



UNIVERSIDAD
COMPLUTENSE
MADRID

Proyecto de Innovación

Convocatoria 2018/2019

Nº de proyecto: 225

Título del Proyecto: Big Data en Educación III: recomendaciones personalizadas de actuación ante estudiantes en riesgo de abandono

Responsable del proyecto: Adolfo Hernández Estrada
Centro: Facultad de Comercio y Turismo
Departamento: Economía Financiera y Actuarial y Estadísticas

1. Objetivos propuestos en la presentación del proyecto

Este proyecto se presentó como continuidad a los llevados a cabo en convocatorias anteriores (conv. 2006: nº 390, conv. 2007: nº 559, conv. 2011: nº 246, conv. 2014: nº 244, conv 2015: nº 379, conv. 2016 83 y conv. 2017 245). En estos proyectos se habían desarrollado materiales para el aprendizaje autónomo del alumno, una herramienta de aprendizaje teórico, dos herramientas de evaluación y autoevaluación, un paquete de vídeos que facilitan el aprendizaje en modelos de aula de aprendizaje inverso, siempre en el entorno del aprendizaje de la Estadística, tanto en el desarrollo de asignaturas convencionales como en el desarrollo de Trabajo Fin de Grado.

En particular, este proyecto es continuación directa del proyecto 245 de la convocatoria 2017: “Big data en educación II: metodologías adaptativas en el proceso de enseñanza-aprendizaje desde el diagnóstico del estudiante”. En este proyecto se cumplió el objetivo principal, “plantear metodologías adaptativas en la práctica docente que permitan ajustarse a las características personales y sociales del estudiante”, y cinco de los seis objetivos secundarios. El único objetivo secundario no alcanzado, “desarrollar modelos estadísticos de clasificación, que permitan detectar de manera previa posibles situaciones de abandono académico antes de que se materialicen”, no llegó a realizarse debido a su gran complejidad. Sin embargo, nos pareció de tal importancia que motivó la presente propuesta.

Los objetivos que se enumeran a continuación fueron los que se incluyeron en el proyecto.

Objetivo principal:

Elaboración de un conjunto de recomendaciones personalizadas para guiar la actuación del profesor con estudiantes que hayan sido detectados en riesgo de abandono de una asignatura o de la titulación en general.

Objetivo secundario:

Desarrollo de métodos estadísticos que permitan la detección del riesgo de abandono de un estudiante. Este proyecto parte del análisis de los datos, tanto de los recogidos a nivel institucional (nota y ruta de acceso, sexo, datos demográficos, etc.) como de los que se generan día a día y que quedan registrados en la plataforma Moodle. Los modelos estadísticos que se desarrollarán pertenecen a la familia de lo que se ha dado en llamar en los últimos años, “Learning Analytics” o “Educational Data Mining”. Estos modelos clasificatorios y predictivos permitirán anticiparse a posibles situaciones de abandono.

2. Objetivos alcanzados

Objetivo secundario:

Totalmente conseguido: a partir de los datos disponibles (tanto del sistema de gestión académica GEA como de la plataforma educativa MOODLE) se ha construido una base de datos relacional y se han testado múltiples modelos clasificatorios y predictivos: métodos clúster (jerárquicos y k-medias), LDA (Análisis Discriminante Lineal), QDA (Análisis Discriminante Cuadrático), Regresión Logística, Árboles Aleatorios (Random Forest) y Redes Neuronales. Todos estos métodos han sido implementados en lenguaje Python.

Objetivo principal:

Parcialmente conseguido. Las recomendaciones que se proponen para prevenir la deserción, están basadas en el análisis de la literatura existente sobre este tipo de hecho, iniciándose estas con el trabajo de revisión realizado por Torrado y Figuera (2019), por compartir, con este estudio, el perfil de estudiantes a quien va dirigido este trabajo. Estos autores analizaron, en un estudio longitudinal, las razones del abandono de estudiantes de Administración y Dirección de empresas (ADE) describiendo

el abandono como un fenómeno multidimensional, en el que inciden factores suprainstitucionales (políticas de ayuda al estudio), factores institucionales (organización del currículum) y personales (falta de competencias y motivación), y proponen realizar una reflexión profunda en cada uno de los niveles de actuación “en un entorno universitario de recortes económicos que ha afectado fundamentalmente al eje clave de la atención a los estudiantes: los servicios y los programas de orientación y tutoría” (p. 414). De todos los posibles factores estudiados (González et al., 2007; Cabrera et al., 2006; Casaravilla et al., 2012;) los relacionados la situación académica son los que más peso tienen en la decisión de abandono del estudiante así como los factores relacionados con la organización y la satisfacción con la docencia, por lo que seleccionaremos las propuestas centradas en la mejora del apoyo académico.

Apoyo mediado por las tecnologías

Algunos autores proponen mejorar y ampliar los sistemas de registro de la información sobre la actividad académica del alumnado, y en especial, la accesibilidad a estos datos por parte del profesorado. Esta mejora de la accesibilidad permitiría diseñar, desde las plataformas virtuales, “sistemas de alerta temprana” que seleccionen a aquellos estudiantes que necesitan acompañamiento, lo que haría posible realizar adaptaciones personalizadas tanto de los materiales didácticos, como de los sistemas de evaluación, es decir generando “itinerarios de aprendizaje diferenciados”. En esta línea, Alamán et al. (2019) defienden el uso de herramientas basadas en analítica prescriptiva. Y exponen dos ejemplos concretos implementados en la Universidad Autónoma, EDX-WS: Warning System y EDX-FAS: Feedback Analysis System. “La primera herramienta genera recomendaciones para los estudiantes y gráficos comparativos para el seguimiento de su evolución, que pueden ser utilizadas por el personal docente y administradores de los MOOCs para orientar a los estudiantes. La segunda es una evolución de la primera y tiene como principal objetivo facilitar el acompañamiento de los alumnos para reforzar su proceso de aprendizaje en función de su situación y necesidades. Este proceso comprende funcionalidades como la detección de patrones (de aprendizaje positivo o negativo) y la proposición de acciones en consecuencia. Para Burgos et al. (2018) ejercer alguna atención específica a aquellos alumnos que no cumplen plazos en entregas de las primeras actividades, tareas o trabajos obligatorios, tales como establecer un protocolo de mensajes o llamadas a esos posibles abandonos, produce efectos positivos. Algunos estudios, van más allá, Cobos, Jurado y Blázquez-Herranz (2019) proponen el análisis de los sentimientos y emociones de los estudiantes, aplicando técnicas NLP para analizar los contenidos de los cursos en línea y las contribuciones de sus alumnos (transcripciones en vídeo, lecturas, preguntas y respuestas de las actividades de evaluación y mensajes de los alumnos en foros, entre otros) para mejorar el material didáctico y el aprendizaje-enseñanza.

Apoyo mediado desde la orientación académica y la tutoría

Para otros expertos, sin embargo, la clave está en la potenciación de las funciones orientadoras y tutoriales. Una acción orientadora dