

**UNIVERSIDAD COMPLUTENSE DE MADRID**  
**FACULTAD DE INFORMÁTICA**



**TESIS DOCTORAL**

**Aportaciones al modelado computacional de trastornos  
emocionales mediante fuentes de información alternativas**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

**Pavél Harold Llamocca Portella**

Directoras

**Matilde Santos Peñas**  
**Victoria López López**

Madrid

**UNIVERSIDAD COMPLUTENSE DE MADRID**  
**FACULTAD DE INFORMÁTICA**



**TESIS DOCTORAL**

Aportaciones al modelado computacional de trastornos emocionales mediante fuentes de información alternativas

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Pavél Harold Llamocca Portella

DIRECTORAS

Matilde Santos Peñas y María Victoria López López

# **Aportaciones al modelado computacional de trastornos emocionales mediante fuentes de información alternativas**

Pavél Harold Llamocca Portella

**Memoria para optar el grado de Doctor en Informática**

Facultad de Informática

Directoras: Matilde Santos Peñas y María Victoria López López

Programa de Doctorado en Ingeniería Informática



UNIVERSIDAD  
**COMPLUTENSE**  
MADRID

**Septiembre 2022**



*A pesar de la distancia y de las adversidades,  
su apoyo y aliento han sido inquebrantables.*

*Quiero dedicar este trabajo a  
mis padres Daniel y Violeta.*

*Quiero también dedicarlo a  
mis directoras Victoria y Matilde  
por todo el tiempo y la infinita  
paciencia que me han tenido.*

# Agradecimientos

No encuentro manera de agradecer a todas las personas que hicieron posible este trabajo sin extenderme demasiado. Sin embargo, no puedo dejar de mencionar a mis directoras Victoria y Matilde por dirigir mi tesis a lo largo de varios años. Es mucho el tiempo que han dedicado en esta tesis, y el tiempo es irrecuperable. Por ello, quiero agradecerles infinitamente por todo el esfuerzo y paciencia que me han tenido.

Tampoco puedo dejar de mencionar al Doctor Diego Urgelés Puertolas por haberme compartido todo el conocimiento necesario para abordar este desafío. Su buena voluntad y disponibilidad son cualidades que tengo muy presentes y han sido fundamentales para el trabajo con los pacientes y los materiales.

Gracias también a la doctora Milena Čukić por haber participado exhaustivamente en algunas de las publicaciones. Su experiencia reforzó en mí el enfoque científico que este tipo de trabajos necesita y me permitió conocer mucho de su trabajo que en parte direccionó el curso de esta tesis.

Agradezco también a todas las personas encargadas del “Programa de Doctorado” de la Facultad de Informática por estar supervisando constantemente el avance de esta tesis y por organizar eventos y charlas que me han aportado capacidad de desenvolvimiento en un alto nivel académico.

Por último y no por eso menos importante, quiero agradecer especialmente a mis padres por ser el motor de todo mi esfuerzo y por enseñarme que aunque las circunstancias pueden no estar de nuestro lado, el derrumbarse nunca es una alternativa. Estoy y estaré eternamente agradecido por todo lo que me siguen enseñando sin escatimar ningún tipo de esfuerzo.

**Pavél Harold Llamocca Portella**

Madrid, Septiembre 2022



# Índice general

Agradecimientos	iii	
Índice general	v	
Índice de figuras	vii	
Índice de tablas	ix	
Acrónimos	xi	
Resumen	xiii	
Abstract	xv	
<b>1</b>	<b>Introducción</b>	<b>1</b>
1.1	Motivación de la tesis y procedimientos . . . . .	3
<b>2</b>	<b>Estado del arte</b>	<b>7</b>
2.1	Materiales: Fuentes de información . . . . .	7
2.2	Métodos: Técnicas de <i>machine learning</i> para diagnóstico de estados emocionales . . . . .	17
2.3	Aportaciones de la tesis en el estado del arte . . . . .	23
<b>3</b>	<b>Datos y herramientas computacionales</b>	<b>25</b>
3.1	Fuentes de información . . . . .	25
3.2	Integración de datos . . . . .	27
3.3	Herramientas computacionales . . . . .	31
<b>4</b>	<b>Desarrollo y contribuciones de esta tesis</b>	<b>35</b>
<b>5</b>	<b>Conclusiones y trabajos futuros</b>	<b>45</b>
	<b>Bibliografía</b>	<b>49</b>
<b>6</b>	<b>Lista de publicaciones</b>	<b>57</b>
	<b>Publicaciones</b>	<b>60</b>
<b>I</b>	<b>Data source analysis in mood disorder research</b>	<b>61</b>
<b>II</b>	<b>Bip4Cast: Some advances in mood disorders data analysis</b>	<b>69</b>

III	Personalized characterization of emotional states in patients with bipolar disorder	77
IV	Weighted dependence of the day of the week in patients with emotional disorders. A mathematical model	97
V	The proposition for bipolar depression forecasting based on wearables data collection	105
	Anexos	115
A	Librerías de R utilizadas	117
B	Procesamiento de ficheros binarios (GGIR)	119
C	Ejecución de algoritmos de <i>machine learning</i>	121
D	Variables del <i>self-report</i>	123
E	Variables del <i>smartwatch</i>	127
F	Variables de las consultas médicas - escala HDRS	137
G	Variables de las consultas médicas - escala YMRS	143
H	Variables de las consultas médicas - estados	149
I	Variables de las consultas médicas - eventos	151
J	Variables generales	153

# Índice de figuras

1.1	Fuentes de información . . . . .	4
1.2	Procedimiento . . . . .	4
3.1	Aplicación Android <i>self-report</i> . . . . .	26
3.2	Aplicación con librería <i>shiny</i> . . . . .	27
3.3	Proceso de integración . . . . .	31
3.4	Arbol de decisión . . . . .	33
4.1	Coefficientes de correlación de <i>Pearson</i> . . . . .	38
4.2	Transiciones entre estados emocionales . . . . .	39
4.3	Comportamiento de variables en distintos estados emocionales . . . . .	39
4.4	Análisis de componentes principales (PCA) . . . . .	40
4.5	Variable <i>D_wakeup_time</i> por semanas . . . . .	41
4.6	Variable <i>D_sleep_duration</i> en cada estado emocional . . . . .	42



# Índice de tablas

3.1	Ficheros utilizados para el registro de datos . . . . .	28
4.1	Precisión de algoritmos de <i>machine learning</i> . . . . .	36
A.1	Librerías utilizadas en la tesis . . . . .	117
D.1	Variables del <i>self-report</i> . . . . .	123
E.1	Variables del <i>smartwatch</i> . . . . .	127
F.1	Variables de las consultas médicas: escala HDRS . . . . .	137
G.1	Variables de las consultas médicas: escala YMRS . . . . .	143
H.1	Variables de las consultas médicas - estados . . . . .	149
I.1	Variables de las consultas médicas - eventos . . . . .	151
J.1	Variables generales del dataset . . . . .	153



# Acrónimos

AUC	Area Under Curve. 20–22
AUROC	Area Under Curve/Receiver Operating Characteristics. 18, 21
BDI	Beck Depression Inventory. 14, 15
BSN	Body Sensors Network. 13
CIDI	Composite International Diagnostic Interview. 18
DT	Decision Tree. 13, 19, 23, 36, 40, 46, 118
EDA	Electrodermal Activity. 10, 11
EEG	Electroencefalografía. 8
EPDS	Edinburgh Postnatal Depression Scale. 14, 21
FS	Feature Selection. 2, 5, 19, 20, 39, 40, 46, 118
GAD7	Generalized Anxiety Disorder-7. 19, 21
HDRS	Hamilton Depression Rating Scale. 16, 19–21, 25, 28, 36, 37
HRV	Variabilidad del ritmo cardiaco. 12
KNN	K-Nearest Neighbors. 13, 19, 20, 23, 118
MADRS	Montgomery-Asberg Depression Rating Scale. 7, 15
mHealth	Mobile Health. 13, 14
ML	Machine Learning. 2, 3, 5, 8, 12, 17, 19–24, 31, 35, 37, 40, 43, 46, 47, 118
MLP	Multi-layer Perceptron. 20, 21, 23
OCT	Optical Coherence Tomography. 9
OWA	Ordered Weighted Average. xiii, xiv, 5, 41, 47
PCA	Análisis de Componentes Principales. 20, 37, 40

## Acrónimos

---

PHQ9	Patient Health Questionnaire. 13, 14, 18, 19
PSG	Polisomnografía. 8, 9
PSQI	Pittsburgh Sleep Quality Index. 15, 16, 21
RF	Random Forest. 17, 19, 21–23, 36, 40, 46, 47, 118
RLG	Regresión Logística. 19, 36, 40, 46, 47, 118
RNFL	Retinal Nerve Fiber Layer. 9
RVM	Relevance Vector Machines. 18, 19, 23
SGDS	Short Geriatric Depression Scale. 11
SVM	Support Vector Machines. 19–21, 23, 36, 40, 46, 47, 118
TB	Trastorno Bipolar. 1–3, 5, 7–11, 14–16, 18–20, 23–25, 35, 45, 46
TC	Tomografía Computarizada. 9
WD	Wearable Devices. 2, 5, 10, 15
WHO	World Health Organization. xiii
YMRS	Young Mania Rating Scale. 15, 16, 20, 25, 28, 36, 37

# Resumen

Según estadísticas de la *World Health Organization (WHO)*, de todos los problemas de salud, los trastornos mentales representan el 12.5%, una cifra superior a la del cáncer. Aproximadamente una de cada 4 personas sufrirá un trastorno mental a lo largo de su vida, y entre el 35 y 50% no recibe o recibe el tratamiento inadecuado, llegando a causar 800 000 suicidios cada año. Estos datos reflejan lo relevante de la detección y el diagnóstico temprano y adecuado. Sin embargo actualmente el sector clínico tiene una limitación en este proceso, ya que las técnicas clásicas para asignar tratamientos se realizan en base a información reportada por el paciente sobre sus emociones (datos subjetivos) y observación médica, siendo esta última apoyada también en gran medida de la información subjetiva.

Para mejorar esta situación es necesario contar con procedimientos que no solo exploten información subjetiva facilitada por el paciente sino que también utilicen datos objetivos, como pueden ser patrones fisiológicos. Actualmente existen en el mercado distintos dispositivos inteligentes capaces de registrar este tipo de información y que pueden ser utilizados por los psiquiatras para mejorar los diagnósticos. Dentro de la clasificación de *wearable devices* los más comunes son los *smartwatches*. Por otra parte, considerando la gran cantidad de indicadores que se pueden obtener a partir de estos dispositivos, el uso de técnicas de *machine learning* y *feature selection* se hacen necesarias.

Entre los objetivos de esta tesis, el más importante es el dar soporte clínico a los tratamientos de pacientes con trastorno bipolar a través de un análisis predictivo del estado emocional, utilizando fuentes de información que provean tanto datos subjetivos como objetivos. Otros objetivos son el análisis de todas las variables registradas para identificar aquellas que tienen mayor impacto en el estado emocional, analizar el impacto del día de la semana así como establecer una relación de los indicadores del sueño con el estado emocional. Se ha realizado un ensayo en el cuál han participado 101 pacientes diagnosticados con trastorno bipolar utilizando como fuentes de información formularios electrónicos y *smartwatches*. Se realizó un análisis personalizado de los datos facilitados aplicando técnicas de *machine learning*, *feature selection*, y operadores *Ordered Weighted Average (OWA)*, entre otros métodos matemáticos y estadísticos, para desarrollar modelos computacionales del estado emocional.

Los resultados han demostrado que la aplicación de técnicas de *machine learning* para la clasificación del estado emocional utilizando datos objetivos consigue una precisión de 74% promedio y que al añadir los datos subjetivos, esa precisión promedio aumenta a un 80%. Además, se ha observado que la precisión de los algoritmos de *machine learning* es mayor cuando las variables de entrada han sido filtradas previamente con técnicas de *feature selection*. También se ha

elaborado un modelo basado en operadores OWA para determinar el impacto que tiene el día de la semana en el estado emocional. Por último, se ha demostrado la importancia de los indicadores de sueño y movilidad para la clasificación del estado emocional en este tipo de enfermedades.

**Palabras clave.** Trastorno bipolar, Wearable devices, smartwatches, machine learning, feature selection, estado emocional, diagnóstico.

# Abstract

According to World Health Organization, mental disorders represent 12.5 % of all health problems, higher than rate for cancer. One in 4 people will suffer a mental disorder throughout its life, and among 35 and 50 % do not receive any treatment or do not receive the proper treatment with the result of 800 000 suicides by year. These statistics highlights the relevance of early and proper detection and diagnosis. However, the clinical attempts has to deal with a limitation in this process since classic techniques to assign treatments are carried out in base of information reported by the patient about its emotions (subjective data) and clinical observations, being the last one also supported by the subjective date in a large degree.

In order to improve this situation, new procedures are needed not only to use subjective information provided by patient but objective data as physiological patterns also. Several smart devices models are currently able in the industry which are able to record this kind of information and can be used by psychiatrist to improve the diagnosis. Within classification of “wearable devices”, the most common are the smartwatches. By other hand, taking into account the large quantity of indicator which can be recorded with those devices, the use of Machine Learning and Feature Selection techniques are needed.

Between the several goals of this thesis, the most important is to give clinical support in treatment to patients with bipolar disorder through a predictive analysis of its emotional state using sources of information that provide subjective and objective data. Other goals are the analysis of all recorded variables in order to identify the ones with the highest impact in the emotional state, analyze the impact of the day of the week as well as establish a relations between sleep indicators and emotional state. An essay was carried out with 101 patients diagnosed with bipolar disorder. The sources information used were: smartwatches, electronic forms and periodic clinical follow-up. It was performed a personalized analysis of data provided. Machine Learning and Feature Selection techniques were applied as well as Ordered Weighted Average (OWA) operators among others mathematical and statistical methods to build computational models of the emotional state.

The results have shown an accuracy of 74 % on average for the classification of the emotional sate by applying Machine Learning techniques over objective data. The accuracy increased to 80 % by adding the objective data. In addition, it was observed that the accuracy of Machine Learning algorithms increases when variables are previously filtered using Feature Selection techniques. A model based in OWA operators was also built in order to calculate the impact of weekday over the emotional state. Finally, it was demonstrated the importance of sleep indicators and mobility in order to determine the emotional state in this

## Abstract

---

kind of illnesses.

**Key words.** Bipolar disorder, wearable devices, smartwatches, machine learning, feature selection, emotional state, diagnosis.

# Capítulo 1

## Introducción

Las enfermedades mentales, si no son diagnosticadas a tiempo, pueden llevar a las personas que las padecen a una carga emocional, social y económica muy considerable. La incidencia de estas enfermedades es del 13% a nivel mundial [66].

El acceso a herramientas de las tecnologías de la información que apoyen un mejor tratamiento y diagnóstico suele ser limitado, en algunos casos por ser caro y en otros por no haber mucha oferta. Por tanto, es una práctica común que los especialistas e investigadores tomen decisiones basadas únicamente en datos subjetivos proporcionados por los mismos pacientes.

Es fundamental contar con métodos ampliamente accesibles que ayuden en la pronta detección de un deterioro mental [32]. La reducción de los costos de estos métodos permite que su uso pueda ser ampliamente extendido. Por otra parte, las técnicas utilizadas deben considerar los factores que actualmente se vienen investigando, por ejemplo, los desórdenes del estado de ánimo se dan por una combinación entre actividad, cognición y emoción que varían en el tiempo [43].

Mathur y Sutton [45] afirman que la calidad del paciente mejora significativamente cuando es correctamente diagnosticado y recibe el tratamiento adecuado. Además, el diagnóstico temprano de enfermedades como el trastorno bipolar puede tener un gran impacto en los resultados del tratamiento ya que las fases iniciales de la enfermedad son más reactivas y se necesitan terapias más ligeras [80].

En el estudio de los trastornos mentales, el diagnóstico de enfermedades es un proceso que termina siendo impreciso en muchos casos. Shen et al. han demostrado que se puede alcanzar un porcentaje de error de 76.8% [71]. En gran medida esto se debe a que para diferentes diagnósticos los pacientes pueden presentar comportamientos similares [56] y el basarse únicamente en evidencias clínicas y datos subjetivos del paciente puede conllevar un diagnóstico equivocado [29].

Dentro de los trastornos mentales, según la Organización Mundial de la salud, el Trastorno Bipolar (TB) tuvo una prevalencia de entre 0.3 y 1.5% a nivel mundial en el 2021 [73], siendo la sexta causa que más discapacidad genera. Además, es importante mencionar que en el TB el riesgo de intento de suicidio en etapa de depresión es del 33.9% [19]. De ahí la importancia de técnicas que ayuden en el correcto diagnóstico y tratamiento de esta enfermedad.

El TB es un cuadro crónico que se manifiesta por cambios bruscos del estado emocional. Se trata de una enfermedad que tiene un complejo proceso de seguimiento por tres motivos: 1) su evolución presenta patrones distintos entre pacientes; 2) presenta una sintomatología muy variada; y 3) los cambios del estado emocional son desencadenados repentinamente. Por tanto se hace

## 1. Introducción

---

especialmente importante la monitorización del paciente a lo largo del tiempo. El estado emocional de los pacientes con TB varía entre: depresión, eutimia (ausencia de cualquier estado), hipomanía (manía ligera) y manía. Fundamentalmente existen tres tipos de este trastorno [59]:

**TB1** Es diagnosticado por la presencia de episodios de manía, hipomanía y depresión con episodios que se denominan mixtos.

**TB2** Es diagnosticado por la presencia de episodios hipomanía y depresión. Estos pacientes sufren a lo largo de su vida muchos más episodios que los pacientes con TB1

**TB Ciclotímico** Es diagnosticado por la presencia de varios episodios hipomaníacos, alternando con periodos con episodios depresivos, pero no suficientes para clasificarse como depresión.

Uno de los errores más frecuentes se da al intentar detectar el TB, ya que éste es en ocasiones incorrectamente diagnosticado como depresión (referido en la literatura como “Unipolar Depression” o “Major Depression”) [72] y tratado como tal. Aunque ambas enfermedades tienen síntomas similares, algunos indicadores de conducta de un paciente con TB pueden ser distintos de aquellos con depresión, por ejemplo, los pacientes con TB tienen una actitud más compleja en cuanto a irritación y enojo [51]. Este error trae como consecuencia que pacientes con TB sean tratados equivocadamente con antidepresivos y psicotrópicos, pudiendo ser este tratamiento contraproducente para su estado clínico [58][71].

Para reducir el porcentaje de error en el diagnóstico se debe disponer no sólo de evidencia clínica e información subjetiva del paciente, sino también de datos objetivos que permitan observar sus indicadores fisiológicos y psicológicos. Una de las alternativas que existen en el mercado actual y que permiten lidiar con la falta de datos objetivos es el uso de Wearable Devices (WD), que son bastante accesibles y útiles para obtener información objetiva sobre sus usuarios. Existe una gran cantidad de indicadores que se pueden extraer con estos dispositivos: actividad física, calidad del sueño, ritmo cardíaco, oxigenación, etc. Una de las ventajas de estos dispositivos es que extraen información sin interferir en la vida cotidiana de la persona que los usa. El uso de estos dispositivos en diagnósticos/tratamientos sobre personas con enfermedades mentales puede llevar a conclusiones relevantes acerca del impacto que tiene una enfermedad sobre un paciente. Hay una fuerte evidencia clínica en la literatura actual de que ciertos indicadores psicológicos y fisiológicos están estrechamente relacionados con los cambios de ánimo [15]. Dado que la variedad de datos obtenidos con estos dispositivos es bastante alta, tenemos que considerar el uso de técnicas de computación como *Machine Learning (ML)* y *Feature Selection (FS)*.

Por otro lado, la medicina personalizada es un nuevo enfoque prometedor que propone adaptar la terapia en base a las características de cada paciente para asegurar los mejores resultados del tratamiento. Esto es especialmente cierto en ramas de la medicina que lidian con enfermedades mentales ya que los patrones de conducta en cada persona son distintos. El tener las capacidades para realizar

un temprano y correcto diagnóstico, asignar un tratamiento adecuado y ser capaces de evaluar los riesgos permite no solo mejorar la calidad de la salud y vida [45] del paciente, sino también disminuir sus costes.

### 1.1. Motivación de la tesis y procedimientos

La monitorización del paciente es fundamental en este tipo de enfermedades. Por esto el uso de dispositivos tecnológicos que ayuden en la monitorización puede aportar un valor muy importante al momento de establecer diagnósticos y asignar tratamientos. Sin embargo, para que estas herramientas puedan dar información relevante, se deben establecer tres puntos: 1) los indicadores más importantes que afecten el estado emocional del paciente; 2) el tipo de dispositivo a utilizar que permitan monitorizar esos indicadores; y 3) técnicas a utilizar sobre los datos registrados y que clasifiquen el estado emocional del paciente. Cada uno de estos puntos serán detallados a lo largo del desarrollo de esta tesis.

Esta tesis presenta una serie de aportaciones para ayudar a la predicción de crisis en el TB y para ello se analizan datos recopilados en pacientes diagnosticados con este trastorno. Se ha estudiado diversos indicadores de conducta del paciente para encontrar una relación entre estos y el estado emocional del paciente. Para el tratamiento de este tipo de indicadores López et al. [42] destacan la importancia de organizar la información en base al tipo y a la frecuencia de actualización. Esta estructura así como las fuentes de información utilizadas se muestran en la figura 1.1.

Los indicadores con los que se ha trabajado se pueden clasificar en dos grandes grupos:

**Indicadores Subjetivos.** Aquellos indicadores reportados por el paciente de acuerdo a preguntas realizadas sobre sus sentimientos. Este tipo de indicadores son los que también se llaman indicadores psicológicos. Entre estos podemos mencionar: la irritación, el ánimo, la motivación, etc.

**Indicadores Objetivos.** Son aquellos que se recogen a partir de una medición. A su vez se pueden clasificar en:

- Aquellos indicadores que son registrados por un *smartwatch*, como por ejemplo: horas de sueño, tiempo de actividad física moderada, ligera, etc.
- Aquellos indicadores que son registrados por el psiquiatra en las sesiones periódicas que mantienen con el paciente, por ejemplo: concentración, la apariencia, el peso, etc.

El procedimiento utilizado consistió en lo siguiente: Se estableció un periodo de recolección de datos en el cual se fueron registrando todos estos indicadores mediante tres fuentes de datos: *smartwatches*, consultas médicas y *self-reports*. A partir de estas fuentes se generó un único dataset mediante un proceso de integración. Sobre este dataset se aplicaron diferentes técnicas estadísticas y computacionales, entre ellas algoritmos de ML. En base a los modelos

# 1. Introducción

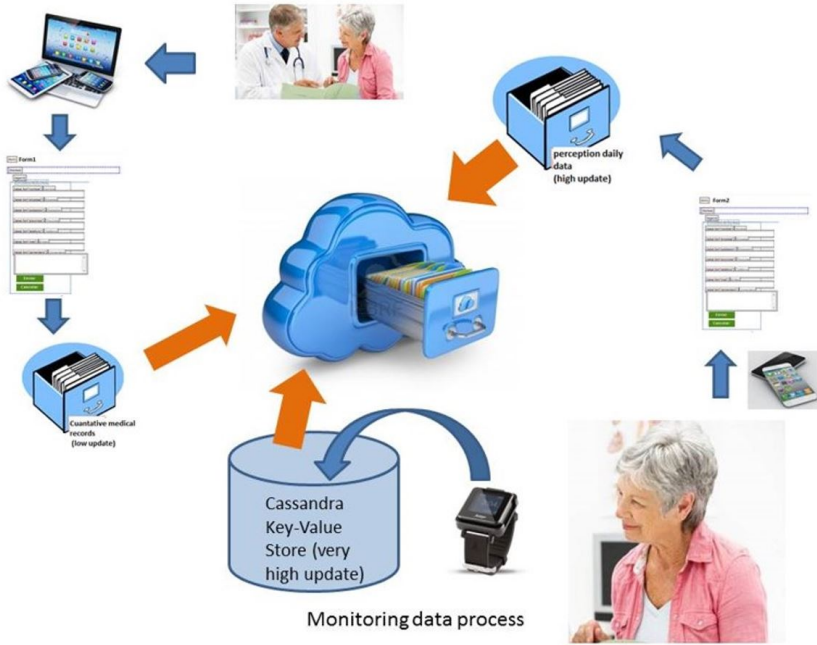


Figura 1.1: Fuentes de información.

matemáticos generados se realizaron clasificaciones del estado emocional del paciente. De manera general, este procedimiento se muestra en la figura 1.2.

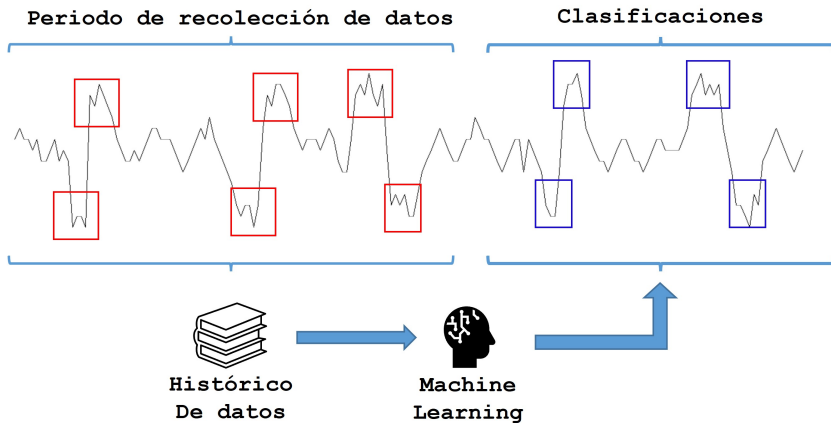


Figura 1.2: Procedimiento General.

Las aportaciones de esta tesis se han realizado en conjunción con un equipo formado por ingenieros, matemáticos, y un médico psiquiatra que fue la persona que aportó las directrices clínicas. La participación de este equipo permitió

validar los resultados obtenidos.

### 1.1.1. Objetivos

El objetivo principal de este trabajo es aportar soluciones tecnológicas que den soporte clínico en la toma de decisiones sobre pacientes con TB.

Muchas de las soluciones propuestas hasta el momento no son aplicables a nivel general debido a la variabilidad de la sintomatología entre pacientes. Para superar esta limitación, en esta tesis proponemos el uso de modelos personalizados que se definen en base a los indicadores más característicos de cada paciente para cada estado emocional. La caracterización del estado eutímico del paciente es uno de los hitos principales de este trabajo. En base a ésta se pueden detectar anomalías y analizar si son lo suficientemente agudas como para desencadenar un cambio de estado emocional. Desde el punto de vista clínico esto servirá como referencia para decidir si se mantiene o se modifica el tratamiento.

Es decir, en este trabajo se pretende ofrecer una nueva alternativa para el diagnóstico del estado emocional de pacientes con TB utilizando modelos matemáticos construidos a partir de un histórico de indicadores fisiológicos y psicológicos que aporten mayor conocimiento cuantitativo al médico especialista.

Entre los objetivos específicos los más importantes son:

- Analizar la viabilidad del uso de WD para el registro de indicadores fisiológicos.
- Generar un dataset en el que se han integrado todas los datos provenientes de todas las fuentes de información utilizadas durante el periodo de recopilación de datos.
- Reducir la dimensionalidad de los datos identificando aquellos indicadores relevantes y los que causen “ruido” y distorsionen los resultados.
- Estudiar del impacto del día de la semana sobre los indicadores, entre otros.

Dentro del alcance de este trabajo se han incluido distintas técnicas informáticas y matemáticas: para la recopilación e integración de los datos, estadísticas para reducir la dimensionalidad, FS, operadores OWA, y técnicas de ML para la clasificación del estado emocional de un paciente. El resultado de esta clasificación puede ser utilizado por los médicos como referencia para el diagnóstico del estado emocional de los pacientes.



## Capítulo 2

# Estado del arte

En los últimos años se han desarrollado herramientas que ayudan al diagnóstico y tratamiento de enfermedades relacionadas con el estado de ánimo como la depresión y la ansiedad. Al realizar un estudio detallado de la literatura relacionada se puede observar claramente la existencia de dos ramas de investigación muy delimitadas. Por un lado, una línea sobre las fuentes de información (tecnología existente para el registro de datos) y, por otro lado, los diversos métodos para explotar estos datos. Existen investigaciones que se centran en una rama o en otra, pero son pocas las que integran ambas. En cualquiera de los casos, una limitación común es el reducido número de pacientes con los que se puede trabajar.

Algunas investigaciones utilizan un número razonablemente grande de pacientes pero tampoco consiguen objetivos de estudio genérico. Constantinides et al. [13] realizaron una comparación entre modelos personalizados y genéricos. El resultado muestra que el modelo personalizado tiene una precisión de casi un 30 % más que el modelo genérico. Hoy en día se ha reconocido que los modelos personalizados, especialmente en medicina, son más exitosos.

Dado que la falta de datos es una desventaja bastante extendida para este tipo de iniciativas, hay investigaciones que publican sus datos a la comunidad científica para que puedan surgir nuevos aportes. Por ejemplo, Garcia-Ceja E. [26] hace público un dataset de 23 pacientes depresivos o con TB en estado de depresión y 32 sujetos de control. Este dataset contiene datos de actigrafos y también de la escala de *Montgomery-Asberg Depression Rating Scale (MADRS)*. Se puede descargar de: <http://datasets.simula.no/depresjon/> o de <https://doi.org/10.5281/zenodo.1219550>.

Este capítulo incluye 3 secciones, la primera es el estado del arte relativo a la tecnología existente para recopilar indicadores útiles para el estudio de los cambios en el estado emocional. La segunda sección es el estado del arte relativo a las técnicas de Inteligencia artificial (IA) utilizadas para explotar estos indicadores y como última sección se incluye las aportaciones de esta investigación dentro del estado del arte.

### 2.1. Materiales: Fuentes de información

Las fuentes de información son herramientas utilizadas para el registro de los datos. En el ámbito médico éstas pueden ser intrusivas respecto a la actividad diaria o la comodidad del paciente, o no intrusivas al no interferir en ellas. En esta sección vamos a realizar un estudio de estas herramientas de acuerdo a la tecnología empleada.

### 2.1.1. Equipos clínicos

Dentro de esta categoría se incluyen todas aquellas pruebas que son realizadas en un centro médico. Estas pruebas requieren el manejo de un especialista y suelen tener un costo elevado. Por norma general son prescritas por un médico y, por tanto, solo es posible registrar datos de los pacientes en determinado momento.

La actividad cerebral de una persona viva nunca se detiene, incluso estando dormido las neuronas siguen en constante comunicación. Esta comunicación se establece mediante impulsos eléctricos. La Electroencefalografía (EEG) es una técnica de exploración del sistema nervioso central. Es una prueba clínica que detecta esta actividad eléctrica en el cerebro. Para realizarla se utiliza unos electrodos que se fijan en el cuero cabelludo y que convierten estos impulsos eléctricos en un gráfico que se conoce como electroencefalograma. Este permite determinar cambios en la actividad del cerebro mediante los cuales se pueden diagnosticar diversos desórdenes mentales.

La EEG se ha aplicado para el estudio y clasificación de varios tipos de enfermedades neuropsiquiátricas y neurodegenerativas. Por ejemplo, la EEG se ha utilizado para predecir la edad cronológica a partir de técnicas de ML. Participaron 468 personas (entre sujetos de control y personas que sufrían algún desorden mental). En los resultados se observó que la correlación entre edad calculada y la edad real alcanzó un 0.6 [3].

Otro ensayo realizado en Shangai, China, [34] tenía como finalidad proponer un método de detección de la depresión basado en señales del EEG. En este ensayo participaron 40 personas (16 depresivos y 14 sujetos de control) que fueron expuestos a imágenes de 6 rostros humanos [63]. Se aplicaron técnicas de ML para la clasificación cuyos resultados mostraron una precisión máxima del 85.7%.

Las investigaciones más recientes buscan proporcionar marcadores más estables y fiables en base a las señales EEG. Entre estas investigaciones podemos mencionar que un estudio [2] concluyó que se puede diferenciar personas sanas de entre paciente con desordenes anímicos y de ansiedad en base a los *EEG-microstate (EEG-ms)*, que es la señal en unas pocas decenas de milisegundos.

La Polisomnografía (PSG) es una técnica utilizada para detectar posibles anomalías del sueño, permite el registro de la respiración, el ritmo cardiaco, la actividad muscular y niveles de oxígeno en la sangre mientras se duerme. Dado que la calidad del sueño es un indicador bastante relevante en el estudio del TB, esta técnica es recomendada para su diagnóstico. Se puede realizar en un centro médico o en casa. Similar al EEG, se utilizan diversos electrodos en el mentón, el cuero cabelludo, los párpados y en el pecho para controlar las frecuencias cardíacas/respiratorias.

Entre las investigaciones en las que se ha utilizado la PSG se puede recalcar un estudio en el que se analizó el efecto del género sobre la vulnerabilidad de una persona ante los desórdenes mentales [31]. En este estudio se utilizó la PSG de 44 personas saludables y los resultados concluyeron que existen diferencias relacionadas con el sexo durante el sueño.

También se estudió el uso de PSG para el diagnóstico de la depresión o ansiedad [20]. En este artículo se intentó demostrar la hipótesis de que personas no diagnosticadas con depresión o ansiedad, pero con una alta probabilidad de padecerlas, muestran unos parámetros de PSG distintos de aquellos que no tienen esa tendencia. Participaron 129 personas no diagnosticadas con altos índices de ansiedad y depresión. Aunque no se llegó a validar dicha hipótesis, los resultados demostraron que las mujeres son ligeramente más vulnerables que los hombres a padecer ansiedad, pero igual de vulnerables para la depresión. Además, se observó que los pacientes depresivos exhibieron menores momentos de sueño *Rapid Eye Movement (REM)* que aquellos que no son depresivos.

La Tomografía Computarizada (TC) es un procedimiento clínico indoloro utilizado para el diagnóstico de enfermedades utilizando imágenes producidas por una fuente de rayos X que gira alrededor del paciente. Entre muchas de las utilidades de la TC está el detectar la concentración de unas proteínas llamadas “amiloides” y “TAU” en el organismo. Los amiloides son unas proteínas anormales que alteran el funcionamiento de distintos sistemas, entre ellos el nervioso. Las TAU son proteínas encontradas usualmente en las neuronas y cuya principal función es la estabilización de los microtúbulos. Estas últimas están bastante relacionadas con muchas enfermedades mentales. La concentración anormal de estas proteínas se conoce como taupatía. Existe un trabajo que investiga si ambas proteínas están relacionadas con la depresión en personas con Alzheimer [5]. Para ello, se utilizó la tomografía por emisión de positrones (PET) en 301 adultos mayores y los resultados mostraron que aquellos que tenían una mayor concentración de tau eran más propensos a estar depresivos. No se encontró ninguna relación del amiloide con la depresión. Adicionalmente se concluyó que los antidepresivos tienen un impacto sobre las TAU incrementando las posibilidades de depresión en adultos mayores.

La *Optical Coherence Tomography (OCT)* es una tomografía que, a diferencia de la anterior, solo examina la zona ocular. El paciente apoya el mentón en una de las bases de un scanner. Este utiliza ondas de luz para tomar imágenes de secciones transversales de la retina. En base a esta herramienta existe una investigación que relaciona las *Retinal Nerve Fiber Layer (RNFL)* y el grosor total de la retina en imágenes OCT con prevalencia del TB. En esta investigación participaron 70 individuos (entre pacientes con TB y sujetos de control). Se concluyó que el TB está significativamente relacionada con un menor grosor de la retina y un menor grosor de las RNFL [28]. Dentro de la literatura revisada, y también relacionada con la OCT, se encontró un artículo que estudia el tamaño del menisco lagrimal (zona que genera y acumula la producción lagrimal) en pacientes con ansiedad y depresión [77]. En este trabajo participaron 40 pacientes con depresión, 35 pacientes con ansiedad y 37 sujetos de control. Se midió la altura (TMH), la profundidad (TMD) y el área (TMA) del menisco lagrimal. En base a los resultados se concluyó que estos parámetros son significativamente menores en los pacientes con ansiedad y depresivos en comparación con los sujetos de control.

### 2.1.2. Dispositivos inteligentes

Dentro de toda la amplia gama de dispositivos inteligentes existentes, los *WD* permiten recopilar indicadores fisiológicos de una persona. Por tanto son los que se han utilizado para este tesis. Los *WD* son aquellos dispositivos que uno lleva consigo en su actividad diaria y están constantemente recopilando información de las mismas con un nivel de intrusión mínimo o ausente.

Muchos indicadores psicológicos y fisiológicos están asociados a enfermedades neuro-psiquiátricas, y estos pueden ser reflejados en señales de actividad, de ubicación, ritmo y variabilidad cardíaca [64]. Son precisamente los *WD* los que permiten la obtención de esta información por lo que terminan siendo bastante útiles para el diagnóstico de este tipo de enfermedades.

La información obtenida varía en según al tipo de dispositivo, pudiendo ser: la cantidad de pasos, calorías quemadas, calidad del sueño, etc. Estos dispositivos únicamente recopilan información y para ser visualizados o analizados tienen que venir integrados con un sistema de comunicación a una terminal que permita procesar los datos. En algunos casos vienen ya directamente con una interfaz gráfica que normalmente es una app de un *smartphone*. Entre los más utilizados tenemos los *smartwatches* pero también hay otro tipo de tecnologías como las E-textiles.

Los *smartwatches* son dispositivos que generalmente se usan en la mano no-dominante y tienen integrado un acelerómetro para detectar el movimiento. En lo que respecta al estudio de los trastornos mentales, existe literatura que evidencia diferencias en los parámetros medidos por los *smartwatches* entre sujetos sanos y sujetos con depresión o con TB [74].

Entre las diversas investigaciones realizadas para la detección de la depresión con *smartwatches*, podemos resaltar el uso del Parkinson's KinetiGraph [62]. Este es un dispositivo utilizado clínicamente para estudiar la bradicinesia, la discinesia y el temblor en pacientes con Parkinson. En este estudio también se incluyó un análisis de los síntomas motores en episodios de depresión. Como resultado se demostró una significativa correlación (0.7) de los estados de melancolía y los síntomas del Parkinson.

En otra investigación se utilizó el dispositivo "Silme W20" para diagnosticar la depresión [75]. Este dispositivo mide la cantidad de pasos, energía gastada, movimiento del cuerpo, tiempo del sueño, ritmo cardíaco, temperatura de la piel y exposición a la luz ultravioleta. En esta investigación se involucraron a 45 pacientes depresivos y 41 sujetos de control. Después del estudio se observó que los parámetros del ritmo cardíaco, cantidad de pasos y tiempo de sueño eran significativamente distintos entre el grupo depresivo y el grupo de control. Se elaboró un modelo matemático que con una precisión de 0.76 identificó los pacientes que tenían síntomas de depresión.

En una publicación reciente, los autores utilizan el dispositivo "E4 wristband" [14]. Este dispositivo tiene como novedad que tienen un sensor *Electrodermal Activity (EDA)*. El objetivo de esta investigación era la predicción del estado de manía y eutimia en pacientes con TB. Participaron 47 pacientes (22 maníacos y 25 eutímicos). Se aplicaron varios métodos de predicción en los que se utilizaron

parcialmente los parámetros que ofrecía el dispositivo usado. La precisión alcanzada fue del 0.91 utilizando las prestaciones completas: acelerómetros y EDA.

Otra publicación [35] presenta una herramienta que clasifica la depresión entre adultos mayores que viven solos en Korea. En la investigación participaron 47 adultos entre 65 y 94 años que alcanzaban un índice superior a 5 en la *Short Geriatric Depression Scale (SGDS)*. Esto indicaba que tiene síntomas de depresión moderados y severos. Los datos fueron recopilados entre mayo del 2017 y enero del 2018 y se utilizó el *smartwatch* “Actiwatch Spectrum PRO”. Esta investigación se caracterizó por incluir los datos de EDA registrados en el *smartwatch*. Estos permiten detectar las experiencias y el estado de ánimo en entornos reales y en diferentes situaciones. Se construyeron dos grupos: un grupo depresivo ( $SGDS > 7$ ) y otro grupo no-depresivo ( $SGDS < 7$ ). Las pruebas de clasificación se realizaron utilizando la regresión logística alcanzando una precisión de 0.91.

Por último, en cuanto al uso de los *smartwatches*, también se ha revisado el proyecto Bip4Cast [38], cuyo objetivo era predecir estados maníacos y depresivos en pacientes ya diagnosticados con TB. Se utilizaron 3 fuentes de información: formularios diarios, sesiones periódicas con el psiquiatra y un *smartwatch*, concretamente se utilizó el dispositivo GeneActive de ActiveInsights. Se realizaron varias clasificaciones obteniendo una precisión máxima de 0.72. Esta información se usó como base para la presente tesis.

Aunque no resulte muy común, las prendas de vestir se están también utilizado para la recolección de información del paciente. Se conoce como e-textiles a aquellas prendas a las que se les incorpora sensores con la finalidad de registrar parámetros de la persona que los usa. Dado que lo normal es llevarlas gran parte del día, si consideramos la capacidad de obtener indicadores de la salud, se consigue una fuente de información bastante fiable. Existen diversos ámbitos en los que estos productos pueden ser útiles, los más importantes son: el médico, deportivo y militar [81].

PSYCHE [33] es un proyecto basado en una plataforma textil y sensores portables cuyo objetivo es el desarrollo de un sistema de monitorización multiparamétrico para la recopilación de datos de pacientes con TB. Estos datos son procesados y analizados para ayudar en el diagnóstico y predicción de cambios relevantes en el estado emocional. Este sistema involucra el uso de una camiseta *Wellness Wearable System (WWS)* y un *smartphone*. La camiseta está elaborada de un material adecuado y totalmente ergonómico para no interferir con la actividad del paciente y al mismo tiempo no obstaculizar el rendimiento de los sensores. Se utilizan 2 sensores a la altura del tórax que permiten el registro de: ritmo y variabilidad cardíaca, ritmo y amplitud de respiración, clasificación de actividad (caminar, correr, echado y parado) y estimación de la energía gastada. La camiseta está conectada a un dispositivo portable *Side Electronic Wearable (SEW)*, y éste a su vez se conecta al *smartphone* vía *bluetooth*. El *smartphone* envía los datos a un servidor de datos de PSYCHE.

Este dispositivo se ha utilizado en un ensayo cuya finalidad fue predecir los cambios de ánimo en pacientes con TB. Fue realizado en el Hospital Universitario

## 2. Estado del arte

---

de Estrasburgo y Ginebra. Durante 14 semanas el estado emocional de 14 pacientes fue evaluado dos veces por semana. Para la clasificación de estos estados se utilizaron técnicas de ML y como resultado promedio entre todos los pacientes se obtuvo una precisión, sensibilidad y especificidad de: 0.69, 0.57 y 0.78 respectivamente. Sin embargo, para un paciente en particular, la precisión alcanzó un 0.83 [79].

A pesar de que los datos obtenidos por PSYCHE pueden conllevar a resultados positivos por la fiabilidad de los datos, las personas que han llegado a probarlo mencionaron que uno de los principales inconvenientes era el uso de la misma prenda todos los días, lo que causaba una gran incomodidad [33].

En otra investigación se estudió la relación de la Variabilidad del ritmo cardíaco (HRV) con los síntomas de desórdenes funcionales [83]. Participaron 133 pacientes con síntomas de desórdenes funcionales bajos (Low FD), 9 con síntomas altos (High FD) y 3 sujetos de control. Para la adquisición de los datos se utilizó el dispositivo “LifeShit”. Este producto es similar a un “chaleco” y está compuesta por diferentes tipos de sensores: sensores de señales cardíacas, sensores de posturas y una banda anatómica que registrar el ritmo respiratorio. Los resultados del estudio mostraron un desbalance en el sistema nervioso autónomo (entre el sistema nervioso simpático y el parasimpático) en pacientes con una alta ocurrencia de síntomas de desórdenes funcionales. Además, para este mismo grupo también se observó unos altos niveles de depresión y ansiedad.

Para validar la fiabilidad en las mediciones de las funciones cardiorrespiratorias se realizó un estudio en los cuales se recopilaron tres parámetros: ritmo cardíaco (HR), ritmo respiratorio (BR) y ventilación (VE) [11]. Se utilizó el dispositivo “Hexoskin”, un producto que tiene sensores corporales entrelazados y pueden rastrear la frecuencia cardíaca, la variabilidad de la frecuencia cardíaca, la recuperación de la frecuencia cardíaca, la frecuencia respiratoria, la ventilación por minuto, el nivel de actividad, la aceleración, las calorías quemadas, la cadencia y el conteo de pasos. En el estudio participaron 28 voluntarios. Se analizaron los datos de cada uno bajo esfuerzo físico utilizando dos sistemas de medidas y en los resultados se observó que el coeficiente de correlación entre HR y BR alcanzó un 0.99. No fueron encontrados resultados relevantes para VE.

### 2.1.3. Body Sensor Networks

Los *Wireless Body Area Network (WBAN o BAN)* son sensores portátiles con diferentes capacidades de detección, almacenamiento, computación y transmisión. Este sistema consiste en el uso de bio-sensores que se colocan en distintas partes del cuerpo mientras se realizan actividades comunes y diarias. Recogen información como: señales de electrocardiograma, electromiografía, electroencefalografía, presión de la sangre, etc. Se pueden clasificar en dos tipos: “in-body”, se trata de sensores implantados y que se comunican directamente con una estación base, y los “on-body”, se trata de sensores no implantados (*wearable sensors*) y que se comunican con un centralizador [17]. Estos sistemas están implementados para enviar la información automáticamente al médico para su análisis a través de Internet, lo que conocemos como Internet de las cosas

(IoT). Por tanto, al utilizar estas herramientas se debe tener especial cuidado con temas de seguridad [53].

Uno de los proyectos de investigación en los que se utiliza esta tecnología tuvo como finalidad la detección de cierto tipo de actividades (sentarse, caminar, pararse, etc) en base a “SensorButton”, un sistema compuesto por sensores de sonido, de luz y un acelerómetro. En este ensayo participaron 9 personas que realizaron estas actividades cada 30 minutos. Los resultados al aplicar los algoritmos de *Decision Tree (DT)* y K-Nearest Neighbors (KNN) alcanzaron una precisión de 0.86 y 0.91, respectivamente. También se plantea aplicar este sistema para la detección de actividades más complejas [41].

Otra publicación hecha por la Universidad de Chiktara en La India presenta el *Heartrate Monitoring System (HMS)*, un sistema que utiliza *Body Sensors Network (BSN)* para analizar el ritmo cardíaco [25]. Participaron 40 personas sanas de entre 18 y 20 años que no tenían registros de haber sufrido enfermedades cardiovasculares. Los parámetros se recopilaron en dos situaciones: en relajación y en intensa actividad física. Se utilizó el electrocardiograma para contrastar los resultados. El ritmo cardíaco resultado del HMS tuvo una correlación de 0.99 con los electrocardiogramas.

#### 2.1.4. Smartphones

Hoy en día la utilización de los *smartphones* es prácticamente imprescindible para nuestras actividades diarias. Si a eso le añadimos las diversas herramientas con las que cuenta un *smartphone* (*bluetooth*, GPS, 3G, apps, etc.), éste se convierte en un dispositivo con un gran potencial para la extracción de información. Las propiedades que vienen integradas en un *smartphone* determinan su capacidad para la recopilación de datos. Sin embargo, son las aplicaciones móviles las que procesan toda esa información y nos la muestran en una interfaz gráfica. Las ventajas que pueda brindar el uso de un *smartphone* pueden ser ampliamente aprovechadas cuando lo que intentamos es recopilar información acerca de la actividad de una persona.

En esta sección vamos a reseñar algunas de las publicaciones más recientes realizadas a partir de datos extraídos por aplicaciones móviles, *Mobile Health (mHealth)* y que tienen como finalidad el diagnóstico o tratamiento de algún desorden mental.

Menezes et al. [50] investigan la efectividad de CONEMO (Control Emocional), un programa que consta de ensayos supervisados en personas depresivas y que han sido diagnosticadas de hipertensión y/o diabetes. La finalidad de estos ensayos fue mejorar el estado anímico de los 66 participantes involucrados: 21 de São Paulo, Brasil, 21 de Lima, Perú, de un primer grupo y 25 también de Lima en un segundo grupo. Los datos fueron recolectados entre el 2015 y 2016. Todos ellos antes de empezar el ensayo tenían un puntaje del índice *Patient Health Questionnaire (PHQ9)* superior a 10, lo que significaba que tenían fuertes síntomas de depresión. En este ensayo se utilizó la app del mismo nombre que tenía la finalidad de reducir los síntomas de depresión mediante 18 sesiones de activación conductual impartidas en 6 semanas (3 sesiones por semana).

## 2. Estado del arte

---

Terminado el ensayo, un 58 % de todos los participantes consiguieron un puntaje PHQ9 inferior a 10 [50].

Un ensayo realizado por Sawyer et al. en Australia durante un año y que tiene un par de publicaciones, [68][69], consiguió, mediante un programa de ayuda, un descenso de los niveles de depresión postnatal en madres en las 4 semanas posteriores al alumbramiento. En este programa de ayuda estuvieron involucradas las enfermeras y las madres. Utilizaron la app “eMums Plus” que ofrece una plataforma que permite a las madres reflejar su estado anímico como así como el de su bebé y ofrece una comunicación entre otras madres y con las enfermeras. Estas últimas tienen una propia configuración de la app mediante la cual pueden rastrear la actividad de la madre (material visto, webs abiertas, el estado emocional, etc.) para que puedan dar tanto soporte material como psicológico. Para este ensayo participaron 133 madres. Entre los criterios de selección se consideró un *Edinburgh Postnatal Depression Scale (EPDS)* superior a 7. Se conformaron 2 grupos, un grupo de intervención al cual se aplicó el programa de ayuda y un grupo de cuidado standard que recibiría las atenciones estándar en Australia. Aunque los resultados no mostraron diferencias significativas entre ambos grupos, las madres del grupo de intervención tuvieron sensaciones positivas acerca del programa, obtuvieron mucha información útil a través de la aplicación y además calificaron el uso como muy sencillo.

Para comprobar la eficacia de los mHealth en los tratamientos que mejoran los desórdenes emocionales, Mehta et al. [49] realizaron un ensayo en el cual se ponía a prueba la app “Youper”, una aplicación diseñada con la finalidad de dar soporte a los tratamientos de depresión, ansiedad, y otros tipos de trastornos. Esta aplicación incluye un agente IA con el que es posible entablar una conversación. Para generar respuestas utiliza algoritmos de árboles de decisión. En este ensayo participaron 4517 usuarios. Para medir la aceptabilidad, los participantes evaluaron la app en una escala de 1-5. Los resultados arrojaron una media de 4.36. En cuanto a la eficacia, los participantes expresaron menores niveles de ansiedad, pero no expresaron diferencias significativas para la depresión [49].

Para el estudio del TB, Dargél et al. [16] realizaron un ensayo en París, Francia, utilizando la aplicación “Toi Mème”, que además de recoger información de actividad: cantidad de pasos, la distancia recorrida, etc., también incluye pruebas cognitivas para registrar la actividad motora, respuesta afectiva, cognición, etc. (mediante la integración de dos juegos diseñados por los propios autores). La finalidad de este ensayo es investigar las correlaciones entre los síntomas clínicos del TB con los resultados de la experimentación con la app. La etapa de recopilación de datos del ensayo finalizó en diciembre del 2021 y participaron 93 pacientes con TB: 31 depresivos, 31 eutímicos y 31 hipomaníacos [16].

Been-Zeev et al. [9] ha conseguido reducir los síntomas de la depresión en un ensayo realizado en Chicago, EEUU, sobre 82 personas con depresión severa y esquizofrenia mayores de 18 años. La severidad de la depresión fue evaluada utilizando la escala de Beck Depression Inventory (BDI) [84]. Para el ensayo se utilizó la app “FOCUS”, que hace uso de un enfoque auto-evaluativo de aspectos como el estado de ánimo, problemas para dormir, ingesta de medicación

y desenvolvimiento social. En base a estas evaluaciones la app muestra una serie de contenido e indicaciones con las que se busca mejorar su estado emocional. En los resultados se observó una reducción en el BDI promedio post-ensayo (-2.72) y una reducción mantenida los 6 meses (-4.03).

En la Universidad de Munich Litvin y Maier realizaron un ensayo [37] que buscaba reducir la depresión en 276 personas mayores de 18 años a los que se evaluó con la escala BDI (137 depresivos y 139 sujetos de control). Para este ensayo se utilizó la app “PsycApps” basada en métodos de intervención como la psicoeducación, la autoevaluación, el diario, la autogestión y el establecimiento de objetivos. A nivel general, los resultados mostraron un importante descenso de los niveles de depresión (media=2.68) respecto a los sujetos de control (media=1.91).

La depresión en estudiantes ha sido evaluada por McCloud et al. [47]. Este ensayo se realizó en cooperación con 4 universidades: University College London, School of Oriental and African Studies University of London, University of Buckingham y University of Roehampton. Fueron 168 los estudiantes que participaron en el ensayo entre marzo y junio del 2016. Tenían una edad superior a 18 años y mostraron un índice superior a 8 en la *Hospital Anxiety and Depression Scale (HADS)* (indicador de que sufrían de depresión y/o ansiedad). Se formaron 2 grupos (de 84 personas): uno al que se le aplicaría una intervención y otro que seguiría los tratamientos comunes. La intervención consistía en el uso de la app “Feel Stress Free” que integraba una terapia conductual cognitiva que ayuda a reducir los niveles de depresión y ansiedad. La aplicación consta de 4 actividades de relajación: respiración tranquila, meditación plena, relajación muscular profunda y autohipnosis. Los resultados del ensayo probaron que el grupo de intervención mostró un descenso continuo en los niveles de depresión: a las 4 semanas del 1.08 y a las 6 semanas un descenso de 1.56.

En otra investigación, Anýž et al. [4] mostraron que los indicadores de la depresión y manía pueden ser evidenciados mediante autoevaluaciones realizadas con una app. En este ensayo, realizado en República Checa, participaron 99 pacientes con TB entre 18 y 60 años. Para esto, los autores utilizaron la app ASERT, que consistía en un cuestionario de 10 preguntas (4 asociadas a depresión, 4 a manía y 2 para síntomas no específicos). Los resultados mostraron que las respuestas para la depresión estaban altamente relacionadas con el índice MADRS ( $P < .001$ ; bootstrap), y las respuestas para la manía estaban altamente relacionadas con el índice *Young Mania Rating Scale (YMRS)* ( $P < .001$ ; bootstrap). También se aplicó una regresión logística para la detección de recaídas y las precisiones para la depresión y manía fueron de 0.85 y 0.87, respectivamente.

Los cambios de los indicadores del sueño en pacientes con TB utilizando WD han sido estudiados por Fellendorf et al. [24]. Los autores realizaron un ensayo en la “Medical University of Graz” (Austria) en el que participaron 22 pacientes y 23 sujetos de control durante 3 semanas. Los indicadores de sueño fueron recopilados mediante la app “UP!”, que registra la hora de dormir y la hora de despertar, además de permitir a los participantes auto evaluar su estado anímico. Se utilizó también un *smartwatch* (Axivity de GeneActive) y el *Pittsburgh Sleep Quality Index (PSQI)*, para recopilar indicadores del sueño. Los resultados mostraron, en cuanto a la hora de dormir, una correlación de 0.77 y

## 2. Estado del arte

---

0.64 con los datos del *smartwatch* y el PSQI, respectivamente. En cuanto a la hora de despertar, la app consiguió una correlación de 0.59 y 0.53 con los datos del *smartwatch* y el PSQI, respectivamente.

En Polonia se realizó un ensayo sobre 51 pacientes con TB durante 208 días [18]. Se evaluó la relación de los datos recopilados con *smartphones* y recopilados con los índices YMRS y *Hamilton Depression Rating Scale (HDRS)*. Los autores utilizaron la app “BDmon” que registraba información relativa al uso del *smartphone* (llamadas y mensajes), información de auto evaluaciones e indicadores de la voz. En los resultados se observó que a medida que los síntomas de manía van aumentando, la duración de las llamadas varían más ( $P=0.03$ ), se envían más mensajes de texto ( $P<0.001$ ) y se consigue un mejor índice del estado anímico en las auto evaluaciones ( $P<0.001$ ), entre otras conclusiones. Por el contrario, a medida que los síntomas de la depresión van aumentando, se reduce el número de llamadas respondidas ( $P<0.001$ ), mayor fracción de llamadas perdidas ( $P<0.001$ ) y se consigue un menor índice del estado anímico en las auto evaluaciones ( $P<0.001$ ).

MONARCA es uno de los proyectos pioneros en el uso de los *smartphones* para el soporte de enfermedades mentales. Se desarrollo entre el 2009 y 2013, fue una iniciativa de la fundación “Monsenso”. Tuvo como objetivo aportar soluciones para el monitoreo multiparamétrico de información conductual para el TB, además de incluir la implementación de un complejo sistema para la gestión y tratamiento de la enfermedad. En este sistema se aprovecha las distintas prestaciones de los *smartphones* (*bluetooth*, GPS y conexión 3G) para, en base a ellos, obtener un alto número de parámetros del paciente. Consta de 2 piezas, una app y un portal web. La app permite al paciente registrar datos de auto evaluación así como almacenar información recopilada por los sensores del *smartphone* (llamadas, ubicaciones, calidad de sueño, etc.). El portal web ofrece una vista histórica de la información del paciente, con una configuración para el paciente y otra para el psiquiatra [85].

Son 5 los pilares sobre los que se diseñó este sistema: auto evaluación (información subjetiva), monitorización de actividades (a través del GPS), vista de datos históricos (en el portal web), auto tratamiento (configuración de alarmas de síntomas de episodios) y compartición de datos (paciente/psiquiatra). Este sistema fue sometido a exhaustivos ensayos clínicos con la finalidad de evidenciar su eficiencia.

En uno de los ensayos realizados con este sistema participaron 61 pacientes diagnosticados con TB. Los síntomas de manía y depresión fueron evaluados con los índices YMRS y HDRS, respectivamente. El ensayo duró 6 meses. Fueron muchos los resultados obtenidos, pero algunos de los más importantes reflejan que hay una correlación negativa entre los parámetros de actividad y el índice HDRS, lo que indica que mientras menos actividad se realiza mayor es el nivel de depresión. También se observa que hay una correlación positiva entre el estado de ánimo y actividades diarias con el índice de la manía, lo que indica que mientras más actividad se realiza se es más propenso a estar maníaco [23].

Sin embargo, entre las conclusiones no clínicas que se obtuvieron se pueden resaltar algunos aspectos que se deben considerar en la utilización de los

*smartphones*, por ejemplo, asegurar que todos los pacientes cuenten con un plan de datos, la necesidad de capacitar a los especialistas en el uso de la aplicación, y el sobrecalentamiento del teléfono al usar el GPS y el *bluetooth* [46].

## 2.2. Métodos: Técnicas de *machine learning* para diagnóstico de estados emocionales

La explotación de los datos comprende todas aquellas técnicas actualmente utilizadas para procesar los datos previamente recolectados y que terminan dándoles valor específico. Dada la dimensionalidad, se debe utilizar técnicas complejas que brinden información no evidente. Concretamente, en este contexto, los datos se irán incrementando en el tiempo, por lo que las técnicas de ML supervisadas son las más apropiadas. Estas permiten que el computador aprenda de los datos históricos para luego aplicar ese aprendizaje sobre datos nuevos. En este caso lo que se persigue es realizar la clasificación el estado emocional del paciente a partir de estos datos nuevos.

El sector clínico, en general, no es ajeno a las ventajas que puede proveer el uso de técnicas de ML para el diagnóstico de enfermedades [65][54]. Actualmente se pueden encontrar soluciones comerciales que las utilicen para diferentes diagnósticos. Sin embargo, concretamente en el área de la salud mental hay un vacío significativo en cuanto a su uso.

En esta sección vamos a detallar algunas de las investigaciones y publicaciones (que no tienen fines comerciales) en las que se utilizan los algoritmos de ML para la clasificación del estado emocional del paciente. Aunque estas investigaciones tengan distintas finalidades, una común a todas ellas es la de aportar una alternativa que permita la clasificación del estado emocional de los pacientes.

### 2.2.1. Detección del estado anímico a partir de ritmos circadianos

Entre las distintas aplicaciones de técnicas de ML podemos encontrar el uso del algoritmo de *Random Forest (RF)* [12]. Se ha utilizado en un ensayo realizado sobre 55 pacientes con alguno de estos tres diagnósticos: depresión, TB1 o TB2. El ensayo duró dos años y tenía como finalidad la predicción del estado anímico del paciente a través de la recolección y explotación de información estrechamente relacionada a los ritmos circadianos. Estos han sido identificados como un factor importante para el tratamiento de los desórdenes del estado de ánimo. Son cambios producidos en nuestro organismo (físicos, mentales y conductuales) que siguen un ciclo diario y que responden principalmente a la luz y la oscuridad. Para este ensayo se reunió información de sueño, ritmo cardíaco y exposición a la luz utilizando la app “eMoodChart”. El estado clínico real de los pacientes fue determinado por los médicos. Entre todos los pacientes se registraron 57 estados de depresión, 11 maníacos y 13 hipomaníacos. La cantidad de observaciones fue de 2003 con 131 variables cada observación. El modelo alcanzó una predicción del 0.65, 0.64 y 0.65; una sensibilidad de 0.57, 0.68 y

0.85; y una especificidad de 0.57, 0.68 y 0.85 para los pacientes con depresión, TB1 y TB2 respectivamente.

### 2.2.2. Diagnóstico del trastorno bipolar

Uno de los errores más frecuentes en el proceso de diagnóstico del TB es diagnosticarlo como depresión. En la universidad de Cambridge se realizó un ensayo cuyo objetivo fue desarrollar una herramienta de diagnóstico que ayude a reducir la frecuencia con la que se diagnostica erróneamente el TB como depresión [76]. Lo interesante de este ensayo fue que parte los datos utilizados se basaban en muestras de sangre que el paciente se extraía periódicamente utilizando el *Dried Blood Spot Kit (DBS)* y que enviaba por correo postal (194 items). Otra parte de los datos se obtuvieron a partir de cuestionarios de salud mental (957 items). En este ensayo participaron 688 pacientes (en dos reclutamientos), entre 18 y 45 años, con un índice de PHQ9 superior a 5 (ligeramente depresivos), no embarazadas, mujeres que no estén en periodo de lactancia y que tampoco tengan antecedentes suicidas. Antes de empezar el ensayo, se determinó el estado clínico del paciente utilizando el *Composite International Diagnostic Interview (CIDI)*. Se utilizó el algoritmo de *Extreme Gradient Boosting (XGBoost)* por su capacidad de manejar los valores vacíos y detectar relaciones no lineales entre las variables. El modelo matemático se alimentó de información de 1151 variables (entre muestras de sangre y cuestionarios) de pacientes con depresión o pacientes cuyo diagnóstico fue cambiado a TB por el CIDI. Como resultado del estudio en el primer reclutamiento se observó: 126 pacientes fueron clasificados con TB pero fueron diagnosticados inicialmente con depresión, 187 fueron clasificados como depresivos e inicialmente tuvieron el mismo diagnóstico. Para el segundo reclutamiento se observó: 98 pacientes fueron clasificados con TB pero fueron diagnosticados inicialmente con depresión, 112 fueron clasificados como depresivos e inicialmente tuvieron el mismo diagnóstico, 165 no tenían diagnósticos previos de trastornos mentales. Los algoritmos entrenados mostraron un *Area Under Curve/Receiver Operating Characteristics (AUROC)* de 0.92 al distinguir los pacientes incorrectamente diagnosticados de los correctamente diagnosticados. Al analizar subgrupos, el AUROC fue más alto en pacientes con TB1 (0.94) que en pacientes con TB2 (0.88).

Sin embargo, también se han utilizado técnicas más intrusivas para estudiar el estado emocional de un paciente. Por mencionar alguna, existe una investigación que identificó a pacientes con TB y sujetos de control en base a neuroimágenes extraídas por escáneres [55]. El total de sujetos analizados fue 256 de los cuales 128 estaban diagnosticados con TB y los otros 128 eran sujetos de control. Las imágenes fueron segmentadas en tres diferentes tipos de tejido (materia gris, materia blanca y líquido cerebrospinal). El número de variables excedía en gran medida el número de observaciones (problema llamado como “small-n-large-p” o “curse-of-dimensionality”). Para resolver este inconveniente se realizó un análisis de varianza (ANOVA) y así poder prescindir de variables redundantes. El algoritmo utilizado fue *Relevance Vector Machines (RVM)*, el cual distinguió pacientes reales de los sujetos de control con una precisión del 70.3%. Esta

investigación explica también las razones por las cuales el RVM es mejor que el algoritmo *Support Vector Machines (SVM)* en estos casos.

El uso del *smartphone* también se ha analizado para distinguir pacientes con TB de pacientes sanos. Existe un estudio realizado en base a información registrada por los *smartphones*: cantidad de llamadas, de SMS's o tiempo de la pantalla del *smartphone* encendida [22]. En este estudio se analizan datos objetivos comparando las acciones con el dispositivo entre 29 pacientes con TB con 37 individuos sanos durante un periodo de un año. Entre los resultados de este estudio se observa que el número de llamadas o de mensajes de texto es mayor en pacientes con TB, o que los pacientes en estados de depresión tienen la pantalla del teléfono encendida un menor tiempo que los individuos sanos. Un ensayo similar en que se utilizaron varias técnicas de ML fue realizado en Pekín, China, entre el 2019 y 2020. El objetivo fue realizar clasificaciones del estado emocional en pacientes diagnosticados con depresión. Participaron 261 pacientes ambulatorios diagnosticados con depresión durante 12 semanas. Se utilizaron varias fuentes de datos: La app de nombre "Mood Mirror" (disponible solo en Android) diseñada por los propios autores, el *smartwatch* "Mi Band 2" (Xiaomi) y las sesiones periódicas con el psiquiatra. La app permitía al paciente hacer una auto evaluación de su estado de ánimo, escogía en una escala de -3 (muy triste) a 3 (muy contento), además de incluir el PHQ9 y *Generalized Anxiety Disorder-7 (GAD7)*. La app permitía recopilar información del uso del teléfono (histórico de llamadas, uso de aplicaciones, GPS y estado de la pantalla). El *smartwatch* se conectaba a la app vía *bluetooth* y recopilaba información de los datos del sueño, cantidad de pasos y ritmo cardíaco. En cada visita, el médico utilizaba la escala HDRS para evaluar el estado clínico de cada paciente. Los estados clínicos de los pacientes se determinaron a partir de el PHQ9 y se establecieron 4 clasificaciones: estable-remisión, estable-depresivo, variación-drástica, variación-moderada. Para encontrar la mejor representación de los datos se realizaron pruebas con distintas combinaciones de datos y para reducir la dimensionalidad de estos se utilizó técnicas de FS. Posteriormente, se aplicaron distintas técnicas de ML: SVM, KNN, DT, Naïve Bayes, RF y Regresión Logística (RLG). En base a los resultados se concluyó que la precisión más alta alcanzada fue de 0.84 con una sensibilidad de 0.97 conseguida por el algoritmo de RF, y se logró utilizando solo los datos de la calidad del sueño para la clasificación entre estable-remisión y variación-moderada. Por el contrario, la precisión más baja fue de 0.66 con una sensibilidad de 0.65 conseguida por el algoritmo DT, y se consiguió utilizando únicamente los datos del *smartphone* para la clasificación entre estable-depresivo y variación-drástica [6].

Aunque el alcance del proyecto PSYCHE se limita a la recolección de datos con la tecnología ya mencionada, existen trabajos que utilizan los datos recolectados por PSYCHE para realizar clasificaciones del estado clínico de pacientes con TB. Uno de estos ensayos tenía como finalidad la clasificación del estado de ánimo con señales del sistema nervioso aprovechando las ventajas ofrecidas por PSYCHE para la recolección de datos [78]. Fue realizado en Pisa, Italia. Como primera fase se analizó el estado emocional de tres pacientes durante un periodo de 90 días. Se utilizaron las escalas: *Bauer Internal Mood Scale (IMS)*, *Profile of Mood States*

## 2. Estado del arte

---

(POMS) y YMRS. Se realizó un preprocesamiento sobre los datos recopilados con PSYCHE, se aplicó la técnica de Análisis de Componentes Principales (PCA) y de FS para la reducción de la dimensionalidad. Posteriormente se aplicaron los siguientes métodos de ML: *Linear Discriminant classifier (LDS)*, *Quadratic Discriminant Classifier (QDC)*, *Mixture of Gaussian (MOG)*, KNN, *Kohonen Self Organizing Map (KSOM)*, *Multi-layer Perceptron (MLP)* y *Probabilistic neural network (PNN)*. Para conocer cuál de estos algoritmos fue el más preciso se realizó el test de ANOVA, siendo el MLP el que alcanzó la máxima precisión de 0.95.

De entre los diversos parámetros utilizados para el estudio del estado de ánimo, el análisis de la voz es un parámetro que puede dar conclusiones importantes. En una investigación realizada en Michigan, EEUU, [27] se analiza la voz de pacientes en llamadas telefónicas intentando clasificar cada llamada dentro de uno de los siguientes grupos: maníaco/eutímico o depresivo/eutímico. En este trabajo se incluyeron 37 pacientes diagnosticados con TB1 y TB2. Para el registro de la voz se utilizó *smartphones* de distintos modelos. A cada participante se le proporcionó un Samsung Galaxy modelo S3, S4 o S5. Estos *smartphones* tenían instalada la app PRIORI que permitía la grabación de las voces. Con estos equipos se registraron 34.380 llamadas en 2436 horas. Se utilizaron las escalas de YMRS y HDRS para evaluar el estado clínico del paciente. Dado que la calidad de la voz durante una llamada telefónica depende de factores externos, se aplican diversos algoritmos para estandarizar los datos (*declipping*, *audio normalization* y *noise-robust segmentation*). Para realizar la clasificación objeto de esta publicación se utilizó el algoritmo SVM. Los resultados generales mostraron una *Area Under Curve (AUC)* de 0.64 para identificar pacientes depresivos/eutímicos y un AUC de 0.57 para pacientes maníacos/eutímicos. Sin embargo, haciendo una clasificación a nivel de grupos, se observó que los modelos S5 tenían un mejor rendimiento para los pacientes maníacos con un AUC de 0.78, lo que podría indicar que la calidad de los registros de voz del S5 influye de manera positiva en los resultados.

Un trabajo de investigación realizado en Madrid, España, [42] estudia el cambio del estado emocional en pacientes diagnosticados con TB. El objetivo de este trabajo es mostrar una especificación para un sistema basado en datos de distintas fuentes. Se utilizan 3 fuentes de información: (1) datos históricos personales del paciente, (2) datos de auto evaluación del paciente obtenidos mediante formularios y (3) datos de un *smartwatch*. Participaron 25 pacientes del hospital “Nuestra Señora de la Paz” y se aplicó un preprocesamiento y estandarización de estos datos. Se trabajó en dos enfoques: uno general que incluía todas las variables excepto los datos personales y un enfoque personal que sí las utilizaba. Los resultados en el segundo enfoque mostraron que el “alineamiento de secuencia” [8] y la “detección de anomalías” fueron las mejores técnicas para clasificar los estados emocionales atípicos de los pacientes.

### 2.2.3. Respuesta a los tratamientos para la depresión

Hay un estudio en el que se predice la respuesta a los tratamientos antidepresivos en pacientes diagnosticados con depresión [36]. En esta investigación se ha trabajado con un enfoque de “Deep Learning Prediction” y se analizan factores clínicos y genéticos como el SNP (Single Nucleotide Polymorphism), sexo, edad, el HDRS, episodios depresivos, estado civil e intentos de suicidio. El estudio incluyó 455 pacientes diagnosticados con depresión. Para definir si el paciente responde o no a los tratamientos se evalúa el HDRS; si se observa un cambio de -50 % se establece que no hay respuesta. Se ha empleado la red neuronal MLP considerando 1, 2 y 3 capas ocultas. En los resultados se observa que con 2 capas ocultas se obtiene un AUC de 0.82 y con 3 capas ocultas un AUC de 0.80, lo que significa que el MLP es también un buen algoritmo que se ajusta a investigaciones de este tipo.

Una investigación realizada en Hunan, China, [86] para la detección de la PPD (Depresión postparto) analiza las ventajas del uso de técnicas de ML, concretamente el uso del algoritmo de RF y SVM. En esta investigación participaron 508 mujeres que fueron reclutadas de un centro de maternidad en Changsha y en Yiyan, ambos en la provincia de Hunan entre septiembre del 2016 y febrero del 2017. Los investigadores recopilaron datos hasta las 6 semanas posteriores al parto. Se utilizaron muchos índices para recopilar información de las participantes: un cuestionario construido a medida que recogía información de la edad, educación, síntomas de menopausia y antecedentes de enfermedades mentales, el EPDS, el *Brief Resilience Scale (BRS)*, el PSQI, el *Social Support Rating Scale (SSRS)* y el GAD7. El criterio para categorizar a una paciente como depresiva era un EPDS superior a 9.5. Un 34.1 % (173) de pacientes alcanzó este puntaje. Se utilizaron 2 métodos para reducir la dimensionalidad de los datos: la consulta con los expertos y el “Random Forest-Based filter feature selection” (FFS-RF). Para el entrenamiento de los modelos matemáticos se utilizó el 75 % (381) de participantes, el restante se utilizó como datos de prueba y verificación. El estudio incluyó un total de 25 variables: 17 fueron seleccionadas en base a la opinión de los expertos y 7 en base al FFS-RF. Se construyeron 4 modelos: (1) E-RF: Modelo RF utilizando las variables seleccionadas por los expertos, (2) E-SVM: Modelo SVM utilizando las variables seleccionadas por los expertos, (3) F-RF: Modelo RF utilizando las variables seleccionadas por el FFS-RF; y (4) F-SVM: Modelo SVM utilizando variables seleccionadas por el FFS-RF. Al comparar estos 4 modelos, en cuanto a sensibilidad, especificidad y AUROC los modelos alcanzaron los siguientes valores: E-RF: 0.48, 0.86 y 0.75, E-SVM: 0.68, 0.87 y 0.81, F-RF: 0.48, 0.86 y 0.70; y F-SVM: 0.69, 0.83 y 0.78 respectivamente. Al observar los resultados se pudo concluir que el modelo F-SVM fue el óptimo. Se pudo concluir también que los modelos SVM tienen una ratio ligeramente menor para la clasificación así como una sensibilidad ligeramente mayor en comparación con los modelos RF. En cuanto al AUROC, los modelos SVM son también ligeramente superiores a los modelos RF.

### 2.2.4. Diagnóstico de la ansiedad, estrés y suicidio

Otra de las investigaciones publicada en el 2020 es la realizada en Amsterdam por el *Netherlands Study of Depression and Anxiety (NESDA)*, cuya finalidad era la clasificación en la recuperación de la ansiedad en 2 años aplicando técnicas de ML. Inicialmente se contó con 2981 participantes pero luego, por falta de datos o datos inconsistentes, se redujo a 887 diagnosticados con ansiedad. Sólo era necesario que los participantes hayan sido diagnosticados alguna vez de ansiedad o estén diagnosticados de depresión o ansiedad (al inicio del ensayo). Los datos se recogieron entre el 2004 y 2007. Se establecieron dos variables a clasificar, la recuperación de ansiedad (que se definía como la ausencia de síntomas de ansiedad) y la recuperación de todas las *Common Mental Disorders (CMD)*, estas son : ansiedad, depresión, distimia y dependencia del alcohol. Se utilizaron indicadores de 5 dominios distintos (clínicos, psicológicos, demográficos, biológicos y estilo de vida) llegando a un total de 569 variables. Para la clasificación se utilizó el algoritmo de RF que separó los datos de entrenamiento (90 %) y de test (10 %) en 100 repeticiones. El rendimiento final se calculó en base a la media de todas éstas. Después de los dos primeros años, 484 (54.6 %) se recuperaron de la ansiedad y 362 (40.8 %) no tuvieron síntomas de ningún CMD. Los resultados del modelo mostraron un AUC entre 0.49 y 0.67 para los 5 dominios. En cuanto a la precisión, se alcanzó la mayor incluyendo la información de los 5 dominios: 0.62 con una sensibilidad de 0.62 y una especificidad de 0.62. También se investigó el rendimiento del modelo en grupos: (1) con comorbilidad (presencia de más de una enfermedad en un mismo paciente) y (2) sin comorbilidad. El AUC obtenido fue de 0.68 en grupo comorbilidad y de 0.64 en el grupo sin ella, concluyéndose que el modelo es ligeramente superior en el grupo con comorbilidad [10].

Las redes neuronales son modelos computacionales que también se han utilizado. Existe un estudio [52] publicado en el 2017 que analiza el nivel del estrés de las personas a partir del movimiento, distancias y ubicaciones obtenidas desde el GPS de sus teléfonos móviles. Este se realizó en New Hampshire, EEUU. Se utilizó el dataset público: StudentLife [82] que contenía información de trazas del GPS, tiempo de conversación, datos de acelerómetros, etc. de 49 estudiantes graduados y no graduados. Son 12 los nodos de entrada de la red neuronal (distancia recorrida en un día, máximo desplazamiento, desviación estándar de la distancia, etc) y 3 los estados de salida que indican el nivel de estrés: “encima de la media”, “debajo de la media”, “estrés medio”. Este estudio se basa inicialmente en el estudio de Mehrotra y Musolesi [48], en el cual se concluye que la información de ubicación obtenida por la mayoría de los teléfonos inteligentes se puede usar exitosamente para monitorizar y predecir niveles de depresión. Aunque se pretende seguir trabajando en mejorar el modelo, los primeros resultados mostraron una precisión de 0.43 y una sensibilidad de 0.45.

Cuando hablamos de trastornos depresivos, también tenemos que hacer mención a las fatales consecuencias que éstos puedan acarrear, en concreto el suicidio. Hay investigaciones que han trabajado sobre este tema. En Texas, EEUU, se ha realizado un estudio en el cual se analiza la probabilidad de intentos de

suicidios en pacientes con trastornos del estado de ánimo [57]. Se incluyeron 144 pacientes con diagnósticos de depresión, TB1 o TB2, y se analizaron variables que fueron consideradas relevantes para esta problemática por parte de especialistas: edad, sexo, raza, años de educación, fobia social, pánico, desorden obsesivo compulsivo, síntomas psicósomáticos, etc. El factor “riesgo de suicidio” se estableció en base a los intentos de suicidios registrados en el *Conte center for the Neuroscience of Mental Disorders (CCNMD)*. Los algoritmos de ML utilizados para realizar el estudio fueron *Least Absolute Selection Shrinkage and Selection Operator (LASSO)*, RVM y SVM. El algoritmo RVM fue el más preciso al predecir correctamente 103 de los 144 pacientes (72 % precisión, 71.1 % de sensibilidad y 71.3 % de especificidad).

Otra investigación que utilizó ML para identificar personas proclives a cometer suicidio se realizó en Santiago de Chile [7]. En este estudio se ha trabajado con 707 sujetos que tienen algún desorden mental (112 con TB), de los cuales 349 tienen actitudes suicidas y 358 no. El total de variables estudiadas es 343 que se obtuvieron aplicando diversas escalas: *Outcome Questionnaire (OQ45.2)*, *State Trait Anger Expression Inventory (STAXI-2)*, *Reasons for Living Scale (RFL)*, *Depressive Experience Questionnaire*, *Family APGAR*, así como variables demográficas, sociales, clínicas, etc. El modelo selecciona 22 variables que definen si la persona está en la zona de riesgo (más proclive a cometer suicidio o está pensando en el suicidio). Se aplicaron diferentes algoritmos de ML para identificar personas con riesgo de suicidio, como entre ellas están: *Classification and Regression Tree (CART)*, KNN, RF, SVM, AdaBoost o MLP. El algoritmo que SVM alcanzó el mayor rendimiento con unos valores para precisión, sensibilidad y especificidad del 0.78, 0.77 y 0.79, seguido por RF (0.78 de media de precisión) y AdaBoost (0.76 de media de precisión).

Existe también una investigación publicada en el 2020 realizada en Malasia. En este trabajo se analizan las actuales técnicas de ML para la predicción de problemas de salud mental en alumnos de educación superior. En este país es una problemática importante dada su prevalencia. En base a datos de la *National health and morbidity survey (NHMS)* del 2017, una de cada 5 personas sufre depresión, dos de cada 5 personas sufren de ansiedad y una de cada 10 personas padece estrés. En esta publicación se hace un resumen de artículos en las que se mencionan distintas técnicas de ML para el diagnóstico de problemas de salud mental. Como conclusión, los algoritmos más utilizados son: SVM, DT y redes neuronales, que alcanzan una precisión sobre el 0.70, por lo que pueden ser una alternativa a considerar para abordar esta problemática [70].

### 2.3. Aportaciones de la tesis en el estado del arte

Una vez revisados los últimos trabajos en las vertientes antes mencionadas, se puede observar que dada la gran dimensionalidad y variedad de datos con los que este ámbito nos fuerza a trabajar, hace falta una herramienta que haga una integración solida de ambas vertientes pero que al mismo tiempo ponga el foco en la caracterización de los pacientes. El análisis personalizado es un

## 2. Estado del arte

---

enfoque que nos va a permitir lidiar con la diferencia existente en los parámetros conductuales entre pacientes [13] y por tanto aportará más valor en el diagnóstico. Además se debe mencionar que muchos de los trabajos revisados se centran en la aplicación de ML para la determinación del estado emocional del paciente. También existen otro tipo de factores que se pueden aprovechar para obtener conclusiones relevantes para el psiquiatra, por ejemplo, el impacto que tiene el día de la semana en ciertos indicadores de conducta.

Esta tesis, busca llenar ese vacío mediante la presentación de una investigación enfocado al estudio personalizado de los cambios de conducta en pacientes con TB mediante el análisis de fuentes de información alternativas (*smartphone* y *smartwatches*) y distintas técnicas de ML. Esta tesis incluye una serie de publicaciones en las que se ha presentado el detalle de este trabajo.

## Capítulo 3

# Datos y herramientas computacionales

Para la realización de esta tesis, se ha trabajado con pacientes ya diagnosticados con TB. En total se han recogido datos de 101 pacientes tratados en el hospital “Nuestra Señora de la Paz” de Madrid. Todas las partes involucradas en esta investigación firmaron todas las políticas y acuerdos de confidencialidad necesarios para la investigación.

Una vez definido el conjunto de pacientes, se estableció un periodo de recolección de datos de un año y medio, empezando a inicios del 2018 y finalizando a mediados del 2019. Se utilizaron diferentes fuentes de información. Por una parte, relojes médicos de monitorización de actividad (*smartwatch*), que proporcionaban datos objetivos de forma automática y no intrusiva. Por otra parte, datos diarios procedentes de formularios que completaban los propios pacientes, para lo cual los pacientes recibieron las indicaciones y las herramientas necesarias. Durante el periodo de monitorización, los pacientes también mantuvieron sesiones de seguimiento con el psiquiatra con una periodicidad promedio de dos semanas, y se añadieron a las bases de datos los resultados anotados en estas sesiones incluyendo los indicadores calculados por el propio psiquiatra.

### 3.1. Fuentes de información

A continuación se describen las tres fuentes de información utilizadas en esta tesis.

#### 3.1.1. Consultas médicas

Dado que se está trabajando con pacientes ya diagnosticados con TB, según tratamiento estándar deben asistir a sesiones periódicas de seguimiento con el psiquiatra. Es el psiquiatra quien después de observar y evaluar al paciente registraba la información, por tanto es información objetiva.

Se estableció como procedimiento utilizar tres hojas Excel para almacenar diferentes tipos de variables. Se registraron las correspondientes a las escalas de YMRS y HDRS. También se registraron variables del tipo de consulta así como observaciones clínicas hechas por el psiquiatra. Además en estas consultas se registró la variable más importante para el objetivo de esta investigación: el estado emocional del paciente. Esta información se registraba en cada sesión, es decir, aproximadamente cada dos semanas.

### 3. Datos y herramientas computacionales

#### 3.1.2. Self-report

Se trata de un formulario electrónico que el paciente debió cumplimentar diariamente durante el periodo de recopilación de datos. Las preguntas fueron definidas por el psiquiatra. Es una fuente utilizada por el paciente para registrar información acerca de sus sentimientos, por tanto, son datos subjetivos.

Existe una versión desarrollada en Android [44]. Para su uso el paciente debe registrarse con un correo electrónico y generar un PIN. La aplicación permite también modificar las respuestas una vez cumplimentado el formulario así como exportar las respuestas en formato “csv” (figura 3.1). Las preguntas son de opción múltiple con una sola respuesta. También existe una versión elaborada con GoogleForm que permite descargar la información en ficheros “csv”. Cada registro de estos ficheros corresponde a un día de un paciente.

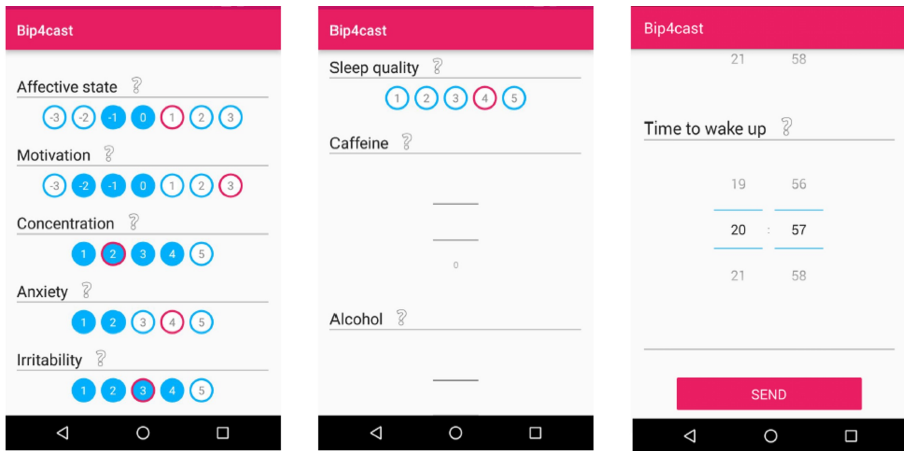


Figura 3.1: Aplicación Android *self-report*

#### 3.1.3. Smartwatch

Actualmente existen diversas marcas y modelos de *smartwatches*. Considerando el objetivo de esta tesis se optó por un dispositivo que se focalice en obtener información del nivel de actividad diaria y la calidad del sueño. Después de evaluar las prestaciones de los distintos modelos existentes, se decidió finalmente optar por el dispositivo GENEActiv v.1.2 del fabricante Activinsights Ltd [1]. Este dispositivo es una fuente de datos objetivos dado que registra variables fisiológicas.

Se adquirió un dispositivo para cada paciente para que fuera usado durante el periodo de recopilación de datos. La información almacenada en los dispositivos se fue extrayendo en cada consulta médica e incluía la información diaria desde la última consulta. La información se extrae en ficheros binarios y por tanto fue necesario una conversión a un formato “csv”.

Muchas de las variables de las tres fuentes de información fueron incluidas en una aplicación interactiva que permitía seleccionar un paciente, un rango de fechas y unas variables para mostrar gráficamente su comportamiento (figura 3.2).

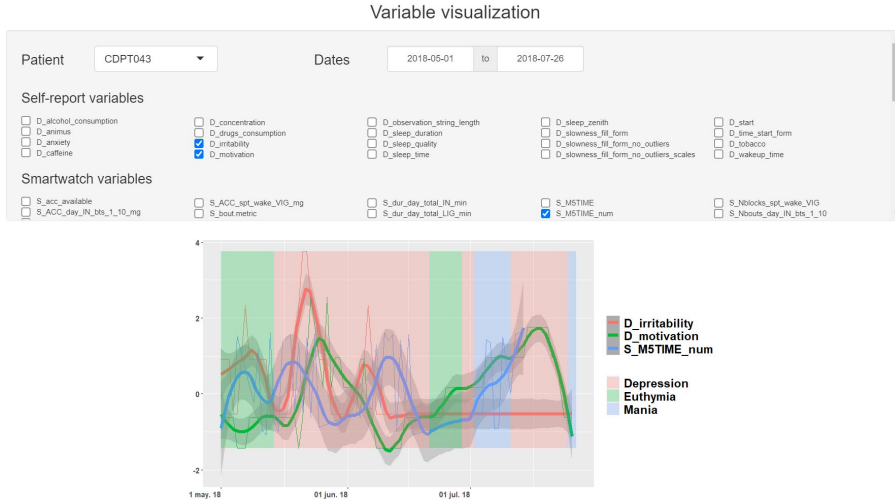


Figura 3.2: Visualización de variables

## 3.2. Integración de datos

Los datos extraídos de cada una de las fuentes de información tienen formato y estructura diferente. Para poder realizar un análisis transversal con todos ellos es necesario unos procesos de carga e integración que permitan construir una única estructura de datos que albergue todas las variables de cada una de las fuentes. Mediante esta estructura se podrá observar el estado emocional de un paciente en un determinado día y los valores cuantitativos de cada una de las variables en ese mismo día. Los procesos de carga de datos así como de integración de los mismos han sido elaborados utilizando scripts del lenguaje de programación *R* (3.3.1).

### 3.2.1. Datos

Para llevar a cabo el proceso de integración previamente es necesario conocer los datos que serán objeto de la integración. En esta sección detallaremos los datos “crudos” extraídos de cada fuente de información, así como la dimensión de cada uno de ellos.

Mediante el *self-report* se registró información subjetiva de los pacientes. Al finalizar el periodo de recolección de datos, la información de todos los pacientes

### 3. Datos y herramientas computacionales

---

se volcó en ficheros *csv*. Para facilitar el manejo de esta información, se fusionaron en un único fichero Excel de nombre *Diario.xlsx*.

En cuanto a la información objetiva registrada durante las consultas médicas se utilizó tres hojas Excel, una para la escala HDRS (*HDRS.xlsx*), otra para la escala YMRS (*YMRS.xlsx*) y una para registrar otros criterios clínicos así como la variable objetivo: el estado emocional del paciente (*Estados.xlsx*). A partir de ésta, una vez finalizado el periodo de recolección, se construyó una hoja Excel adicional (*Eventos.xlsx*) en la que se registró el estado emocional del paciente para cada uno de los días del periodo de recolección de datos.

La información objetiva de los *smartwatches* se exportó en ficheros binarios manualmente por el psiquiatra en cada consulta médica. Por tanto se necesitó de una transformación a formato *csv* con una herramienta especial que detallaremos más adelante. Una vez finalizado el periodo de recolección de datos se tuvo disponibles un total de 111 ficheros binarios pertenecientes a distintos pacientes.

Todos los ficheros utilizados como entrada para el proceso de integración los podemos observar en la tabla 3.1.

Cuadro 3.1: Ficheros utilizados para el registro de datos

Fuente	Fichero	Descripción	Variables	Volumen
Self-report	Diario.xlsx	Datos subjetivos	18	2634 registros
Consulta	HDRS.xlsx	escala HDRS	24	2499 registros
Consulta	YMRS.xlsx	escala YMRS	21	2507 registros
Consulta	Estados.xlsx	Criterios clínicos y estado emocional del paciente	25	2696 registros
Consulta	Eventos.xlsx	Estado emocional del paciente por día	5	4268 registros
Smartwatch	*.bin	Variables de actividad y sueño	113	111 ficheros

#### 3.2.2. Normalización de los datos

Previo al proceso de integración hemos aplicado algunas técnicas de normalización. La finalidad de éstas es conseguir que los datos se asemejen para facilitar su enlace y evitar problemas en la integración. Algunos de los motivos por los que fue importante estandarizar los datos en esta investigación son:

1. **Disminución la cantidad de anomalías causadas por el registro manual de los datos.** El registro manual de la información conlleva posibles errores de digitación, el registro de outliers o duplicados. También se ha encontrado comentarios escritos por el psiquiatra en columnas que no son variables.

2. **Codificación de variables cualitativas.** La transformación de un valor cualitativo a un valor cuantitativo permite utilizar esa variable para su posterior análisis matemático. Por ejemplo, para la variable estado emocional del paciente podemos reemplazar los valores “depresivo”, “eutímico” y “maníaco”, por valores 1,2 y 3, respectivamente.
3. **Establecer un formato único para variables comunes entre los datos.** Las variables comunes en todas las fuentes de información son: la fecha *date* y el id del paciente *id*. Sin embargo no en todas las fuentes han utilizado el mismo formato para la fecha ni tampoco para el id.
4. **Escalar variables cuyo rango sea muy distinto a los demás.** Para la generación de gráficos de dos variables o más, éstas deben tener una escala similar. Esto sucedió bastante al graficar variables de distintas fuentes, por ejemplo, consultas y *smartwatches*. La escalabilidad de los valores es configurable en nuestro proceso.
5. **Interpolación de datos nulos.** La existencia de vacíos en los datos fue una problemática que hemos solventado mediante la interpolación y “arrastre” de datos. Esta es una alternativa configurable en nuestro proceso.

### 3.2.3. Procesos de carga de datos

Cada uno de las plantillas cumplimentadas así como los ficheros binarios de los *smartwatches* fueron cargados con scripts independientes. Dentro de cada script se aplicaron ciertas transformaciones a las variables y se crearon algunas variables adicionales. La salida de cada uno de estos procesos fue un dataset que tenía la variable que identifica al paciente  $A\_id$ , la fecha  $A\_date$  y las variables propias de cada plantilla. Los scripts utilizados son los siguientes:

1. **01\_01\_carga\_diario.r:** Este script carga las variables de la plantilla “Diario.xlsx” para almacenarlas en el dataset *diario*. Además, en base a éstas, calcula las siguientes variables adicionales:  $D\_sleep\_zenith$ ,  $D\_slowness\_fill\_form$  y  $D\_sleep\_duration$ . Las variables cargadas con este script se pueden observar en el anexo D.
2. **01\_02\_carga\_pulsera.r:** Este script se encarga de procesar los ficheros binarios de los *smartwatches* para transformarlos en ficheros “csv”. Los ficheros “csv” contienen información de la actividad física del paciente en cada día. Esta transformación se consigue mediante una función especial de *R* llamada *g.shell.GGIR* y será detallada en la sección 3.3.1. Todas las variables obtenidas se almacenan en el dataset *pulsera*. Las variables cargadas con este script se pueden observar en el anexo E.
3. **02\_01\_carga\_hdrs.r:** Este script carga las variables de la plantilla “HDRS.xlsx” para almacenarlas en el dataset *hdrs*. Además, en base a éstas calcula las siguientes variables adicionales:  $H\_score$  y  $H\_hrsd\_dx$ . Las variables cargadas con este script se pueden observar en el anexo F.

### 3. Datos y herramientas computacionales

---

4. **02\_02\_carga\_ymrs.r:** Este script carga las variables de la plantilla “YMRS.xlsx” para almacenarlas en el dataset *ymrs*. Además, en base a éstas calcula las siguientes variables adicionales: *Y\_score* y *Y\_ymrs\_dx*. Las variables cargadas con este script se pueden observar en el anexo G.
5. **03\_01\_carga\_estados.r:** Este script carga las variables de la plantilla “Estados.xlsx” para almacenarlas en el dataset *estados*. En base a éstas calcula las siguientes variables adicionales: *I\_dx\_short* y *Z\_dx\_numeric*. Las variables cargadas con este script se pueden observar en el anexo H.
6. **03\_02\_carga\_eventos.r:** Este script carga variables de la plantilla “Eventos.xlsx” para almacenarlas en el dataset *eventos*. Las variables cargadas con este script se pueden observar en el anexo I.

#### 3.2.4. Procesos de integración

Los procesos de carga previamente detallados fueron invocados desde unos scripts de integración. Cada uno de estos scripts realizaron operaciones *joins* sobre datos de la misma periodicidad. Se utilizó el identificador del paciente: *A\_id* y la fecha *A\_date* como variables clave para realizar los *joins*. Estos scripts fueron:

1. **01\_carga\_continuos.r:** Este script invoca a los scripts: *01\_01\_carga\_diario.r* y *01\_02\_carga\_pulsera.r* para generar los datasets *diario* y *pulsera* respectivamente. A partir de estos se generó el dataset *continuos* que contenían datos de periodicidad diaria.
2. **02\_carga\_discretos.r:** Este script invoca a los scripts: *02\_01\_carga\_hdrs.r* y *02\_02\_carga\_ymrs.r* para generar los datasets *hdrs* y *ymrs* respectivamente. A partir de estos se generó el dataset *discretos* que contenían los valores de las índices registrados en las consultas médicas.
3. **03\_carga\_diagnosticos.r:** Este script invoca a los scripts: *03\_01\_carga\_estados.r* y *03\_02\_carga\_eventos.r* para generar los datasets *estados* y *eventos*, respectivamente. A partir de estos se generó el dataset *diagnosticos* que contenían información de los estados emocionales de los pacientes por cada día.
4. **00\_iniciar.r:** Este script es el encargado de iniciar todo el proceso de integración. Invoca a los tres anteriores para generar los tres datasets: *continuos*, *discretos* y *diagnosticos*. Construye el dataset *dftotal* haciendo un *join* de *continuos* y *discretos* y posteriormente *diagnosticos*. Una vez hecha la integración realiza un renombrado de todas las variables para anteponerles los literales “D\_”, “S\_”, “H\_”, “Y\_”, “I\_” y “E\_” si provienen de los datasets *diario*, *pulsera*, *hdrs*, *ymrs*, *estados* y *eventos*, respectivamente. La variable objetivo es renombrada a *Z\_dx\_numeric*.

Las variables objetivo, como las variables  $A\_id$  y  $A\_date$ , se pueden observar en el anexo J. El esquema del proceso de integración se muestra en la figura 3.3.

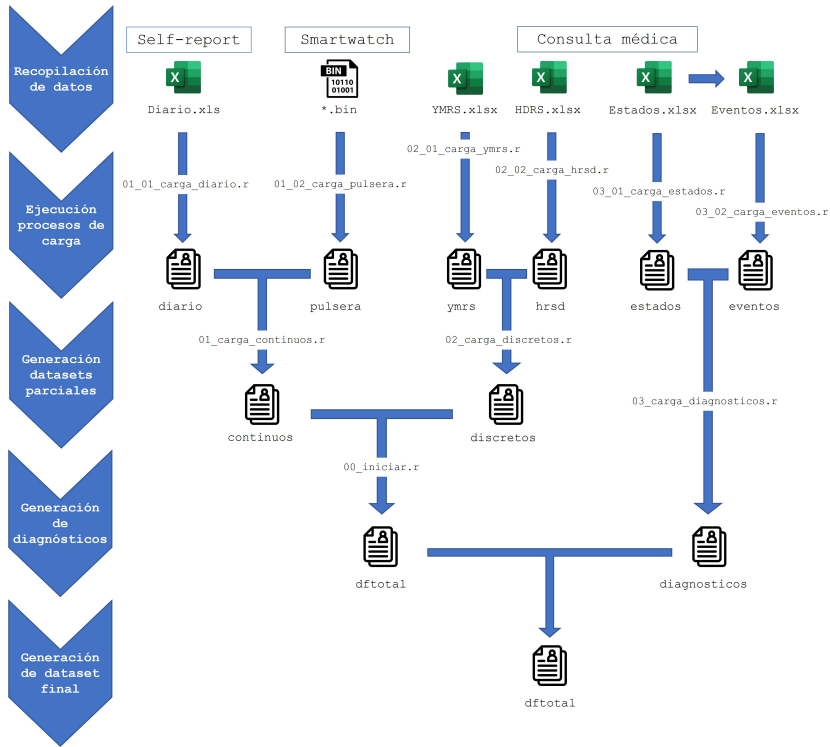


Figura 3.3: Proceso de integración.

### 3.3. Herramientas computacionales

En esta sección se va a detallar el software así como las técnicas utilizadas en esta investigación.

#### 3.3.1. Software

El lenguaje de programación utilizado fue *R* versión 3.6.3. Esta es una herramienta *open source* que dispone de gran cantidad de librerías útiles para el análisis estadístico, generación de gráficos así como para la aplicación de técnicas de ML. El IDE utilizado fue *R Studio* versión 1.4.1106 también de disponibilidad gratuita. Utilizando este lenguaje, sus librerías y con este IDE se ha realizado: la carga de datos, las técnicas de normalización, el análisis estadístico y matemático, la construcción de interfaces gráficas para la visualización de indicadores, la aplicación de técnicas de clasificación, etc.

### 3. Datos y herramientas computacionales

---

El proceso de carga de los *smartwatches* necesitó un proceso especial que consistía en una conversión de los ficheros binarios (ficheros en los que se almacena la información del *smartwatch*) a ficheros “csv”. Esta transformación se ha realizado con la librería *GGIR*, concretamente la función *g.shell.GGIR*.

Esta función aplica 5 fases sobre el fichero de entrada. La primera fase carga y pre procesa los ficheros en base a unos parámetros establecidos por el usuario. La segunda fase analiza y resume la salida de la fase anterior en ficheros *RData* que servirán para las siguientes fases. La tercera fase, identifica los periodos de inactividad que servirán de entrada para la fase 4. La cuarta fase recoge los periodos de inactividad de la fase anterior para clasificarlos como “SPT” (sleep period time) o no. Y la quinta fase reúne la salida de la segunda y cuarta fases para generar reportes de la actividad diaria y nocturna. Cada una de estas fases es parametrizable. La invocación a la función *g.shell.GGIR* que hemos utilizado se puede observar en el anexo B.

Para la integración de los datos se utilizaron librerías de *R* con diversas finalidades (la lista de librerías utilizadas se muestra en la tabla A.1), Esto nos permitió participar en la “Jornada de usuarios de R 2019” en la que, entre otras cosas presentamos nuestra aplicación interactiva (figura 3.2) que se construyó con la librería *shiny*.

#### 3.3.2. Técnicas utilizadas

Entre las técnicas de machine learning utilizadas en esta investigación, las que mejores resultados han producido se explican a continuación.

##### 3.3.2.1. Árboles de decisión

Es una técnica de aprendizaje supervisado que toma decisiones mediante reglas de inferencia. Estas técnicas son utilizadas tanto para clasificación como para predicción y se pueden utilizar tanto para variables categóricas como cuantitativas. Concretamente para predecir una variable categórica (como ha sido el caso de esta investigación) se utiliza la moda de la región de entrenamiento a la que pertenece una observación. El objetivo del entrenamiento es lograr el nodo más “puro”, es decir, el nodo que cubre las observaciones de una sola clase.

Para identificar la región en la que se dividirán los datos, usualmente se utiliza la técnica de *gain information* (formula 3.1) que nos proporciona el atributo de división en base a la información requerida para una mejor descripción del árbol. En términos generales, la ganancia de información esperada es el cambio de la entropía de información entre nodos. En la figura 3.4 se puede observar una forma muy simplificada de árbol de decisión utilizado en esta investigación.

$$Gain(A) = entropy(parent\ node) - [Average\ (entropy(children))] \quad (3.1)$$

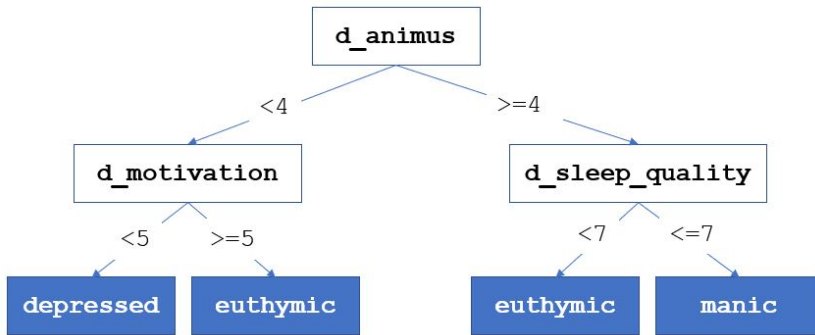


Figura 3.4: Arbol de decisión.

### 3.3.2.2. Random Forest

Esta técnica generaliza la técnica anterior construyendo varios arboles de decisión individuales y agrupando los resultados en una única predicción final. La colección de resultados para tomar una decisión final se categoriza como *ensemble technique*. El algoritmo selecciona aleatoriamente de entre todos los datos un cierto número de observaciones y variables para construir un árbol de decisión a partir de ellas. Este proceso se repite muchas veces y se construyen árboles diferentes. El número de árboles es un factor relevante ya que cuanto mayor sea este número menor será el error de clasificación.

El entrenamiento de machine learning utiliza un indicador para decidir la mejor división. El indicador utilizado es el índice de *Gini* mostrando en la ecuación 3.2.

$$Gini = 1 - \sum_{i=1}^c (p_i)^2 \quad (3.2)$$

Donde,  $p_i$  representa la frecuencia relativa de la clase observada y  $c$  representa el número de clases. Sin embargo, la entropía es también una alternativa para decidir la mejor división de datos, esta se expresa en la formula 3.3 donde las variables son las mismas que en el índice *Gini*.

$$Entropy = \sum_{i=1}^c -(p_i \cdot \log(p_i)) \quad (3.3)$$

### 3.3.2.3. Máquinas de vectores de soporte

Es una técnica de de entrenamiento supervisado muy popular para solucionar problemas de clasificación. La idea principal del algoritmo en el que se basa es separar clases en los datos de entrenamiento. La manera más sencilla de separar dos grupos es usando una línea recta (para una dimensión), un plano (para dos dimensiones) o un hiperplano multi dimensional (para tres o más dimensiones). Cuando una región no puede separarse mediante hiperplanos lineales se utiliza

### 3. Datos y herramientas computacionales

---

un función kernel no lineal. Por lo tanto cada clase se acota mediante hiperplanos. La técnica de Máquinas de Vector Soporte (svm) es una técnica de optimización buscando los mejores hiperplanos. Los vectores de soporte corresponden con las observaciones más cercana a los hiperplanos. Estos son considerados críticos ya que si se eliminasen, podrían alterar la posición del hiperplano divisor. Un hiperplano se representa como el conjunto de puntos  $\vec{x}$  que satisfacen la formula 3.4.

$$\vec{w} \cdot \vec{x} + b = 0 \quad (3.4)$$

$$m = \frac{2}{\|\vec{w}\|} \quad (3.5)$$

Tal que se maximice el margen  $m$  como muestra la ecuación 3.5. Donde  $\|\vec{w}\|$  es el modulo del vector  $\vec{w}$ .

#### 3.3.2.4. Regresión lineal y logística

La regresión lineal es una de las técnicas más utilizadas en machine learning. Sirve para calcular la relación entre una variable dependiente y una o varias variables independientes. Para efectos de esta tesis, dado que tenemos varias variables independientes, usaremos la regresión lineal múltiple. La idea principal es construir la ecuación 3.6 que se ajuste a nuestros datos. Los coeficientes de la ecuación se optimizan mediante la técnica de mínimos cuadrados.

$$Y_i = \beta_0 + \beta_1 \cdot x_{i1} + \beta_2 \cdot x_{i2} + \dots + \beta_p \cdot x_{ip} + e \quad (3.6)$$

Donde  $p$  es la cantidad de variables independientes,  $Y_i$  es la variable dependiente,  $\beta_i$  es el coeficiente que pondera a cada variable,  $\beta_0$  es una traslación y  $e$  representa el error que se comete al tratar de estimar la variable dependiente como una transformación lineal de las variables independientes. Es decir, el error que se comete en el ajuste.

La Regresión Logística (RLG) es una generalización de la regresión lineal: si una variable categórica es codificada como uno o cero, se puede ajustar el modelo a una regresión lineal por coeficientes de mínimos cuadrados. La regresión logística resuelve mejor el ajuste cuando hay valores atípicos. Para ello utiliza una función sigmoide (ecuación 3.7).

$$\text{sigmoide} = \sigma(x) = \frac{1}{1 + e^{-x}} \quad (3.7)$$

La ecuación de ajuste para la regresión logística se muestra en la ecuación 3.8.

$$p(Y) = \frac{e^{\beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_i \cdot x_i}}{1 + e^{\beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_i \cdot x_i}} \quad (3.8)$$

El conjunto total de librerías utilizadas para este investigación se muestra en el anexo A.

## Capítulo 4

# Desarrollo y contribuciones de esta tesis

Este capítulo tiene como objetivo reflejar la cronología de las publicaciones realizadas describiendo paso por paso los avances y resultados de la investigación.

- La investigación tiene como punto de partida un análisis realizado en el 2018 acerca de los métodos existentes para el registro de indicadores psicológicos y fisiológicos en pacientes con algún trastorno mental. Los resultados se publicaron en el Congreso de la Asociación Española para la Inteligencia Artificial (CAEPIA) celebrado en Granada en 2018 (Publicación I [61]). El análisis realizado fue bastante general dada la etapa temprana de la investigación. También se describieron proyectos relacionados como PSYCHE y MONARCA, en los que los autores realizaron ensayos con resultados muy interesantes a pesar de presentar limitaciones en su aplicación.

En este punto de la investigación existía ya un convenio de colaboración con el hospital psiquiátrico Nuestra Sra. de la Paz (Madrid), y se acordó con los médicos del programa las fuentes de datos de interés. Además de los datos recabados en consulta médica se decidió la recolección de datos monitorizados con relojes inteligentes (*smartwatch*) y con formularios diarios (*self-reports*), que los pacientes rellenarían a mano mediante formularios Google y/o una aplicación móvil específica que se diseñó exclusivamente para esta investigación. Una vez que todas las partes involucradas firmaron las políticas de protección de datos, y que los pacientes recibieron la capacitación para la utilización del *smartwatch* y los *self-reports*, el proceso de recolección de datos se inició.

Los datos se empezaron a recabar de un conjunto de pacientes diagnosticados con TB en un periodo inicialmente programado de un año, iniciándose a comienzos del 2018. Paralelamente se comenzaron también los trabajos de integración y procesado. Debido al diagnóstico de los pacientes, los datos eran registrados de manera supervisada por el psiquiatra en su propio equipo informático de forma que también en ese momento él mismo establecía el estado emocional de cada paciente (Eutimia, Depresión o Manía). Con los datos registrados hasta el momento de la redacción de la primera publicación se realizaron las primeras pruebas de la aplicación de técnicas de ML, utilizando los datos de cada fuente por separado y también integrando algunas de ellas. En ese momento, la intención era realizar un estudio grupal que permitiera la identificación de un patrón de

#### 4. Desarrollo y contribuciones de esta tesis

---

comportamiento en las personas diagnosticadas. Sin embargo los primeros resultados no fueron muy prometedores (62% de precisión en promedio, tabla 4.1). Al analizar el problema se observó que el conjunto de variables características del comportamiento de un paciente podía ser muy distinto al de otro paciente. Como resultado positivo de esta primera etapa de la investigación, se confirmó con los datos la relación directa entre el trastorno bipolar, el factor de la calidad del sueño y la actividad diurna.

Cuadro 4.1: Precisión de algoritmos de *machine learning*

<b>Datasets/Algoritmos</b>	<b>DT</b>	<b>RF</b>	<b>SVM</b>	<b>RLG</b>	<b>Promedio</b>
YMRS	0.36	0.38	0.38	0.76	0.47
HDRS	0.78	0.55	0.36	0.62	0.58
diario	0.70	0.75	0.68	0.67	0.69
estados	0.44	0.65	0.57	0.51	0.54
YMRS-HDRS	0.63	0.63	0.70	0.33	0.57
diario-estados	0.67	0.17	0.42	0.83	0.52
<b>Promedio</b>	<b>0.60</b>	<b>0.52</b>	<b>0.52</b>	<b>0.62</b>	

- A partir de estos resultados la investigación dio un giro hacia la definición de un conjunto de características que estableciera el perfil de cada paciente particular. Esta decisión también se vio reforzada por las limitaciones que otras investigaciones relacionadas presentaron en la recolección de datos suficientes de distintos pacientes y por las diferentes propuestas encontradas hacia la personalización de los análisis y los tratamientos consecuentes,

Siendo conscientes, por tanto, de la importancia de realizar análisis personalizado, se empezó a reformular el modelo y las técnicas de esta investigación para mejorar los resultados obtenidos. Estos resultados se publicaron en la *International Alan Turing Conference on Decision Support and Recommender Systems*, celebrada en Londres en noviembre de 2019 (Publicación II [38]). A diferencia de la anterior publicación, aquí se diseñó un proceso de integración de los datos de todas las fuentes de información. La idea fue construir una única estructura que contuviese la información de todos los pacientes, siendo cada observación correspondiente a los datos diarios del periodo de recolección. A partir de este nuevo dataset y con la intención de seguir un enfoque personalizado, se realizó un filtrado por paciente antes de realizar cualquier análisis posterior. Aislado los datos de cada paciente confirmamos que no era factible realizar un análisis en general porque cada paciente presentaba unas dinámicas de estado emocional distintas y un perfil estadístico diferente.

Para resolver la ausencia de información se aplicaron técnicas para completar datos. También se aplicó normalización para resolver la heterogeneidad de algunas variables. Sin embargo, la dimensionalidad de la estructura integrada era muy alta, en ese momento 187 variables. Este elevado número de variables requirió un análisis de ruido y dependencias que

---

permitiese una reducción de la dimensionalidad. Analizamos las variables que tienen poco o ningún impacto en el estado emocional y también otras variables que ocasionan ruido o redundancia. Por ejemplo, las variables indicadoras de consumo de cafeína y tabaco. La cantidad de café consumido diariamente es una variable descartable desde el punto de vista estadístico. La explicación más plausible es que el consumo de café corresponde a un hábito realizado de forma recurrente como una costumbre y no como un acto impulsivo. Los actos impulsivos son precisamente los que se asocian a las emociones, así el consumo de tabaco, aunque es una variable que sólo es significativa en los pacientes fumadores, es muy determinante en los modelos predictivos personalizados.

En la Publicación II [38] se incluyen los primeros pasos en esta dirección. Además en este artículo se realiza un PCA para buscar el número adecuado de componentes necesarias para cubrir un alto nivel de variabilidad de los datos. También se evalúa la correlación entre todas las variables, y se visualizan los coeficientes de correlación de Pearson para una evaluación rápida de la correlación. De esta forma podríamos descartar las variables con coeficiente de correlación absoluta muy alta. Sin embargo, en este punto de la investigación las matrices se generaron por cada fuente de información de forma separada y no mostraron coeficientes demasiado significativos (figura 4.1).

Dentro del dataset integrado, las variables recogidas mediante formularios diarios (*self-report*) son de tipo cualitativo aunque se incluye su valor discretizado con el fin de poder generar analíticas de tipo cuantitativo. Para efectos de análisis con técnicas de ML, descartamos las variables cualitativas reduciéndose el número de variables prácticamente a la mitad. Aún así, la cantidad total de variables se mantenía en 173. Para facilitar el manejo de todas ellas se decidió modificar el nombre anteponiéndole un literal que identifique la fuente de procedencia: “D\_”, “S\_”, “H\_”, “Y\_”, “I\_” y “E\_” para formularios diarios, monitorización por *smartwatch*, escala HDRS, escala YMRS, intervenciones y eventos, respectivamente (la variable objetivo, estado emocional del paciente, fue también renombrada a *Z\_dx\_numeric*).

- Conociendo ya la importancia de la personalización y de la reducción de la dimensionalidad, empezamos a estudiar los cambios de estado emocional de los pacientes, poniendo especial atención en la eutimia (o normalidad del paciente particular). La idea era aislar un conjunto de variables que mejor definen el estado de eutimia en cada paciente. De esta forma podríamos también determinar las variables que van mostrando alguna alteración cuando se presenta un cambio desde el estado eutímico al estado depresivo o al estado maníaco. La caracterización del estado eutímico es importante por ser el estado a partir del cual van a acontecer otros estados. En la figura 4.2, se observa las diferentes transiciones desde el estado eutímico. En base a esta conclusión comenzamos el proceso de caracterización del perfil eutímico de cada paciente. Para ello filtramos las observaciones separando

#### 4. Desarrollo y contribuciones de esta tesis

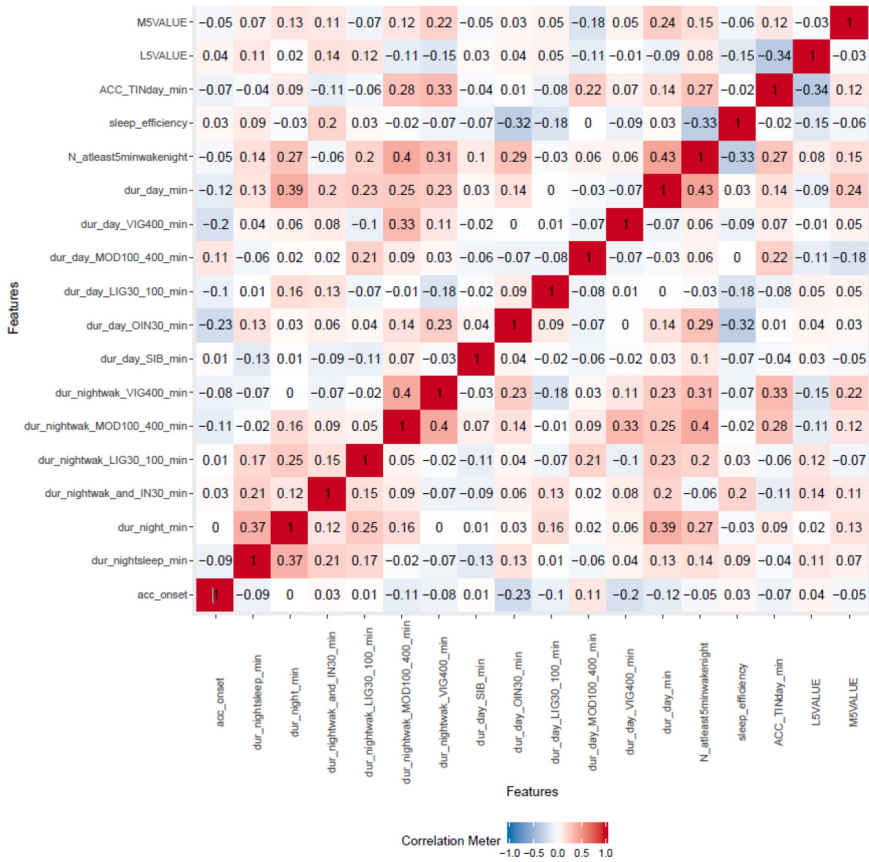


Figura 4.1: Primeros coeficientes de correlación de *Pearson*

los episodios de manía y depresión. El perfil eutímico se caracteriza a partir de las variaciones en las ventanas de tiempo con diagnóstico de no crisis. Construido este perfil se pudo observar la alteración de ciertas variables cuando ocurre un cambio de estado emocional, siendo diferentes las variables que afectan a los estados diagnosticados como depresión y los estados diagnosticados como manía. Se comparó el comportamiento de estas variables en estado eutímico con su comportamiento en un estado no-eutímico (figura 4.3).

Los resultados de esta investigación se publicaron en la revista *Mathematics* (perteneciente al primer cuartil del ranking JCR) y los resultados más importantes de esta publicación (Publicación III [40]) se describen a continuación.

Se realizó una selección de los pacientes que participaron más activamente en el estudio para que los resultados fuesen acordes con el análisis personalizado pretendido. Se presentó una tabla resumen de los estados

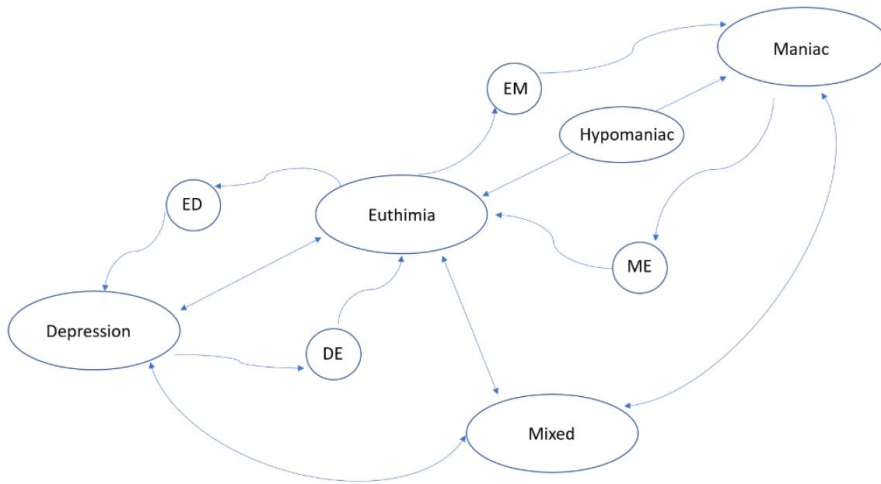


Figura 4.2: Transiciones desde el estado eutímico

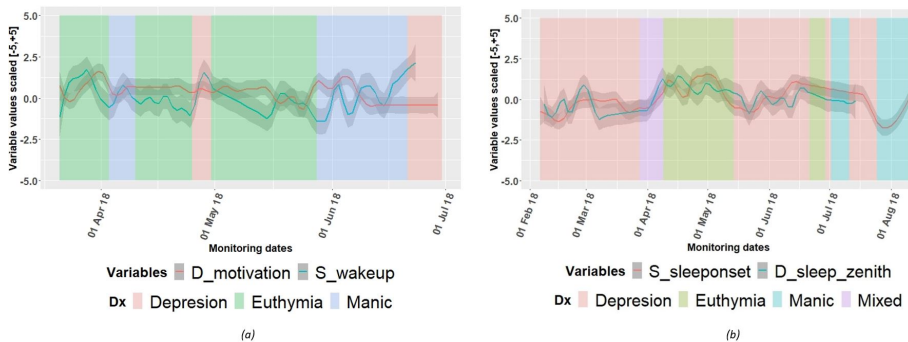


Figura 4.3: Comportamiento de las variables en distintos estados emocionales (a)  $D\_motivation$  y  $S\_wakeup$ . (b)  $S\_sleeponset$  y  $D\_sleep\_zenith$

emocionales registrados para cada uno de los pacientes durante el periodo de recolección de datos. A partir de esta tabla se puede también inferir el tipo de trastorno bipolar (TB1 ó TB2) de cada paciente. Para reducir la dimensionalidad, además de las técnicas mencionadas en la publicación anterior, también se incluyó un ranking de importancia de variables en base a varias técnicas de FS (considerando el estado emocional del paciente como variable objetivo). Los resultados probaron las hipótesis de las primeras publicaciones donde se manifestaba la importancia de las variables relacionadas con el sueño y la actividad diaria para la determinación del estado emocional. En cuanto a la correlación, se calcularon nuevas matrices pero a diferencia de la publicación anterior, éstas fueron generadas por cada paciente y considerando únicamente las variables seleccionadas por las

## 4. Desarrollo y contribuciones de esta tesis

técnicas de FS aplicadas. En estas nuevas matrices se observó que algunos pares de variables alcanzaban ya un coeficiente de correlación bastante alto (0.97, 0.85) lo que indicaba la posibilidad de mejorar el modelo mediante la reducción de variables. Además, con la finalidad de descartar variables muy parecidas, calculamos el coeficiente de correlación de cada variable con la variable objetivo. Dentro de la heterogeneidad de los datos, encontramos que estos coeficientes en algunos casos alcanzaron valores como 0.92 y 0.95. Otra de las técnicas utilizadas para reducir la dimensionalidad fue el PCA. La finalidad de esta técnica es calcular el menor número de componentes con los que se cubra el mayor porcentaje de varianza de los datos. Al aplicar este análisis para uno de los pacientes se concluyó que los tres primeros componentes son los que contienen la mayor información y que con 7 componentes se llega a acumular el 99% de la varianza (figura 4.4).

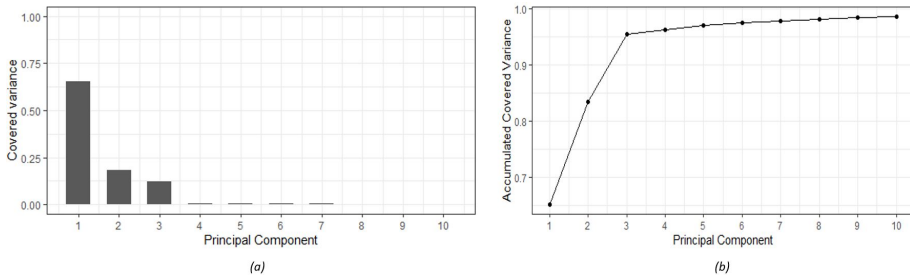


Figura 4.4: (a) Varianza cubierta.(b) Varianza acumulada

A partir del análisis anterior se aplicaron diversas técnicas de ML (SVM, RF, DT y RLG) para caracterizar el estado emocional del paciente. El pseudocódigo utilizado para estas ejecuciones se pueden ver en el anexo C. Estos resultados fueron muy concluyentes pues se demostró que utilizando únicamente las variables seleccionadas se alcanza en promedio una mejora del 12% en la precisión respecto a utilizar todas las variables del dataset. El algoritmo de RF fue el óptimo en todos los casos. También se comprobó que utilizando las variables de los *smartwatch* y los *self-report* conjuntamente, los algoritmos de ML alcanzan una mayor precisión (en un 80%) que al utilizar únicamente los variables de los *smartwatch*. Para esta última conclusión, dada la heterogeneidad de los datos y para efectos de la publicación, se consideró únicamente variables relacionadas con la calidad del sueño y la actividad diaria tanto del *smartwatch* como del *self-report* para aplicar las técnicas de ML bajo las mismas condiciones.

- Poco después de esta publicación recibimos nuevos datos de monitorización de los pacientes por parte del psiquiatra a cargo del proyecto. La nueva información fue fácilmente incorporada mediante el proceso de integración preparado con antelación, lo que también demostró la escalabilidad de nuestro sistema en este sentido. En ese momento de la investigación planteamos la realización del análisis personalizado basado en técnicas de

series temporales bajo la premisa de que ciertos estados emocionales, en especial la depresión, están muy condicionados por la estacionalidad. Sin embargo, a principios de 2021 descartamos esta idea tras un análisis sencillo en ventanas de tiempo amplias, pues la gran mayoría de los pacientes presentan crisis bipolar en distintas estaciones y los resultados estadísticos no mostraban relación con la estacionalidad ni ningún patrón observable. La explicación desde el punto de vista médico es que estos pacientes están muy condicionados por variables exógenas de su entorno inmediato y no sólo por variables endógenas que podrían justificar una estacionalidad en la crisis. Por dicho motivo se descartó el estudio con series temporales al observar que los datos presentan una aleatoriedad diferente y que el cambio de estado emocional no respondía a ningún periodo en particular.

Al tratarse de pacientes bipolares, el estado anímico es proclive a cambios imprevisibles causados por eventos cotidianos pero significantes para un paciente particular. No obstante sí observamos fluctuaciones en el ciclo semanal (figura 4.5). Por ello se decidió realizar un estudio de los datos teniendo en cuenta distintos momentos en la semana: días laborables - fines de semana - cada día de la semana por separado. Se consideró

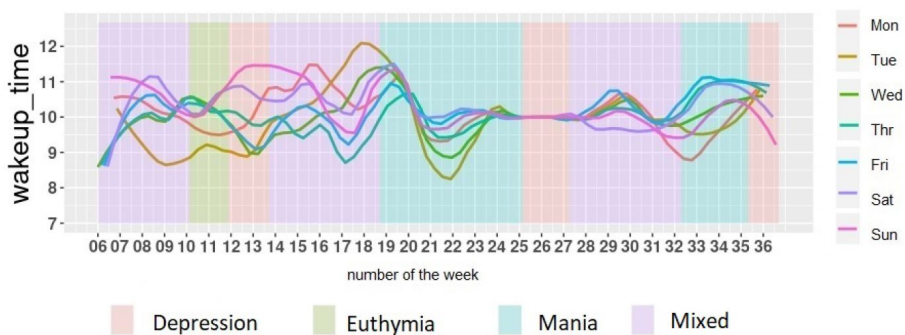


Figura 4.5: Variable  $D\_wakeup\_time$  por semanas

oportuno realizar un análisis de las variables sobre ventanas semanales de información ya que el día de la semana era un indicador importante de actividad. Es decir, por lo general los fines de semana se descansa o se hace más actividad física a diferencia de los días laborables. Nos pareció interesante añadir a las variables con las que habíamos trabajado el factor día de la semana en nuestro análisis para mejorar el perfil eutímico de cada paciente en particular. Para ello utilizamos agregadores ponderados de información, concretamente operadores OWA de dimensión 7. El criterio para determinar el peso de estos operadores fue una normalización en base a la desviación estándar de las variables. Estos operadores permitieron ponderar la importancia de cada día de la semana para cada paciente. Los resultados de esta investigación se presentaron en el *IEEE International Conference on Progress in Informatics and Computing* (Publicación IV

#### 4. Desarrollo y contribuciones de esta tesis

[60]). En este artículo se incluye el análisis realizado para un paciente concreto utilizando las variables que mejor lo caracterizan: motivación, irritabilidad, duración del sueño y hora de despertar. Se presenta un modelo matemático que define el estado emocional particular como una función que puede estimarse mediante un estadístico obtenido como una transformación lineal de las variables seleccionadas, que a su vez se estiman mediante los datos recabados por monitorización. Las fluctuaciones moderadas acotan la función modelizada en una banda de confianza representativa del estado eutímico. Cuando la estimación de la función se escapa de dicha banda de confianza, el modelo indica una tendencia a estado de crisis por depresión (si la se escapa de la cota inferior) o por manía (si supera la cota superior). En este artículo no sólo se hace una particularización para cada paciente sino que también se detalla el desarrollo del último y más elaborado modelo en esta investigación.

Como resultado se pudo observar que los episodios de manía suelen producirse en días de semana, mientras que en fines de semana existe una menor fluctuación de las variables. Esta información permite mejorar la estimación de la función del estado emocional del paciente así como de la banda de confianza que se le asigna de forma particular.

- A finales del año 2021, paralelamente a nuestro artículo anterior, también empezamos a redactar una disertación de la investigación realizada en los últimos años para publicarla en la revista *Frontiers in Physiology* (Publicación V [39]) por invitación del editor (Q1 en el ranking JCR).

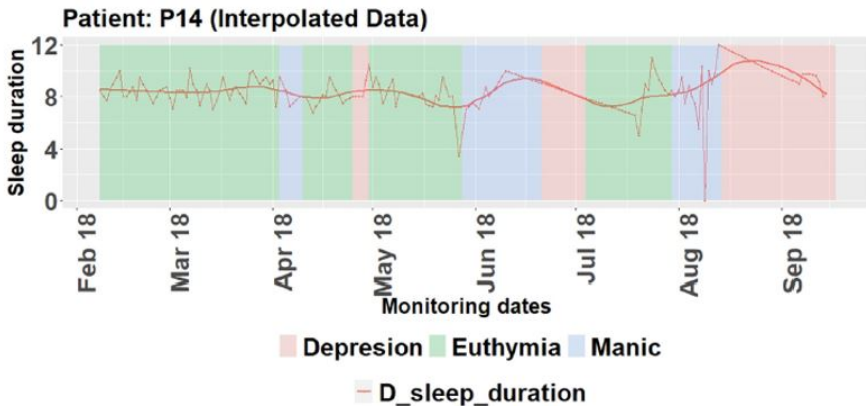


Figura 4.6: Variable  $D\_sleep\_duration$  en cada estado emocional

En esta publicación se resumen las técnicas aplicadas, los resultados, conclusiones y propuestas de mejora para trabajos futuros así como también las limitaciones encontradas en este tipo de investigaciones. Además también se incluyen resultados novedosos que no fueron publicados anteriormente, especialmente comparaciones de resultados de aplicación en

---

pacientes diagnosticados sólo con depresión. En esta publicación se hace especial énfasis en la importancia de la personalización para evitar asignar tratamientos erróneos (lo que puede en ocasiones empeorar el estado clínico del paciente) y la importancia de la selección de variables que caractericen a cada paciente. De manera similar a otras publicaciones, se hace mención al impacto de las variables relacionadas con el sueño en el estado emocional. En la figura 4.5, se observa el comportamiento de la variable de duración del sueño en cada estado emocional de un paciente.

Por otra parte, los resultados obtenidos mostraron que al usar técnicas de ML supervisadas sobre datos de distintas fuentes de información se consigue una importante precisión en la caracterización del estado emocional. Además, en comparación con la depresión, la manía es un estado más sencillo de prevenir ya que los factores que lo indican no sólo son más distinguibles visualmente por parte del psiquiatra, sino que a nivel de datos tienen una mayor acentuación y, por lo tanto, más variabilidad estadística.



## Capítulo 5

# Conclusiones y trabajos futuros

Además de tener una prevalencia importante, el TB es una enfermedad crónica para la cuál no existe cura hasta la fecha [30]. El riesgo de intento de suicidio es de un 33.9% [19] y la mortalidad es de hasta un 20% [21] [67]. Estas estadísticas evidencian la relevancia de contar con técnicas más eficaces para el diagnóstico y tratamiento. Varios autores afirman que el TB es una enfermedad que con cierta frecuencia es erróneamente diagnosticada como depresión; esto trae como consecuencia la asignación de un tratamiento incorrecto que puede llegar a agudizar el estado clínico del paciente con TB. Uno de los motivos de estos desaciertos es la complejidad para encontrar diferencias sintomáticas entre la depresión y el TB en estado de depresivo. Otra causa importante es el tipo de información con la que usualmente se realizan los diagnósticos y además se asignan tratamientos. Los procesos clínicos de diagnóstico y tratamiento estándares se suelen basar principalmente en información subjetiva que el paciente proporciona acerca de sus emociones.

Esta investigación, al igual que otras incluidas en el estado del arte, ponen de manifiesto que para el estudio de los estados emocionales en el trastorno bipolar se tiene que poner especial atención a cada individuo en particular (medicina personalizada) y a los indicadores de la actividad diurna y la calidad del sueño. La presencia de alteraciones en estos indicadores puede mostrar un cambio de estado emocional estaría por ocurrir. La alternativa que planteamos en esta tesis es incorporar datos objetivos y subjetivos. Por un lado, la información subjetiva es registrada mediante formularios cumplimentados por el paciente diariamente; y por otro lado, información objetiva relacionada con indicadores de la calidad del sueño y la actividad diaria mediante el uso de *smartwatches*. Estas fuentes de información tienen un pequeño grado de intrusión en la actividad del paciente. Los formularios debido a que los pacientes tienen que programarse diariamente cumplimentarlo y los *smartwatches* debido a que, en lo posible, deben ser usados las 24 horas del día. Existen otros dispositivos con un nivel mayor o menor de intrusión; sin embargo, se optó por los *smartwatches* dado que son los más accesibles por su coste y además el tipo de indicadores que registran es muy variado.

El efecto que pueda tener nuestro aporte viene también asociado al nivel de participación de los pacientes. Mientras mayor sea el número de días que hayan cumplimentado el formulario, usado el *smartwatch*, o hayan asistido a sus consultas médicas programadas, más datos podrán ser incorporados en el análisis. En este aspecto el entorno social del paciente, especialmente el familiar, podría tener un rol importante. Debido a que el estado emocional de los pacientes es proclive a cambios bruscos, el registro de información (formularios y *smartwatches*) o la asistencia a las consultas médicas se pueden convertir en

## 5. Conclusiones y trabajos futuros

---

actividades no prioritarias para ellos, y por tanto, pueden prescindir de ellos. Es en estas situaciones cuando los familiares y/o amigos podrían ayudar a los pacientes a tener una continuidad en el registro de información. Del mismo modo se debe incentivar la asistencia a las consultas médicas. Esto es fundamental porque es en estas consultas cuando el psiquiatra clasifica el estado emocional del paciente, el indicador más importante en todo nuestro análisis.

En concordancia con otros autores, hemos observado que los pacientes con TB presentan dinámicas muy diferentes entre ellos. Por tanto, un estudio general no es factible y debe aplicarse un estudio personalizado.

La fase de análisis de esta tesis consiste en la aplicación de técnicas supervisadas de ML sobre datos de cada paciente. Esto implica el disponer de un histórico de datos. Mientras más amplio sea el histórico de datos, mayor será la precisión alcanzada por los algoritmos de ML. Concretamente, para los resultados de esta tesis, el histórico de datos incluía información de aproximadamente 6 meses en promedio por cada paciente. Esto supone que las técnicas de ML o los modelos computacionales utilizados pueden mejorarse al ir incorporando más datos con los que se puedan entrenar estos modelos.

Es muy importante recalcar que la aplicación de estos métodos permite la caracterización del estado emocional del paciente. La métrica que evalúa su eficacia es la precisión. A pesar de que los resultados conseguidos muestran altos índices de precisión, en ningún caso se ha conseguido una precisión máxima. Esto indica que la propuesta presentada puede ser utilizada por el psiquiatra como una referencia adicional para asignar tratamientos más efectivos, pero nunca para reemplazar el criterio humano.

La actividad diaria y la calidad del sueño han mostrado tener un fuerte efecto en la determinación de los estados emocionales de personas con trastornos mentales. Esto no solo es afirmado por muchos autores en la literatura revisada, sino que también coincide con resultados obtenidos al aplicar técnicas para la reducción de la dimensionalidad de los datos (FS). Por otra parte se ha observado que el consumo del tabaco tiene especial importancia para la detección de cambios de estado emocional en pacientes fumadores. Uno de los hitos más importantes que hemos conseguido para la detección de cambios del estado emocional fue la elaboración de perfiles eutímicos de los pacientes. Con esto se estableció unos umbrales de las variables bajo los cuales los pacientes no presentaban crisis alguna permitiéndonos distinguir aquellos periodos de depresión o manía.

Se han aplicado diversas técnicas de ML (SVM, RF, DT, RLG, etc.). En una primera aplicación de ML sobre los datos de todos los pacientes sin realizar ningún análisis previo de las variables se alcanzó una precisión media de 56.3% a nivel fuente de información y una precisión máxima de 62% utilizando RLG. Sin embargo, al aplicar posteriormente un enfoque personalizado e integrando ya todas las fuentes de información se logró incrementar esta precisión a una media de 79.4%. Incluso al utilizar únicamente variables seleccionadas (FS) esta media se incrementó a 87.5%. Además los resultados mostraron que se alcanzó una precisión media de 74.1% utilizando únicamente variables seleccionadas de los *smartwatches*, pero ésta se incrementó a 79.9% al incluir variables seleccionadas de los formularios.

---

Otra conclusión importante es el impacto del día de la semana en el estado emocional de los pacientes. Ciertas variables presentan un comportamiento diferente entre días laborales y fines de semana. Para llegar a esta conclusión hemos construido un modelo matemático que utiliza operadores OWA de 7 dimensiones con el cuál a cada variable le asignamos un peso en cada día de la semana. Todas las variables utilizadas en este modelo, así como en los modelos detallados anteriormente, se han podido representar gráficamente con la librería *ggplot* de *R*. Además se construyó con la librería *shiny* una herramienta interactiva que permitía ir incluyendo diversas variables en un mismo gráfico para un paciente seleccionado previamente.

Cabe destacar que la efectividad de estos modelos está limitada al nivel de participación de los pacientes. Mientras mayor sea la rigurosidad en el registro de los datos (en las tres fuentes de información), mas exactos serán los resultados obtenidos.

Como resumen se puede decir que el registro continuo de los datos y el uso de técnicas supervisadas de ML permitirían alimentar los modelos matemáticos y caracterizar el estado emocional del paciente.

Uno de los propósitos bajo el cual fue diseñado el proceso de integración fue la inclusión de nuevas fuentes de información. Con la finalidad de tener una mejor caracterización del estado emocional del paciente, a mediano plazo nos proyectamos la integración de datos de electrocardiogramas o datos electrofisiológicos en actividad y durante la noche.

Dadas las diferentes dinámicas del trastorno bipolar, los datos presentan una heterogeneidad muy significativa. A partir de esta variabilidad, como trabajo futuro nos planteamos también el cálculo de la entropía de Shannon para identificar aquellos indicadores con una alta probabilidad de desencadenar cambios en el estado emocional.

Esta investigación se ha centrado en la utilización de técnicas de aprendizaje supervisadas (SVM, RF, RLG, etc.). Sin embargo el ML incluye también técnicas no supervisadas por ejemplo: *Subgroup Discovery*, *Association Rule Discovery* o clustering predictivo. Como siguiente paso de investigación nos planteamos calcular la relación entre los grupos generados por las técnicas no supervisadas y la variable objetivo utilizada en las técnicas supervisadas.



# Bibliografía

- [1] Activinsights. *Geneactiv original - wrist-worn actigraphy device: Geneactiv accelerometers*. URL: <https://www.activinsights.com/actigraphy/geneactiv-original/> (visitado 20-06-2021).
- [2] Al Zoubi, O. y col. «EEG microstates temporal dynamics differentiate individuals with mood and anxiety disorders from healthy subjects». En: *Frontiers in human neuroscience* vol. 13 (2019), pág. 56.
- [3] Al Zoubi, O. y col. «Predicting age from brain EEG signals—A machine learning approach». En: *Frontiers in aging neuroscience* vol. 10 (2018), pág. 184.
- [4] Anýž, J. y col. «Validity of the Aktibipo Self-rating Questionnaire for the Digital Self-assessment of Mood and Relapse Detection in Patients With Bipolar Disorder: Instrument Validation Study». En: *JMIR mental health* vol. 8, n.º 8 (2021), e26348.
- [5] Babulal, G. M. y col. «Depression is associated with tau and not amyloid positron emission tomography in cognitively normal adults». En: *Journal of Alzheimer's Disease* vol. 74, n.º 4 (2020), págs. 1045-1055.
- [6] Bai, R. y col. «Tracking and monitoring mood stability of patients with major depressive disorder by machine learning models using passive digital data: prospective naturalistic multicenter study». En: *JMIR mHealth and uHealth* vol. 9, n.º 3 (2021), e24365.
- [7] Barros, J. y col. «Suicide detection in Chile: proposing a predictive model for suicide risk in a clinical sample of patients with mood disorders». En: *Brazilian Journal of Psychiatry* vol. 39 (2016), págs. 1-11.
- [8] Becerro, Ó. S., López, V. L. y Pérez, B. G. «Performance improvement in the sequence alignment algorithms using dynamic libraries and multi-agent system». En: *XXXV Congreso Nacional SEIO: IX Jornadas de Estadística Pública: Universidad Pública de Navarra, Pamplona, del 26 al 29 de mayo de 2015*. Departamento de Estadística e Investigación Operativa. Universidad de Navarra. 2015, págs. 53-53.
- [9] Ben-Zeev, D. y col. «Transdiagnostic mobile health: smartphone intervention reduces depressive symptoms in people with mood and psychotic disorders». En: *JMIR mental health* vol. 6, n.º 4 (2019), e13202.
- [10] Bokma, W. A. y col. «Predicting the naturalistic course in anxiety disorders using clinical and biological markers: a machine learning approach». En: *Psychological medicine* vol. 52, n.º 1 (2022), págs. 57-67.
- [11] Cherif, N. H. y col. «Physiological data validation of the hexoskin smart textile». En: (2018).

- [12] Cho, C.-H. y col. «Mood prediction of patients with mood disorders by machine learning using passive digital phenotypes based on the circadian rhythm: prospective observational cohort study». En: *Journal of Medical Internet Research* vol. 21, n.º 4 (2019), e11029.
- [13] Constantinides, M. y col. «Personalized versus generic mood prediction models in bipolar disorder». En: *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. 2018, págs. 1700-1707.
- [14] Côté-Allard, U. y col. «Long-Short Ensemble Network for Bipolar Manic-Euthymic State Recognition Based on Wrist-worn Sensors». En: *arXiv preprint arXiv:2107.00710* (2021).
- [15] Čukić, M. y col. «Nonlinear analysis of EEG complexity in episode and remission phase of recurrent depression». En: *International journal of methods in psychiatric research* vol. 29, n.º 2 (2020), e1816.
- [16] Dargél, A. A. y col. «Toi Même, a mobile health platform for measuring bipolar illness activity: protocol for a feasibility study». En: *JMIR Research Protocols* vol. 9, n.º 8 (2020), e18818.
- [17] Devaruppala, M. y Potharaju, V. «ARM7 and IoT based modern health care system using body sensor network». En: *Journal of Engineering Sciences* vol. 11, n.º 2 (2020), págs. 316-320.
- [18] Dominiak, M. y col. «Behavioral and Self-reported Data Collected From Smartphones for the Assessment of Depressive and Manic Symptoms in Patients With Bipolar Disorder: Prospective Observational Study». En: *Journal of Medical Internet Research* vol. 24, n.º 1 (2022), e28647.
- [19] Dong, M. y col. «Prevalence of suicide attempts in bipolar disorder: a systematic review and meta-analysis of observational studies». En: *Epidemiology and psychiatric sciences* vol. 29 (2020).
- [20] Doshi, N. y Auerbach, S. «0898 Sleep Polysomnogram Biomarkers for Identifying Anxiety and Depression». En: *Sleep* vol. 42 (2019), A361.
- [21] Evans, D. L. «Bipolar disorder: diagnostic challenges and treatment considerations». En: *Journal of Clinical Psychiatry* vol. 61 (2000), págs. 26-31.
- [22] Faurholt-Jepsen, M. y col. «Objective smartphone data as a potential diagnostic marker of bipolar disorder». En: *Australian & New Zealand Journal of Psychiatry* vol. 53, n.º 2 (2019), págs. 119-128.
- [23] Faurholt-Jepsen, M. y col. «Smartphone data as an electronic biomarker of illness activity in bipolar disorder». En: *Bipolar disorders* vol. 17, n.º 7 (2015), págs. 715-728.
- [24] Fellendorf, F. T. y col. «Monitoring Sleep Changes via a Smartphone App in Bipolar Disorder: Practical Issues and Validation of a Potential Diagnostic Tool». En: *Frontiers in Psychiatry* vol. 12 (2021), pág. 218.

- 
- [25] Gandhi, V. y Singh, J. «WBSN based safe lifestyle: A case study of heartrate monitoring system». En: *International Journal of Electrical and Computer Engineering* vol. 10, n.º 3 (2020), pág. 2296.
- [26] Garcia-Ceja, E. y col. «Depresjon: a motor activity database of depression episodes in unipolar and bipolar patients». En: *Proceedings of the 9th ACM multimedia systems conference*. 2018, págs. 472-477.
- [27] Gideon, J., Provost, E. M. y McInnis, M. «Mood state prediction from speech of varying acoustic quality for individuals with bipolar disorder». En: *2016 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE. 2016, págs. 2359-2363.
- [28] Gokcinar, N. B. y col. «Optical coherence tomography neurodegenerative findings in patients with bipolar disorder». En: *Asia-Pacific Psychiatry* vol. 12, n.º 4 (2020), e12394.
- [29] Han, K.-M. y col. «Differentiating between bipolar and unipolar depression in functional and structural MRI studies». En: *Progress in Neuro-Psychopharmacology and Biological Psychiatry* vol. 91 (2019), págs. 20-27.
- [30] Healthdirect. *Bipolar disorder*. URL: <https://www.healthdirect.gov.au/bipolar-disorder/> (visitado 20-03-2022).
- [31] Hein, M. y col. «Alterations of neural network organization during REM sleep in women: implication for sex differences in vulnerability to mood disorders». En: *Biology of sex differences* vol. 11, n.º 1 (2020), págs. 1-12.
- [32] Jacobson, N. C., Weingarden, H. y Wilhelm, S. «Digital biomarkers of mood disorders and symptom change». En: *NPJ digital medicine* vol. 2, n.º 1 (2019), págs. 1-3.
- [33] Javelot, H. y col. «Telemonitoring with respect to mood disorders and information and communication technologies: overview and presentation of the PSYCHE project». En: *BioMed research international* vol. 2014 (2014).
- [34] Jiang, C. y col. «Enhancing EEG-based classification of depression patients using spatial information». En: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* vol. 29 (2021), págs. 566-575.
- [35] Kim, H. y col. «Depression prediction by using ecological momentary assessment, actiwatch data, and machine learning: observational study on older adults living alone». En: *JMIR mHealth and uHealth* vol. 7, n.º 10 (2019), e14149.
- [36] Lin, E. y col. «A deep learning approach for predicting antidepressant response in major depression using clinical and genetic biomarkers». En: *Frontiers in psychiatry* vol. 9 (2018), pág. 290.
- [37] Litvin, M. y Maier, M. «How mHealth programmes can treat depression: a randomized controlled trial». En: *Psychreg J Psychol* vol. 3, n.º 2 (2019), págs. 65-78.

- [38] Llamocca, P. y col. «Bip4Cast: Some advances in mood disorders data analysis». En: *Proceedings of the 1st International Alan Turing Conference on Decision Support and Recommender Systems, London*. 2019.
- [39] Llamocca, P., López, V. y Čukić, M. «The Proposition for Bipolar Depression Forecasting Based on Wearable Data Collection». En: *Frontiers in Physiology* (2022), pág. 2333.
- [40] Llamocca, P. y col. «Personalized characterization of emotional states in patients with bipolar disorder». En: *Mathematics* vol. 9, n.º 11 (2021), pág. 1174.
- [41] Lombriser, C. y col. «On-Body Activity Recognition in a Dynamic Sensor Network». En: *Proceedings of the ICST 2nd International Conference on Body Area Networks*. BodyNets '07. Florence, Italy: ICST (Institute for Computer Sciences, Social-Informatics y Telecommunications Engineering), 2007.
- [42] López, V. y col. «Specification of a cad prediction system for bipolar disorder». En: *Uncertainty Modelling in Knowledge Engineering and Decision Making: Proceedings of the 12th International FLINS Conference*. World Scientific. 2016, págs. 162-167.
- [43] Malhi, G. S. y col. «Modelling mood disorders: an ACE solution?» En: *Bipolar disorders* vol. 20 (2018), págs. 4-16.
- [44] Martínez Gómez, A. M. «System to predict bipolar disorder crises analysing massive data». Universidad Complutense de Madrid, 2016.
- [45] Mathur, S. y Sutton, J. «Personalized medicine could transform healthcare». En: *Biomedical reports* vol. 7, n.º 1 (2017), págs. 3-5.
- [46] Mayora, O. y col. «Mobile health systems for bipolar disorder: the relevance of non-functional requirements in monarca project». En: *E-health and telemedicine: Concepts, methodologies, tools, and applications*. IGI Global, 2016, págs. 1395-1405.
- [47] McCloud, T. y col. «Effectiveness of a mobile app intervention for anxiety and depression symptoms in university students: randomized controlled trial». En: *JMIR mHealth and uHealth* vol. 8, n.º 7 (2020), e15418.
- [48] Mehrotra, A. y Musolesi, M. «Designing effective movement digital biomarkers for unobtrusive emotional state mobile monitoring». En: *Proceedings of the 1st Workshop on Digital Biomarkers*. 2017, págs. 3-8.
- [49] Mehta, A. y col. «Acceptability and Effectiveness of Artificial Intelligence Therapy for Anxiety and Depression (Youper): Longitudinal Observational Study». En: *Journal of Medical Internet Research* vol. 23, n.º 6 (2021), e26771.
- [50] Menezes, P. y col. «Use of a mobile phone app to treat depression comorbid with hypertension or diabetes: a pilot study in Brazil and Peru». En: *JMIR mental health* vol. 6, n.º 4 (2019), e11698.

- 
- [51] Mesbah, R. y col. «Anger and cluster B personality traits and the conversion from unipolar depression to bipolar disorder». En: *Depression and anxiety* vol. 38, n.º 6 (2021), págs. 671-681.
- [52] Mikelsons, G. y col. «Towards deep learning models for psychological state prediction using smartphone data: Challenges and opportunities». En: *arXiv preprint arXiv:1711.06350* (2017).
- [53] Momin, M. A. y col. «Smart body monitoring system using IoT and machine learning». En: (2021).
- [54] Murnane, E. L. y col. «Self-monitoring practices, attitudes, and needs of individuals with bipolar disorder: implications for the design of technologies to manage mental health». En: *Journal of the American Medical Informatics Association* vol. 23, n.º 3 (2016), págs. 477-484.
- [55] Mwangi, B. y col. «Individualized prediction and clinical staging of bipolar disorders using neuroanatomical biomarkers». En: *Biological psychiatry: cognitive neuroscience and neuroimaging* vol. 1, n.º 2 (2016), págs. 186-194.
- [56] Ozten, M. y Erol, A. «Impulsivity differences between bipolar and unipolar depression». En: *Indian journal of psychiatry* vol. 61, n.º 2 (2019), pág. 156.
- [57] Passos, I. C. y col. «Identifying a clinical signature of suicidality among patients with mood disorders: A pilot study using a machine learning approach». En: *Journal of affective disorders* vol. 193 (2016), págs. 109-116.
- [58] Patel, R. y col. «Do antidepressants increase the risk of mania and bipolar disorder in people with depression? A retrospective electronic case register cohort study». En: *BMJ open* vol. 5, n.º 12 (2015), e008341.
- [59] Perrotta, G. «Bipolar disorder: definition, differential diagnosis, clinical contexts and therapeutic approaches». En: *J Neuroscience and Neurological Surgery* vol. 5 (2019).
- [60] Portella, P. L., López, V. y Santos, M. «Weighted Dependence of the Day of the Week in Patients with Emotional Disorders: A Mathematical Model». En: *2021 IEEE International Conference on Progress in Informatics and Computing (PIC)*. IEEE. 2021, págs. 420-425.
- [61] Portella, P. L. y col. «Data source analysis in mood disorder research». En: *XVIII Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA 2018) 23-26 de octubre de 2018 Granada, España*. Asociación Española para la Inteligencia Artificial (AEPIA). 2018, págs. 893-898.
- [62] Powell, A. y col. «Wearable technology to assess bradykinesia and immobility in patients with severe depression undergoing electroconvulsive therapy: a pilot study». En: *Journal of psychiatric research* vol. 130 (2020), págs. 75-81.
- [63] Qian, K. y col. «Teaching machines to know your depressive state: On physical activity in health and major depressive disorder». En: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. 2019, págs. 3592-3595.

- [64] Reinertsen, E. y Clifford, G. D. «A review of physiological and behavioral monitoring with digital sensors for neuropsychiatric illnesses». En: *Physiological measurement* vol. 39, n.º 5 (2018), 05TR01.
- [65] Sahoo, A. K., Pradhan, C. y Das, H. «Performance evaluation of different machine learning methods and deep-learning based convolutional neural network for health decision making». En: *Nature inspired computing for data science*. Springer, 2020, págs. 201-212.
- [66] Saloni Dattani, H. R. y Roser, M. «Mental Health». En: *Our World in Data* (2021). <https://ourworldindata.org/mental-health>.
- [67] Sankar, A. y col. «Graph theory analysis of whole brain functional connectivity to assess disturbances associated with suicide attempts in bipolar disorder». En: *Translational psychiatry* vol. 12, n.º 1 (2022), págs. 1-10.
- [68] Sawyer, A. y col. «The effectiveness of an app-based nurse-moderated program for new mothers with depression and parenting problems (eMums Plus): pragmatic randomized controlled trial». En: *Journal of medical Internet research* vol. 21, n.º 6 (2019), e13689.
- [69] Sawyer, A. C. y col. «Evaluating the effectiveness of an app-based nurse-moderated program for new mothers with depression and parenting problems (eMums Plus): protocol for a pragmatic randomized controlled trial». En: *JMIR Research Protocols* vol. 8, n.º 1 (2019), e11549.
- [70] Shafiee, N. S. M. y Mutalib, S. «Prediction of mental health problems among higher education student using machine learning». En: *International Journal of Education and Management Engineering (IJEME)* vol. 10, n.º 6 (2020), págs. 1-9.
- [71] Shen, H. y col. «Analysis of misdiagnosis of bipolar disorder in an outpatient setting». En: *Shanghai archives of psychiatry* vol. 30, n.º 2 (2018), pág. 93.
- [72] Singh, T. y Rajput, M. «Misdiagnosis of bipolar disorder». En: *Psychiatry (Edgmont)* vol. 3, n.º 10 (2006), pág. 57.
- [73] Stephen Soreff, M. *Bipolar disorder*. Feb. de 2022.
- [74] Tazawa, Y. y col. «Actigraphy for evaluation of mood disorders: A systematic review and meta-analysis». En: *Journal of affective disorders* vol. 253 (2019), págs. 257-269.
- [75] Tazawa, Y. y col. «Evaluating depression with multimodal wristband-type wearable device: screening and assessing patient severity utilizing machine-learning». En: *Heliyon* vol. 6, n.º 2 (2020), e03274.
- [76] Tomasik, J. y col. «A machine learning algorithm to differentiate bipolar disorder from major depressive disorder using an online mental health questionnaire and blood biomarker data». En: *Translational psychiatry* vol. 11, n.º 1 (2021), págs. 1-12.

- 
- [77] Ulusoy, M. O., Işık-Ulusoy, S. y Kıvanç, S. A. «Evaluation of dry eye disease in newly diagnosed anxiety and depression patients using anterior segment optical coherence tomography». En: *Eye and Vision* vol. 6, n.º 1 (2019), págs. 1-6.
- [78] Valenza, G. y Scilingo, E. P. «Autonomic nervous system dynamics for mood and emotional-state recognition: Significant advances in data acquisition, signal processing and classification». En: (2013).
- [79] Valenza, G. y col. «Predicting mood changes in bipolar disorder through heartbeat nonlinear dynamics». En: *IEEE journal of biomedical and health informatics* vol. 20, n.º 4 (2016), págs. 1034-1043.
- [80] Vieta, E. y col. «Early intervention in bipolar disorder». En: *American Journal of Psychiatry* vol. 175, n.º 5 (2018), págs. 411-426.
- [81] Viswanathan, S. «E-Textile and its Applications». En: *International Journal of Engineering Science* vol. 20174 (2019).
- [82] Wang, R. y col. «StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones». En: *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*. 2014, págs. 3-14.
- [83] Werner, F. «Autonomic Function in Individuals With Symptoms of Functional Disorders». Tesis doct. California State University, Northridge, 2019.
- [84] Williams, Z. J., Everaert, J. y Gotham, K. O. «Measuring depression in autistic adults: Psychometric validation of the Beck Depression Inventory-II». En: *Assessment* vol. 28, n.º 3 (2021), págs. 858-876.
- [85] Wylich-Muxoll, B. v. *The Monarca system for bipolar disorder - the foundation of monsense*. Feb. de 2020.
- [86] Zhang, W. y col. «Machine learning models for the prediction of postpartum depression: application and comparison based on a cohort study». En: *JMIR medical informatics* vol. 8, n.º 4 (2020), e15516.



## Capítulo 6

# Lista de publicaciones

### Publicación I

Llamocca, P., Junestrand, A., Čukić M., Urgelés, D. y López, V. «Data source analysis in mood disorder research». En: *XVIII Conferencia de la Asociación Española para la Inteligencia Artificial 23-26 de octubre de 2018 Granada, España*. (2018), pp. 893–898.

### Publicación II

Llamocca, P., Urgelés, D., Čukić M y López, V. «Bip4Cast: Some advances in mood disorders data analysis». En: *Proceedings of the 1st International Alan Turing Conference on Decision Support and Recommender Systems, London*. (2019, Noviembre), pp. 5–10.

### Publicación III

Llamocca, P., López, V., Santos M. y Čukić M. «Personalized characterization of emotional states in patients with bipolar disorder». En: *Mathematics*. Vol. 9, no. 11:1174, (2021a), DOI: 10.3390/math9111174, ISSN. 2227-7390.

### Publicación IV

Llamocca, P., López, V. y Santos M. «Weighted dependence of the day of the week in patients with emotional disorders. A mathematical model». En: *2021 IEEE International Conference on Progress in Informatics and Computing (PIC)*. (2021b), pp. 420–425. DOI: 10.1109/PIC53636.2021.9687019.

### Publicación V

Llamocca, P., López, V. y Čukić M. «The proposition for bipolar depression forecasting based on wearables data collection». En: *Frontiers in Physiology*. Vol. 12, (2022), DOI: 10.3389/fphys.2021.777137, ISSN. 1664-042X.



# Publicaciones



Publicación I

## **Data source analysis in mood disorder research**

**Llamocca, P., Junestrand, A., Čukić M., Urgelés, D., López, V.**

Publicado en *XVIII Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA 2018)*, Granada-España, Octubre 2018, pp. 893–898.



# *Data source analysis in mood disorder research*

*Pavél Llamocca Portella*  
Universidad Complutense de Madrid  
[pavel.llamocca@hotmail.com](mailto:pavel.llamocca@hotmail.com)

*Milena Čukić*  
University of Belgrade  
[milena.cukic@gmail.com](mailto:milena.cukic@gmail.com)

*Victoria López López*  
Universidad Complutense de Madrid  
[vlopezlo@ucm.es](mailto:vlopezlo@ucm.es)

*Axel Junestrand*  
Universidad Complutense de Madrid  
[axel.junestrand@hotmail.com](mailto:axel.junestrand@hotmail.com)

*Diego Urgelés*  
Hospital Ntra. Sra. De la Paz  
[diego.urgeles@gmail.com](mailto:diego.urgeles@gmail.com)

**Abstract**—Mood disorders have been a relevant topic for the last few years. Nowadays, there are projects in the mental health area which are supported by technological devices that improve the efficiency of treatments by effortlessly allowing the gathering of biological and psychological indicators from patients. One of the goals of this document is to describe the most common methods for collecting most of those indicators and to study which of them can be applied to the Bip4Cast project. The purpose of this article is to analyze the sources of information that have been used successfully in the study of emotional disorders as well as alternative sources of information from the monitoring of movement and sounds in the patient's environment. This article shows the results of the analysis of traditional information sources. The results show a lack of precision in the data on fundamental variables such as sleep quality and motor activity. Therefore, the study demonstrates the need to include new sources of information to increase the quality of the data before applying crisis prediction algorithms. The need to monitor the sleep and movement of patients in order to achieve a sufficient quality in the source data from the evolutionary analysis of patients is concluded.

**Keywords**—*bipolar disorder; mood disorder; data gathering; machine learning; data analysis.*

## I. INTRODUCTION

The quality of treatments in mood disorders has acquired a high attention for researchers in the mental health field. However, despite current efforts, there is still plenty of room for improvement. Many practitioners agree that knowing how patients react to treatments in advance and predicting when their mood could vary significantly are two of the most important issues to solve in order to ensure the quality of new treatments.

The proposal of the Bip4Cast project is to keep using the current monitoring of personal sessions between patients and psychiatrists, but also to add new data sources and their analysis to improve the prediction of Bipolar Disorder crises in patients. The main idea is to get advantage of the new developments in data gathering, data cleaning, and Machine Learning to monitor a set of patients and make a new approach

with these data. The patients are encouraged to follow certain methods for gathering psychological and biological indicators during a particular period of time. The goal is to analyze the data gathered in order to find some common patterns that could trigger a crisis. For the process of pattern detection, some Machine Learning tools and mathematical models are being used.

The goal of this document is to cover a discussion about some methods for gathering indicators and the feasibility of their usage in this project. Apart from this introduction, this document includes the following sections: section 2 presents the state of the art related researches in mood changes and patient monitoring. In section 3, several methods for gathering psychological and biological parameters are described. Section 4 covers the preliminary analytics on the Bip4Cast data sets and, finally, section 5 includes the conclusions.

## II. STATE OF THE ART

Several studies about Bipolar Disorder state that a relationship exists between the different behaviors of the patients before the occurrence of a crisis [1-2]. For example, during a manic or a depressive crisis, some of these studies agree that sleeping rates are very important indicators. Vocal features as well as the rate of speech are other important indicators and there are some studies stating that the pitch is lower in a depressed state [3]. Also, parameters like the time of exposure to dark or sunny places and the physical activity are considered.

In [4], the author introduces a mobile health system using several sensors for mood detection. In [5], the author presents a research that includes Machine Learning models in a mobile application in order to estimate the mood in depressive patients. However, no objective psychological or psychiatric markers are considered due to the recording of the data being done manually by patients. Furthermore, there is an interesting study which describes the use of electroencephalography for the gathering of brain signals. It also uses non-linear features like Higuchi's Fractal Dimension and Sample Entropy to feed different Machine Learning methods [6]. In [7], a mobile

application is presented for supporting the treatments of patients with Bipolar Disorder. Its key is to compare objective and subjective data. It records objective data using some features given by mobile phones like accelerometers and phone call rates. This information is used for predicting trends in the mood of the patients. However, the focus of this application is to record subjective data using a self-reporting approach.

At the present time, there are some projects within the scope of mental health which have similar approaches and try to obtain certain markers from which a common pattern can be inferred to help with the treatments. One of these projects is PSYCHE [8], whose main idea is for the patients to use a special garment made of a proper material, of which the main goal is to collect parameters from the patient in his/her daily life. The outcomes of PSYCHE are positive. However, patients said that the main inconvenient was to use the same clothes all the time, which implicates a high discomfort for the patient and therefore its non-use. Another really interesting project is MONARCA [9]. It emphasizes the use of mobile phones for the electronic monitoring of patients. The number of parameters that can be obtained through the use of a mobile phone is really high, but nevertheless, none of them are physiological parameters. Furthermore, after 3 years of activity with this application, an analysis of some non-functional requirements for the treatment of patients with Bipolar Disorder concluded that, for new developments, some details have to be taken into consideration in order to improve the ease of use, e.g.: ensuring that the patients have a data plan for their 3G connection, the need for teaching the clinicians how to operate the system, and the overheating of the smart phone from the use of an application that requires both GPS and Bluetooth.

Since few years ago, Body Sensor Networks (BSN) have made an appearance, which are a branch of wireless sensor networks (WSNs) that conform one of the core technologies of IoT developments in the healthcare system [10]. Its purpose is to provide an integrated hardware and software platform which facilitates the future development of pervasive monitoring systems. BSN allow the monitoring of patients by using a collection of tiny-powered and lightweight wireless sensor nodes. These nodes are placed on the skin and sometimes integrated with different garments, so that the patient's health-related data can be collected and transferred to the healthcare staff in real time. However, the development of this new technology in healthcare applications without considering security makes patient privacy vulnerable. For this reason, several research projects are currently being carried out to try to cover this vulnerability [11].

### III. DATA SOURCE ANALYSIS

All the research that is currently being conducted suggests a wide variety of indicators for taking mood disorder treatments into account. The indicators themselves and the way in which they are collected are strongly related. In this section, several data sources for gathering those indicators are described. Depending on the indicator they gather or the type of device, they are classified into 6 groups as shown below.

**1. Smart wristband/smart band.** These devices include a set of sensors which measure daily activity by means of

accelerometers. They create variables such as an activity tracker (resting, moving or sleeping), a pedometer (steps taken and distance traveled), the calories burned, a sleep monitor (awake, slight and deep sleeping), the heart rate and the blood oxygen level. The most relevant variable measured from this type of device is the sleep indicator. Almost all researchers agree that sleep quality is the best indicator in Bipolar Disorder treatments. The CHOICE [12] study states that lower levels of depression are correlated with improvements in insomnia treatments, and by the other hand, high levels of mania are correlated with less need for sleep. Furthermore, a pilot randomized controlled trial demonstrated that sleep disturbance appears to be an important pathway contributing to Bipolar Disorder [13]. These data can easily be gathered from popular apps. Fig. 1 (a) shows one of the monitors used in the Bip4Cast project.

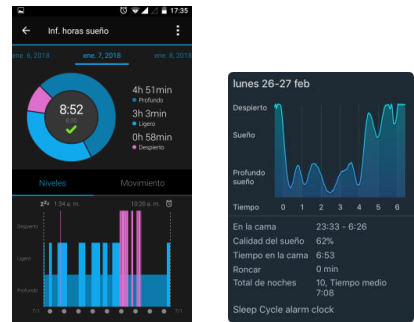


Fig. 1. Two sleep monitors in Bip4Cast (a- Garmin Vivofit3; b- Sleep Cycle for iPhone)

**2. Medical bands.** They allow measuring more parameters because they usually include more hardware and better features like atmospheric pressure (barometer), GPS location and magnetometer. There is some research which links episodes to disturbances in circadian rhythms and lifestyle regularity. Those indicators can be collected through these devices using the activity tracker or gyroscope. Furthermore, this research suggests that methods for tracking behavior, nutrition, blood pressure and lipid profile as well as physical/social activity and sleep-awake routines may improve treatments. In the Bip4Cast project we are using GENEActiv v.1.2 for a total of 25 patients, (see [14] for more details about their use in Bip4Cast).

**3. Mobile Sensors.** In this group, any other kind of sensors that can be worn by the user on any other part of the body is included, e.g. for wearing on the leg, ActivPAL is a kind of device used to investigate the correlation between physical behaviors and chronic disease [15]. For using as a necklace, LeafUrban is an option (there are some versions for wearing on the wrist or attached to the clothes). It is a device designed for women and what it makes different from other devices is the tracking of the menstrual cycle, the fertility and the breathing [16]. For wearing on the head, ELF Emmitt is a headband that helps the user improve the state of both mind and body by using pulsed electromagnetic stimulation (PEMS) [17]. Relevant variables include skin and breath changes,

electrocardiogram and respirogram data, stress level and menstrual cycle among others.

**4. Sleep Activity Recording Devices.** In this group, any device specialized in recording sounds and activities during sleep is included. There are hundreds of mobile applications that record sounds during sleep. One of the main objectives is to detect snoring, for which four of the most popular applications at the moment are SleepGenius, SleepCycle (see Fig. 1 (b)), SleepBot and SleepTime. However, there are other kinds of devices with different non-invasive designs, e.g. devices attached to the mattress which can track sounds as well as heart rate, breathing, movement, etc. The reason for using these methods is to improve sleeping conditions. Almost all of these methods have one parameter in common: “breathing”, which allows the detection of snoring. Habitual snoring is a prevalent condition that is not only a marker for Obstructive Sleep Apnea (OSA) but can also lead to vascular risks [18]. Some researchers have found a relationship between OSA and Major Depressive Disorder/Bipolar Disorder.

**5. Forms and Questionnaires.** This group contains any method which uses a questionnaire or a form for the self-reporting of mood. In current literature, these methods were designed by psychiatrists and are presented as scales. There are several scales for detecting the risk of a euphoria episode outbreak: Altman Self-Rating Mania Scale (ASRM) [19], the Clinician-Administered Rating Scale for Mania (CARS-M), the Internal State Scale (ISS), the Self-Report Manic Inventory (SRMI), etc. For depression episodes, there are several scales, like the Patient Health Questionnaire (PHQ-9) [20]. All of them consist in questionnaires which can be performed by patients. This presents the opportunity of developing digital forms based on these patients in order to facilitate their use.

Scales for detecting the risk of euphoria or mania episodes, like the Young Mania Rating Scale (YMRS) and the Bech-Rafaelsen Mania Scale (MAS), or the Hamilton Depression Rating Scale (HDRS), which detects the risk of depression, are not included because they are performed by the clinician (however, for the scope of this project, these scales are included in normal monitoring sessions). All of these questionnaires collect variables from which it is possible to measure the presence and severity of mania, depression, affective, psychological and somatic symptoms. Fig. 2 shows the interface of an application developed for collecting these data. All the details about this work are in [21].

**6. Mobile Apps / Time Consumption.** This group includes mobile applications that support BPD treatments and/or record smart phone use. For the aim of this project, these mobile applications were classified into two subclasses: the first one, named Bipolar Disorder Apps (BPDA) in this document, includes any applications that have been developed for supporting the treatment itself, and the second one, named Time Consumption Apps (TCA) in this document, includes any application that is able to measure the time spent using a specific mobile application.

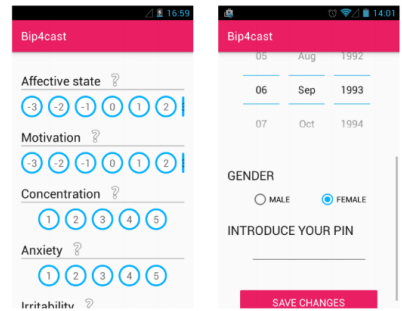


Fig. 2. First version of the app for collecting daily personal data

**7. Conventional methods.** Finally, patients are also being assessed periodically through interviews. This evaluation is done by psychiatrists in medical centers. For the patient, this does not imply any kind of alteration in the current treatment. However, the procedure will need the psychiatrist to send the collected data from those sessions to the data server. It is important to mention that in this phase, the Young Mania Rating Scale (YMRS) and the Hamilton Depression Rating Scale (HDRS) are included for detecting mania or depression episodes.

Also, psychiatrists can take advantage of these interviews for downloading the data recorded from wristbands and medical bands in order to later send them to a dedicated server (just in case these devices are not able to send the recorded data automatically).

#### IV. THE BIP4CAST PROJECT

Patients with Bipolar Disorder are characterized by a behavior which is difficult to predict. There is great deal of information which can be retrieved from biological, physiological and physical signals in order to detect episodes. Knowing which variables are correlated and which features or parameters are important is essential to build a model that will successfully predict the target of a study. The aim of this study is to investigate which features have the highest importance in health. In order to achieve this, Machine Learning algorithms and techniques are used for feature ranking.

The data used for this project is anonymized patient data gathered by psychiatrists at Clínica Nuestra Señora de la Paz in Madrid. All the data were available in an Excel file with different sheets. Even though 25 patients are already wearing a medical band (GENEActiv 1.2) and we have developed an application for gathering their daily activity, for this study most of the data have been gathered in a supervised manner during medical appointments with four different patients that suffer from Bipolar Disorder. The goal for the future is that these data are both recorded by the psychiatrists in appointments and with the help of mobile applications. This way, the patients can actively participate in their own diagnosis. The data consist of 4 data sets: Episodes, which represents different episode periods in the patients (depression/mania) from a total of four patients; YMRS data set, which contains Young Mania Rating

Scale [22] data (to assess mania symptoms) from a total of 48 days; HDRS data set, which contains Hamilton Depression Rating Scale [23] data (for depression), also from a total of 48 days; Interview data set, which contains 728 registers about physical and psychological items, the latter including variables like anxiety, irritability or concentration problems, and the former including more objective data, as could be the number of cigarettes smoked by the patient or the time in which the patient woke up or went to bed. The last data set used in the study is Interventions. It includes data about all the medical interventions that different doctors have had with the patients, in a total of 92 registers. For the gathering of data included in the Interview data set, a mobile application [21] has been developed, which patients can use daily to store quantitative data (number of cigarettes, menstruation, etc.) and qualitative data (feeling of stress, anxiety, etc.). In the project, we have also included studies with data from a medical bracelet (GENEActiv 2.1 for 25 patients) and an application for recording night sounds. The data collected by these last two exercises will be included in subsequent studies.

The programming language used in this project is Python 2.7, which has a lot of libraries that make data cleaning and visualization less complicated, as well as applying Machine Learning algorithms. The environment used is Jupyter Notebook [24]. Scikit-learn [25] is the Machine Learning library of choice for this project because it includes preprocessing and cross-validation tools as well as all the known baseline Machine Learning algorithms. This project is shared in a public GitHub repository, which can be found at [26].

#### A. Data Cleaning

The first step of the project consisted in the data cleaning which included the gathering of the data that we would be working with. In order to gather the data, we exported each sheet, from the Excel file that was given to us by the hospital, to csv format. The initial Excel file was divided into five different sheets: Episodes, YMRS, HDRS, Interview Data Set (IDS) and Interventions. In order to export them to a format readable by Pandas [27], we saved each sheet as CSV UTF-8 in Microsoft Excel. Some other improvement was done in relation to data cleaning: filling the empty values, converting them from Float to Integer and data type revision.

#### B. Exploratory Data Analysis

After the data cleaning, we performed an Exploratory Data Analysis, in order to visualize how the data behaved. We used histograms, heatmaps and scatterplots in this part. For instance, the YMRS data set correlation heatmap showed that aggressiveness and verbal expression were correlated. This could mean that if the patient talks a lot (excessive speech rate), this behavior would probably be accompanied by excessive energy or hyperactivity (Disruptive-Aggressive Behavior). The scatterplot matrix from the YMRS data showed that hyperactivity and irritability have a similar distribution as well as a correlation between verbal expression and euphoria. The HDRS data set analysis showed a similar distribution between suicide and precocious insomnia (difficulty of sleep early in the night). The HDRS data set correlation heatmap showed a high correlation between depressed mood and work: the less a

patient is willing to work or do other activities, the more depressed he or she will probably feel. At this stage of the research, the best data set regarding both size and accuracy was the Interview Data Set (IDS). During its analysis, we found a clear linear relationship between the variables mood and motivation. The Intervention data set presented a lack of correlation between the level of relief in a patient and the GAF (Global Assessment of Functioning). A good summary of all these results can be read at [26].

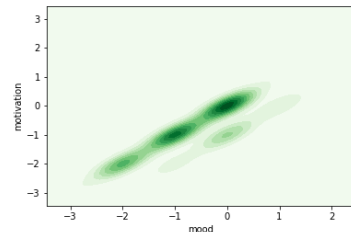


Fig. 3. 2D kernel density plot of mood and motivation in Interview data set

#### C. Data Combination

The goal of this phase was to find data set combinations that had enough data for the algorithms to process so that we could later see which data sets returned the highest accuracy. In order to get the combinations right, we defined a function that obtained the date of each entry and compared it with the different episodes of depression and mania in the Episode data set, which was the target of the prediction. For the entries or rows that were not recorded in the Episode data set, we assumed that the patient was in a euthymic state. This way, we got three possible states that a patient could be in: Depression (D), Mania (M) and Euthymic (N). For instance, with the HDRS and Episode combination we could see that when the patients had a depression episode, the value of depressed mood was much higher, almost always between 2 and 3, which meant that they either spontaneously reported feeling depressed or they communicated feeling depressed in a non-verbal way, judging by the rating items from the Hamilton Depression Rating Scale (HDRS). We could also see that when the patients were in a depression state (because of the predominance of green points on higher values of the work axis, where green represented patients in a state of depression), they started feeling loss of interest in activities they usually performed or there was a decrease in the time spent on work and other activities, which made perfect sense according to this rating scale.

#### D. Application of the Algorithms

The data sets on which we tested the algorithms were: YMRS (Young Mania Rating Scale data), HDRS (Hamilton Depression Rating Scale data), Interviews (interview data, IDS), Interventions (intervention data), YMRS-HDRS (combination of the YMRS and HDRS data) and Interventions-Interventions (combination of the Interview and Intervention data). Fig.4 shows the diagram of the process followed during this study.

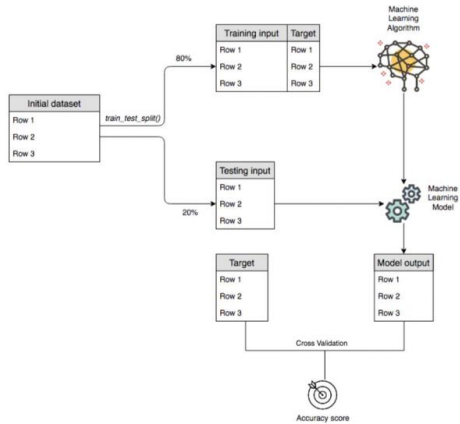


Fig. 4. Diagram of the Machine Learning algorithm application process

The algorithms that we used for this part were: Decision Tree [27], Random Forest [28], Support Vector Machines [29] and Logistic Regression [30]. The reason why these algorithms have been chosen for this project is explained in [26], in the section belonging to each algorithm.

The fact that these algorithms have been applied in this project does not mean that they are the best option for the classification of Bipolar Disorder states, but rather that they are the most suitable ones given the amount of data and the number of features used for the project. In future studies that make use of this project, if the data sets are larger it could be interesting to apply other algorithms too, like the Naïve Bayes [31] algorithm or any kind of Boosting algorithm [32], as to see how they perform on this particular classification problem.

Before applying each algorithm, as is necessary for every classification problem, the original data needed to be split into training and testing sets. Later, the training data would be used to train the prediction models and the testing data would be used to compare the output of the model with the real targets by cross-validation, a technique presented first by M. Stone in 1974 [33], that is used widely in Machine Learning for algorithm performance comparison.

The testing set is used for obtaining the accuracy of the model, as mentioned above, which is done by comparing the output obtained from the testing input and the real output of the testing set. In order to divide the original data sets into training and testing sets, we used the `train_test_split()` function from the scikit-learn library [25], where the test size represents the percentage of the data that is used for the testing set. After the algorithms were applied, the `cross_val_score()` function, which is also included in the scikit-learn library, was called in order to evaluate the score by k-fold cross-validation [33].

The best way to compare the accuracies obtained with the different algorithms on all the data sets was to make an algorithm performance matrix, which is shown in Table 1. This matrix showed that, in average, the data set that returned the best prediction accuracy (69%) with the algorithms was the Interview data set, as seen on the right column. The algorithm

that performed the best, also in average (62%), was the Logistic Regression algorithm, as seen on the column farthest down.

Even though the algorithm that had the best accuracy average was the Logistic Regression, we stated that the Random Forest algorithm made the most accurate predictions in the sense that they were very reasonable given the behavior of the patients, which we tested with randomized data. These predictions made possible the implementation of a small program with the Random Forest classifier that we obtained, and which can be seen in [26].

TABLE I. ACCURACIES WITH DIFFERENT ALGORITHMS

Algorithms/Datasets	Decision Tree	Random Forest	SVM	Logistic Regression	Average
YMRS	36%	38%	38%	76%	47%
HDRS	78%	55%	36%	62%	58%
Interviews	70%	72%	68%	67%	69%
Interventions	44%	65%	57%	51%	54%
YMRS-HDRS	63%	63%	70%	33%	57%
Interviews-Interventions	67%	17%	42%	83%	52%
Average	60%	52%	52%	62%	

## V. CONCLUSIONS AND FUTURE WORK

Having a deep understanding of the data is essential in any Machine Learning project focused on a branch of medical science like psychiatry, where knowing which behaviors are normal and abnormal in the patients can help us create much more precise prediction models. The amount of data used and the nature of the data source are very important factors because with a larger suitable amount of data we will be able to get prediction accuracies with a much higher level of confidence. In the same way that understanding the data is important, having a deep perception of the theory behind each algorithm used, as well as their many implementations, is crucial in order to get the models to perform in the best possible way.

In this project, several groups of data collected in a supervised way have been analyzed and a set of Machine Learning algorithms has been applied. The results allow us to make decisions about the new sources of relevant information to be incorporated in consequent studies. It is concluded that the data from sleep and daily activity, measured both by movement and sounds, are relevant for improving the prediction of a crisis in patients with Bipolar Disorder. Therefore, a future project that includes group 1 bracelets instead of the current medical wristbands is proposed, because the latter are too expensive and invasive, and the development of a new mobile application that, in addition to the daily data, includes sensitization data and sounds. Future work will also analyze the EEG data collected during supervised monitoring for the purpose of performing a comparative analysis. The implementation with Jupyter will also allow us to perform the same studies on larger databases when the number of patients in the experiment is higher.

The most immediate use of the results obtained in this project would be to train the same algorithms used but with larger amounts of data, in order to see if they perform in a similar way. Gathering objective data from devices like phones or wristbands is something that can be accomplished quite easily according to the work already done in this sense. The

goal of this task would be to compare the performance of different algorithms on the objective data gathered from these devices with the performance results obtained on the subjective data used in this project.

As for other more indirect applications of the results obtained during this project, the implementation of a drug recommending system for patients with Bipolar Disorder could be made by predicting the states in which the patients are during a certain period of time. These predictions could be stored in a database which also contains the medicine that these patients have been prescribed with during the same period of time, thus providing the possibility of seeing how each patient reacts to the different types of drugs used.

## REFERENCES

- [1] P. Vanello et al., "Speech analysis for mood state characterization in bipolar patients," Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, pp. 2104-2107, 2012
- [2] T. Beiwinkel, et al. "Using Smartphones to Monitor Bipolar Disorder Symptoms: A Pilot Study". Eysenbach G, ed. JMIR Mental Health. 2016.
- [3] MN. Burns, et al., "Harnessing Context Sensing to Develop a Mobile Intervention for Depression". J Med Internet Res 2011;13(3):e55.
- [4] M. Cukic, et al. "EEG machine learning with Higuchi fractal dimension and Sample Entropy as features for successful detection of depression" CoRR abs/1803.05985 (2018): n. pag..
- [5] M. Frost, A. Doryab, M. Faurholt-Jepsen, L. Vedel Kessing, and J. E. Bardram. "Supporting disease insight through data analysis: refinements of the monarca self-assessment system". In Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing (UbiComp '13). ACM, New York, NY, USA, pp. 133-142, September-2013
- [6] H. Javelot, et al., "Telemonitoring with respect to Mood Disorders and Information and Communication Technologies: Overview and Presentation of the PSYCHE Project", BioMed Research International, vol. 2014, Article ID 104658, 12 pages, 2014.
- [7] O. Mayora, et. al., "Mobile Health Systems for Bipolar Disorder: The relevance of Non-Functional Requirements in MONARCA Project". IGI International Journal of Handheld Computing Research (IJHCR). 10.4018/978-1-4666-8756-1.ch070.
- [8] Z. Guan, T. Yang, X. Du and M. Guizani, "Secure data access for wireless body sensor networks," 2016 IEEE Wireless Communications and Networking Conference, Doha, pp. 1-6. 2016.
- [9] P. Gope and T. Hwang, "BSN-Care: A Secure IoT-Based Modern Healthcare System Using Body Sensor Network," in IEEE Sensors Journal, vol. 16, no. 5, pp. 1368-1376, March, 2016.
- [10] L. Sylvia, et. al., "Sleep disturbance may impact treatment outcome in bipolar disorder: A preliminary investigation in the context of a large comparative effectiveness trial", Journal of Affective Disorders, Volume 225, 2018, Pages 563-568, ISSN 0165-0327.
- [11] AG. Harvey, et. al., "Treating Insomnia Improves Mood State, Sleep, and Functioning in Bipolar Disorder: A Pilot Randomized Controlled Trial". Journal of consulting and clinical psychology. 83. 10.1037/a0038655.
- [12] J. Anchiraco, "Diseño de una Arquitectura Big Data para la Predicción de Crisis en el Trastorno Bipolar", Trabajo Fin de Master en Ingeniería Informática, ePrints-UCM, Madrid, 2016.
- [13] ActivePAL (2018). Retrieved from <http://www.palt.com/>.
- [14] Bellabeat – Urban Collection, Health Trackers (2018). Retrieved from <https://webshop.bellabeat.com/pages/leaf-urban>
- [15] ELF Emmit Device from NewMed (2018). Retrieved from <https://www.news-medical.net/ELF-Emmit-Device-from-NewMed>
- [16] H. Nakano, et al. "Monitoring Sound To Quantify Snoring and Sleep Apnea Severity Using a Smartphone: Proof of Concept." Journal of Clinical Sleep Medicine : JCSM : Official Publication of the American Academy of Sleep Medicine 10.1 (2014): pp. 73–78, PMC. Web. 12 June 2018.
- [17] The Altman Self-Rating Mania Scale (2018). Retrieved from <https://psychology-tools.com/altman-self-rating-mania-scale/>
- [18] The Patient Health Questionnaire (PHQ-9) (2018). Retrieved from <https://patient.info/doctor/patient-health-questionnaire-phq-9>
- [19] A. Martínez, "Introduction to Big Data and First Steps in a Big Data Project", Trabajo Fin de Doble Grado Matemáticas-Informática, ePrints-UCM, Madrid, 2016.
- [20] R.C. Young et al., "A rating scale for mania: reliability, validity and sensitivity", British Journal of Psychiatry, 1978
- [21] M. Hamilton, "A rating scale for depression", Journal of Neurology, Neurosurgery, and Psychiatry, 1960.
- [22] F. Perez and B. E. Granger, "IPython: A System for Interactive Scientific Computing," in Computing in Science & Engineering, vol. 9, no. 3, pp. 21-29, May-June 2007.
- [23] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python", Journal of Machine Learning Research 12 (2011), pp. 2825-283. November, 2011.
- [24] A. Junestrand, "Application of Machine Learning Algorithms for Bipolar Disorder Crisis Prediction", Trabajo Fin de Grado Ingeniería Informática, ePrints-UCM, Madrid, 2018.
- [25] W. McKinney, "Data Structures for Statistical Computing in Python", Proceedings of the 9th Python in Science Conference, 2010.
- [26] T. Hastie, R. Tibshirani and J. Friedman, "The Elements of Statistical Learning", Springer Series in Statistics, 2009.
- [27] Garima, H. Gulati and P. K. Singh, "Clustering techniques in data mining: A comparison", New Delhi: 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), 2015.
- [28] J. R. Quinlan, "Induction of Decision Trees", Hingham, Massachusetts: Kluwer Academic Publishers, 1986.
- [29] L. Breiman, "Random Forests", Berkeley, California: Statistics Department Technical Report, 2001.
- [30] C. W. Hsu, C. C. Chang and C. J. Lin, "A Practical Guide to Support Vector Classification", Taipei: National Taiwan University, 2003.
- [31] C. Y. J. Peng, K. L. Lee and G. M. Ingersoll, "An Introduction to Logistic Regression Analysis and Reporting", Bloomington, Indiana: The Journal of Educational Research, 2002.
- [32] T. Patil and S. Sherekar, "Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification", International Journal Of Computer Science And Applications, 2013.
- [33] P. Bühlmann and T. Hothorn, "Boosting Algorithms: Regularization, Prediction and Model Fitting", Zürich: Seminar für Statistik, ETH Zürich, 2007.
- [34] M. Stone, "Cross-validatory choice and assessment of statistical predictions", London: Royal Statistical Society, 1974.

Publicación II

## **Bip4Cast: Some advances in mood disorders data analysis**

**Llamocca, P., Urgelés, D., Čukić M, López, V**

Publicado en *Proceedings of the 1st International Alan Turing Conference on Decision Support and Recommender Systems*, Londres-Reino Unido, Noviembre 2019, pp. 5–10.





# Bip4Cast: Some advances in mood disorders data analysis

P Llamocca<sup>1</sup>, D Urgelés<sup>2</sup>, M Cukic<sup>3</sup>, V Lopez<sup>4</sup>

<sup>1,4</sup>Complutense University, Madrid, SPAIN

<sup>2</sup>Hospital Ntra. Sra. de la Paz, Madrid, SPAIN

<sup>3</sup>3EGA, Amsterdam, NEDERLAND

[1pavellam@ucm.es](mailto:pavellam@ucm.es), [2diego.urqeles@sjd.es](mailto:diego.urqeles@sjd.es), [3cukic@3ega.nl](mailto:cukic@3ega.nl), [4vlopezlo@ucm.es](mailto:vlopezlo@ucm.es)

<sup>1,4</sup>[www.ucm.es](http://www.ucm.es), <sup>2</sup>[www.sjd.es](http://www.sjd.es), <sup>3</sup>[www.3ega.nl](http://www.3ega.nl)

## ABSTRACT

Mood disorders have been a relevant topic for the last decade. According to the World Health Organization, the cost of mood disorders and anxiety in the EU is about €170 billion per year. Bip4Cast is a project aiming at crisis prediction for patients with bipolar disorder. Traditional indicators like Hamilton and Yang are insufficient to predict and avoid crises. This study adds up to 149 new variables from data gathered from different sources as wearable devices and smartphones. The analytics include correlation between all the accelerometer variables and linear regression between variables that come from different sources. The results show the existence of a relationship between biological, psychological, physical indicators with the appearance of a crisis of depression or mania. These relations are the base of the predictive analytics that clinicians need in order to make better decisions on the future treatment plans.

## 1. INTRODUCTION

Bipolar disorder and depression are chronic and severe mental disorders and a major mental health problem in Europe. The Regional Office for Europe of the World Health Organization, in 2019, reports that 25% of the population suffer from depression or anxiety and that neuropsychiatric disorders account for 26% of the burden of disease in European Union countries. They have a lifetime prevalence rate of 0.6% (0.4%-1.4%) and a high suicide index. These mental disorders can be monitored by mobile devices. They can be useful to improve the efficiency of treatments because they allow the gathering of physical, biological and psychological indicators from patients. The importance of these indicators is discussed in recent publication (Vanello, 2012). As a continuation of this, the project Bip4Cast ([bip4cast.org](http://bip4cast.org)) focus on discovering the relationship among the indicators provided by the data gathered and developing a CAD tool (Lopez, 2016) to support the clinical decisions. These studies seek to find common patterns that could trigger a crisis of mania or depression. Bip4Cast is a project developed together with clinicians and patients from the hospital Ntra. Sra. de la Paz in Spain. Data from patients are collected from wearable medical devices (Actigraph, [www.activinsights.com](http://www.activinsights.com)) and integrated into the dataset that psychiatrists collect during a regular consultation. The patients wear a smart band and periodically send the data to the system. In addition, the patients fill in a daily form on smartphones. All these data are collected in a 188-variable data set. This data set is the input of a data integration process that returns a structure with normalized variables and crisis indicators that facilitate a preliminary analysis and the application of a future Machine Learning analysis. This analysis shows how to detect some dependencies and reduce the number of variables after the data integration stage. Besides, data integration is carried out to complement the Young and Hamilton indexes on the mania and depression crisis (Llamocca, 2018; Junestrand, 2018) and to provide clinicians with additional information.

In addition to this introduction, section 2 shows the procedure of the research. Section 3 contains a description of the data and sources of information. In section 4, the integration of the data collected is explained. Section 5 is about the analysis and results and finally, section 6 shows the conclusions and the future work on the project.

## 2. PROCEDURE

To accomplish the project and gather the prototype data, 12 patients from the partner hospital are involved in the project. The participation of these patients requires the compliance with all privacy policies and prior informed consent. After getting all permissions and starting on September 2017, a period of collecting the data was scheduled. Data from patients were gathered using different methods and sources. The next section explains the most relevant sources. In addition, a process of cleaning and integration of the data was also developed by executing algorithms in R. One of the resulting data set, `data2.csv`, consist of 2328 observations of 188 variables. This data set contains the most relevant variables in the study and also several other indicators (calculated variables). The "crisis indicator" is a categorical variable that indicates the stage of the patient: Mania, Depression or Euthymia (the normal state of a patient). Only in a few observations this indicator is different than NA (not available) value. The set of observations in which the real value is known (set by the doctor during the consultations) will be used to train the system when the number of observations become large enough. Meanwhile the data set is used as a prototype to be utilized for experiments. Junstrand (2018) performed a comparison of the behaviour of several machine learning algorithms on the dataset. The following step is to model the euthymic state of each particular patient. Comparisons between the real state and the euthymic state of the patient are relevant for the prediction of crisis. This study facilitates the decision making for the clinician.

## 3. SOURCES OF INFORMATION

The most relevant sources of information in this project are described as follows.

- *Interviews.* The patients are observed periodically by a psychiatrist in an interview session. During the session, the psychiatrist registers a total of 38 variables. One of these variables is the "crisis indicator" according to his/her diagnosis. Also relevant HDRS scale (Hamilton Depression Rating Scale) and YMRS scale (Young Mania Rating Scale) are added in these interview sessions.
- *Smartwatches/bands.* Each patient involved in the project has its own medical smart band. This device gathers a set of important indicators (heart rate, physical activity, sleep quality...) and a total of 108 variables. Data are automatically recorded by the device and during the appointment with psychiatrist are periodically collected. (Bellivier, 2014) and (Anchiriaco, 2017) are publications on actigraphy as data source, the latter containing the R code that Bip4Cast has created to import data from the medical bands to the server.
- *Fill-in Forms.* Patients complete an electronic form daily: The Bip4Cast app (Martinez, 2016) or a Google form as alternative. This form was designed as part of the Bip4Cast project for gathering quantitative and qualitative data. The source provides them with 41 variables. Some of them are objective data, e.g. coffee consumption, tobacco or alcohol consumption, periods of menstruation, etc. Other data are subjective data, e.g. anxiety, concentration, comfort, etc.

## 4. INTEGRATION

Every source output yield data with different structure/formats. This phase deals with centralizing all the variables in a single structure. Some packages of *RStudio* as *tidyverse* are being used for data transformation. The following two goals are already achieved.

- *Building a single structure.* A single structure capable of storing any variable from any source. The new structure is also scalable for allowing to store variables coming from new sources (as voice recording, images and video). Currently, there are 187 variables extracted from sources already incorporated. Table 1 provides information about this structure. One record in this structure corresponds to one observation day for a specific patient.
- *Building an integration process.* This process is capable of treating all formats/structures from any incorporated source. To achieve this goal, several tasks are performed: cleaning, formatting, standardization, interpolation, etc. For categorical variables we use the factor data type in R which is a very useful way to make the analysis shorter in combination with the non-categorical variables.

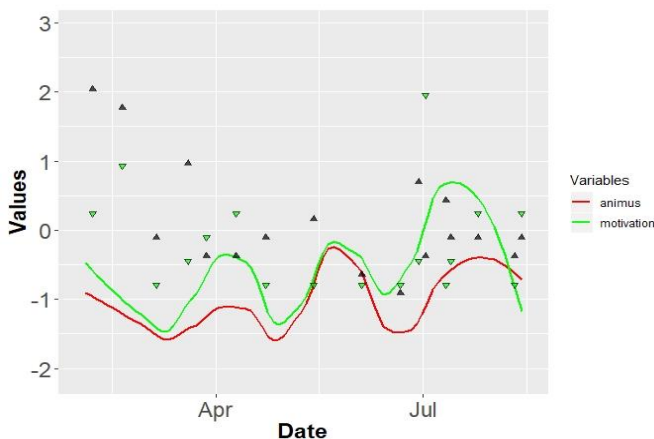
<sup>6</sup> Iván Palomares (Ed.): Proc. 1st International 'Alan Turing' Conference on Decision Support and Recommender Systems (DSRS-Turing'19) The Alan Turing Institute, London, United Kingdom, 21-22nd November 2019 ©DSRS-Turing'19 ; The Alan Turing Institute. ISBN: 978-1-5262-0820-0

**Table 1. Integration Structure**

Group of Variables	Type	Nº of Variables	Source	Frequency
Patient	CHAR	1	All	Daily
Date	DATE	1	All	Daily
Fill-in Forms Parameters	NUMERIC	39	Fill-in Forms	Daily
Smartwatch Parameters	NUMERIC	108	Smartwatches	Daily
Interviews Parameters	NUMERIC	37	Interviews	Each 2 weeks
Crisis Indicator	NUMERIC	1	Interviews	Each 2 weeks
“New Source” Parameters	TBD	TBD	Voice, video	TBD

## 5. ANALYSIS AND RESULTS

In this study patient’s data are individually analysed. As the structure contains all the data from all patients, filtering data per patient before analysing is necessary. One of the first steps is to build the correlation matrix. To remove noise, a Principal Component Analysis (PCA) is carried out. The similarity between variables can be also visualized in graphics using "ggplot" R-package. For example, the relation between variables *animus* and *motivation* (Figure 1) is established from the fill-in form data. The correlation coefficient between them reaches 85% average (for all patients). This result can be useful to discard one of the variables on the base of high correlation. In relation to the accelerometer features, Pearson’s correlations are computed (Figure 2). The results show that the features related to sleep quality are the most relevant for crisis prediction: the maximum positive correlation (0.43) exists between features ‘at least 5min wake at night’ and ‘duration\_day\_minutes’. Strong positive correlations exist between features ‘dur\_night\_min’ and ‘dur\_day\_min’(0.39); and ‘dur\_night\_sleep’ and ‘ACC\_tinday\_min’ (0.37). Maximal values in negative correlation are between ‘acc\_onset’ and ‘during\_day\_LiG300\_min’ (-0.23); between ‘during\_day\_OIN30\_min’ and ‘acc\_onset’(-0.23); and (-0.32) between ‘during\_day\_OIN30\_min’ and ‘sleep\_efficiency’; even (-0.33) between ‘at least 5 min wake at night’ and ‘sleep\_efficiency’.



**Figure 1. Similarity between variables: animus (red) and motivation (green)**

Also that it is an implication for decorrelation of the features for further data mining, and especially a practical meaning for clinical psychiatrist. For instance, they could draw additional conclusions about the features most important to understand the dynamics in every particular patient. From data mining point of view, decorrelation is important to decrease the dimensionality of a problem leading to an easier further

performance in machine learning sense. The data show an example of a good classification based on a smaller number of features once internal relations and connections are revealed. At this point, the study of the entropy measures of the data can be relevant. Cukic et al. (2018, 2019) relate presence of depression of mania crisis to a low entropy levels in de data of mood.

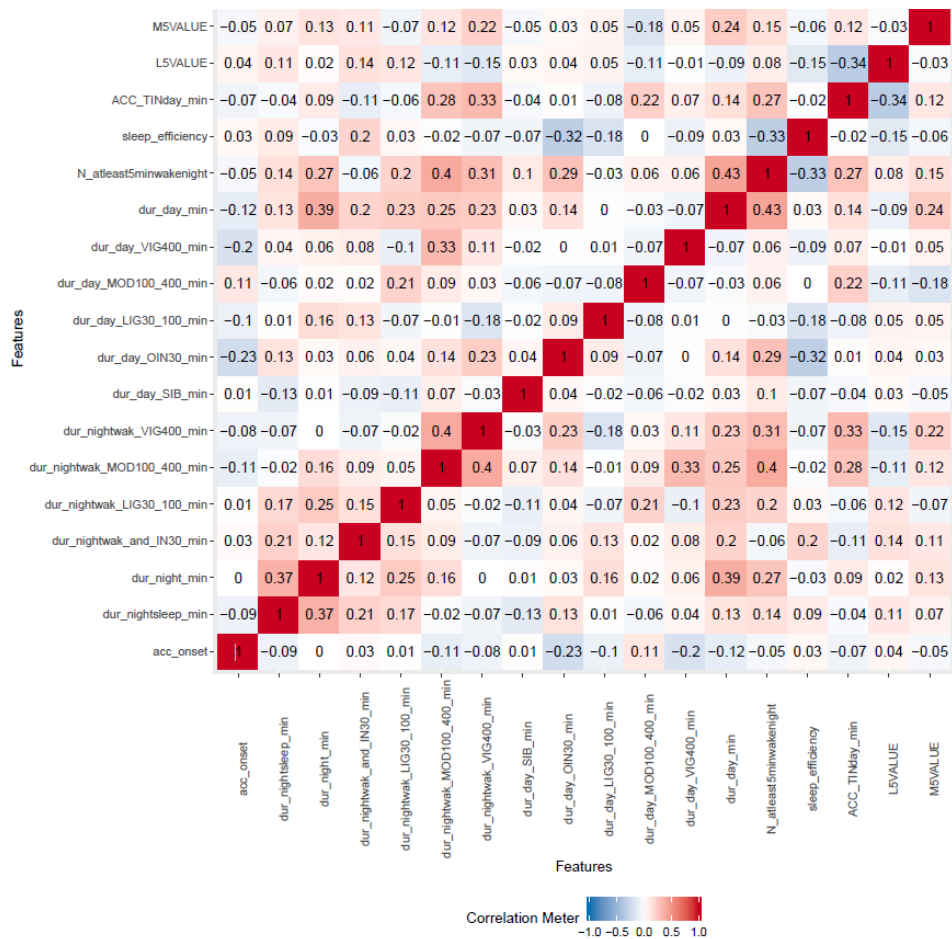


Figure 2. Correlation between accelerometer features.

### 6. CONCLUSIONS AND FUTURE WORK

This work describes the Bip4Cast project on mood disorders analytics. The project is in data analysis phase within real data. Several sources of data from patients are being collected and the behaviour and evolution of relevant variables are in process of analysis to provide predictions that supports the decision making in further medical treatments. As a future work, a prediction model will be developed and some machine learning algorithms will be used. In addition, Bayesian Analysis will be use to aggregate future data and debug the model where previous low mood fluctuations indicate the presence of depression or mania crisis.

**Acknowledgements:** This work is supported by the Project H2020-MSCA-RISE-2015-690874 (2016-2020).

## 7. REFERENCES

- J. Anchiraico (2017), Design of a Big Data architecture for predicting crises in bipolar disorder, Master Thesis, *Eprints Complutense* (<https://eprints.ucm.es/41633/>)
- F. Bellivier et al. (2014), Sleep in remitted bipolar disorder: A naturalistic case-control study using actigraphy, *Journal of Affective Disorders*, 158, DOI: 10.1016/j.jad.2014.01.012, pp. 1-7
- M. Cukic et al. (2018), EEG machine learning with Higuchi fractal dimension and Sample Entropy as features for successful detection of depression, *Journal ArXiv Cornell University*, pp. 1-34
- M. Čukić Radenković (2019). Novel Approaches in Treating Major Depressive Disorder (Depression), ISBN: 978-1-53614-382-9, NOVA Scientific Publishers Ltd. July 2019.
- M. Čukić Radenković and V. Lopez Lopez (2019). Machine Learning Approaches for Detecting the Depression from Resting-State Electroencephalogram (EEG): A Review Study. Cornell Repository Arxiv.org. <https://arxiv.org/abs/1909.03115>
- A. Junestrand (2018) Application of Machine Learning Algorithms for Bipolar Disorder Crisis Prediction. Bachelor Thesis, *Eprints Complutense* (<https://eprints.ucm.es/48866/>)
- P. Llamocca, A. Junestrand, M. Cukic, D. Urgeles, V. Lopez (2018). Data Source Analysis in mood disorder research, *Proceedings of the XVIII Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA)*, Granada, Spain, pp. 893-898.
- V. Lopez, G. Valverde, J. Anchiraico and D. Urgeles (2016) Specification of a Cad Prediction System for Bipolar Disorder, *Uncertainty Modelling in Knowledge Engineering and Decision Making*, World Scientific Proceedings Series on Computer Engineering and Information Science, vol. 10, pp. 162-167
- A. Martinez, (2016) System to predict bipolar disorder crises analysing massive data, Bachelor Thesis, *Eprints Complutense* (<https://eprints.ucm.es/38722/>)
- N. Vanello et al. (2012), Speech analysis for mood state characterization in bipolar patients, *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, CA, pp. 2104-2107.



Publicación III

# Personalized characterization of emotional states in patients with bipolar disorder

**Llamocca, P., López, V., Santos M., Čukić M.**

Publicado en *MDPI Mathematics*, June 2021, volumen 9, Número 11:1174, pp. 123–456. DOI: 10.3390/math9111174., ISSN. 2227-7390





Article

# Personalized Characterization of Emotional States in Patients with Bipolar Disorder

Pavel Llamocca <sup>1</sup>, Victoria López <sup>2,\*</sup>, Matilde Santos <sup>3</sup>  and Milena Čukić <sup>4</sup><sup>1</sup> Computer Sciences Faculty, Complutense University of Madrid, 28040 Madrid, Spain; pavellam@ucm.es<sup>2</sup> Quantitative Methods Department, Cunef University, 28040 Madrid, Spain<sup>3</sup> Institute of Knowledge Technology, Complutense University of Madrid, 28040 Madrid, Spain; msantos@ucm.es<sup>4</sup> AHTI, Amsterdam Health and Technology Institute, HealthInc, 1062KS Amsterdam, The Netherlands; micu@3ega.nl

\* Correspondence: victoria.lopez@cunef.edu; Tel.: +34-62-997-5771

**Abstract:** There is strong clinical evidence from the current literature that certain psychological and physiological indicators are closely related to mood changes. However, patients with mental illnesses who present similar behavior may be diagnosed differently, which is why a personalized study of each patient is necessary. Following previous promising results in the detection of depression, in this work, supervised machine learning (ML) algorithms were applied to classify the different states of patients diagnosed with bipolar depressive disorder (BDD). The purpose of this study was to provide relevant information to medical staff and patients' relatives in order to help them make decisions that may lead to a better management of the disease. The information used was collected from BDD patients through wearable devices (smartwatches), daily self-reports, and medical observation at regular appointments. The variables were processed and then statistical techniques of data analysis, normalization, noise reduction, and feature selection were applied. An individual analysis of each patient was carried out. Random Forest, Decision Trees, Logistic Regression, and Support Vector Machine algorithms were applied with different configurations. The results allowed us to draw some conclusions. Random Forest achieved the most accurate classification, but none of the applied models were the best technique for all patients. Besides, the classification using only selected variables produced better results than using all available information, though the amount and source of the relevant variables differed for each patient. Finally, the smartwatch was the most relevant source of information.

**Keywords:** decision making; machine learning; classification; bipolar disorder; mental healthcare



**Citation:** Llamocca, P.; López, V.; Santos, M.; Čukić, M. Personalized Characterization of Emotional States in Patients with Bipolar Disorder. *Mathematics* **2021**, *9*, 1174. <https://doi.org/10.3390/math9111174>

Academic Editors: Giancarlo Consolo and Vitaly Volpert

Received: 8 April 2021

Accepted: 19 May 2021

Published: 22 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Personalized medicine is a new promising approach that proposes the tailoring of therapy according to the patient's personal characteristics to ensure the best treatment outcomes. By enabling each patient to receive early (correct) diagnosis, risk assessments, and optimal treatments, personalized medicine holds promise for not only improving health care but also lowering its costs and improving patients' quality of life [1]. This is especially true in some branches of medicine such as those dealing with mental illnesses. Indeed, there has been a large increase in the numbers of patients with mental health issues amplified by the current pandemic [2].

There is strong clinical evidence from the current literature that certain psychological and physiological indicators are closely related to mood changes [3]. However, patients with mental illnesses who present similar behavior may be diagnosed differently, which is why a personalized study of each patient is necessary. For instance, some behavioral indicators shown by patients diagnosed with bipolar depressive disorder (BDD) may be

quite different from those with major depressive disorders (MDDs), being similar illnesses. BDD patients have a more complex attitude.

A current issue regarding mental health is that based solely on standard clinical inquiry (self-report and observation by a trained clinician), BDD is often misdiagnosed as unipolar depression and is treated as such. Bipolar depression patients may be treated wrongly with antidepressants for that reason, even when this therapy may aggravate their state [4,5]. However, it is possible to distinguish more than one phase in patients with BDD [6].

Therefore, in this work, supervised machine learning (ML) algorithms were applied to the classification of the different states of patients diagnosed with bipolar disorder (BD). Random Forest (RF), Decision Trees (DT), Logistic Regression (LR), and Support Vector Machine (SVM) were applied under different configurations. The final goal was to provide relevant information to medical staff and the relatives to help them deal with these situations and to make decisions on the best way to treat these patients.

The information used was collected from 17 BDD patients through wearable devices (smartwatches), daily self-reports, and medical observation at regular appointments. The variables (a total of 173) were processed and then statistical techniques for data analysis, normalization, noise reduction, and features selection were applied. An individual analysis of each patient was carried out.

The data presented in this study are available upon request from the corresponding author. Although the data have been completely anonymized and provided together with the corresponding certificate, it is not publicly available, due to the requirements of the provider, the Hospital Nuestra Señora de la Paz, in Madrid (Spain).

The results with all the variables and with only the best ranked variables by feature-selected methods were obtained and compared. Besides, some selected variables from the smartwatch were used to confirm the medical diagnosis of these patients, with good results. Moreover, adding two specific variables from the self-reporting source improved the classification. Although Random Forest gave a slightly better accuracy, none of the techniques were best for all of the patients. This result confirmed that each patient requires a personalized analysis to be able to obtain a better classification of their mood states.

The rest of the paper is structured as follows. Section 2 summarizes some related works. Section 3 details the materials used, and the machine learning techniques and methods applied. Section 4 describes the data analysis carried out, including the correlation and principal component analysis. Section 5 presents the application of some machine learning techniques to the data and discusses the results of the models. The paper ends with conclusions and future work.

## 2. Previous Studies

Disease diagnosis and prediction using health databases have always been a potential application area for machine learning methods [7]. They are usually applied to public databases of patients or to historical medical data collected by a certain healthcare institution. They have been used for different diseases, from cancer diagnosis [8–10] to diabetes [11,12].

According to [13], the Support Vector Machine (SVM) algorithm is applied most frequently, followed by the Naïve Bayes algorithm and Random Forest (RF) algorithm, the latter of which usually shows superior accuracy comparatively. Nevertheless, the application of these machine learning techniques to identify mental illnesses is not so frequent, due to the difficulty of collecting data from these patients [14]. Cho et al. [15] found that, based on a literature review, SVM, Gradient Boosting Machine (GBM), RF, Naïve Bayes, and K-Nearest Neighborhood (KNN) are the most frequently employed in the mental health area. However, they stated that, though every ML algorithm has its own advantages, researchers should be aware of the limitations of the results obtained under restricted data conditions in the practice of mental health. In the survey by Thieme [16], the authors reflected on the current state-of-the-art of ML work for mental health, and they

suggested more consideration of the potential implications that ML models can have in real-world mental health contexts, as proved by other works [17].

In Costello et al. [18], an extended version of the conventional Tree-Augmented Naive Bayesian Network was applied to find the main causes leading to depression in Korea. The goal of this proposal was to help healthcare decision-makers comprehend changes in depression status. The same database was used in [19], to estimate factors affecting depression using deep learning and machine learning algorithms.

Particularly, bipolar disorder has always been considered especially difficult to diagnose accurately, as stated by Phillips and Kupfer [20]. In [21], a review of the existing literature on the use of machine learning techniques in the assessment of subjects with bipolar disorder was presented. They concluded that, given the clinical heterogeneity of samples of patients with BD, machine learning techniques may provide clinicians and researchers with important insights in fields such as diagnosis, personalized treatment, and prognosis orientation. In a more recent literature review, some advice for the proper applications of ML techniques to BDD was provided [22]. Highlighting the paramount goal of identifying etiological markers of MDD in order to help the prevention and treatment of these patients, in [23], a systematic review was conducted of all published studies reporting on the prospective relationship between positive and negative emotionality and the longitudinal course of MDD in diagnosed adults. The authors concluded that emotions are promising predictors of MDD course. In a similar way, Lima et al. [24] carried out medical research on cognitive deficits in bipolar patients. They focused on two key predictors: cognition and emotionality. They provided evidence that problems in cognition and emotionality are prominent among those diagnosed with bipolar disorder.

Regarding the sources of information for the detection of BP, they may be very different. For instance, [25] used a purpose-built online mental health questionnaire and dried blood spot samples for biomarker analysis from nine individuals with depressive symptoms. Extreme Gradient Boosting and nested cross-validation were used to train and validate diagnostic models. In [26], the design of a sensor based on physiological predictors for the continuous and autonomous monitoring of bipolar patients was proposed. It uses a pulse rate sensor to record heart rate, and an electrodermal activity sensor to trace the emotional and cognitive state changes. The extracted features from the data were then used to build machine learning models to predict the given psychological change in response to the emotional stimuli. In [27], the authors reviewed some papers on classification models applied to electroencephalographic recordings of patients with a depression diagnosis. In [28], BDD patients' data collected from different sources were analyzed. The results showed that data from sleep and daily activity, measured both by movement and sounds, are relevant for improving the prediction of crises in patients with bipolar disorder.

In the paper by M'Bailara et al. [29], the authors used self-completed questionnaires, showing that normothymic bipolar patients displayed higher levels of emotional lability and intensity than the controls did. They carried out an emotional induction experiment based on the viewing of a set of some positive, negative, and neutral pictures. They then analyzed the reaction of normothymic bipolar patients and compared them with the control group.

In [30], the authors explored speech analysis based on phone recordings collected for the detection of mood in individuals with BDD. They classified every call into the following groups: mania/euthymic or depressive/euthymic, by applying SVM. They also proved that data preprocessing, feature extraction, and data modeling improve significantly the classifier performance. Similarly, in [31], the dynamic analysis of data from voice monitoring with accelerometers was proved to be an adequate complement to other kinds of data for predicting crises in BDD patients with a combination of machine learning algorithms.

There has been an increasing interest in using mobile sensing technologies for mental health monitoring. The use of the mobile phone was analyzed in Faurholt-Jepsen et al., [32]. In this paper, the authors claimed that objective smartphone data reflecting behavioral activities allow the classification of patients with bipolar disorder compared with healthy

individuals. The results showed that the number of text messages/day and the duration of phone calls/day were increased in patients with bipolar disorder. In a similar way, Antosik-Wójcińska et al. [33] concluded that objective data automatically collected using smartphones (voice and smartphone-usage data) are valid markers of a mood state.

However, in mental health, despite the numerous mobile applications available, there are few smartwatch solutions on the market. By contrast, smartwatches, which are small and unobtrusive, and can collect data continuously, are able to monitor users' activities and detect behavioral patterns [34]. Nevertheless, the use of wearable technology for people diagnosed with bipolar disorder and, potentially, other affective disorders, is growing [35]. As an example, in [36], contextual and sensor data were collected using smartphone and smartwatch sensors under free-living conditions. Personalized models for detecting emotional transitions and states were built using a wide range of supervised ML algorithms.

New lines of research are arising regarding this topic. For instance, the paper by Perri et al. [37] linked the disruptions in structural, functional, and effective connectivity that have been documented in BD with the impact on circuits that support emotional processes, cognitive control, and executive functions. That is, those at high risk for BD show patterns of connectivity that differ from both matched control populations, so it is possible to define those neurobiological markers.

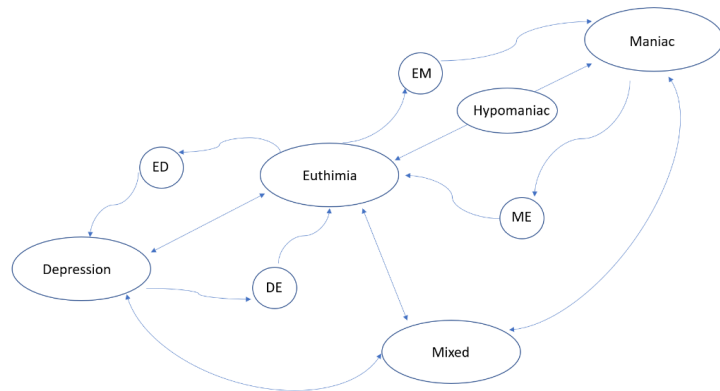
### 3. Materials and Methods

One of the main contributions of this article is the analysis of both objective and subjective variables related to BDD patients. This section starts with a brief description of this mental illness to better understand some relevant indicators and variables. The database consisted of real data collected during the Bip4Cast project [31]. For this research, a proper approval of the local Ethic committee was obtained, as well as a written informed consent from all the participants. Different indicators of 17 patients diagnosed with BDD were taken from three sources: daily forms (daily self-reports filled out by the patient via a mobile app), smartwatch measurements, and clinical interviews (173 in total). Although the previous Bip4Cast research analyzed a similar dataset, further data processing is necessary in order to differentiate the specific characteristics of each patient. Indeed, in [38], it has been shown that personalized analysis is 30% more accurate than generic models.

Based on [6], four types of Bipolar Depression Disorder were identified:

- Bipolar Disorder I, characterized by depressed episodes, frequent manic episodes, or mixed episodes;
- Bipolar Disorder II, typified by full depressed and hypomanic episodes;
- Cyclothymic Disorder or Cyclothymia. It is a chronically unstable mood state in which people experience hypomania and moderate depressive episodes for at least two years;
- Unspecified Bipolar Disorder, when a patient does not meet any of the above criteria but has experienced periods of clinically significant abnormal mood elevation.

Based on these types of BDD disorders, the psychiatrist diagnoses the situation of the patient corresponding to one of the following states: Depressed, Hypomanic, Manic, or Mixed. Clinically, a crisis is defined as the appearance of one of those states. The absence of any of them is a state known as Euthymic (i.e., the patient does not suffer from any crisis). In this study, the variable "diagnosis" takes one of these 5 states (Depressed, Hypomanic, Manic, Mixed, or Euthymic). These states can be described according to clinical indicators, mood changes, or phases of the disease already described in clinical practice. The clinician modifies the diagnosis (or state), after observing significant symptoms that characterize a different new state. Figure 1, elaborated by the authors, shows the general state transitions: the nodes are the five main BDD states plus 4 more nodes for intermediate states (ED, DE, EM, and ME). The arrows show the transition between states.



**Figure 1.** Transitional state-directed graph illustrating bipolar disorder phases (nodes represent the states and arrows represent the transition).

3.1. Model Fundamentals

A patient P is characterized by a set of variables,  $X_1, \dots, X_n$ , where some are more explanatory than others regarding the patient’s emotional state. The value of these variables determines the mood state of the patient. However, many of these variables are correlated and, therefore, one of the first steps in the processing of the information is to determine the minimum set of variables that best represents the patient’s state.

The formalization of the Algorithm 1 that characterizes each patient is as follows.

---

**Algorithm 1** Characterization Algorithm

---

1: Initialization. Let C be the set of candidate variables to be part of the patient characterization, and let W be the characterization set. Initially, vector  $C = \{X_1, \dots, X_n\}$  contains all possible variables collected, and  $W = \emptyset$  (empty set). The vector  $CV = \bar{X}/s_x$  that represents the variation in the variables is calculated as the quotient between the mean and the standard deviation of each variable.

The correlation matrix  $COR \in M_{n \times n}$  between these variables,  $COR_{ij} = COR(X, Y) = \frac{s_{xy}}{s_x s_y}$ ,  $X = x_i \wedge Y = x_j$ , is obtained.

The parameter  $Max = Max(P)$  is the maximum number of variables for each patient P. This number will be used as Halting criteria.

2: First selection. The variable  $X_k \in C = \{X_1, \dots, X_n\}$  with the highest CV value (variation) is selected (Initial Criterion). Then,  $W_1$  (the winner 1) is updated as  $W_1 = X_k$  and  $W = \{W_1\}$ . Remove  $X_k$  from the candidates’ set:  $C = C \setminus \{W_1\}$ . The index is updated,  $i = 1$ .

3: General selection. Select the variable  $X_k$  with the lowest correlation of the set of winners W; in case of a tie, select the one with the highest variation coefficient (General Criterion). That is:

$$\sum_i COR(X_k, W_i) \leq \sum_i COR(X_j, W_i), \forall X_j \in C$$

4: Update the characterization set  $W_{i+1} = X_k$ ;  $W = W \cup \{W_{i+1}\}$ ,  $C = C \setminus W$ ;  $i = i + 1$ ;

5: Halting criterion. If  $|W| < Max$  (or another Halting criterion is met), stop. Otherwise, go to Step 2.

---

After executing the Characterization Algorithm for all patients, a winner variable set  $W(P) = \{W_1, \dots, W_p\}$  is obtained for patient P. This is the set of variables that characterizes the mood state of that patient. The algorithm can use any feature selection procedure. For the Halting criterion, we have used principal component analysis (PCA), in order to obtain the number of loops the algorithm must perform. This way, we try to cover 99% of the variance.

In addition, as the variables come from different sources, an integration of all the variables is necessary. For instance, the variables from the daily report are taken every 24 h. However, the variables from the accelerometer (smartwatch) are measured every 5 min. Other values of variables are obtained only when the patient goes to the doctor’s appointment. In this work, we have considered the day as the sample time unit, so the final dataset has  $m$  observations that correspond to days of the patient’s monitoring.

A patient dataset consists of an  $m \times (p + 2)$  matrix where the first column represents the day, the last column contains the diagnosis of the patient (target), and the rest of the columns are the variables  $W_j$  at  $t_i$  recording date.

$$dataset(P) = \begin{pmatrix} t_1 & w_{11} & \cdots & w_{1p} & d_1 \\ t_2 & w_{21} & \cdots & w_{2p} & d_2 \\ \cdots & \cdots & \ddots & \cdots & \cdots \\ t_m & w_{m1} & \cdots & w_{mp} & d_m \end{pmatrix}$$

There are more than 365 observations available for each patient, taken in a time period from October 2017 to January 2019. The monitoring time interval of a patient  $t_1, \dots, t_m$ , is divided into slots that correspond to the different emotional states,  $d_i$  (Figure 1). Thus,  $state(P, t_{i1}, \dots, t_{im}, d_i)$  is a submatrix of  $dataset(P)$ ,

$$\forall t_{ik} \in \{t_{i1}, \dots, t_{im}\} : t_{ik} = t_{ik-1} + 1,$$

$$d_i \in \{ED, DE, EM, ME\},$$

$$state(P, t_{i1}, \dots, t_{im}, d_i) = \begin{pmatrix} t_{i1} & w_{i11} & \cdots & w_{i1p} & d_i \\ t_{i2} & w_{i21} & \cdots & w_{i2p} & d_i \\ \cdots & \cdots & \ddots & \cdots & \cdots \\ t_{im} & w_{im1} & \cdots & w_{imp} & d_i \end{pmatrix}$$

Machine learning methods were applied to these sub-matrices, after initial cleaning and other processing operations explained in the following sections.

### 3.2. Data Collection

The Bip4Cast data came from a study carried out in collaboration with a public hospital, Nuestra Sra. de la Paz (Madrid, Spain). The BDD patients were evaluated and diagnosed by an experienced psychiatrist. All of the patients satisfied both ICD-11 (International Classification of Diseases. 11th Revision) and DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition) criteria.

The variables were collected from different information sources:

- Smartwatches/bands. Each patient has their own medical smart band. This device gathers a set of objective variables (physical activity, sleep quality, etc.) and up to 110 activity variables. The actigraphy as a data source has been proved to be very useful [39].
- Fill-in forms. Patients complete an electronic form on a daily basis. This form was designed as part of the Bip4Cast project for gathering a total of 20 variables. Some of them are quantitative, e.g., coffee, tobacco, or alcohol consumption, and periods of menstruation. Others are qualitative variables such as anxiety, concentration, and comfort. When there are missing data in self-daily reports, interpolation is applied. These are subjective variables as they are directly reported by the patient.
- Periodical interviews during medical consultation. In these sessions, the psychiatrist registers a total of 40 variables. HDRS (Hamilton Depression Rating Scale) and YMRS (Young Mania Rating Scale) items are included. Patients are diagnosed more than once monthly (usually every two weeks). Therefore, the variable “diagnosis” is also collected. According to the psychiatrist, a diagnosis should be maintained until changes in the patient’s symptoms occur, that is, until the next diagnosis. The

diagnosis is directly related to the HDRS and YMRS tests. This source of data consists of subjective and objective variables.

The data from these three sources have a different structure/format. Thus, data processing was necessary to integrate all this information into a single data structure. This work used the preprocessing functions offered by RStudio in the specific-use library “tidyverse” [40]. This library contains the necessary functions for information fusion from different sources and to apply some data processing such as cleaning, formatting, normalization, standardization, and interpolation.

In this project, patients used a specific smartwatch, GENEACTIV 1.1 [41] <https://www.activinsights.com/technology/geneactiv/>, accessed on 21 May 2021), whose accuracy has been measured [42]. We did not use any specific application other than the R language, particularly, the R CGIR library. GGIR is an R-package to process multi-day raw accelerometer data for physical activity and sleep research [43].

The new single data structure integrates 173 variables, where two of them are key variables (ID of the patient and timestamp). A record of this data structure corresponds to one day of a specific patient. The new structure is scalable and open to growing with other variables and new sources (as voice recording, images, and video).

Table 1 shows the data structure, with 6 groups of variables, and the type, source of information, and frequency of each variable. The different frequency of sampling imposes escalation of the data.

**Table 1.** Data structure of all the variables.

Group of Variables	Type	No. of Variables	Source	Frequency
Patient	CHAR	1	All	Daily
Date	DATE	1	All	Daily
Fill-in Forms	NUMERIC	20	Fill-in Forms	Daily
Smartwatch	NUMERIC	110	Smartwatches	Daily
Interviews	NUMERIC	40	Interviews	Each 2 weeks
Diagnosis	NUMERIC	1	Interviews	Each 2 weeks

The dataset contains diagnosis observations from 17 patients who have been monitoring with smartwatches, and who have filled the self-daily reports during a significant period (more than one month in a row). The personal information of the patients was previously anonymized. In this work, we referred to them as P01 to P17. Table 2 shows the states of the 17 patients between October 2017 and January 2019. All of them have had one or more crisis episodes within the period of observation.

To better identify the source of each variable, all of them have been named by adding a character at the beginning: “S\_” for smartwatches, “D\_” for daily forms (self-daily reports), “I\_” for interviews, “H” for HRDS scale, and “Y\_” for Young scale. Although the same variables have been collected for all patients, according to the characteristics of each one, only some of them will be relevant for each particular patient. For example, the variable tobacco does not provide any useful information about a nonsmoking patient.

To obtain the relevant variables for a specific patient, 8 Feature Selection (FS) methods were applied (A–H in Table 3): Boruta from Random Forest algorithm (A), Variable Importance from machine learning algorithms (B), LASSO (Least Absolute Shrinkage and Selection Operator) (C), STEPWISE (D), Recursive Feature Elimination (E), Genetic Algorithm (F), Simulated Annealing (G), and Descriptive Machine Learning Explanation (DALEX) (H).

**Table 2.** States of the patients during the recollection period.

Patient	Depressed	Euthymic	Hypomanic	Manic	Mixed
P01	x	x	x	x	x
P02		x		x	
P03	x	x	x	x	x
P04	x	x		x	x
P05	x	x	x	x	x
P06	x	x	x	x	
P07	x	x	x	x	x
P08	x	x		x	x
P09		x		x	x
P10	x	x		x	x
P11		x			x
P12	x	x			
P13	x	x			x
P14	x	x		x	x
P15		x			
P16	x	x		x	x
P17	x	x		x	x

**Table 3.** Selected variables by the Feature Selection methods for patient P07.

Variable	Description	Feature Selection Procedures								Rank
		A	B	C	D	E	F	G	H	
D_sleep_zenith	Middle point of the sleep	1	0	1	0	1	0	1	1	5
S_ACC_day_LIG_unbt_mg	Avg. acceleration in light state	1	0	1	1	0	1	0	1	5
S_Nblocks_spt_wake_MOD	Times of waking up in moderate state	1	1	0	0	0	1	1	1	5
D_sleep_time	Time to go to rest	1	0	1	0	0	1	0	1	4
D_wakeup_time	Time to wake up	1	0	0	0	0	1	1	1	4
S_Nblocks_spt_wake_LIG	Times of waking up in light state	1	0	0	1	0	1	1	0	4
S_ACC_day_total_MOD_m	Avg. acceleration total in moderate	1	0	1	0	0	1	0	1	4
S_dur_day_total_VIG_min	Total time spent in vigorous state	1	1	0	0	1	0	1	0	4
S_Nblocks_spt_sleep	Times of waking up during sleep	1	0	1	1	0	0	1	0	4
S_dur_spt_wake_LIG_min	Duration of waking up time in light activity	1	1	0	0	1	0	0	1	4

Each of those FS algorithms returns a list of variables that have more impact in order to analyze the target variable (diagnosis). However, not all FS procedures gave the same group of variables. For this reason, we ranked each variable by the number of FS procedures that returned it. Table 3 shows, as an example, the top 10 ranked variables for patient P07.

As it is possible to see in Table 3, the 10 first variables of patient P07 were selected by at least 4 different methods. The smartwatch was the most relevant source of information in all cases (7 out of 10 in the list). Nevertheless, the amount and nature of the relevant variables differed for each patient.

### 3.3. Classification Methods

Different machine learning (ML) algorithms were applied to classify the observations. Random Forest (RF), Decision Trees (DT), Logistic Regression (LR), and Support Vector Machine (SVM) were used under different configurations. The code has been written in R language with RStudio. Table 4 shows the algorithms and the specific libraries applied in this study.

**Table 4.** Machine learning algorithms and the corresponding R libraries.

Algorithm	R Libraries
Decision Tree	rpart
Logistic Regression	glmnet:glm
Random Forest	RandomForest
SVM	e1071::svm

A brief description of each ML technique that was applied is given below.

Decision Tree (DT) is an algorithm that predicts values of responses by learning the decision rules derived from the features. There are different ways to construct decision trees; the one used here was the Classification and Regression Trees (CART). This is a supervised learning technique where there is a target variable (which is dependent) and the goal is to obtain a function that procures the value of the target variable for unknown cases, from predictor variables (which are independent). The particular implementation of CART we used was Recursive Partitioning and Regression Trees (RPART).

Random Forest (RF) is an algorithm that uses supervised training. It builds individual decision trees for training. Predictions from all the trees are aggregated to achieve one final prediction. The algorithm randomly selects a certain number of columns and rows from entire data for building a decision tree from them. This process is repeated many times and different trees are built. The higher the number of trees constructed, the smaller the classification errors.

Support Vector Machine (SVM) is a supervised algorithm used for classification and regression problems. The main idea of SVM is to separate classes with a straight line (1 dimension), flat plane (2 dimensions), or an N-dimensional hyperplane. SVM uses kernel (nonlinear) functions to separate classes when it is necessary according to the data. SVM finds the best hyperplane by defining a decision boundary with the margin between the groups as large as possible.

Logistic regression (LG) is a regression model (also known as logit regression) for estimating the parameters of a logistic model, a binary model that has a dependent variable with two possible values, usually 0 and 1. In this model, the logarithm of the odds (or 1) is a linear combination of the independent (predictors) variables. In the logistic regression model, the dependent variable is a logistic function (Equation (1)). This function varies between (0, 1) for an input Y in  $(-\infty, +\infty)$  and, thus, the logistic function is ideal for modeling any binary response. The inverse of the logistic function is logit (Equation (2)). In logistic regression, the variable Y is treated as a linear function of a vector variable  $\vec{X}$ . That is, the logistic regression model can be defined as Equation (3).

$$\sigma(Y) = \frac{e^Y}{1 + e^Y} \tag{1}$$

$$g(\sigma(Y)) = \ln \left[ \frac{\sigma(Y)}{1 - \sigma(Y)} \right] = Y \tag{2}$$

$$\ln \left[ \frac{\sigma(Y)}{1 - \sigma(Y)} \right] = \theta_1 b_1(\vec{X}) + \theta_2 b_2(\vec{X}) + \dots + \theta_m b_m(\vec{X}) \tag{3}$$

Detailed information of these algorithms and applications can be found in [44].

### 4. Data Analysis

This section describes the kind of personalized analysis applied to each patient’s data. First, the correlation obtained between the variables is presented and the necessity to reduce dimensionality is justified. Then, principal component analysis (PCA) is carried out and the results are discussed.

#### 4.1. Correlation Analysis

The first step of the data analysis was to determine the correlation between the variables for each individual patient. The different degrees of correlation justified the feature selection process that was carried out as a second step. In this article, patient P07 has been selected as an illustration for the results obtained.

Figure 2 shows the correlation matrix of the first 15 ranked variables obtained for patient P07 from the feature selection procedures. The colour code is red, maximum value of correlation; white, minimum correlation; reddish, positive correlation; bluish, negative correlation. The results showed some highly dependent variables, such as variables related to nighttime activity (*S\_dur\_spt\_wake\_LIG\_min*, *S\_Nblocks\_spt\_wake\_MOD* and *S\_Nblocks\_spt\_wake\_LIG*). It can also be concluded that this patient (P07) engaged in light and moderate physical activities in the nights of recordings. The high correlation (0.83) between *D\_sleep\_zenith* and *D\_wakeup\_time* means that, regardless of the time patient P07 fell asleep, the number of sleeping hours was almost the same every day.

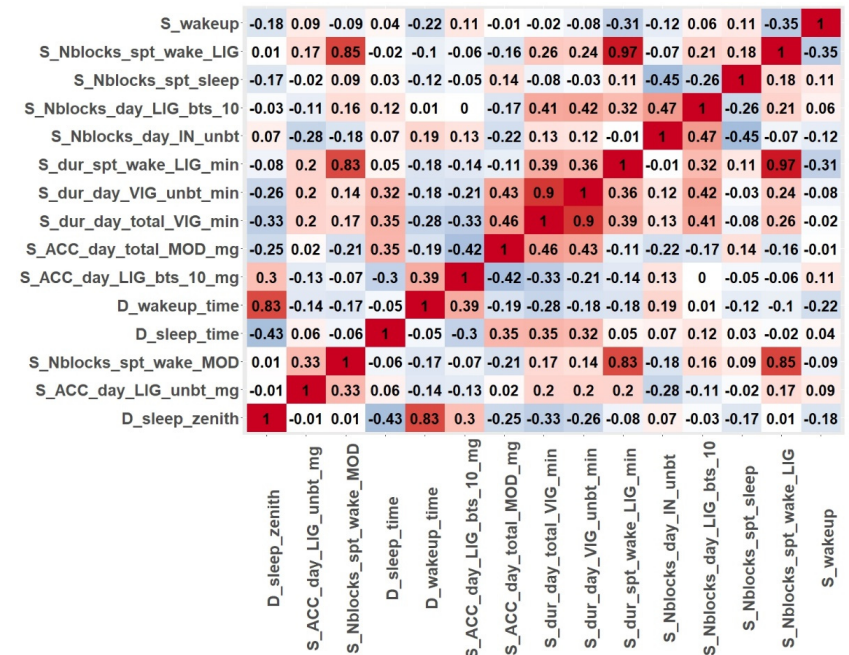
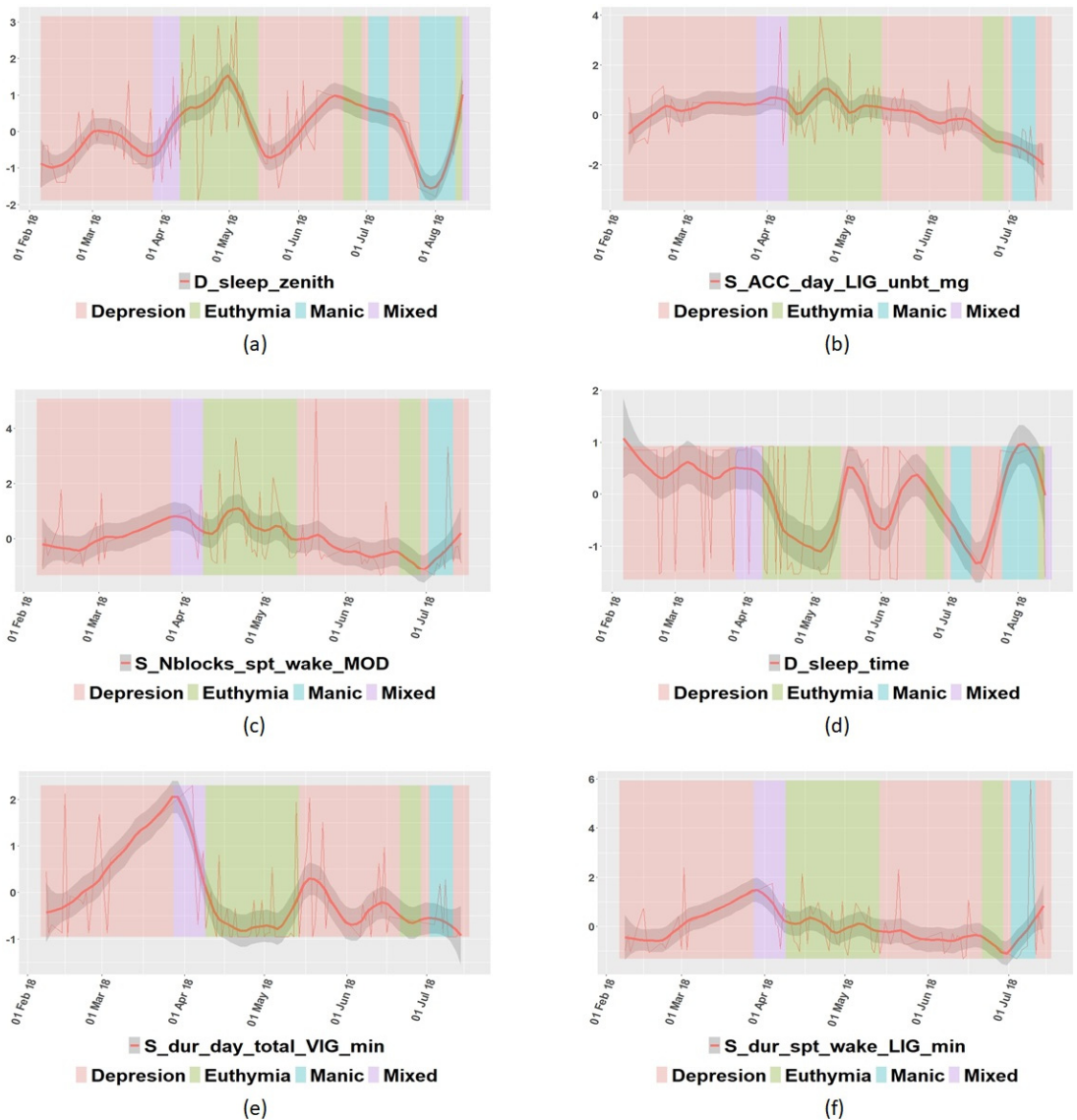


Figure 2. Correlation matrix for the features selected for patient P07.

The study of the correlation between variables together with the large number of variables of the original dataset clearly justified the need for reducing the dimensionality of the problem.

In addition to the correlation among features, some specific variables were further analyzed. In Figure 3, some variables from patient P07’s data are represented, and some interesting insights can be drawn. For example, Figure 3a shows that P07 tended to sleep

better in the euthymic state, and according to Figure 3b, this patient engaged in less light physical activity when in the depressed state. In Figure 3c, it is shown that patient P07 usually woke up more often in the euthymic state than in any other state. Figure 3d shows that this patient fell asleep earlier in the euthymic state than in any other case. In Figure 3e, it is possible to observe that the level of vigorous physical activity was unstable when depressed. Finally, Figure 3f allows us to conclude that patient P07 engaged in the same light physical activity in the sleep period times in any state. This information can help to establish some rules to predict the state of each particular individual.



**Figure 3.** Graphical representation of some variables of patient P07: (a) sleep zenith; (b) acceleration in light state; (c) times of waking up in moderate state; (d) time to go to rest; (e) time spent in vigorous state; (f) time spent in light state in sleep period time.

The correlation coefficient was also calculated for all variables with the target variable (*diagnosis*) for every patient. Table 5 shows some of the variables for only those patients whose correlation coefficient was greater than 0.5 or lower than  $-0.5$ .

**Table 5.** Variables with the highest correlation with variable diagnosis for some patients.

Patient	Source	Variable	Correlation (Variable vs. Diagnosis)
P06	Smartwatch	S_dur_spt_min	0.7427
P06	Smartwatch	S_ACC_day_IN_unbt_mg	-0.7310
P03	Young	Y_irritability	0.7888
P03	Young	Y_young	0.7828
P05	HRSD	H_mid_insonmia	0.6363
P05	HRSD	H_psychic_anxiety	0.8721
P05	HRSD	H_late_insonmia	-0.5960
P05	HRSD	H_general_somatic_symptoms	-0.5960
P06	Smartwatch	S_daysleeper	-0.8967
P06	Smartwatch	S_ACC_day_VIG_unbt_mg	0.8935
P06	Young	Y_sleep	0.8367
P06	Young	Y_young	0.9539
P07	HRSD	H_early_insonmia	0.5837
P10	Young	Y_content	0.8783
P10	Young	Y_young	0.8886
P10	Smartwatch	S_sleeponset	-0.6876
P10	HRSD	H_late_insonmia	0.6325
P16	Self-reports	D_sleep_time	0.6168
P16	HRSD	H_depression	0.7319
P16	HRSD	H_guilt	0.8660
P16	HRSD	H_activity	0.9299

It is possible to see in Table 5 that the most correlated variables with the diagnosis were those that belong to HRSD and Young scales, as expected. They are directly related with the diagnosis itself and the specialists use them to diagnosis the patient.

#### 4.2. Principal Component Analysis

Principal component analysis (PCA) is a statistical technique that is applied to reduce the dimension of a dataset. It finds the components that define the basic structure of the dataset. In our case, the nonnumeric variables were removed before applying PCA. In addition, NA values were interpolated, and the complete dataset was normalized.

The PCA procedure results in a list of variables (“principal components”-PCx) that covers the variance of the original dataset. The list is sorted by the level of variance covered. That is, the first components cover the majority of the variance of the original dataset. Figure 4 shows, as an example, the covered variance: (a) relative and (b) cumulative, of the patient P05 dataset. It is possible to see how the first three principal components covered together more than 95% of the variance, and the first seven covered more than 99%. That means that the dimensionality of the study for patient P05 could be reduced to seven variables.

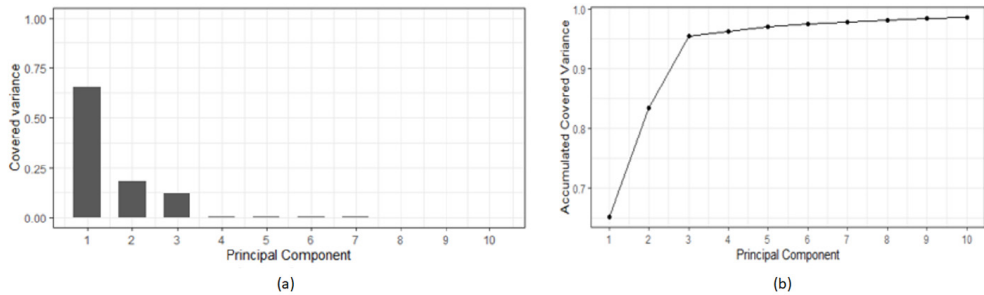


Figure 4. Principal component analysis results of patient P05 data: (a) Covered variance; (b) accumulated variance.

4.3. General Settings and Configuration of the ML Techniques

Supervised ML algorithms were applied to the available data. The dataset was split into training and testing datasets. The training set consisted of 70% observations randomly selected, and the testing data consisted of the 30% remaining observations. Several mathematical models were obtained using the diagnosis as the target variable. The accuracy of these models was calculated through the confusion matrix. The accuracy was calculated by adding all the values in the diagonal of the confusion matrix and dividing the result by the total number of observations.

$$Accuracy = \frac{\sum_{i=A}^C True\ Positives(i)}{\sum All\ observations} \tag{4}$$

The code used to generate each model and the configuration of each ML method are shown in Table 6. For SVM, a classification machine was used; thus, “type parameters” is “C-classification,” with a linear kernel. Random Forest used 500 as the number of trees to grow, and the minimum size of the nodes for classification was one. The “method” selected for the Decision Trees technique was “class,” to perform a classification. The default parameter method of Logistic Regression was “glm.fit,” which uses iteratively weighted least squares (IWLS). The “family parameter” was binomial. The binomial family accepts different links such as logit, probit, and cauchit, which correspond to logistic, normal, and Cauchy CDFs, respectively.

Table 6. Parametrization used for the generation of machine learning models.

Algorithm	Code Statements
Support Vector Machine	<code>formula &lt;- as.formula(paste0(var_diagnosis,"~.")) modelo_svm &lt;- svm(formula,data = datos.entreno.alg,type = 'C-classification',kernel = 'linear')</code>
Random Forest	<code>formula &lt;- as.formula(paste0(var_diagnosis,"~.")) modelo_rf &lt;- randomForest(formula,data = datos.entreno.alg)</code>
Decision Tree	<code>formula &lt;- as.formula(paste0(var_diagnosis,"~.")) modelo_dt &lt;- rpart(formula,data = datos.entreno.alg,method = 'class')</code>
Logistic Regression	<code>formula &lt;- as.formula(paste0(var_diagnosis,"~.")) modelo_glm &lt;- glm(formula,data = datos.entreno.alg,family = binomial)</code>

5. Results and Discussion

The application of ML techniques was intended to help confirm the diagnosis of the psychiatrist. ML methods acquire the diagnosis based on learning from collected variables

automatically. All the techniques applied performed a supervised classification process, and, thus, they assigned a new observation to a specific mood state. First, the four selected ML techniques were applied to all the variables. Table 7 shows the obtained percentage of accuracy of each model for some of the patients.

**Table 7.** Accuracy (%) of every ML model obtained with all the variables for some of the patients.

Algorithm	P03	P05	P06	P07	P10	P16
Support Vector Machine	80.00	57.14	100.00	69.23	80.00	66.67
Random Forest	100.00	100.00	80.00	69.23	100.00	80.00
Decision Tree	80.00	100.00	60.00	46.15	80.00	100.00
Logistic Regression	80.00	85.71	100.00	38.46	100.00	53.33

From Table 7, it is possible to conclude that, though Random Forest gave the best results, for some patients (P06, P16), other techniques worked better. This means that none of the algorithms were best for all patients. Selecting a specific one for each patient gave the best results.

The same ML techniques were applied to the variables selected by the feature selection algorithms. Variables whose rank was over 3 were used, i.e., those that were selected by at least three FS procedures. Table 8 shows the obtained percentage of accuracy of each model for the same patients of Table 7.

**Table 8.** Accuracy (%) of every ML model obtained with only the feature-selected variables.

Algorithm	P03	P05	P06	P07	P10	P16
Support Vector Machine	100.00	85.71	100.00	61.54	100.00	73.33
Random Forest	100.00	100.00	100.00	84.62	100.00	93.33
Decision Tree	80.00	100.00	60.00	46.15	80.00	100.00
Logistic Regression	100.00	85.71	80.00	76.92	100.00	93.33

The results from Table 8 show that using feature selection before applying machine learning techniques, in almost all cases, provided better results than using all the variables for classification, without selecting the most relevant ones.

Comparing the results with previous studies on a similar dataset, in [28], the best algorithm in terms of accuracy was Random Forest (mean of 72%). The present work reached a higher accuracy, with an average of 80% with that technique. The main improvement was due to the addition of objective variables (smartwatch variables) and by the application of feature selection algorithms prior to classification.

Finally, the classification results using only three smartwatch variables (Table 9) and adding two variables from the self-daily report (Table 10) were obtained for comparison purposes. These accelerometer variables were control sleep time, wake-up time, and daytime activity; the chosen self-report variables were irritability and motivation.

**Table 9.** Accuracy (%) of the classification of mood states based on 3 selected smartwatch variables.

Algorithm	P01	P03	P05	P06	P07	P08	P10	P13
Support Vector Machine	68.08	53.33	67.44	59.09	60.41	70.00	80.00	82.60
Random Forest	100.00	83.33	88.37	86.36	58.33	77.50	90.00	100.00
Decision Tree	97.87	66.66	76.74	68.18	54.16	65.00	80.00	82.60
Logistic Regression	61.70	46.66	72.09	72.72	62.5	70.00	90.00	82.60

**Table 10.** Accuracy (%) of the classification of mood states based on 3 smartwatch variables and 2 daily form variables preselected.

Algorithm	P01	P03	P05	P06	P07	P08	P10	P13
Support Vector Machine	87.23	76.67	76.74	77.27	64.58	62.50	80.00	86.95
Random Forest	100.00	100.00	88.37	95.45	64.58	65.00	90.00	100.00
Decision Tree	93.61	86.67	81.39	77.27	58.33	62.50	80.00	86.95
Logistic Regression	93.61	73.33	79.07	90.91	62.50	62.50	80.00	73.91

From Tables 9 and 10, it can be observed that in 80% of the cases, the classification improved when the variables of the daily self-report were added. The best model was, most prominently, RF. The best-classified cases corresponded to patients with two states (Euthymia-Mixed and Euthymia-Mania) or three states (Euthymia-Mania-Depression). The worse results could have been a consequence of the lack of collaboration of the patients in completing the daily self-report (such as reporting inaccurate information about their present state). It is also shown that some of the patients collaborated when they were in the euthymic state only (that is, in absence of any crisis). This was another limitation of such a methodology pipeline and should be considered in the interpretation of the results.

**6. Conclusions and Future Work**

This study has allowed us to gain another perspective of ML application to detect BDD states from relevant factors extracted from both objective and subjective data, which can facilitate future patients’ treatments. Currently, there are many people who suffer from this mental illness, which also affects their immediate environment and their skills in ordinary life.

In this work, heterogeneous sources of information have been combined. Data from patients diagnosed with BDD have been obtained through smartwatches, questionnaires, and standard medical consultations with the specialist. Several ML techniques have been applied with different configurations to classify the different states present in BDD patients, and to validate the most relevant factors for future detection.

The present results, regardless of the level of accuracy in successfully confirmed diagnosis, have allowed us to draw the following conclusions.

- The study required the active participation of the patients to collect the data from at least two sources of information. However, it was found that the information given by some of the patients in self-reports are sometimes not reliable, especially when the patient is in a specific mood state (mania). In this sense, the people who live with or take care of them can have a crucial role in objectifying some of the factors that they are asked about in the questionnaires.
- Many BDD patients do not use smartwatches frequently, despite being noninvasive devices. They forget to put them on at times when it is important to collect information (sleep and activity). Again, relatives or caregivers can help to remind them to wear those devices.
- There are some patients who actively participate (reporting accurate data) when they are in a state of euthymia only.
- The analysis of the data of each patient must be carried out individually, and it is only possible to compare and detect the changes in mood state in relation to each person. This is due to the fact that each patient manifests different levels of intensity of the symptoms that correspond to a particular emotional state. In addition, the manifestations of a certain state can change for a specific patient over time as a consequence of the changing nature of this illness.
- It is difficult to obtain the data from patients that are monitored and followed by a specialist, hence the low number of samples in the databases of this type of patient, and in general, of all those suffering from mental illness.

Regarding the methods and the results, a reduced set of smartwatch variables have been proven to be good enough to obtain good results in the classification of the different states of the disease. Furthermore, these results are improved when reliable information provided by the patient via self-reports are available.

This study confirms that a simple, low-cost, and noninvasive device could help to assess the correct diagnosis for BDD and to relate symptoms with different mood changes in a feasible way. In order to acquire a deeper understanding of this mental disorder with very complex dynamics, much larger datasets are necessary.

Therefore, further works include acquiring larger datasets, and to apply this methodology to patients with other psychiatric disorders. Besides, inspired by [7,45,46], other classification methods, e.g., the C4.5 classification algorithm or neural network, could be tested on the dataset. This will allow us to design a computational tool for health decision making.

**Author Contributions:** Conceptualization, V.L and M.S.; methodology, V.L. software, P.L.; validation, P.L., V.L., M.S. and M.Č.; formal analysis, V.L.; investigation, P.L.; writing—original draft preparation, P.L. and V.L.; writing—review and editing, M.S. and M.Č.; supervision, M.S. and M.Č.; funding acquisition, V.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** The APC was funded by CUNEF University, Spain.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Project Bip4Cast Ref. A-83-4155904.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors thank the Hospital Nuestra Sra. de la Paz (Madrid) for the data.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Mathur, S.; Sutton, J. Personalized medicine could transform healthcare. *Biomed. Rep.* **2017**, *7*, 3–5. [[CrossRef](#)] [[PubMed](#)]
2. Inkster, B.; Digital Mental Health Data Insights Group (DMHDIG). Early Warning Signs of a Mental Health Tsunami: A Coordinated Response to Gather Initial Data Insights from Multiple Digital Services Providers. *Front. Digit Health* **2021**, *2*, 578902. [[CrossRef](#)]
3. Čukić, M.; Stokić, M.; Radenković, S.; Ljubisavljević, M.; Simić, S.; Savić, D. Nonlinear analysis of EEG complexity in episode and remission phase of recurrent depression. *Int. J. Res. Methods Psychiatry* **2019**, *2019*, e1816. [[CrossRef](#)]
4. Singh, T.; Rajput, M. Misdiagnosis of Bipolar Disorder. *Psychiatry Edmont* **2006**, *3*, 57–63. [[PubMed](#)]
5. Patel, R.; Reiss, P.; Shetty, H.; Broadbent, M.; Stewart, R.; McGuire, P.; Taylor, M. Do antidepressants increase the risk of mania and bipolar disorder in people with depression? A retrospective electronic case register cohort study. *BMJ Open* **2015**, *5*. [[CrossRef](#)]
6. Murnane, E.L.; Cosley, D.; Chang, P.; Guha, S.; Frank, E.; Gay, G.; Matthews, M. Self-monitoring practices, attitudes, and needs of individuals with bipolar disorder: Implications for the design of technologies to manage mental health. *J. Am. Med. Inform. Assoc.* **2016**, *23*, 477–484. [[CrossRef](#)]
7. Sahoo, A.K.; Pradhan, C.; Das, H. Performance evaluation of different machine learning methods and deep-learning based convolutional neural network for health decision making. In *Nature Inspired Computing for Data Science*; Springer: Cham, Switzerland, 2020; pp. 201–212.
8. Farias, G.; Santos, M.; López, V. Making decisions on brain tumor diagnosis by soft computing techniques. *Soft Comput.* **2010**, *14*, 1287–1296. [[CrossRef](#)]
9. Chugh, G.; Kumar, S.; Singh, N. Survey on Machine Learning and Deep Learning Applications in Breast Cancer Diagnosis. *Cogn. Comput.* **2021**, 1–20. [[CrossRef](#)]
10. Ak, M.F. A comparative analysis of breast cancer detection and diagnosis using data visualization and machine learning applications. *Healthcare* **2020**, *8*, 111. [[CrossRef](#)]
11. Manrique-Cordoba, J.; Romero-Ante, J.D.; Vivas, A.; Vicente, J.M.; Sabater-Navarro, J.M. Mathematical modeling of food intake and insulin infusion in a patient with type 1 Diabetes in closed loop. *Rev. Iberoam. Autom. E Inform. Ind.* **2020**, *17*, 156–168.
12. Rghioui, A.; Lloret, J.; Sendra, S.; Oumnad, A. A Smart Architecture for Diabetic Patient Monitoring Using Machine Learning Algorithms. *Healthcare* **2020**, *8*, 348. [[CrossRef](#)] [[PubMed](#)]
13. Uddin, S.; Khan, A.; Hossain, M.E.; Moni, M.A. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med. Inform. Decis. Mak.* **2019**, *19*, 1–16. [[CrossRef](#)] [[PubMed](#)]

14. Graham, S.; Depp, C.; Lee, E.E.; Nebeker, C.; Tu, X.; Kim, H.C.; Jeste, D.V. Artificial intelligence for mental health and mental illnesses: An overview. *Curr. Psychiatry Rep.* **2019**, *21*, 1–18. [[CrossRef](#)] [[PubMed](#)]
15. Cho, G.; Yim, J.; Choi, Y.; Ko, J.; Lee, S.H. Review of machine learning algorithms for diagnosing mental illness. *Psychiatry Investig.* **2019**, *16*, 262. [[CrossRef](#)]
16. Thieme, A.; Belgrave, D.; Doherty, G. Machine learning in mental health: A systematic review of the HCI literature to support the development of effective and implementable ML systems. *ACM Trans. Comput. Hum. Interact. Tochi* **2020**, *27*, 1–53. [[CrossRef](#)]
17. Chekroud, A.M.; Zotti, R.J.; Shehzad, Z.; Gueorguieva, R.; Johnson, M.K.; Trivedi, M.H.; Cannon, T.D.; Krystal, J.H.; Corlett, P.R. Cross-trial prediction of treatment outcome in depression: A machine learning approach. *Lancet Psychiatry* **2016**, *3*, 243–250. [[CrossRef](#)]
18. Costello, F.J.; Kim, C.; Kang, C.M.; Lee, K.C. Identifying High-Risk Factors of Depression in Middle-Aged Persons with a Novel Sons and Spouses Bayesian Network Model. *Healthcare* **2020**, *8*, 562. [[CrossRef](#)] [[PubMed](#)]
19. Oh, J.; Yun, K.; Maoz, U.; Kim, T.S.; Chae, J.H. Identifying depression in the National Health and Nutrition Examination Survey data using a deep learning algorithm. *J. Affect. Disord.* **2019**, *257*, 623–631. [[CrossRef](#)] [[PubMed](#)]
20. Phillips, M.L.; Kupfer, D.J. Bipolar disorder diagnosis: Challenges and future directions. *Lancet* **2013**, *381*, 1663–1671. [[CrossRef](#)]
21. Librenza-Garcia, D.; Kotzian, B.J.; Yang, J.; Mwangi, B.; Cao, B.; Lima, L.N.P.; Passos, I.C. The impact of machine learning techniques in the study of bipolar disorder: A systematic review. *Neurosci. Biobehav. Rev.* **2017**, *80*, 538–554. [[CrossRef](#)] [[PubMed](#)]
22. Claude, L.A.; Houenou, J.; Duchesnay, E.; Favre, P. Will machine learning applied to neuroimaging in bipolar disorder help the clinician? A critical review and methodological suggestions. *Bipolar Disord.* **2020**, *22*, 334–355. [[CrossRef](#)]
23. Moris, B.H.; Bylsma LMRottenberg, J. Does emotion predict the course of major depressive disorder? A review of prospective studies. *Br. J. Clin. Psychol.* **2009**, *48*, 255–273. [[CrossRef](#)]
24. Lima, I.M.; Peckham, A.D.; Johnson, S.L. Cognitive deficits in bipolar disorders: Implications for emotion. *Clin. Psychol. Rev.* **2018**, *59*, 126–136. [[CrossRef](#)] [[PubMed](#)]
25. Tomasik, J.; Han, S.Y.S.; Barton-Owen, G.; Mirea, D.M.; Martin-Key, N.A.; Rustogi, N.; Bahn, S. A machine learning algorithm to differentiate bipolar disorder from major depressive disorder using an online mental health questionnaire and blood biomarker data. *Transl. Psychiatry* **2021**, *11*, 1–12. [[CrossRef](#)]
26. Anwar, Y. Bipolar Disorder Predictive Model: A Study to Analyze and Predict Emotional Change Using Physiological Signals. Master's Thesis, University of Minnesota, Minneapolis, MI, USA, 2019.
27. Čukić, M.; López, V.; Pavón, J. Classification of Depression through Resting-State Electroencephalogram as a Novel Practice in Psychiatry. *J. Med. Internet Res.* **2020**, *22*, e19548. [[CrossRef](#)] [[PubMed](#)]
28. Llamocca Portella, P.; Čukić, M.; Junestrand, A.; Urgelés, D.; López, V.L. Data source analysis in mood disorder research. In Proceedings of the XVIII Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA 2018), Granada, Spain, 23–26 October 2018; pp. 893–898.
29. M'Bailara, K.; Demotes-Mainard, J.; Swendsen, J.; Mathieu, F.; Leboyer, M.; Henry, C. Emotional hyper-reactivity in normothymic bipolar patients. *Bipolar Disord.* **2009**, *11*, 63–69. [[CrossRef](#)] [[PubMed](#)]
30. Gideon, J.; Provost, E.M.; McInnis, M. Mood State prediction from speech of varying acoustic quality for individuals with Bipolar Disorder. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP, Shanghai, China, 20–25 March 2016; pp. 2359–2363.
31. López, V.; Valverde, G.; Anchiraico, J.C.; Urgeles, D. Specification of a Cad Prediction System for Bipolar Disorder. In *Uncertainty Modelling in Knowledge Engineering and Decision Making, Proceedings of the 12th International FLINS Conference, Istanbul, Turkey, 26–29 August 2016*; World Scientific: Singapore, 2016; pp. 162–167.
32. Faurholt-Jepsen, M.; Busk, J.; Þórarinsdóttir, H.; Frost, M.; Bardram, J.E.; Vinberg, M.; Kessing, L.V. Objective smartphone data as a potential diagnostic marker of bipolar disorder. *Aust. N. Z. J. Psychiatry* **2019**, *53*, 119–128. [[CrossRef](#)] [[PubMed](#)]
33. Antosik-Wójcińska, A.Z.; Dominiak, M.; Chojnacka, M.; Kaczmarek-Majer, K.; Opara, K.R.; Radziszewska, W.; Świącicki, Ł. Smartphone as a monitoring tool for bipolar disorder: A systematic review including data analysis, machine learning algorithms and predictive modelling. *Int. J. Med. Inform.* **2020**, *138*, 104131. [[CrossRef](#)] [[PubMed](#)]
34. Motti, V.G. Smartwatch Applications for Mental Health: A Qualitative Analysis of the Users' Perspectives. *JMIR Ment. Health* **2018**. [[CrossRef](#)]
35. Tran, T.; Nathan-Roberts, D. Design Considerations in Wearable Technology for Patients with Bipolar Disorder. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Philadelphia, PA, USA, 1–5 October 2018; Volume 62, pp. 1187–1191.
36. Sultana, M.; Al-Jefri, M.; Lee, J. Using Machine Learning and Smartphone and Smartwatch Data to Detect Emotional States and Transitions: Exploratory Study. *JMIR Mhealth Uhealth* **2020**, *8*, e17818. [[CrossRef](#)] [[PubMed](#)]
37. Perry, A.; Roberts, G.; Mitchell, P.B.; Breakspear, M. Connectomics of bipolar disorder: A critical review, and evidence for dynamic instabilities within interoceptive networks. *Mol. Psychiatry* **2019**, *24*, 1296–1318. [[CrossRef](#)]
38. Constantinides, M.; Busk, J.; Matic, A.; Faurholt-Jepsen, M.; Kessing, L.V.; Bardram, J.E. Personalized versus generic mood prediction models in bipolar disorder. In Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, Singapore, 9–11 October 2018; pp. 1700–1707.
39. Geoffroy, P.A.; Boudebessé, C.; Bellivier, F.; Lajnef, M.; Henry, C.; Leboyer, M.; Scott, J.; Etain, B. Sleep in remitted bipolar disorder: A naturalistic case-control study using actigraphy. *J. Affect. Disord.* **2014**, *158*, 1–7. [[CrossRef](#)]

40. Wickham, H.; Averick, M.; Bryan, J.; Chang, W.; McGowan, L.D.A.; François, R.; Yutani, H. Welcome to the Tidyverse. *J. Open Source Softw.* **2019**, *4*, 1686. [[CrossRef](#)]
41. Geneactiv. Available online: <https://www.activinsights.com/technology/geneactiv/> (accessed on 20 May 2021).
42. Pavey, T.G.; Gomersall, S.R.; Clark, B.K.; Brown, W.J. The validity of the GENEActiv wrist-worn accelerometer for measuring adult sedentary time in free living. *J. Sci. Med. Sport* **2016**, *19*, 395–399. [[CrossRef](#)] [[PubMed](#)]
43. Migueles, J.H.; Rowlands, A.V.; Huber, F.; Sabia, S.; van Hees, V.T. GGIR: A Research Community–Driven Open Source R Package for Generating Physical Activity and Sleep Outcomes from Multi-Day Raw Accelerometer Data. *J. Meas. Phys. Behav.* **2019**, *2*. [[CrossRef](#)]
44. Flach, P. *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*; Cambridge University Press: Cambridge, UK, 2012; ISBN 9781107422223.
45. Paoletti, M.E.; Haut, J.M.; Plaza, J.; Plaza, A. A Comparative study of ML techniques for hyperspectral image classification. *Rev. Iberoam. Autom. E Inform. Ind.* **2019**, *16*, 129–137. [[CrossRef](#)]
46. Guevara, C.; Santos, M. Intelligent models for movement detection and physical evolution of patients with hip surgery. *Log. J. IGPL.* **2020**, jzaa032. [[CrossRef](#)]

Publicación IV

# **Weighted dependence of the day of the week in patients with emotional disorders. A mathematical model**

**Llamocca, P., López, V., Santos M.**

Publicado en *2021 IEEE International Conference on Progress in Informatics and Computing (PIC)*, 2021, pp. 420–425. DOI: 10.1109/PIC53636.2021.9687019.



# Weighted dependence of the day of the week in patients with emotional disorders. A mathematical model

Pavél Llamocca Portella  
CUNEF University and  
Complutense University of Madrid  
Madrid, Spain

pavellam@ucm.es, pavel.llamocca@cunef.edu

Victoria López  
Quantitative Methods Department  
CUNEF University  
Madrid, Spain

victoria.lopez@cunef.edu

Matilde Santos  
Institute of Knowledge Technology  
Complutense University of Madrid  
Madrid, Spain

msantos@ucm.es

**Abstract**—People who suffer from depression or bipolar disorder have very different and complex indicators of their emotional state. The use of wearable smart devices can help to characterize the behaviour of these people and therefore allows the psychiatrist to decide the best treatment. In addition, those devices are able to extract a great amount of data from patients that can be analyzed with computer techniques. However, most patients experience fluctuations in mood according to a weekly cycle. The day of the week is a factor that influences a set of characteristics that describe the emotional state, like irritability or motivation. In this work, we analyze this factor and its influence on a set of mood variables gathered daily and their relation with the medical diagnostic of the patient. The analysis of the information is personalized since the data presents variations due to factors that affect the emotional state of each patient according to different ways and intensities. This work presents an improved mathematical model on the diagnosis by including the factor described before.

**Index Terms**—data analysis, bipolar disorder, depression, emotional state, mathematical modelling

## I. INTRODUCTION

One of the main challenges in the study of mental illnesses is the lack of objective data, that usually results in a wrong diagnosis. The techniques to get accurate and objective data requires high technology that is usually expensive and not affordable. Thus, specialized staff and researchers have to count only on subjective information given by patients themselves in most of the cases. A clear example of easily misdiagnosed illness is the bipolar disorder which is often treated as unipolar disorder [1] [2]. A research by [3] states that the quality of life of a patient is significantly improved when a patient is well diagnosed and receive the proper treatment. Besides, since the COVID pandemic, the cases of mental illnesses have shown an increase [4].

The analysis of clinical data using compute techniques helps in the diagnosis of some illnesses, for example in [5] [6] the different types of brain tumors are identified by using some classification methods. In [7] A mathematical model is designed to simulate the glucose - insulin dynamics for patients with Type 1 Diabetes. Current literature relates psychological and physiological indicators to mood changes

[8]. Big data tools emerge as new alternatives to find patterns in the data from patients and this way help prevent and better diagnose these mental illnesses. Some studies have drawn interesting conclusions using machine learning (ML) algorithms to analyze those indicators. Following this line, the main goal of this work is to analyse episodes of patients diagnosed with bipolar disorder using historical data. Not only subjective but also objective data are gathered using wearable devices as well. Data come from patients who participated in the Bip4Cast Project [9].

The main problem found in the reliable application of machine learning techniques to data from patients diagnosed with any mental illness is the lack of data and the untrustworthy information [10]. In addition, different data techniques deal with data differently and get various results. That is why several ML techniques are usually applied and compared. For instance, statistical comparison among supervised machine learning algorithms such as SVM (Support Vector Machine), Naive Bayes, and Random Forest (RF) have been applied in [11], being Random Forest the one that showed the highest accuracy. Particularly in the mental health area, the most frequent techniques are : SVM, Gradient Boosting Machine, RF, Naive Bayes and K-Nearest Neighborhood (KNN) [12].

Bipolar Disorder is especially difficult to diagnose [13]. Thus, some studies applied different techniques to assess the diagnosis. For instance, deep learning (DL) and Bayesian Networks have been used to find the main causes of depression in Korea [14] [15], to help in the clinical decision-making. But most of the works accept that it is difficult to apply a single technique to the data, and that it is not easy to consider a specific algorithm better than others in general. Thus, personalized medicine now states that a study must be done for each patient. This is even more necessary due to the heterogeneity of data that are usually available [16]. A comparison between personalized and generic models is performed in [17]. It concludes that personalized models are 30% more accurate.

There are a wide variety of sources used to collect the information from bipolar patients. In [18] the authors show a

technique that complements the mental-health questionnaires with the analysis of dried blood samples. In [19] it is proposed the use of a sensor based on physiological predictors for the monitoring of bipolar patients. Other authors used electroencephalographic recordings from depressed patients [20]. All these innovative techniques are used in combination with other algorithms.

Among the variables used, some authors found that data from sleeping and daily activity are relevant for the prediction of crisis in patients with bipolar disorder. In addition, the study performed in [21] suggests that there are differences between weekdays and weekends. There are also evidence on how the mood varies on weekly cycles for bipolar patients [22]. The work here presented goes deeper in this evidence and relates variables such as sleep and activity to the day of the week. We will perform a personalized analyze for each variable and for each day in order to show if there is any relationship that may help better diagnose the mood state of the patient.

The structure of the work is as follows. Section 2 presents materials and data. The mathematical model is described in Section 3. Section 4 presents the results for some specific patients as an example of how the model includes some particular information. The paper ends with the conclusions and future works.

## II. MATERIALS: DATA AND INFORMATION SOURCES

The data were collected from patients diagnosed with Bipolar Disorder who participated in the Bip4Cast project [23]. In this work, we do not use data from healthy patients since the goal is to analyze the evolution of diagnosed patients to determine whether or not they require intervention in their treatment. Three different ways were used to collect information from those patients.

**Daily Forms.** Every day the patients filled an online form with 20 questions related to their mood state (self reporting). Those forms could be filled via web or by the mobile phone. The answers were stored in "csv" files.

**Smartwatches.** All the patients were requested to wear a smartwatch (wearable device) most of the time. It gathered information about the sleep and physical activity in 110 variables (objective information). This information was extracted from the smartwatch at each clinical session. The device used was GeneActive from ActiveInsights [24]. The quality of the measures and information was evaluated in [25]. This device generates a file which all the information that is processed with R language and the GGIR package [26].

**Clinical Observation.** The patients attend clinical sessions weekly. In those sessions, the psychiatrist interviewed the patient and collected up to 40 variables from the subjective point of view of an expert. In addition, the scales: YMRS (Young mania rating scale) and HDRS (Hamilton mania rating scale) were also filled by the psychiatrist at each session. All this information is presented in Excel files. In those clinical sessions the "crisis indicator" or diagnosis that says whether the patient was on depressed, manic, mixed or euthymic state was also determined.

Those data were collected along one year and a half (from the beginning of 2018 to mid 2019). It contains information for 101 patients. But some of them were not included in this study from the beginning while others did not complete the collection period. In addition, some of them were not using the smartwatch all the time, while others did not fill the form in a daily basis and others did not assist to the clinical evaluations. For the analysis purposes, there are 16 patients whose data is considered enough complete. Those patients were named P01,P02,...,P16.

The programming language R was used to integrate all the different sources of information. The .csv files from the daily forms as well as the .bin files from the smartwatches and the Excel files from the interviews were merged in a unique dataset. The common variables in all of them are the patient and the date. Once all the sources are integrated, the dataset has the structure shown in Table I

Variables	Type	Quantity	Source	Frequency
Patient id	CHAR	1	All	Daily
Date	DATE	1	All	Daily
Fill-in Forms	NUMERIC	20	Fill-in Forms	Daily
Smartwatch	NUMERIC	110	Smartwatches	Daily
Interviews	NUMERIC	40	Interviews	2 weeks
Diagnosis	NUMERIC	1	Interviews	2 weeks

TABLE I: Data Structure of the variables.

## III. MATHEMATICAL MODEL

The emotional state of an individual  $p$  can be modeled as a real function  $m_p(t)$  of the real variable time,  $t$ , where  $p$  means a specific patient, that reflects when sudden events cause drastic alterations in the behavior of such patient. For example, the death of a family member may cause a rapid decline to a depressed state ( $m_p(t) \gg m_p(t + \epsilon)$ ), while winning the lottery may generate a quick rise to a state of euphoria ( $m_p(t) \ll m_p(t + \epsilon)$ ). Knowledge of the  $m_p(t)$  function of each individual would represent a great advance in personalized medicine. However,  $m_p(t)$  is a very complex function and any approach will be subjected to random events such as those mentioned before. The  $m$ -function is therefore a continuous indicator of emotional state. In practice, psychiatrists use discrete indicators such as the Young's indicator, used to measure manic emotional state and the Hamilton indicator, used to measure depressive emotional state. However, there are no continuous indicators of emotional state that are applied in practice. Young and Hamilton indicators are obtained from a set of measurements recorded directly by the psychiatrist during the consultation with the patient. Through fuzzy logic and other approaches, continuous functions that combine both indicators can be developed, but it is not easy to automate it as the data collection is manual. The biggest problem in estimating a continuous indicator (such as the  $m$ -function) is the correct choice and computing of continuous variables that actually determine the status of each particular patient.

Nevertheless, the benefits of just a simple estimation ( $\hat{m}_p(t)$ ) of this function are particularly important in patients

with emotional disorders such as depression or bipolar disorder. Rough models can help the psychiatrist analyze the trend of  $m$ -function, predict a relapse, and modify the patient's medication before it occurs. In this way,  $m_p(t) \in (\alpha_p, \beta_p)$  indicates a state of normality while  $m_p(t) < \alpha_p$  is an indicator of depression and  $m_p(t) > \beta_p$  is a mania crisis indicator.

However, decision-making must be carried out on the estimation ( $\hat{m}_p(t)$ ) that is all that we can obtain. This must be calculated based on a set of measurable characteristics of the individual and the environment. We denote  $\{X_1, X_2, \dots, X_N\}$  the set of control-variables that characterize the mood state. Those variables have been previously studied in [27] and they are the inputs of the function

$$m(t) = T(X_1(t), X_2(t), \dots, X_N(t)) \quad (1)$$

where  $T$  is a transformation of  $R^N$  into  $R$ . That is, it is a statistic from which we can obtain its distribution,  $F_T(t)$ , which it is also the distribution of the  $m$ -function.

As said, in [27], a set of decision-making variables have been isolated for patients in general: motivation, irritability (recognized by the patient itself), duration of the nocturnal sleep period, time of awakening (data collected automatically by a smart device) and diagnosis (mania, hypomania, depression, euthymia, noted by the psychiatrist). In the aforementioned study, the values are collected daily, so although the emotional state represented by these variables is not strictly continuous, a continuous function can be inferred by interpolation. In the present work, the dependence of each of the mentioned variables with respect to the day of the week is analyzed, under the assumption that the patients can present substantial changes in those variables depending on the day of the week or, at least, that it will be possible to distinguish between weekdays and weekends. Under this hypothesis, the set of independent variables can be weighted according to the average variability observed on the different days of the week.

Ideally, the diagnosis function (or emotional state)  $m_p(t)$  can be calculated as a weighted aggregation of the independent variables that characterize the patient's condition. However, the reality presents several problems: variables generally have dependencies. Eliminating these dependencies implies in many cases removing relevant information. Another problem is redundancy. There are periods of time and situations that do not add any relevant information. The aggregation of these data only increases the redundancy of cases with little information.

Emotional state  $m_p(t)$  is estimated by the weighted aggregation of a selected set of variables (minimal dependence and statistical sufficiency). Thus,  $m_p(t)$  will be a function of  $n$  variables  $(X_1, \dots, X_n)$  where  $X_1, \dots, X_n$  is a subset of  $X_1, \dots, X_N$ . Therefore,  $N > n$ .

The values of each of the variables are calculated from the observed data  $x_1^i, \dots, x_n^i \forall i$ . In many cases, these data show little or no variability, so they only add noise to the study. Specifically, in data from patients with bipolar disorder, characteristics present different intensity depending on the day of the week. For example, the characteristic "irritability" can be very different on weekends than on weekdays, since there

are factors associated with the work environment that modify its value.

In the model, the characteristics are aggregated every week by the weighted aggregation of the daily average data. Weighting is done using OWA (ordered weighted averaging aggregation) operators. OWA operators were introduced by R. Yager in 1988 [28]. An OWA operator of dimension  $n$  is a function  $F$  that is associated with a vector of weights  $W = (w_1, \dots, w_n)$  so that

$$F(a_1, \dots, a_n) = \sum_{i=1}^n w_i a_{(n-i+1)} \quad (2)$$

where

$$\forall i, 0 \leq w_i \leq 1, \quad \sum_{i=1}^n w_i = 1$$

and

$$\{a_{(1)}, \dots, a_{(n)}\}$$

is the reordered increased data, that is  $a_{(k)} \leq a_{(k+1)}$

A model of the weekly estimation of the discretized diagnosis is described below. As said, values are grouped weekly:

$$Week_k = [N(\hat{X}_1^k), \dots, N(\hat{X}_n^k)] \quad (3)$$

where  $N(\hat{X}_i^k)$  is the normalization of the discretized estimate of the variable and it is defined as follows.

$$\hat{X}_i^k = \sum_{j=1}^7 \omega_j^i \cdot x_{(7-j+1)}^{i,k} = OWA_i(x_1^{i,k}, \dots, x_7^{i,k}) \quad (4)$$

where each term  $x_d^{i,k}$  is the value of the variable  $X_i$  on the week day  $d \in \{1..7\}$ . Therefore,  $\hat{X}_i^k$  is an estimation of the variable  $X_i$  on each week  $k$ . The OWA vector  $(\omega_1, \dots, \omega_7)$  sets the weight of each variable  $X_i$ . This vector makes it possible to configure the aggregation of the variables, giving different weight to the data according to their relevance for a diagnosis. A high value for the characteristic "irritability" is an indicator of a state of mania, while a low value can be an indicator of a state of depression. The model allows associating a different vector of weights for each variable.

The vector  $Week_k = [N(\hat{X}_1^k), \dots, N(\hat{X}_n^k)]$  shows an image of the emotional state that can be interpreted by the psychiatrist. But the aggregation of these variables allows us to approximate the emotional state of the patient raised at the beginning of the model. This last aggregation can be defined by means of a weighted average as follows.

$$m_p(\hat{t}_k) = \sum_{i=1}^n (w_i \times N(\hat{X}_i^k)) \quad (5)$$

This equation is an estimation of the emotional state on week  $k$ . In the following section, a configuration of the OWA operators is tested over some of the available data.

#### IV. RESULTS AND DISCUSSION

After applying a feature selection process, a small set of variables have been considered: Motivation (degree of motivation that the patient manifests on a specific day, with a variation range of -3 to 3), irritability (degree of irritability within a range of 1 to 5), sleep duration and wake-up time. Figure 1 illustrates a summary of the statistics of each variable using boxplots for every day of the week for patient P03. The boxplots show the differences among the variance of the different variables. This kind of visualization and analytic help decide the weight of the variables for a particular patient.

The variability of these variables was also analyzed for each day of the week and different dependencies were observed for each patient. In fact, patient state information is enriched by analyzing data by weekly time windows. Figure 2 shows the variation of the wake-up-time variable for one of the patients with the day of the week, which justifies the selection of the week day factor as an additional factor in the analysis of a patient's data. In this figure, the mood state of the patient is represented by a color code (pink-depression, green-euthymia, blue-mania, and purple-mixed).

It is possible to see how the first episode of mania takes place on weekdays (Monday to Thursday) and it is quite different from the rest of the week, especially Saturday and Sunday, where the variable presents a lower variability. Something similar is observed when focusing attention on the different episodes. According to the model proposed in the previous section, a vector of weights suitable for the weekly representation of this variable should give more importance to the working days where there is more variability.

In this work, a normalization based on standard deviation of the variables is the criteria used to determine the coefficients of the dim-7 OWA operator when predicting episodes of mania, hypomania or mixed. This specific decision criteria is made as follows.

- Step 1. Euthymic parameters: with the data from euthymia episodes, compute mean ( $\mu_p$ ) and standard deviation ( $\sigma_p$ ) per each variable.
- Step 2. Comparison: Let  $(d_1, \dots, d_7)$  be the values of the variable on an specific week. Compute  $K = \#(d_i > \mu_p + \sigma_p)$ .
- Step 3. OWA coefficients: If  $K > 0$ , compute  $w_i = 1/K, \forall i \in \{1..K\}$ ,  $w_i = 0, \forall i \in \{(K+1)..7\}$ . If  $K = 0$ ,  $w_i = 0, \forall i \in \{1..6\}$  and  $w_7 = 1$ .

The calculation turns to 0 the NA values. This process can be adapted in the prediction of the depressive state in a future work.

In order to evaluate the model, patients with different types of crisis has been considered: Patient P03. To illustrate the way of working of the model, the specific week 2018-04-09 to 2018-04-15 is selected. Table II shows the details of the parametrization. Each variable is weighted by a scaled of its standard deviation, that is shown in the second column. The third column shows the values (Monday to Sunday) of the specific week in which the patient had been diagnosed with

Mixed state. The 4th column shows the OWA weight vector computed according to the specific decision criteria described before, as well as parameter  $K$  in column 5th. Last column shows the weighted average value of each variable, regarding formula (4).

For example, patient P03 presents irritability average of 1.35 with standard deviation of 0.62 along the euthymia period. During this week, the patient got the following records of irritability: 3, 4, NA, 3, NA, NA, 5.  $K = 4$  out of the 7 are greater than  $1.35 + 0.65$ . That is why the OWA vector of coefficients become 0.25, 0.25, 0.25, 0.25, 0, 0, 0. The OWA-average for Patient P03 on this week will be 3.75 for the evaluation of the variable *irritability*.

The *Motivation* variable presents very low values in that week, none of them exceeds the established limit (sum of the mean and the variance in euthymia). Therefore the resulting OWA operator is the minimum. Almost the opposite occurs with the variable *Sleep duration*, where only the value of Monday exceeds the limit, therefore the resulting OWA operator is the maximum. The *Wakeuptime* variable splits the weight between the two maximum values, since  $k = 2$  in that case.

#### V. CONCLUSIONS AND FUTURE WORK

The emotional disorder manifests itself in very different ways for each individual. External factors may modify the behavior of a patient and the values that can be measure may be very different. Indeed, the emotional state is determined by the characteristics of the individual and their interaction with the environment. Hence, it is necessary to create customized models for each individual. In this work, a model is proposed that collects the relationship of the emotional state weighted with respect to the control variables and the days of the week. The proposed model can be applied to any patient regardless of the type of emotional disorder. As future work, it is proposed to validate the model with a larger set of patients and various types of emotional disorders, as well as to mature the aggregation model of the weighted estimates of the variables.

#### ACKNOWLEDGMENT

Partially supported by the grant PID2020-113192GB-I00 (Mathematical Visualization: Foundations, Algorithms and Applications) from the Spanish MICINN.

#### REFERENCES

- [1] T. Singh and M. Rajput, "Misdiagnosis of bipolar disorder," *Psychiatry (Edmont)*, vol. 3, no. 10, p. 57, 2006.
- [2] R. Patel, P. Reiss, H. Shetty, M. Broadbent, R. Stewart, P. McGuire, and M. Taylor, "Do antidepressants increase the risk of mania and bipolar disorder in people with depression? a retrospective electronic case register cohort study," *BMJ open*, vol. 5, no. 12, p. e008341, 2015.
- [3] S. Mathur and J. Sutton, "Personalized medicine could transform health-care," *Biomedical reports*, vol. 7, no. 1, pp. 3-5, 2017.
- [4] B. Inkster *et al.*, "Early warning signs of a mental health tsunami: A coordinated response to gather initial data insights from multiple digital services providers," *Frontiers in Digital Health*, vol. 2, p. 64, 2021.
- [5] G. Farias, M. Santos, and V. López, "Making decisions on brain tumor diagnosis by soft computing techniques," *Soft Computing*, vol. 14, no. 12, pp. 1287-1296, 2010.

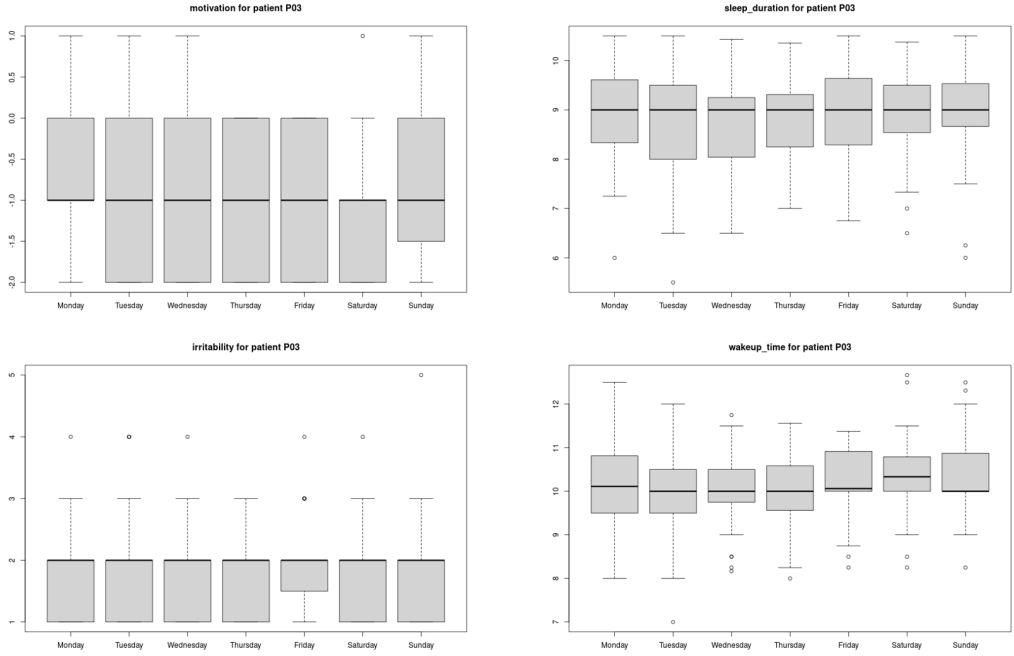


FIG. 1: Variability of different features for patient P03

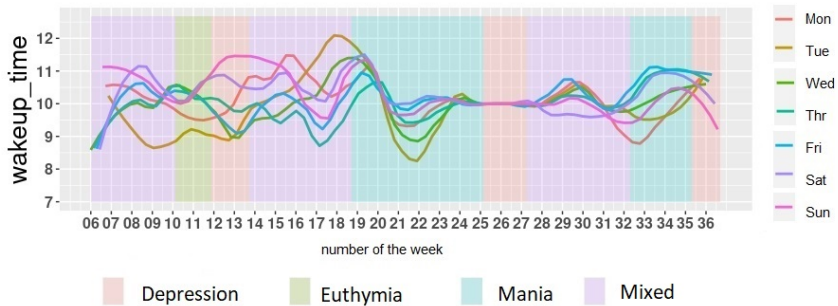


FIG. 2: Wake-up time for each day, patient P03

- [6] G. Farias, M. Santos, and V. Lopez, "Brain tumour diagnosis with wavelets and support vector machines," in *2008 3rd International Conference on Intelligent System and Knowledge Engineering*, vol. 1. IEEE, 2008, pp. 1453–1459.
- [7] J. Manrique-Cordoba, J. Romero-Ante, A. Vivas, J. Vicente, and J. Sabater-Navarro, "Mathematical modeling of food intake and insulin infusion in a patient with type 1 diabetes in closed loop," *REVISTA IBEROAMERICANA DE AUTOMATICA E INFORMATICA INDUSTRIAL*, vol. 17, no. 2, pp. 156–168, 2020.
- [8] M. Ćukić, M. Stokić, S. Radenković, M. Ljubisavljević, S. Simić, and D. Savić, "Nonlinear analysis of eeg complexity in episode and remission phase of recurrent depression," *International journal of methods in psychiatric research*, vol. 29, no. 2, p. e1816, 2020.
- [9] P. Llamocca, D. Urgelés, M. Cukic, and V. Lopez, "Bip4cast: Some advances in mood disorders data analysis," in *Proceedings of the 1st International Alan Turing Conference on Decision Support and Recommender Systems, London*, 2019.
- [10] S. Graham, C. Depp, E. E. Lee, C. Nebeker, X. Tu, H.-C. Kim, and D. V. Jeste, "Artificial intelligence for mental health and mental illnesses: an overview," *Current psychiatry reports*, vol. 21, no. 11, pp. 1–18, 2019.
- [11] S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, "Comparing different supervised machine learning algorithms for disease prediction,"

TABLE II: Parametrization for patient P03 week 2018/04/09 - 2018/04/15

Variable Name ( $\mu_i, \sigma_i$ )	Weight ( $w_i$ )	Week values	OWA vector	K value	Weighted Average ( $X_i^k$ )
Motivation (-0.475, 0.98)	0.239	(0, 0, 0, 0, 0, 0, -1)	(0, 0, 0, 0, 0, 0, 1)	0	-1
Irritability (1.35, 0.65)	0.225	(3,4,NA,3,NA, NA, 5)	(0.25, 0.25, 0.25, 0.25, 0, 0, 0)	4	3.75
Sleep duration (8.52, 1.03)	0.271	(10.5, 8.5, NA, 8.5, 8.5, NA, 9)	(1, 0, 0, 0, 0, 0, 0)	1	10.5
Wakeup time (10.13, 1.14)	0.265	(13, 10.5, NA, 11, NA, NA, 12)	(0.5, 0.5, 0, 0, 0, 0, 0)	2	12.5

*BMC medical informatics and decision making*, vol. 19, no. 1, pp. 1–16, 2019.

[12] G. Cho, J. Yim, Y. Choi, J. Ko, and S.-H. Lee, “Review of machine learning algorithms for diagnosing mental illness,” *Psychiatry investigation*, vol. 16, no. 4, p. 262, 2019.

[13] M. L. Phillips and D. J. Kupfer, “Bipolar disorder diagnosis: challenges and future directions,” *The Lancet*, vol. 381, no. 9878, pp. 1663–1671, 2013.

[14] F. J. Costello, C. Kim, C. M. Kang, and K. C. Lee, “Identifying high-risk factors of depression in middle-aged persons with a novel sons and spouses bayesian network model,” in *Healthcare*, vol. 8, no. 4. Multidisciplinary Digital Publishing Institute, 2020, p. 562.

[15] J. Oh, K. Yun, U. Maoz, T.-S. Kim, and J.-H. Chae, “Identifying depression in the national health and nutrition examination survey data using a deep learning algorithm,” *Journal of affective disorders*, vol. 257, pp. 623–631, 2019.

[16] D. Librenza-Garcia, B. J. Kotzian, J. Yang, B. Mwangi, B. Cao, L. N. P. Lima, M. B. Bermudez, M. V. Boeira, F. Kapczynski, and I. C. Passos, “The impact of machine learning techniques in the study of bipolar disorder: a systematic review,” *Neuroscience & Biobehavioral Reviews*, vol. 80, pp. 538–554, 2017.

[17] M. Constantinides, J. Busk, A. Matic, M. Faurholt-Jepsen, L. V. Kessing, and J. E. Bardram, “Personalized versus generic mood prediction models in bipolar disorder,” in *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, 2018, pp. 1700–1707.

[18] J. Tomasik, S. Y. S. Han, G. Barton-Owen, D.-M. Mirea, N. A. Martin-Key, N. Rustogi, S. G. Lago, T. Olmert, J. D. Cooper, S. Ozcan *et al.*, “A machine learning algorithm to differentiate bipolar disorder from major depressive disorder using an online mental health questionnaire and blood biomarker data,” *Translational psychiatry*, vol. 11, no. 1, pp. 1–12, 2021.

[19] Y. Anwar, “Bipolar disorder predictive model: A study to analyze and predict emotional change using physiological signals,” 2019.

[20] M. Čukić, V. López, and J. Pavón, “Classification of depression through resting-state electroencephalogram as a novel practice in psychiatry,” *Journal of Medical Internet Research*, vol. 22, no. 11, p. e19548, 2020.

[21] R. M. Ryan, J. H. Bernstein, and K. W. Brown, “Weekends, work, and well-being: Psychological need satisfactions and day of the week effects on mood, vitality, and physical symptoms,” *Journal of social and clinical psychology*, vol. 29, no. 1, pp. 95–122, 2010.

[22] J. Proudfoot, A. E. Whitton, G. Parker, V. Manicavasagar, J. Nicholas, and M. Smith, “Evidence of weekly cyclicality in mood and functional impairment in those with a bipolar disorder,” *Psychiatry research*, vol. 218, no. 3, pp. 290–294, 2014.

[23] V. López, G. VALVERDE, J. C. Anchiraico, and D. URGELES, “Specification of a cad prediction system for bipolar disorder,” in *Uncertainty Modelling in Knowledge Engineering and Decision Making: Proceedings of the 12th International FLINS Conference*. World Scientific, 2016, pp. 162–167.

[24] Activinsights, “Geneactive,” 2021, last accessed 20 May 2021. [Online]. Available: <https://www.activinsights.com/technology/geneactiv/>

[25] T. G. Pavey, S. R. Gomersall, B. K. Clark, and W. J. Brown, “The validity of the geneactiv wrist-worn accelerometer for measuring adult sedentary time in free living,” *Journal of science and medicine in sport*, vol. 19, no. 5, pp. 395–399, 2016.

[26] J. H. Migueles, A. V. Rowlands, F. Huber, S. Sabia, and V. T. van Hees, “Ggir: a research community-driven open source r package for generating physical activity and sleep outcomes from multi-day raw accelerometer data,” *Journal for the Measurement of Physical Behaviour*, vol. 2, no. 3, pp. 188–196, 2019.

[27] P. Llamocca, V. López, M. Santos, and M. Čukić, “Personalized characterization of emotional states in patients with bipolar disorder,” *Mathematics*, vol. 9, no. 11, p. 1174, 2021.

[28] R. Yager, “On ordered weighted averaging aggregation operators in multi-criteria decision making,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 18, pp. 183–190, 1988.

Publicación V

# The proposition for bipolar depression forecasting based on wearables data collection

**Pavél Llamocca Portella, Victoria López López, Milena Čukić**

Publicado en *Frontiers in Physiology*, 2022, volumen 12, DOI: 10.3389/fphys.2021.777137, ISSN 1664-042X.





# The Proposition for Bipolar Depression Forecasting Based on Wearable Data Collection

Pavel Llamocca<sup>1</sup>, Victoria López<sup>2</sup> and Milena Čukić<sup>3,4,5\*</sup>

<sup>1</sup> Computer Architecture Department, Complutense University of Madrid, Madrid, Spain, <sup>2</sup> Quantitative Methods Department, Cuneif University, Madrid, Spain, <sup>3</sup> Institute for Technology of Knowledge, Complutense University of Madrid, Madrid, Spain, <sup>4</sup> 3EGA, Amsterdam, Netherlands, <sup>5</sup> Department for General Physiology and Biophysics, Belgrade University, Belgrade, Serbia

## OPEN ACCESS

### Edited by:

Carlo Massaroni,  
Campus Bio-Medico University, Italy

### Reviewed by:

Beth Lewandowski,  
Glenn Research Center, United States  
Frank Pernett,  
Mid Sweden University, Sweden  
Fernando Marmolejo-Ramos,  
University of South Australia, Australia

### \*Correspondence:

Milena Čukić  
micukic@ucrm.es;  
micu@3ega.nl

### Specialty section:

This article was submitted to  
Physio-logging,  
a section of the journal  
Frontiers in Physiology

**Received:** 14 September 2021

**Accepted:** 29 November 2021

**Published:** 25 January 2022

### Citation:

Llamocca P, López V and Čukić M  
(2022) The Proposition for Bipolar  
Depression Forecasting Based on  
Wearable Data Collection.  
*Front. Physiol.* 12:777137.  
doi: 10.3389/fphys.2021.777137

Bipolar depression is treated wrongly as unipolar depression, on average, for 8 years. It is shown that this mismedication affects the occurrence of a manic episode and aggravates the overall condition of patients with bipolar depression. Significant effort was invested in early detection of depression and forecasting of responses to certain therapeutic approaches using a combination of features extracted from standard and online testing, wearables monitoring, and machine learning. In the case of unipolar depression, this approach yielded evidence that this data-based computational psychiatry approach would be helpful in clinical practice. Following a similar pipeline, we examined the usefulness of this approach to foresee a manic episode in bipolar depression, so that clinicians and family of the patient can help patient navigate through the time of crisis. Our projects combined the results from self-reported daily questionnaires, the data obtained from smart watches, and the data from regular reports from standard psychiatric interviews to feed various machine learning models to predict a crisis in bipolar depression. Contrary to satisfactory predictions in unipolar depression, we found that bipolar depression, having more complex dynamics, requires personalized approach. A previous work on physiological complexity (complex variability) suggests that an inclusion of electrophysiological data, properly quantified, might lead to better solutions, as shown in other projects of our group concerning unipolar depression. Here, we make a comparison of previously performed research in a methodological sense, revisiting and additionally interpreting our own results showing that the methodological approach to mania forecasting may be modified to provide an accurate prediction in bipolar depression.

**Keywords:** bipolar depression, detection, forecasting, wearables, telehealth, physiological complexity

**Abbreviations:** EEG, Electroencephalogram; ECG, Electrocardiogram; HRV, Heart rate variability; CVC, Cardio-vagal control; HFD, Higuchi Fractal Dimension; DFA, Detrended Fluctuation Analysis; ROC, Receiver operating characteristic; AUC, Area under the curve; PCA, Principal component analysis; GA, Genetic algorithm; FDA, Functional data analysis; LASO, the name of the algorithm; a type of linear regression that uses shrinkage; ANOVA, Analysis of variance; SVM, Support vector machines.

## INTRODUCTION

Those who suffer from bipolar depressive disorder (BDD) are often misdiagnosed with unipolar depression and treated as such in average for 8 years (Singh and Rajput, 2006; Lloyd et al., 2011). In addition, there are findings suggesting that antidepressant medication can aggravate their condition (Patel et al., 2015; Robillard et al., 2021). Bipolar disorder in its various forms affects 2.4% of the population of the world (Merikangas et al., 2011; World Health Organization [WHO], 2017). It is a recurrent mood disorder that produces everything from extreme euphoria to severe depression. It is accompanied by alterations in thought and behavior and can produce psychotic symptoms, such as delusions and hallucinations. People who suffer from it have a high risk of suicide, 20 times more than general population (Baldessarini et al., 2020). Even with treatment, more than a third of patients will suffer at least one relapse in the first year after diagnosis and more than 60% will have a new crisis in the first 2 years. It is a disease that typically appears during adolescence or early adulthood, affecting the person throughout his/her entire life (World Health Organization [WHO], 2018). Pharmacological treatment is the main pillar in the approach to this debilitating disease. It aims to shorten crises and prevent their occurrence but the medication has serious side effects, especially at high doses. It is therefore particularly important to detect the onset of a crisis as soon as possible. Rapid treatment of a new crisis can make a big difference in the overall effectiveness. However, this early detection is very difficult from a current standardized clinical approach. At the beginning of a crisis, the symptoms and changes can be very subtle, almost impossible to notice. It is very challenging to differentiate between unipolar and bipolar depression. We showed that the detection of unipolar depression is possible by combination of machine learning and non-linear characterization of electroencephalographic (EEG) signals (Čukić et al., 2020a,b,c; Čukić and Lopez, 2020). Additionally, we demonstrated that with the same methodological approach, it is possible to differentiate between two phases of the disease, episode, and remission (Čukić et al., 2019), which can have immense significance for clinical decisions.

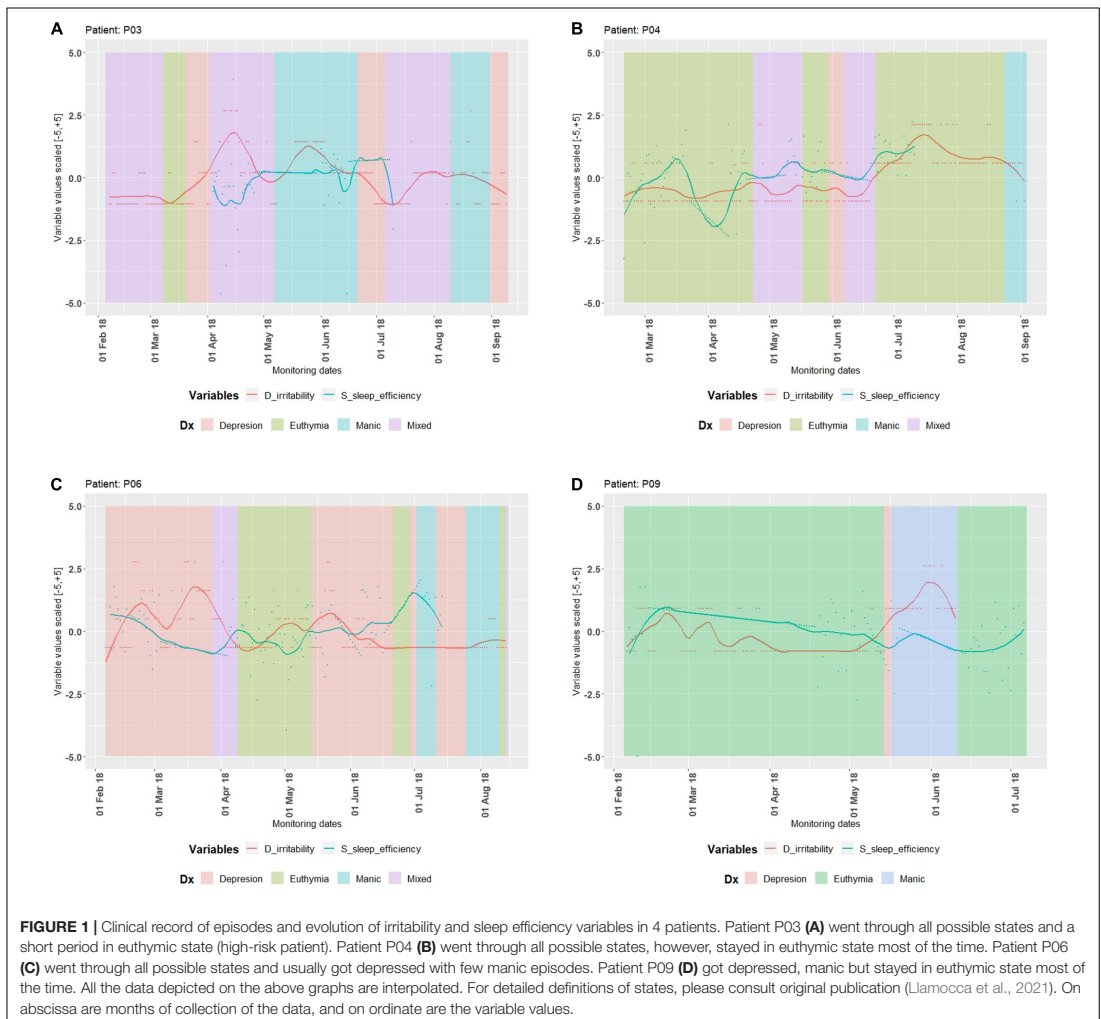
A common denominator at the onset of crises is the change in sleep and activity pattern. Weeks before the crisis, there are always changes in these variables (Llamocca et al., 2021). The early detection of these changes would allow for the improved possibility of social and occupational integration of the patients and would also allow for the decrease of the dose of drug needed for stabilization.

In our previous work, we dealt with prediction of the occurrence of crisis in BDD based on actigraphy measurements combined with standard reports from psychiatrists and self-report data obtained from outpatients *via* a mobile application (Llamocca et al., 2018, 2019, 2021). We used a number of methods for feature selection and a number of machine learning models that were previously applied in similar detection tasks (Llamocca et al., 2021). In conclusion, we stated that this methodology led to a real precision medicine application. It was shown that non-linear analysis of electrophysiological data could be used for monitoring state of patients with bipolar depression

(Pincus, 2006; Migliorini et al., 2012; Moon et al., 2013; Nardelli et al., 2017; Byun et al., 2019). Spectral and non-linear biomarkers extracted from ECG are corresponding to the aberrations of the autonomous nervous system (ANS) of patients, but also to the severity of the disease. The relation between variability of heart rate (VHR) and depression is well described (Kemp et al., 2010, 2011, 2012). Based on the non-linear analysis of ECG (as a robust marker of vagal control), it is possible to differentiate between comorbid disorders (Kemp et al., 2012) or subtypes of depression (Kemp et al., 2014), and to point to the unreported suicidal ideation (Khandoker et al., 2017), an information of enormous significance for accurate diagnosis and effective treatment. We argue here that electrophysiological data (ECG measured by portable monitoring device) as a source of detection and forecast, properly characterized by non-linear measures, can be a game changer. We revisited and additionally interpret some of our already published data, important for developing an accurate warning system for the proximity of the crisis, allowing timely and appropriate action.

## COMPARATIVE ANALYSIS AND DISCUSSION

Our main aim in the most recent publication was to isolate relevant variables for BDD (irritability and duration of sleep turned out to be the most significant) and discover the relations between them (Llamocca et al., 2021). Being successful in the detection of unipolar depression states/phases, we applied the same method to BDD and revealed quite different dynamics of the disease with more phases than in unipolar depression (Llamocca et al., 2021). According to our results (based on accumulated clinical observations and advanced analytics), there are five distinct states with as many intermediary (bidirectional) states in BDD dynamics, described by directed graph approach (for more details of our methodology, please consult the original publication, Llamocca et al., 2021). We could not discuss all aspects of our results, due to the scope and the limitations of the journal. In this retrospective analysis together with additional interpretation of those results, we are discussing suggestions for improvement of the future methodology that might lead to a simpler solution, more attractive to clinicians. Due to very complex dynamics of bipolar depression, the personal analysis of every single case is still required, as in the classical personalized approach (Llamocca et al., 2021). Other research aiming at forecasting for BDD, also concluded that the time series extracted from similarly collected data are not possible to generalize since they are very *heterogenous*; this is actually preventing the automated mood forecasting in BDD (Moore et al., 2012). Moore and colleagues reported that for some patients the mania scores were always zero during the monitoring period, which is probably the effect of medication. **Figure 1** shows the periods for defined states (depression, euthymia, manic, or mixed) in which some patients were, as well as the evolution of self-report variable D irritability and the actigraph variable S sleep efficiency. For detailed definitions of states, please consult original publication (Llamocca et al., 2021). From **Figure 1**,



**FIGURE 1 |** Clinical record of episodes and evolution of irritability and sleep efficiency variables in 4 patients. Patient P03 (A) went through all possible states and a short period in euthymic state (high-risk patient). Patient P04 (B) went through all possible states, however, stayed in euthymic state most of the time. Patient P06 (C) went through all possible states and usually got depressed with few manic episodes. Patient P09 (D) got depressed, manic but stayed in euthymic state most of the time. All the data depicted on the above graphs are interpolated. For detailed definitions of states, please consult original publication (Liamocca et al., 2021). On abscissa are months of collection of the data, and on ordinate are the variable values.

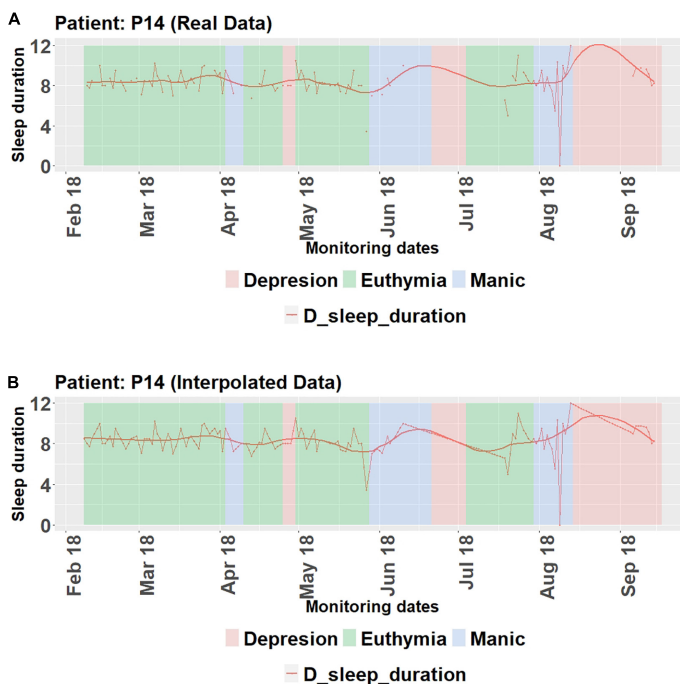
we can see different dynamics in four different patients; P03 exhibited mania and mixed state, P04 experienced euthymia and mixed state, P06 exhibited all possible states in the same period, while P09 was in the phases of long euthymia and mania, with a brief phase of depression. Although these four persons are all diagnosed with *the same clinical entity*, it is difficult to compare their dynamics as they are so different. Knowing that the mood (or states, as we labeled them) is the outcome of many complex physiological processes (that generate series of sequential data), the problem of forecasting seems to be more complicated than previously thought [in various artificial intelligence (AI) applications]. Addition of physiological complexity (fractal and non-linear) analysis to this methodology, based on our interpretation coming from Information theory,

may improve the characterization of their states leading to better crisis prediction.

One of the first authors to write about the quantitative assessment strategies in mood disorders, Steven M. Pincus, introduced a novel understanding of physiological complexity, based on his rich experience with deciphering hormonal dynamics (Pincus et al., 1996). Pincus argues that we should pay closer attention to time series that reflect essential physiological information, for there is very important history of the data, i.e., the order of samples in the time series. Pincus is the author of the Approximate Entropy algorithm (ApEn), which is a model-independent quantification of the regularity (complexity) of the data (Pincus, 1991, 1995; Pincus and Huang, 1992; Pincus and Viscarello, 1992). The fundamental difference between

regularity statistics (such as, ApEn) and conventional variability measures is that the conventional approach is focusing on tasks of quantifying the degree of spread about the central value, while the order of the input data is irrelevant; whereas in irregularity statistics, ApEn tracks changes from random to very regular and the order of samples is *essential* to the algorithm (Pincus, 2003). If we shuffle the data, the intrinsic dynamics is lost, since time series reflect the essential physiological information (Pincus, 1994). Since the sequential order of mood data is relevant to diagnosis, we must use something beyond SD and means currently used in medicine, to adequately quantify the serial nature of those data (Pincus, 2003). Since ApEn and other similar entropy measures (Shannon entropy, sample entropy, multiscale entropy, etc.) started gathering attention, various research results confirmed that they are indispensable for detecting the slightest changes in the complex physiological systems, that cannot be discovered by conventional methods. In our most recent research, we detected that among those entropy-based measures, Shannon entropy yields the best result, overperforming any previously reported conventional heart rate variability (HRV) analysis (Čukić and Savić, 2021). That has sense since Shannon entropy reflects the amount of information generated by the signal (process), which can lead to discerning the system (and its states) that is functioning in a different way than the healthy

one (Vajapeyam, 2014). ApEn can detect subclinical changes (the patterns that mostly remain undetected), unlike conventional time series analyses (Pincus et al., 1993; Pincus and Goldberger, 1994). In addition, Pincus advises a combined approach of non-linear analysis of atypical heart rate (HR) dynamics and/or EEG, given the hereditary nature of bipolar disorder (Pincus, 2003, 2006). ApEn showed to be capable of detecting changes not in peaks or amplitudes, but in underlying episodic behavior, corresponding to subsystem anatomy, feedback, or coupling (Pincus and Keefe, 1992; Pincus, 1994). It can be therefore useful to predict subsequent clinical changes, such as in mood disorders. Cook used this kind of quantification (irregularity statistics) (Cook et al., 2002) to show that patients with bipolar depression exhibit changes in EEG as a reaction to antidepressant therapy. Glenn and her colleagues managed to distinguish an episode of mania or depression, in 49 patients with bipolar disorder, from the 60 days of prior euthymia, 60 days prior the change by using ApEn algorithm on time series of self-reported state (Glenn et al., 2006). Their research showed that the larger ApEn value suggests that the 60 days prior to manic episode are more disordered (*irregular*) than the 60 days prior to a depressive episode. They argued that non-linear and linear techniques of analysis may measure different underlying components of mood changes capturing patterns that are embedded in the order of the



**FIGURE 2 |** Evolution of sleep duration variable and the states the patient P14 went through. **(A)** Real data. **(B)** Interpolated data. For detailed definitions of states, please consult the original publication (Llamocca et al., 2021). On abscissa are months of collection of the data, and on ordinate are the variable values.

data. Their research suggested that non-linear techniques should complement traditional measures to better delineate the onset (and extent) of an episode, preventing the costly hospitalization, but also the recovery from the crisis. Moore et al. (2012) noted that the quality that seems to vary among the patients with BDD they observed is so-called *roughness*, which they addressed by application of Detrended Fluctuation Analysis (DFA), a fractal methodology belonging to the family of non-linear methods of analysis. Another important study by Migliorini et al. (2012), used portable ECG sensor embedded in T-shirts, so the patients could sleep without restraints while the constant monitoring of heart dynamics was performed. The rationale here is that underlying aberrated dynamics of ANS or cortico-vagal control (known to be disrupted in mood disorders (Rothenberg, 2007), could be used for the detection and forecasting. Again, non-linear measures showed to be superior to conventional ones (see also Gottschalk et al., 1995), and the ratings extracted from signals recorded during the whole 4 nights were more accurate than the result of the standard diagnostic procedure performed before sleep (they used ML models to differentiate between BDD and healthy controls). Faurholt-Jepsen et al. (2014) showed that self-reported (labeled “subjective”) assessment was more efficient in identifying BDD states, as in using mobile technology

(smartphones) or online platforms, such as Mechanical Turk (Gillan and Whelan, 2017). Hence, some electrophysiological recording could significantly improve the chances of BDD mania prediction. Irregularity data would probably act as much more reliable features accurately representing underlying physiological information leading to better predictions. It is also important to distinguish between *detection* and *forecasting* since the latter is a much more demanding task. Having in mind that the symptoms of mood disorders are the consequences of cortico-vagal control or better, the lack of it (Rothenberg, 2007; Willner et al., 2013; Van der Kolk, 2014), non-linear measures as indicators of intrinsic dynamics (provided their sensitive quantification power) of the system are the optimal choice. Kim et al. (2013) showed that in bipolar depression, based on network analysis of EEG, there is an underlying disruption of functional connectivity.

Here, we propose two classes of methodology improvements that can result in more feasible solution for forecasting of manic episodes.

The first one is to add to the method the recording of ECG from the patients with BDD, with portable monitoring devices with medical-grade quality of signal. There are plenty of solutions, such as recording from the fingers, or from the wrists; to perform sufficiently accurate analysis, the recording

**TABLE 1** | Methods of detection and prediction of bipolar depression, used in the literature and recommended, with practical explanations and citation.

	Methods used	Methods recommended	Practical explanation
<b>Detection</b>			
1	Patient's medical history, scales, epidemiological data	Electrophysiological signals (EEG, ECG.)	Coley et al. (2021) showed that epidemiological data cannot help in prediction; physiological dynamics can
2	EEG based detection of depression	ECG based detection of depression	Portable monitoring devices for EEG are still few and expensive, those for ECG are more accessible
3	Sub-bands analysis	Broad-band analysis	There is no physiological explanation of support for importance of sub-bands
4	Small sample sizes	Larger (collaborative) sample sizes	Existing effect can be better detected with decent effect size, demonstrating practically useful results
5	Big number of variables per person	Keep the ratio under 10	Unwarranted optimism (Ng, 1997; Whelan et al., 2013)
6	ECG detected from fingers or wrist	ECG detected from the chest	Medical-grade quality of signal leads to higher accuracies of detection/prediction
7	Conventional time and frequency measures of HRV	Fractal and non-linear measures of HRV (HFD, DFA, entropy based measures, Poincare plots.)	Effect sizes for non-linear detection overperform conventional measures detection for a whole magnitude on scale (corrected Cohen's $d \sim 0.2$ vs. 7.7, Ćukić and Savić, 2021)
8	Aggressive pre-processing of electrophysiological signals	Using artifact free unfiltered signals, or Deep Learning of raw signal to correct for artifacts	By overly filtering and Fourier's decomposition (reductionistic approach) important information about history of data (sequentionality important for regularity statistics) is lost
<b>Prediction/Forecasting</b>			
1	Frequentist statistics	Bayesian approach	Improved accuracy for real life use
2	Historical medical data	Non-linear measures as feature extraction	Features based on complex systems dynamics approach lead to realistic results
3	Variation around mean values	Complex variability (physiological complexity)	Irregularity statistics is much better suitable for quantifying physiological dynamics which is non-stationary, non-linear and noisy
4	SVM and other popular ML models	LASO embedded regularisation, unsupervised learning, clustering, FDA	Practically useful prediction
5	Outliers removal	Deep learning on raw data (ECG)	Keeping the intrinsic structure of the data intact
6	Feature extraction based on t (ANOVA)	PCA, GA or FDA	Much better sensitivity and specificity
7	Non-existing external validation	ROC curve application (AUC)	More realistic results

from the chest is required. The signal should be analyzed by some of the abovementioned non-linear methods, irregularity statistics (entropy-based) and some form of fractal analysis. This kind of characterization of signal would eventually lead to much better prediction. The aim is to connect the values of certain measures/variables to certain diagnostic entities and their phases.

We are proposing recording of portable ECG, and not EEG (that was used for many EEG based depression detection in literature, among others, Alimardani and Boostani, 2018), aware of the problems in acquisition of the signal that can jeopardize the whole project. Telemedicine [with internet of things (IoT)] is gradually entering homes; outpatients are already using mobile applications, and the collection of data is easier than before. It is already shown that non-linear measures of ECG make it possible to differentiate between comorbid disorders (as shown in refs. Kemp et al., 2010, 2011; Kemp, 2011), to delineate melancholic from non-melancholic depression (Kemp et al., 2014), or to detect the suicide ideation (Khandoker et al., 2017), which is particularly important in BDD where the risk of suicide is high (20-fold risk in comparison with controls, Baldessarini et al., 2020). Those are all immensely important for the clinician to make effective treatment decisions. In addition, since sleep is disrupted in BDD, it would make sense to measure ECG during sleep (Migliorini et al., 2012). Pincus (2003) predicted that some form of sleep recording of ECG would be the most sufficient for this task (see also Saad et al., 2019).

An example from our publication (Llamocca et al., 2021) is illustrating how variables connected to sleep (sleep duration as the most significant one) are changing in relation to the state defined for that day (either as self-reported or pronounced by a clinician, since both are included in dataset). **Figure 2** shows how real data from patient P14 differ in respect to the interpolated data. We can conclude that P14 usually sleeps about 8 h in euthymic state, but this time-period varies when P14 is about to enter crisis or is already suffering from one. **Figure 2** shows the periods for states in which patient P14 was, as well as the evolution of self-report variable D sleep duration (duration of their sleep).

The second part of our proposition for further improvement of approach to prediction would be in connection to ML models. We were using various forms of supervised learning to learn from the data. The authors who are dealing with more theoretical approach to computational psychiatry (Whelan and Garavan, 2014) are advocating avoidance of 'unwarranted optimism' by collecting more data and lowering the number of variables per person (Kohavi, 1995; Tibshirani, 1996; Ng, 1997). Relying on Bayesian approaches is recommendable. Special care should be given to dimensionality problem, which our group addressed entirely (Llamocca et al., 2021). In addition, support vector machines (SVM) might be one of the most popular models, but other methods could be used, such as embedded regularization (Whelan et al., 2013). Knowing the heterogeneity problem, we suggest introducing some of unsupervised learning methods, such as subgroup discovery (which is a binary classifier and works on labeled data) and association rule discovery (which is unsupervised ML Model), or predictive and descriptive clustering (distance-based models) (Flach, 2012). The problem

with clusters can be 2-fold: either you have a trivial solution (which corresponds to overfitting in linear models, let us say clustering overfitting) that can be resolved if we penalize the large K, or if we fix the number of clusters K in advance; the problem cannot be solved for large datasets (but a typical dataset is not large). Soft clustering generalizes the notion of partition, in the same way that a probability estimator generalizes a classifier (Flach, 2012). With the abovementioned suggestion, the algorithm can learn from the (properly characterized) data. We can conclude what the subgroups are and what the relations between present instances are, so we can try to interpret them in the light of information theory approach to physiological processes.

Busk et al. (2020) used a similar manner of collecting the data, with different items in the questionnaire. They tested the feasibility of forecasting daily subjective mood scores based on daily self-assessment from 84 patients with bipolar disorders via smartphone in a randomized clinical trial. Combined historic data and currently collected data improved forecasting and used Hierarchical Bayesian approach, a multi-task learning method. They used data from different subjects as additional cases to learn. Ordinal regression (or ordinal classification) is a method of predicting a discrete variable that has a relative ordering of the possible outcomes. First, they started with 1-day forecast with several scenarios (two time-series cross-validation experiments) and applied best model to evaluate 7-day forecast. When increasing the forecast horizon, forecast errors also increased and the forecast regression shifted toward the mean of data distribution; the best model used a 4-day history of self-assessment. Interestingly, authors used similar organization of the dataset that is usually used for entropy-based analysis of physiological data, discussed above; maybe the historicity of the data would be the key for successful forecasting. Besides, some shift in ML models used for much needed realistic forecasting includes much preferred unsupervised learning or functional data analysis (Wang et al., 2016). **Table 1** is offering some recommended techniques with our justification.

We hope that an improved research methodology, based on abovementioned comparison and analysis, would eventually lead to a much better therapeutic and improve the quality of life of patients.

## AUTHOR CONTRIBUTIONS

VL developed the idea for research. PL, VL, and MČ performed the research, wrote the manuscript, and reviewed the manuscript. PL and VL collected and analyzed the data. PL generated figures. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was partially supported by the grant PID2020-113192GB-I00 (Mathematical Visualization: Foundations, Algorithms, and Applications) from the Spanish MICINN.

## REFERENCES

- Alimardani, F., and Boostani, R. (2018). DB-FFR: a modified feature selection algorithm to improve discrimination rate between bipolar mood disorder (BMD) and schizophrenic patients. *Iran. J. Sci. Technol. Trans. Electric. Eng.* 42, 251–260. doi: 10.1007/s40998-018-0060-x
- Baldessarini, R. J., Vázquez, G. H., and Tondo, L. (2020). Bipolar depression: a major unsolved challenge. *Int. J. Bipolar Disord.* 8:1. doi: 10.1186/s40345-019-0160-1
- Busk, J., Faurholt-Jepsen, M., Frost, M., Bardram, J. E., Kessing, L. V., and Winther, O. (2020). Forecasting mood in bipolar disorder from smartphone self-assessments: hierarchical bayesian approach. *JMIR Mhealth Uhealth* 8:e15028. doi: 10.2196/15028
- Byun, S., Kim, A. Y., Jang, E. H., Kim, S., Choi, K. W., Yu, H. Y., et al. (2019). Entropy analysis of heart rate variability and its application to recognize major depressive disorder: a pilot study. *Technol. Health Care* 27, 407–424. doi: 10.3233/THC-199037
- Coley, R. Y., Boggs, J. M., Beck, A., and Simon, G. E. (2021). Predicting outcomes of psychotherapy for depression with electronic health record data. *J. Affect. Disord. Rep.* 6:100198. doi: 10.1016/j.jadr.2021.100198
- Cook, I. A., Leuchter, A. F., Morgan, M., Witte, E., Stubbeman, W. F., Abrams, M., et al. (2002). Early changes in prefrontal activity characterize clinical responders to antidepressants. *Neuropsychopharmacology* 27, 120–131. doi: 10.1016/S0893-133X(02)00294-4
- Čukić, M., and Lopez, V. (2020). “On mistakes we made in prior computational psychiatry data driven approach projects and how they jeopardize translation of those findings in clinical practice,” in *IntelliSys Conference, Amsterdam 3-5 September 2020. “Advances in Intelligent Systems and Computing”*, eds K. Arai, S. Kapoor, and R. Bhatia (Berlin: Springer Verlag), 493–510. doi: 10.1007/978-3-030-55190-2\_37
- Čukić, M., and Savić, D. (2021). Another Godot who is still not coming: more on biomarkers for depression. *Revista de Psiquiatria y Salud Mental* doi: 10.1016/j.rpsm.2021.12.006 (in press)
- Čukić, M., Stokić, M., Radenković, S., Ljubisavljević, M., Simić, S., and Savić, D. (2019). Nonlinear analysis of EEG complexity in episode and remission phase of recurrent depression. *Int. J. Methods Psychiatry Res.* 29:e1816. doi: 10.1002/MPR.1816
- Čukić, M., López, V., and Pavón, J. (2020a). Classification of depression through resting-state electroencephalogram as a novel practice in psychiatry. *J. Med. Internet Res.* 22:e19548. doi: 10.2196/19548
- Čukić, M., Stokić, M., Radenković, S., Ljubisavljević, M., Simić, S., and Savić, D. (2020b). Nonlinear analysis of EEG complexity in episode and remission phase of recurrent depression. *Int. J. Methods Psychiatry Res.* 29:e1816. doi: 10.1002/MPR.1816
- Čukić, M., Stokić, M., Simić, S., and Pokrajac, D. (2020c). The successful discrimination of depression from EEG could be attributed to proper feature extraction and not to a particular classification method. *Cogn. Neurodyn.* 14, 443–455. doi: 10.1007/s11571-020-09581-x
- Faurholt-Jepsen, M., Vinberg, M., Frost, M., Christensen, E. M., Bardram, J., and Kessing, L. V. (2014). Daily electronic monitoring of subjective and objective measures of illness activity in bipolar disorder using smartphones—the MONARCA II trial protocol: a randomized controlled single-blind parallel-group trial. *BMC Psychiatry* 14:309. doi: 10.1186/s12888-014-0309-5
- Flach, P. (2012). *Machine Learning: The Art And Science Of Algorithms That Make Sense Of Data*. Cambridge: Cambridge University Press.
- Gillan, C. M., and Whelan, R. (2017). What big data can do for treatment in psychiatry. *Curr. Opin. Behav. Sci.* 18, 34–42.
- Glenn, T., Whybrow, P. C., Rasgon, N., Grof, P., Alda, M., Baethge, C., et al. (2006). Approximate entropy of self-reported mood prior to episodes in bipolar disorder. *Bipolar Disord.* 8, 424–429. doi: 10.1111/j.1399-5618.2006.00373.x
- Gottschalk, A., Bauer, M. S., and Whybrow, P. C. (1995). Evidence of chaotic mood variation in bipolar disorder. *Arch. Gen. Psychiatry* 52, 947–959. doi: 10.1001/archpsyc.1995.03950230061009
- Kemp, A. (2011). Depression, antidepressant treatment and the cardiovascular system. *Acta Neuropsychiatr.* 23, 82–83. doi: 10.1111/j.1601-5215.2011.00535.x
- Kemp, A. H., Quintana, D. S., Felmingham, K. L., Matthews, S., and Jelinek, H. F. (2012). Depression, comorbid anxiety disorders, and heart rate variability in physically healthy, unmedicated patients: implications for cardiovascular risk. *PLoS One* 7:e30777. doi: 10.1371/journal.pone.0030777
- Kemp, A. H., Quintana, D. S., Gray, M. A., Felmingham, K. L., Brown, K., and Gatt, J. M. (2010). Impact of depression and antidepressant treatment on heart rate variability: a review and meta-analysis. *Biol. Psychiatry* 67, 1067–1074. doi: 10.1016/j.biopsych.2009.12.012
- Kemp, A. H., Quintana, D. S., and Malhi, G. S. (2011). Effects of serotonin reuptake inhibitors on heart rate variability: methodological issues, medical comorbidity, and clinical relevance. *Biol. Psychiatry* 69, e25–e26. doi: 10.1016/j.biopsych.2010.10.0365
- Kemp, A. H., Quintana, D. S., Quinn, C. R., Hopkinson, P., and Harris, A. W. (2014). Major depressive disorder with melancholia displays robust alterations in resting state heart rate and its variability: implications for future morbidity and mortality. *Front. Psychol.* 5:1387. doi: 10.3389/fpsyg.2014.01387
- Khandoker, A. H., Luthra, V., Abouallaban, Y., Saha, S., Ahmed, K. I., Mostafa, R., et al. (2017). Predicting depressed patients with suicidal ideation from ECG recordings. *Med. Biol. Eng. Comput.* 55, 793–805. doi: 10.1007/s11517-016-1557-y
- Kim, D. J., Bolbecker, A. R., Howell, J., Rass, O., Sporns, O., Hetrick, W. P., et al. (2013). Disturbed resting state EEG synchronization in bipolar disorder: a graph-theoretic analysis. *NeuroImage Clin.* 2, 414–423. doi: 10.1016/j.nicl.2013.03.007
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *IJCAI* 14, 1137–1145.
- Llamocca, P., Čukić, M., Junestrand, A., Urgelés, D., and López, V. L. (2018). “Data source analysis in mood disorder research,” in *In XVIII Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA 2018) 23-26 de octubre de 2018 Granada* (España: Asociación Española para la Inteligencia Artificial (AEPIA)), 893–898.
- Llamocca, P., López, V., Santos, M., and Čukić, M. (2021). Personalized characterization of emotional states in patients with bipolar disorder. *Mathematics* 9:1174. doi: 10.3390/math911174
- Llamocca, P., Urgelés, D., Cukic, M., and Lopez, V. (2019). “Bip4Cast: some advances in mood disorders data analysis,” in *Proceedings of the 1st International Alan Turing Conference on Decision Support and Recommender Systems, London* (Berlin: Springer), 5–10.
- Lloyd, L. C., Giaroli, G., Taylor, D., and Tracy, D. K. (2011). Bipolar depression: clinically missed, pharmacologically mismanaged. *Ther. Adv. Psychopharmacol.* 1, 153–162. doi: 10.1177/2045125311420752
- Merikangas, K. R., Jin, R., He, J. P., Kessler, R. C., Lee, S., Sampson, N. A., et al. (2011). Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Arch. Gen. Psychiatry* 68, 241–251. doi: 10.1001/archgenpsychiatry.2011.12
- Migliorini, M., Mendez, M. O., and Bianchi, A. M. (2012). Study of heart rate variability in bipolar disorder: linear and non-linear parameters during sleep. *Front. Neuroeng.* 4:22. doi: 10.3389/fneng.2011.00022
- Moon, E., Lee, S. H., Kim, D. H., and Hwang, B. (2013). Comparative study of heart rate variability in patients with schizophrenia, bipolar disorder, post-traumatic stress disorder, or major depressive disorder. *Clin. Psychopharmacol. Neurosci.* 11, 137–43. doi: 10.9758/cpn.2013.11.3.137
- Moore, P. J., Little, M. A., McSharry, P. E., Geddes, J. R., and Goodwin, G. M. (2012). Forecasting depression in bipolar disorder. *IEEE Trans. Biomed. Eng.* 59, 2801–2807. doi: 10.1109/TBME.2012.2210715
- Nardelli, M., Lanata, A., Bertschy, G., Scilingo, E. P., and Valenza, G. (2017). Heartbeat complexity modulation in bipolar disorder during daytime and nighttime. *Sci. Rep.* 7:17920. doi: 10.1038/s41598-017-18036-z
- Ng, A. Y. (1997). Preventing “Overfitting” of Cross-Validation Data. Presented at the 14th International Conference on Machine Learning (ICML) (1997). Available online at: <https://www.semanticscholar.org/paper/Preventing-%22Overfitting%22-of-Cross-Validation-Data-Ng/ead7bec02e696b602705d4e73dfdd13aa4c6d57> (accessed December 30, 2021).
- Patel, R., Reiss, P., Shetty, H., Broadbent, M., Stewart, R., McGuire, P., et al. (2015). Do antidepressants increase the risk of mania and bipolar disorder in people with depression? A retrospective electronic case register cohort study. *BMJ Open* 5:e008341. doi: 10.1136/bmjopen-2015-008341
- Pincus, S. (1995). Approximate entropy (ApEn) as a complexity measure. *Chaos* 5, 110–117. doi: 10.1063/1.166092

- Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. U. S. A.* 88, 2297–2301. doi: 10.1073/pnas.88.6.2297
- Pincus, S. M. (1994). Greater signal regularity may indicate increased system isolation. *Math. Biosci.* 122, 161–181. doi: 10.1016/0025-5564(94)90056-6
- Pincus, S. M. (2003). Quantitative assessment strategies and issues for mood and other psychiatric serial study data. *Bipolar Disord.* 5, 287–294. doi: 10.1034/j.1399-5618.2003.00036.x
- Pincus, S. M. (2006). Approximate entropy as a measure of irregularity for psychiatric serial metrics. *Bipolar Disord.* 8, 430–440. doi: 10.1111/j.1399-5618.2006.00375.x
- Pincus, S. M., Cummins, T. R., and Haddad, G. G. (1993). Heart rate control in normal and aborted-SIDS infants. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 264, R638–R646. doi: 10.1152/ajpregu.1993.264.3.R638
- Pincus, S. M., and Goldberger, A. L. (1994). Physiological time-series analysis: what does regularity quantify? *Am. J. Physiol. Heart Circ. Physiol.* 266, H1643–H1656. doi: 10.1152/ajpheart.1994.266.4.H1643
- Pincus, S. M., and Huang, W. M. (1992). Approximate entropy: statistical properties and applications. *Commun. Stat. Theory Methods* 21, 3061–3077. doi: 10.1080/03610929208830963
- Pincus, S. M., and Keefe, D. L. (1992). Quantification of hormone pulsatility via an approximate entropy algorithm. *Am. J. Physiol. Endocrinol. Metab.* 262, E741–E754. doi: 10.1152/ajpendo.1992.262.5.E741
- Pincus, S. M., Mulligan, T., Iranmanesh, A., Gheorghiu, S., Godschalk, M., and Veldhuis, J. D. (1996). Older males secrete luteinizing hormone and testosterone more irregularly, and jointly more asynchronously, than younger males. *Proc. Natl. Acad. Sci. U. S. A.* 93, 14100–14105. doi: 10.1073/pnas.93.24.14100
- Pincus, S. M., and Viscarello, R. R. (1992). Approximate entropy: a regularity measure for fetal heart rate analysis. *Obstet. Gynecol.* 79, 249–255.
- Robillard, R., Saad, M., Ray, L. B., Bujáki, B., Douglass, A., Lee, E. K., et al. (2021). Selective serotonin reuptake inhibitor use is associated with worse sleep-related breathing disturbances in individuals with depressive disorders and sleep complaints: a retrospective study. *J. Clin. Sleep Med.* 17, 505–513. doi: 10.5664/jcsm.8942
- Rothenberg, J. (2007). Cardiac vagal control in depression: a critical analysis. *Biol. Psychol.* 74, 200–211. doi: 10.1016/j.biopsycho.2005.08.010
- Saad, M., Ray, L. B., Bujáki, B., Parvaresh, A., Palamarchuk, I., De Koninck, J., et al. (2019). Using heart rate profiles during sleep as a biomarker of depression. *BMC Psychiatry* 19:168. doi: 10.1186/s12888-019-2152-1
- Singh, T., and Rajput, M. (2006). Misdiagnosis of bipolar disorder. *Psychiatry (Edgmont)* 3, 57–63.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B (Methodol.)* 58, 267–288. doi: 10.1111/j.2517-6161.1996.tb02080.x
- Vajapeyam, S. (2014). Understanding Shannon's entropy metric for information. *arXiv [Preprint]*. Available Online at: <https://arxiv.org/abs/1405.2061> (accessed December 30, 2021).
- Van der Kolk, B. (2014). *The Body Keeps The Score: Mind, Brain And Body In The Transformation Of Trauma*. New York: Penguin.
- Wang, J.-L., Chiou, J.-M., and Müller, H.-G. (2016). Review of functional data analysis. *Annu. Rev. Stat. Appl.* 3, 257–295. doi: 10.1146/annurev-statistics-041715-033624
- Whelan, R., and Garavan, H. (2014). When optimism hurts: inflated predictions in psychiatric neuroimaging. *Biol. Psychiatry* 75, 746–748. doi: 10.1016/j.biopsych.2013.05.014
- Willner, P., Scheel-Krüger, J., and Belzung, C. (2013). The neurobiology of depression and antidepressant action. *Neurosci. Biobehav. Rev.* 37, 2331–2371. doi: 10.1016/j.neubiorev.2012.12.007
- World Health Organization [WHO] (2017). *Depression And Other Common Mental Disorders*. Available Online at: <http://apps.who.int/iris/bitstream/10665/254610/1/WHO-MSD-MER-2017.2-eng.pdf> (accessed December 30, 2021).
- World Health Organization [WHO] (2018). *International Suicide Rates, 2018*. Available Online at: [https://www.who.int/health-topics/suicide#tab=tab\\_1](https://www.who.int/health-topics/suicide#tab=tab_1)

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Llamocca, López and Čukić. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# **Anexos**



## Anexo A

# Librerías de R utilizadas

Cuadro A.1: Librerías de R utilizadas en la investigación

Nº	Librería	Descripción	Aporte en la tesis
1	<i>read_xls</i>	Carga de ficheros excel	Cargar todas las plantillas en excel
2	<i>impute_ts</i>	Manejo de valores vacíos	Realizar interpolaciones de datos
3	<i>outliers</i>	Manejo de valores atípicos	Remover outliers de los self-report
4	<i>scales</i>	Generación de vectores escalados	Escalar variables para la generación de gráficos
5	<i>lubridate</i>	Manejo de fechas	Calcular el número de semana de cada observación
6	<i>data.table</i>	Tratamiento complejo de data.frames	Cargar en bloque los ficheros csv generados a partir de los bin (smartwatches)
7	<i>zoo</i>	Manejo de valores vacíos	Realizar el “arrastre” de datos
8	<i>boot</i>	Técnica de bootstrapping	Ejecutar bootstrapping
9	<i>dplyr</i>	Utilidades varias para manejo de datos	Modificar los datasets (mutate) previo a la generación de gráficos
10	<i>tidyr</i>	Manipulación de dataframes	Realizar transposiciones de data.frames
11	<i>ggir</i>	Procesos para analizar datos de acelerómetros	Procesamiento de los ficheros binarios de smartwatches
12	<i>ggplot2</i>	Generación de gráficos	Generación de todos los gráficos de la tesis
13	<i>shiny</i>	Desarrollo de aplicaciones	Desarrollo de aplicación interactiva para visualizar todas las variables
14	<i>rmarkdown</i>	Conversión de scripts en reportes de diversos formatos	Generación de reportes visuales en formato pdf

*Continúa en la página siguiente*

Cuadro A.1 – *Continuada de la página anterior*

Nº	Librería	Descripción	Aporte en la tesis
15	<i>e1071</i>	ML: SVM	Ejecución del algoritmo de ML: Máquina de vectores de soporte
16	<i>glmnet</i>	ML: RLG FS: LASSO	Ejecución del algoritmo de ML: Regresión logística Ejecución del algoritmo de FS: Least Absolute Shrinkage and Selection Operator
17	<i>randomForest</i>	ML: RF	Ejecución del algoritmo de ML: Random Forest
18	<i>rpart</i>	ML: DT	Ejecución del algoritmo de ML: Árboles de decisión
19	<i>class</i>	ML: KNN	Ejecución del algoritmo de ML: K-nearest neighbors
20	<i>relaimpo</i>	FS : Relative Importance	Ejecución de algoritmo de FS: Relative importance
21	<i>caret</i>	FS : Variable Importance	Ejecución de algoritmo de FS: Variable importance
22	<i>boruta</i>	FS : Boruta	Ejecución de algoritmo de FS: boruta
23	<i>dalex</i>	FS : Dalex	Ejecución de algoritmo de FS: Descriptive mAchine Learning EXplanations

## Anexo B

# Procesamiento de ficheros binarios (GGIR)

```
1 library(GGIR)
2
3 # Path where the bin inputfiles are stored
4 dirtemp_in_bin="C:/Datos/input/.."
5 # Path where the csv outputfiles will be stored
6 dirtemp_out_csv="C:/Datos/output/.."
7
8 g.shell.GGIR(# INPUT NEEDED:
9     mode=c(1,2,3,4,5),
10    datadir=dirtemp_in_bin,
11    outputdir=dirtemp_out_csv,
12    f0=1, f1=2,
13    #-----
14    # Part 1:
15    #-----
16    do.enmo = TRUE,                do.anglez=TRUE,
17    chunksize=1,                  printsummary=TRUE,
18    #-----
19    # Part 2:
20    #-----
21    strategy = 3,                  ndayswindow=300,
22    hrs.del.start = 0,             hrs.del.end = 0,
23    maxdur = 900,                  includedaycrit = 16,
24    winhr = c(5,10),
25    qllevels = c(c(1380/1440),c(1410/1440)),
26    qwindow=c(0,24),
27    ilevels = c(seq(0,400,by=50),8000),
28    mvpathreshold =c(100,120),
29    bout.metric = 4,
30    closedbout=FALSE,
31    #-----
32    # Part 3:
33    #-----
34    timethreshold= c(5),           anglethreshold=5,
35    ignorenonwear = TRUE,
36    #-----
37    # Part 4:
38    #-----
39    excludefirstlast = TRUE,       includenightcrit = 16,
40    def.noc.sleep = c(),           outliers.only = FALSE,
41    criterror = 4,                 relyonsleeplog = FALSE,
42    sleeplogidnum = TRUE,          colid=1,
43    coln1=2,                       do.visual = TRUE,
44    nnight = 200 ,
45    #-----
46    # Part 5:
47    #-----
```

## B. Procesamiento de ficheros binarios (GGIR)

---

```
48     threshold.lig = c(30), threshold.mod = c(100), threshold.vig = c(400),
49     boutcriter = 0.8,      boutcriter.in = 0.9,      boutcriter.lig = 0.8,
50     boutcriter.mvpa = 0.8, boutdur.in = c(1,10,30), boutdur.lig = c(1,10),
51     boutdur.mvpa = c(1),  timewindow = c("WW"),
52     #-----
53     # Report generation
54     #-----
55     do.report=c(2,4,5),
56     visualreport=TRUE,    dofirstpage = TRUE,
57     viewingwindow=1)
```

## Anexo C

# Ejecución de algoritmos de *machine learning*

```
1 library(dplyr)
2 library(arm)
3 library(DataExplorer)
4 library(class)
5 library(rpart)
6 library(glmnet)
7 library(e1071)
8 library(nnet)
9
10 # El dataset dftotal es el que contiene todas las fuentes integradas
11 dftotal<- "data.RData"
12 # Variable objetivo
13 targetvariable<- "Z_dx_numeric"
14 formula<-as.formula(paste0(targetvariable, "~. "))
15
16 # Especificamos que un 80% de los datos seran para entrenar
17 # Especificamos que el 20% restante seran para probar
18 split <- round(nrow(dftotal) * 0.80)
19 datos.entreno.alg<-dftotal[1:split, ]
20 datos.test.alg<-dftotal[(split + 1):nrow(dftotal), ]
21
22
23 #####
24 # SVM Support Vector Machines
25 #####
26 modelo_ml<-svm(formula,data=datos.entreno.alg,type='C-classification',kernel='
  linear')
27 predictions <- predict(modelo_ml, datos.test.alg)
28
29 #####
30 # Random Forest
31 #####
32 modelo_ml<-randomForest(formula,data=datos.entreno.alg)
33 predictions <- predict(modelo_ml, datos.test.alg)
34
35 #####
36 # Decission Trees
37 #####
38 modelo_ml<-rpart(formula,data=datos.entreno.alg,method='class')
39 predictions <- predict(modelo_ml, datos.test.alg)
40
41 #####
42 # Regresion Logistica
43 #####
44 modelo_ml<-glm(formula,data=datos.entreno.alg,family=binomial)
45 predictions <- predict(modelo_glm, test,type="response")
46 predictions<-round(predictions,0)
```

## C. Ejecución de algoritmos de *machine learning*

---

```
47
48 #####
49 # KNN k-nearest neighbors
50 #####
51 train<-datos.entreno.alg[!(names(datos.entreno.alg) %in% targetvariable)]
52 test<-datos.test.alg[!(names(datos.test.alg) %in% targetvariable)]
53
54 train_category <- datos.entreno.alg[[targetvariable]]
55 test_category <- datos.test.alg[[targetvariable]]
56
57 modelo_ml<-knn(train,test,cl=train_category,k=5)
58 predictions<-modelo_knn
59
60
61
62 # Mostramos la matriz de confusion (falsos positivos, falsos negativos,
63   verdaderos positivos, verdaderos negativos)
64 (mc<-table(predictions,datos.test.alg[[targetvariable]]))
65
66 # Mostramos la confianza
67 (confidence<-100 * sum(diag(mc)) / sum(mc))
```

## Anexo D

# VARIABLES DEL *self-report*

Cuadro D.1: Variables del *self-report*

Nº	Fuente	<i>Dataset</i>	<i>Nombre Variable</i>	Descripción o posibles valores
1	Self-report	<i>diario</i>	<i>D_alcohol_consumption</i>	Consumo de alcohol
2	Self-report	<i>diario</i>	<i>D_animus</i>	-3 Grado máximo de depresión -2 Grado intermedio de depresión -1 Depresión leve, apenas perceptible 0 Eutimia o normalidad 1 Euforia o irritabilidad leve, apenas perceptible 2 Grado intermedio de euforia o irritabilidad 3 Grado máximo de euforia o irritabilidad
3	Self-report	<i>diario</i>	<i>D_anxiety</i>	1 Ausente 2 Leve y muy ocasional 3 Moderada 4 Grave y continua 5 Muy grave, incapacitante
4	Self-report	<i>diario</i>	<i>D_caffeine</i>	Consumo de café

*Continúa en la página siguiente*

Cuadro D.1 – *Continúa de la página anterior*

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
5	Self-report	<i>diario</i>	<i>D_concentration</i>	1 Buena atención y concentración
				2 Leves problemas apenas perceptibles
				3 Problemas moderados, entre el punto 2 y el 4
				4 Dificultades importantes, es difícil seguir una película o leer un texto
				5 Alteración grave, imposibilidad de mantener una conversación
6	Self-report	<i>diario</i>	<i>D_drugs_consumption</i>	Consumo de drogas
7	Self-report	<i>diario</i>	<i>D_irritability</i>	1 Muy tranquilo, inalterable, ninguna discusión
				2 Leve tendencia a enfadarse, apenas perceptible
				3 Moderada irritabilidad, claramente apreciable
				4 Muy irritable, dificultades para contenerse
				5 Irritabilidad extrema, incluso agresividad física
8	Self-report	<i>diario</i>	<i>D_motivation</i>	-3 grado mínimo de motivación
				-2 Grado intermedio de motivación
				-1 Leve disminución en la motivación, apenas perceptible
				0 Motivación o energía media, normal
				1 Motivación o ganas de hacer cosas un poco aumentadas
				2 Grado intermedio de motivación
3 Grado máximo de energía, actividad y motivación				
9	Self-report	<i>diario</i>	<i>D_observation_string_length</i>	Longitud de comentarios

*Continúa en la página siguiente*

Cuadro D.1 – Continúa de la página anterior

Nº	Fuente	<i>Dataset</i>	<i>Nombre Variable</i>	Descripción o posibles valores
10	Self-report	<i>diario</i>	<i>D_sleep_duration</i>	Duración del sueño
11	Self-report	<i>diario</i>	<i>D_sleep_quality</i>	Calidad del sueño
12	Self-report	<i>diario</i>	<i>D_sleep_time</i>	Hora de sueño
13	Self-report	<i>diario</i>	<i>D_sleep_zenith</i>	Hora de la mitad del periodo de sueño
14	Self-report	<i>diario</i>	<i>D_slowness_fill_form</i>	Tiempo para llenar el formulario
15	Self-report	<i>diario</i>	<i>D_slowness_fill_form_no_outliers</i>	<i>D_slowness_fill_form</i> sin outliers
16	Self-report	<i>diario</i>	<i>D_slowness_fill_form_no_outliers_scales</i>	<i>D_slowness_fill_form</i> escalada entre 1 y 5
17	Self-report	<i>diario</i>	<i>D_start</i>	Hora de inicio del formulario (timestamp)
18	Self-report	<i>diario</i>	<i>D_time_start_form</i>	Hora de inicio del formulario expresado en números
19	Self-report	<i>diario</i>	<i>D_tobacco</i>	Consumo de tabaco
20	Self-report	<i>diario</i>	<i>D_wakeup_time</i>	Hora de despertar



## Anexo E

# Variables del *smartwatch*

Cuadro E.1: Variables del *smartwatch*

Nº	Fuente	<i>Dataset</i>	<i>Nombre Variable</i>	Descripción o posibles valores
21	Smartwatch	<i>pulsera</i>	<i>S_acc_available</i>	(TRUE/FALSE) datos de smartwatch disponibles
22	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_- IN_bts_1_10_mg</i>	Aceleración media en estado de INACTIVIDAD (1-10 mg)
23	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_- IN_bts_10_30_- mg</i>	Aceleración media en estado de INACTIVIDAD (10-30 mg)
24	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_- IN_bts_30_mg</i>	Aceleración media en estado de INACTIVIDAD (más de 30 mg)
25	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_- IN_unbt_mg</i>	Aceleración media en estado de INACTIVIDAD (unbouted)
26	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_- LIG_bts_1_10_- mg</i>	Aceleración media en estado de ACTIVIDAD LIGERA (1-10 mg)
27	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_- LIG_bts_10_mg</i>	Aceleración media en estado de ACTIVIDAD LIGERA (más de 10 mg)

*Continúa en la página siguiente*

Cuadro E.1 – *Continúa de la página anterior*

Nº	Fuente	<i>Dataset</i>	<i>Nombre Variable</i>	Descripción o posibles valores
28	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_LIG_unbt_mg</i>	Aceleración media en estado de ACTIVIDAD LIGERA (unbouted)
29	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_mg</i>	Aceleración media total
30	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_MOD_unbt_mg</i>	Aceleración media en estado de ACTIVIDAD MODERADA (unbouted)
31	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_MV-PA_bts_1_mg</i>	Aceleración media en estado de ACTIVIDAD MODERADA o VIGOROSO (más de 1 mg)
32	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_spt_mg</i>	Aceleración media en “sleep period time” de día
33	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_total_IN_mg</i>	Aceleración media total en estado de INACTIVIDAD
34	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_total_LIG_mg</i>	Aceleración media total en estado de ACTIVIDAD LIGERA
35	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_total_MOD_mg</i>	Aceleración media total en estado de ACTIVIDAD MODERADA
36	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_total_VIG_mg</i>	Aceleración media total en estado de ACTIVIDAD VIGOROSA
37	Smartwatch	<i>pulsera</i>	<i>S_ACC_day_VIG_unbt_mg</i>	Aceleración media en estado de ACTIVIDAD VIGOROSA (unbouted)
38	Smartwatch	<i>pulsera</i>	<i>S_ACC_spt_mg</i>	Aceleración media en “sleep period time” total

*Continúa en la página siguiente*

Cuadro E.1 – Continúa de la página anterior

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
39	Smartwatch	<i>pulsera</i>	<i>S_ACC_spt_sleep_mg</i>	Aceleración media en “sleep period time” durante el sueño
40	Smartwatch	<i>pulsera</i>	<i>S_ACC_spt_wake_IN_mg</i>	Aceleración media en estado de INACTIVIDAD en periodos “waking up”
41	Smartwatch	<i>pulsera</i>	<i>S_ACC_spt_wake_LIG_mg</i>	Aceleración media en estado de ACTIVIDAD LIGERA en periodos “waking up”
42	Smartwatch	<i>pulsera</i>	<i>S_ACC_spt_wake_MOD_mg</i>	Aceleración media en estado de ACTIVIDAD MODERADA en periodos “waking up”
43	Smartwatch	<i>pulsera</i>	<i>S_ACC_spt_wake_VIG_mg</i>	Aceleración media en estado de ACTIVIDAD VIGOROSA en periodos “waking up”
44	Smartwatch	<i>pulsera</i>	<i>S_bout.metric</i>	Métrica para calcular periodos de actividad
45	Smartwatch	<i>pulsera</i>	<i>S_boutcriter.in</i>	Fracción de periodos para llegar al umbral del estado de INACTIVIDAD
46	Smartwatch	<i>pulsera</i>	<i>S_boutcriter.lig</i>	Fracción de periodos para llegar al umbral del estado de ACTIVIDAD LIGERA
47	Smartwatch	<i>pulsera</i>	<i>S_boutcriter.mvpa</i>	Fracción de periodos para llegar al umbral del estado de ACTIVIDAD MODERADA o VIGOROSA
48	Smartwatch	<i>pulsera</i>	<i>S_boutdur.in</i>	Duración de los estados de INACTIVIDAD
49	Smartwatch	<i>pulsera</i>	<i>S_boutdur.lig</i>	Duración de los estados de ACTIVIDAD LIGERA
50	Smartwatch	<i>pulsera</i>	<i>S_boutdur.mvpa</i>	Duración de los estados de ACTIVIDAD MODERADA o VIGOROSO
51	Smartwatch	<i>pulsera</i>	<i>S_cleaningcode</i>	Código de problemas para la evaluación del sueño

Continúa en la página siguiente

Cuadro E.1 – *Continúa de la página anterior*

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
52	Smartwatch	<i>pulsera</i>	<i>S_daysleeper</i>	0: La persona duerme de noche; 1: La persona duerme de día
53	Smartwatch	<i>pulsera</i>	<i>S_daytype</i>	WD: día laborable; WE: fin de semana
54	Smartwatch	<i>pulsera</i>	<i>S_dur_day_IN_bts_1_10_min</i>	Duración total de los periodos de INACTIVIDAD entre 1 y 10 minutos
55	Smartwatch	<i>pulsera</i>	<i>S_dur_day_IN_bts_10_30_min</i>	Duración total de los periodos de INACTIVIDAD entre 10 y 30 minutos
56	Smartwatch	<i>pulsera</i>	<i>S_dur_day_IN_bts_30_min</i>	Duración total de los periodos de INACTIVIDAD de más de 30 minutos
57	Smartwatch	<i>pulsera</i>	<i>S_dur_day_IN_unbt_min</i>	Duración total de los periodos de INACTIVIDAD (unbouted)
58	Smartwatch	<i>pulsera</i>	<i>S_dur_day_LIG_bts_1_10_min</i>	Duración total de los periodos de ACTIVIDAD LIGERA entre 1 y 10 minutos
59	Smartwatch	<i>pulsera</i>	<i>S_dur_day_LIG_bts_10_min</i>	Duración total de los periodos de ACTIVIDAD LIGERA de más de 10 minutos
60	Smartwatch	<i>pulsera</i>	<i>S_dur_day_LIG_unbt_min</i>	Duración total de los periodos de ACTIVIDAD LIGERA (unbouted)
61	Smartwatch	<i>pulsera</i>	<i>S_dur_day_min</i>	Duración de las horas despiertas durante el día
62	Smartwatch	<i>pulsera</i>	<i>S_dur_day_MOD_unbt_min</i>	Duración total de los periodos de ACTIVIDAD MODERADA (unbouted)

*Continúa en la página siguiente*

Cuadro E.1 – *Continúa de la página anterior*

Nº	Fuente	<i>Dataset</i>	<i>Nombre Variable</i>	Descripción o posibles valores
63	Smartwatch	<i>pulsera</i>	<i>S_dur_day_MV-PA_bts_1_min</i>	Duración total de los periodos de ACTIVIDAD MODERADA o VIGOROSA de más de 1 minuto
64	Smartwatch	<i>pulsera</i>	<i>S_dur_day_spt_min</i>	Duración de periodos de “waking up” y “sleep period time”
65	Smartwatch	<i>pulsera</i>	<i>S_dur_day_total_IN_min</i>	Duración total en periodos de INACTIVIDAD
66	Smartwatch	<i>pulsera</i>	<i>S_dur_day_total_LIG_min</i>	Duración total en periodos de ACTIVIDAD LIGERA
67	Smartwatch	<i>pulsera</i>	<i>S_dur_day_total_MOD_min</i>	Duración total en periodos de ACTIVIDAD MODERADA
68	Smartwatch	<i>pulsera</i>	<i>S_dur_day_total_VIG_min</i>	Duración total en periodos de ACTIVIDAD VIGOROSA
69	Smartwatch	<i>pulsera</i>	<i>S_dur_day_VIG_unbt_min</i>	Duración de periodos de ACTIVIDAD VIGOROSA (unbouted)
70	Smartwatch	<i>pulsera</i>	<i>S_dur_spt_min</i>	Duración del “sleep period time”
71	Smartwatch	<i>pulsera</i>	<i>S_dur_spt_sleep_min</i>	Duración del periodo de sueño en el “sleep period time”
72	Smartwatch	<i>pulsera</i>	<i>S_dur_spt_wake_IN_min</i>	Duración del periodo de “waking up” en INACTIVIDAD durante el “sleep period time”
73	Smartwatch	<i>pulsera</i>	<i>S_dur_spt_wake_LIG_min</i>	Duración del periodo de “waking up” en ACTIVIDAD LIGERA durante el “sleep period time”

*Continúa en la página siguiente*

Cuadro E.1 – *Continúa de la página anterior*

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
74	Smartwatch	<i>pulsera</i>	<i>S_dur_spt_wake_MOD_min</i>	Duración del periodo de “waking up” en ACTIVIDAD MODERADA durante el “sleep period time”
75	Smartwatch	<i>pulsera</i>	<i>S_dur_spt_wake_VIG_min</i>	Duración del periodo de “waking up” en ACTIVIDAD VIGOROSA durante el “sleep period time”
76	Smartwatch	<i>pulsera</i>	<i>S_filename</i>	Nombre del fichero de datos
77	Smartwatch	<i>pulsera</i>	<i>S_guider</i>	Guía para diferenciar periodos de inactividad entre día y noche
78	Smartwatch	<i>pulsera</i>	<i>S_L10TIME</i>	Momento de las 10 horas menos activas (timestamp)
79	Smartwatch	<i>pulsera</i>	<i>S_L10TIME_num</i>	Momento de las 10 horas menos activas expresado en números
80	Smartwatch	<i>pulsera</i>	<i>S_L10VALUE</i>	Valor de aceleración de las 10 horas menos activas
81	Smartwatch	<i>pulsera</i>	<i>S_L5TIME</i>	Momento de las 5 horas menos activas (timestamp)
82	Smartwatch	<i>pulsera</i>	<i>S_L5TIME_num</i>	Momento de las 5 horas menos activas expresado en números
83	Smartwatch	<i>pulsera</i>	<i>S_L5VALUE</i>	Valor de aceleración de las 5 horas menos activas
84	Smartwatch	<i>pulsera</i>	<i>S_M10TIME</i>	Momento de las 10 horas más activas (timestamp)
85	Smartwatch	<i>pulsera</i>	<i>S_M10TIME_num</i>	Momento de las 10 horas más activas expresado en números
86	Smartwatch	<i>pulsera</i>	<i>S_M10VALUE</i>	Valor de aceleración de las 10 horas más activas
87	Smartwatch	<i>pulsera</i>	<i>S_M5TIME</i>	Momento de las 5 horas más activas (timestamp)
88	Smartwatch	<i>pulsera</i>	<i>S_M5TIME_num</i>	Momento de las 5 horas más activas expresado en números
89	Smartwatch	<i>pulsera</i>	<i>S_M5VALUE</i>	Valor de aceleración de las 5 horas más activas

*Continúa en la página siguiente*

Cuadro E.1 – *Continúa de la página anterior*

Nº	Fuente	<i>Dataset</i>	<i>Nombre Variable</i>	Descripción o posibles valores
90	Smartwatch	<i>pulsera</i>	<i>S_N_- atleast5minwakenight</i>	Número de periodos de “waking up” superiores a 5 minutos
91	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_day_- IN_bts_1_10</i>	Número de bloques de INACTIVIDAD de entre 1 y 10 minutos
92	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_day_- IN_bts_10_30</i>	Número de bloques de INACTIVIDAD de entre 10 y 30 minutos
93	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_day_- IN_bts_30</i>	Número de bloques de INACTIVIDAD de más de 30 minutos
94	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_day_- IN_unbt</i>	Número de bloques de INACTIVIDAD (unbouted)
95	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_day_- LIG_bts_1_10</i>	Número de bloques de ACTIVIDAD LIGERA de entre 1 y 10 minutos
96	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_day_- LIG_bts_10</i>	Número de bloques de ACTIVIDAD LIGERA de más de 10 minutos
97	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_day_- LIG_unbt</i>	Número de bloques de ACTIVIDAD LIGERA (unbouted)
98	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_day_- MOD_unbt</i>	Número de bloques de INACTIVIDAD (unbouted)
99	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_day_- MVPA_bts_1</i>	Número de bloques de ACTIVIDAD MODERADA o VIGOROSO de más de 1 minuto

*Continúa en la página siguiente*

Cuadro E.1 – *Continúa de la página anterior*

Nº	Fuente	<i>Dataset</i>	<i>Nombre Variable</i>	Descripción o posibles valores
100	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_day_ - total_IN</i>	Número de bloques de INACTIVIDAD
101	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_day_ - total_LIG</i>	Número de bloques de ACTIVIDAD LIGERA
102	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_day_ - total_MOD</i>	Número de bloques de ACTIVIDAD MODERADA
103	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_day_ - total_VIG</i>	Número de bloques de ACTIVIDAD VIGOROSA
104	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_day_ - VIG_unbt</i>	Número de bloques de ACTIVIDAD VIGOROSA (unbouted)
105	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_spt_ - sleep</i>	Número de bloques de sueño dentro del “sleep period time”
106	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_spt_ - wake_IN</i>	Número de bloques en INACTIVIDAD dentro del “sleep period time”
107	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_spt_ - wake_LIG</i>	Número de bloques en ACTIVIDAD LIGERA dentro del “sleep period time”
108	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_spt_ - wake_MOD</i>	Número de bloques en ACTIVIDAD MODERADA dentro del “sleep period time”
109	Smartwatch	<i>pulsera</i>	<i>S_Nblocks_spt_ - wake_VIG</i>	Número de bloques en ACTIVIDAD VIGOROSA dentro del “sleep period time”
110	Smartwatch	<i>pulsera</i>	<i>S_Nbouts_day_ - IN_bts_1_10</i>	Número de periodos de INACTIVIDAD de entre 1 y 10 minutos

*Continúa en la página siguiente*

Cuadro E.1 – Continúa de la página anterior

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
111	Smartwatch	<i>pulsera</i>	<i>S_Nbouts_day - IN_bts_10_30</i>	Número de periodos de INACTIVIDAD de entre 10 y 30 minutos
112	Smartwatch	<i>pulsera</i>	<i>S_Nbouts_day - IN_bts_30</i>	Número de periodos de INACTIVIDAD de más de 30 minutos
113	Smartwatch	<i>pulsera</i>	<i>S_Nbouts_day - LIG_bts_1_10</i>	Número de periodos de ACTIVIDAD LIGERA de entre 1 y 10 minutos
114	Smartwatch	<i>pulsera</i>	<i>S_Nbouts_day - LIG_bts_10</i>	Número de periodos de ACTIVIDAD LIGERA de más de 10 minutos
115	Smartwatch	<i>pulsera</i>	<i>S_Nbouts_day - MVPA_bts_1</i>	Número de periodos de ACTIVIDAD MODERADA o VIGOROSO de más de 1 minuto
116	Smartwatch	<i>pulsera</i>	<i>S_night_number</i>	Número de noche
117	Smartwatch	<i>pulsera</i>	<i>S_nonwear_perc - day</i>	Porcentaje del día que no se usa el dispositivo
118	Smartwatch	<i>pulsera</i>	<i>S_nonwear_perc - day_spt</i>	Porcentaje del día y noche que no se usa el dispositivo
119	Smartwatch	<i>pulsera</i>	<i>S_nonwear_perc - spt</i>	Porcentaje de la noche que no se usa el dispositivo
120	Smartwatch	<i>pulsera</i>	<i>S_quantile_mos-tactive30min_mg</i>	Cuantil más activo de 30 minutos
121	Smartwatch	<i>pulsera</i>	<i>S_quantile_mos-tactive60min_mg</i>	Cuantil más activo de 60 minutos
122	Smartwatch	<i>pulsera</i>	<i>S_sleep_efficiency</i>	Eficiencia del sueño
123	Smartwatch	<i>pulsera</i>	<i>S_sleeplog_used</i>	(TRUE/FALSE) se utilizó el sleep log

Continúa en la página siguiente

Cuadro E.1 – *Continúa de la página anterior*

Nº	Fuente	<i>Dataset</i>	<i>Nombre Variable</i>	Descripción o posibles valores
124	Smartwatch	<i>pulsera</i>	<i>S_sleeponset</i>	Momento de inicio del sueño expresado en números
125	Smartwatch	<i>pulsera</i>	<i>S_sleeponset_ts</i>	Momento de inicio del sueño (timestamp)
126	Smartwatch	<i>pulsera</i>	<i>S_wakeup</i>	Momento del despertar expresado en números
127	Smartwatch	<i>pulsera</i>	<i>S_wakeup_ts</i>	Momento del despertar (timestamp)
128	Smartwatch	<i>pulsera</i>	<i>S_week</i>	Número de semana
129	Smartwatch	<i>pulsera</i>	<i>S_weekday</i>	Día de semana
130	Smartwatch	<i>pulsera</i>	<i>S_window_number</i>	Número de día de registro

## Anexo F

# VARIABLES DE LAS CONSULTAS MÉDICAS - ESCALA HDRS

Cuadro F.1: Variables de las consultas médicas: escala HDRS

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
131	Consultas Médicas	<i>hdrs</i>	<i>H_activity</i>	0 No hay dificultad
				1 Ideas y sentimientos de incapacidad, fatiga o debilidad (trabajos, pasatiempos)
				2 Pérdida de interés en su actividad (disminución de la atención, indecisión y vacilación)
				3 Disminución del tiempo actual dedicado a actividades o disminución de la productividad
				4 Dejó de trabajar por la presente enfermedad. Solo se compromete en las pequeñas tareas, o no puede realizar estas sin ayuda
132	Consultas Médicas	<i>hdrs</i>	<i>H_agitation</i>	0 Ninguna
				1 Juega con sus dedos
				2 Juega con sus manos, cabello, etc.
				3 No puede quedarse quieto ni permanecer sentado
				4 Retuerce las manos, se muerde las uñas, se tira de los cabellos, se muerde los labios

*Continúa en la página siguiente*

Cuadro F.1 – *Continúa de la página anterior*

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
133	Consultas Médicas	<i>hdrs</i>	<i>H_depression</i>	0 Ausente 1 Estas sensaciones las expresa solamente si le preguntan como se siente 2 Estas sensaciones las relata espontáneamente 3 Sensaciones no comunicadas verbalmente (expresión facial, postura, voz, tendencia al llanto) 4 Manifiesta estas sensaciones en su comunicación verbal y no verbal en forma espontánea
134	Consultas Médicas	<i>hdrs</i>	<i>H_early_insonmia</i>	0 No tiene dificultad 1 Dificultad ocasional para dormir, por ej. más de media hora el conciliar el sueño 2 Dificultad para dormir cada noche
135	Consultas Médicas	<i>hdrs</i>	<i>H_gastrointestinal_somatic_symptoms</i>	0 Ninguno 1 Pérdida del apetito, pero come sin necesidad de que lo estimulen. Sensación de pesadez en el abdomen 2 Dificultad en comer si no se le insiste. Solicita o necesita laxantes o medicación intestinal o para sus síntomas gastrointestinales
136	Consultas Médicas	<i>hdrs</i>	<i>H_general_somatic_symptoms</i>	0 Ninguno 1 JPesadez en la extremidades, espalda o cabeza. Dorsalgias, cefalalgias, mialgias. Fatigabilidad y pérdida de energía 2 Cualquiera de los síntomas anteriores se puntúa como 2 si está muy bien definido

*Continúa en la página siguiente*

Cuadro F.1 – *Continúa de la página anterior*

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
137	Consultas Médicas	<i>hdrs</i>	<i>H_genitals_symptoms</i>	0 Ausentes
				1 Débiles
				2 Graves
				3 Incapacitantes
138	Consultas Médicas	<i>hdrs</i>	<i>H_guilt</i>	0 Ausente
				1 Se culpa a si mismo, cree haber decepcionado a la gente
				2 Tiene ideas de culpabilidad o medita sobre errores pasados o malas acciones
				3 Siente que la enfermedad actual es un castigo
				4 Oye voces acusatorias o de denuncia y/o experimenta alucinaciones visuales de amenaza
139	Consultas Médicas	<i>hdrs</i>	<i>H_hypochondria</i>	0 Ausente
				1 Preocupado de si mismo (corporalmente)
				2 Preocupado por su salud
				3 Se lamenta constantemente, solicita ayuda
140	Consultas Médicas	<i>hdrs</i>	<i>H_illness_awareness</i>	0 Se da cuenta que esta deprimido y enfermo
				Se da cuenta de su enfermedad pero atribuye la causa
				1 a la mala alimentación, clima, exceso de trabajo, virus, necesidad de descanso, etc.
				2 No se da cuenta que está enfermo

*Continúa en la página siguiente*

Cuadro F.1 – *Continúa de la página anterior*

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
141	Consultas Médicas	<i>hdrs</i>	<i>H_inhibition</i>	0 Palabra y pensamiento normales
				1 Ligero retraso en el habla
				2 Evidente retraso en el habla
				3 Dificultad para expresarse
				4 Incapacidad para expresarse
142	Consultas Médicas	<i>hdrs</i>	<i>H_late_insonmia</i>	0 No hay dificultad
				1 Se despierta a primeras horas de la madrugada, pero se vuelve a dormir
				2 No puede volver a dormirse si se levanta de la cama
143	Consultas Médicas	<i>hdrs</i>	<i>H_mid_insonmia</i>	0 No hay dificultad
				1 Esta desvelado e inquieto o se despierta varias veces durante la noche
				2 Esta despierto durante la noche, cualquier ocasión de levantarse de la cama se clasifica en 2 (excepto por motivos de evacuar)
144	Consultas Médicas	<i>hdrs</i>	<i>H_psychic_anxiety</i>	0 No hay dificultad
				1 Tensión subjetiva e irritabilidad
				2 Preocupación por pequeñas cosas
				3 Actitud aprensiva en la expresión o en el habla
				4 Expresa sus temores sin que le pregunten
145	Consultas Médicas	<i>hdrs</i>	<i>H_somatic_anxiety</i>	0 Ausente
				1 Ligera
				2 Moderada
				3 Severa
				4 Incapacitante

*Continúa en la página siguiente*

Cuadro F.1 – *Continúa de la página anterior*

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
146	Consultas Médicas	<i>hdrs</i>	<i>H_suicide</i>	0 Ausente
				1 Le parece que la vida no vale la pena ser vivida
				2 Desearía estar muerto o tiene pensamientos sobre la posibilidad de morirse
				3 Ideas de suicidio o amenazas
				4 Intentos de suicidio (cualquier intento serio)
147	Consultas Médicas	<i>hdrs</i>	<i>H_weigth</i>	0 Pérdida de peso inferior a 500 gr. en una semana
				1 Pérdida de más de 500 gr. en una semana
				2 Pérdida de más de 1 Kg. en una semana
148	Consultas Médicas	<i>hdrs</i>	<i>H_score</i>	Puntaje en la escala de HDRS
149	Consultas Médicas	<i>hdrs</i>	<i>H_hrsd_dx</i>	Diagnóstico según la escala. 0-7: Eutimia, 8-13: Depresión Ligera, 14-18: Depresión Moderada, 19-22: Depresión severa y >23: Muy severa



## Anexo G

# VARIABLES DE LAS CONSULTAS MÉDICAS - ESCALA YMRS

Cuadro G.1: Variables de las consultas médicas: escala YMRS

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
150	Consultas Médicas	<i>hdrs</i>	<i>H_activity</i>	0 No hay dificultad
				1 Ideas y sentimientos de incapacidad, fatiga o debilidad (trabajos, pasatiempos)
				2 Pérdida de interés en su actividad (disminución de la atención, indecisión y vacilación)
				3 Disminución del tiempo actual dedicado a actividades o disminución de la productividad
				4 Dejó de trabajar por la presente enfermedad. Solo se compromete en las pequeñas tareas, o no puede realizar estas sin ayuda
151	Consultas Médicas	<i>hdrs</i>	<i>H_agitation</i>	0 Ninguna
				1 Juega con sus dedos
				2 Juega con sus manos, cabello, etc.
				3 No puede quedarse quieto ni permanecer sentado
				4 Retuerce las manos, se muerde las uñas, se tira de los cabellos, se muerde los labios

*Continúa en la página siguiente*

Cuadro G.1 – Continúa de la página anterior

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
152	Consultas Médicas	<i>hdrs</i>	<i>H_depression</i>	0 Ausente
				1 Estas sensaciones las expresa solamente si le preguntan como se siente
				2 Estas sensaciones las relata espontáneamente
				3 Sensaciones no comunicadas verbalmente (expresión facial, postura, voz, tendencia al llanto)
				4 Manifiesta estas sensaciones en su comunicación verbal y no verbal en forma espontánea
153	Consultas Médicas	<i>hdrs</i>	<i>H_early_insonmia</i>	0 No tiene dificultad
				1 Dificultad ocasional para dormir, por ej. más de media hora el conciliar el sueño
				2 Dificultad para dormir cada noche
154	Consultas Médicas	<i>hdrs</i>	<i>H_gastrointestinal_somatic_symptoms</i>	0 Ninguno
				1 Pérdida del apetito, pero come sin necesidad de que lo estimulen. Sensación de pesadez en el abdomen
				2 Dificultad en comer si no se le insiste. Solicita o necesita laxantes o medicación intestinal o para sus síntomas gastrointestinales
155	Consultas Médicas	<i>hdrs</i>	<i>H_general_somatic_symptoms</i>	0 Ninguno
				1 JPesadez en la extremidades, espalda o cabeza. Dorsalgias, cefalalgias, mialgias. Fatigabilidad y pérdida de energía
				2 Cualquiera de los síntomas anteriores se puntúa como 2 si está muy bien definido

Continúa en la página siguiente

Cuadro G.1 – *Continúa de la página anterior*

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
156	Consultas Médicas	<i>hdrs</i>	<i>H_genitals_symptoms</i>	0 Ausentes
				1 Débiles
				2 Graves
				3 Incapacitantes
157	Consultas Médicas	<i>hdrs</i>	<i>H_guilt</i>	0 Ausente
				1 Se culpa a si mismo, cree haber decepcionado a la gente
				2 Tiene ideas de culpabilidad o medita sobre errores pasados o malas acciones
				3 Siente que la enfermedad actual es un castigo
				4 Oye voces acusatorias o de denuncia y/o experimenta alucinaciones visuales de amenaza
158	Consultas Médicas	<i>hdrs</i>	<i>H_hypochondria</i>	0 Ausente
				1 Preocupado de si mismo (corporalmente)
				2 Preocupado por su salud
				3 Se lamenta constantemente, solicita ayuda
159	Consultas Médicas	<i>hdrs</i>	<i>H_illness_awareness</i>	0 Se da cuenta que esta deprimido y enfermo
				Se da cuenta de su enfermedad pero atribuye la causa
				1 a la mala alimentación, clima, exceso de trabajo, virus, necesidad de descanso, etc.
				2 No se da cuenta que está enfermo

*Continúa en la página siguiente*

Cuadro G.1 – Continúa de la página anterior

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
160	Consultas Médicas	<i>hdrs</i>	<i>H_inhibition</i>	0 Palabra y pensamiento normales
				1 Ligero retraso en el habla
				2 Evidente retraso en el habla
				3 Dificultad para expresarse
				4 Incapacidad para expresarse
161	Consultas Médicas	<i>hdrs</i>	<i>H_late_insonmia</i>	0 No hay dificultad
				1 Se despierta a primeras horas de la madrugada, pero se vuelve a dormir
				2 No puede volver a dormirse si se levanta de la cama
162	Consultas Médicas	<i>hdrs</i>	<i>H_mid_insonmia</i>	0 No hay dificultad
				1 Esta desvelado e inquieto o se despierta varias veces durante la noche
				2 Esta despierto durante la noche, cualquier ocasión de levantarse de la cama se clasifica en 2 (excepto por motivos de evacuar)
163	Consultas Médicas	<i>hdrs</i>	<i>H_psychic_anxiety</i>	0 No hay dificultad
				1 Tensión subjetiva e irritabilidad
				2 Preocupación por pequeñas cosas
				3 Actitud aprensiva en la expresión o en el habla
				4 Expresa sus temores sin que le pregunten
164	Consultas Médicas	<i>hdrs</i>	<i>H_somatic_anxiety</i>	0 Ausente
				1 Ligera
				2 Moderada
				3 Severa
				4 Incapacitante

Continúa en la página siguiente

Cuadro G.1 – Continúa de la página anterior

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
165	Consultas Médicas	<i>hdrs</i>	<i>H_suicide</i>	0 Ausente
				1 Le parece que la vida no vale la pena ser vivida
				2 Desearía estar muerto o tiene pensamientos sobre la posibilidad de morirse
				3 Ideas de suicidio o amenazas
				4 Intentos de suicidio (cualquier intento serio)
166	Consultas Médicas	<i>hdrs</i>	<i>H_weigth</i>	0 Pérdida de peso inferior a 500 gr. en una semana
				1 Pérdida de más de 500 gr. en una semana
				2 Pérdida de más de 1 Kg. en una semana
167	Consultas Médicas	<i>hdrs</i>	<i>H_score</i>	Puntaje en la escala de HDRS
168	Consultas Médicas	<i>hdrs</i>	<i>H_hrsd_dx</i>	Diagnóstico según la escala. 0-7: Eutimia, 8-13: Depresión Ligera, 14-18: Depresión Moderada, 19-22: Depresión severa y >23: Muy severa



## Anexo H

# VARIABLES DE LAS CONSULTAS MÉDICAS - ESTADOS

Cuadro H.1: Variables de las consultas médicas - estados

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
169	Consultas Médicas	<i>estados</i>	<i>I_alleviation</i>	1:Muchísimo mejor, 2:Mucho mejor, 3:Un poco mejor, 4:Ningun cambio, 5:Un poco peor, 6:Mucho peor, 7:Muchísimo peor
170	Consultas Médicas	<i>estados</i>	<i>I_assist_ind</i>	1:Asistió a la consulta, 2:No asistió
171	Consultas Médicas	<i>estados</i>	<i>I_dx_large</i>	Literales utilizados por el psiquiatra para establecer el estado del paciente (24 niveles)
172	Consultas Médicas	<i>estados</i>	<i>I_dx_short</i>	Agregación de <i>I_dx_large</i> en las categorías: Depresión, Eutimia, Manía, Mixto y No informado
173	Consultas Médicas	<i>estados</i>	<i>I_intervencion_ty-pe</i>	1:Presencial, 2:Online, 3:Teléfono, 4:Correo
174	Consultas Médicas	<i>estados</i>	<i>I_scheduled_intervention</i>	1:Consulta programada, 2:Consulta no programada
175	Consultas Médicas	<i>estados</i>	<i>I_significant_-change</i>	0:No presenta cambios significativos, 1:Si presenta cambios significativos
176	Consultas Médicas	<i>estados</i>	<i>I_therapist</i>	0:El psiquiatra no es DU, 1:El psiquiatra es DU



## Anexo I

# Variables de las consultas médicas - eventos

Cuadro I.1: Variables de las consultas médicas - eventos

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores
177	Consultas Médicas	<i>eventos</i>	<i>E_recorded_dx</i>	(Si/No) Existe o no información del estado del paciente
178	Consultas Médicas	<i>eventos</i>	<i>E_dx_large</i>	Literales utilizados por el psiquiatra para establecer el estado del paciente (12 niveles)
179	Consultas Médicas	<i>eventos</i>	<i>E_dx_short</i>	Agregación de <i>E_dx_large</i> en las categorías: Depresión, Eutimia, Hipomanía, Mixto, Ingresado u Otros
180	Consultas Médicas	<i>eventos</i>	<i>E_dx_numeric</i>	Codificación de <i>E_dx_short</i> . -1:Depresión, 0:Eutimia, 1:Hipomanía, 2:Mixto, 3:Ingresado, 4:Otros
181	Consultas Médicas	<i>eventos</i>	<i>E_crisis</i>	(Si/No) <i>E_dx_numeric</i> es o no es Eutimia (0)



## Anexo J

# Variables generales

Cuadro J.1: Variables generales del dataset

Nº	Fuente	Dataset	Nombre Variable	Descripción o posibles valores										
182	Todas	<i>Todos</i>	<i>A_id</i>	Identificador del paciente										
183	Todas	<i>Todos</i>	<i>A_date</i>	Fecha de registro										
184	Consultas médicas	<i>estados</i>	<i>Z_dx_numeric</i>	<p><b>Variable Objetivo</b> Codificación de <i>I_dx_short</i></p> <table border="1"> <tr><td>-1</td><td>Depresión</td></tr> <tr><td>0</td><td>Eutimia</td></tr> <tr><td>1</td><td>Manía</td></tr> <tr><td>2</td><td>Mixto</td></tr> <tr><td>3</td><td>NA</td></tr> </table>	-1	Depresión	0	Eutimia	1	Manía	2	Mixto	3	NA
-1	Depresión													
0	Eutimia													
1	Manía													
2	Mixto													
3	NA													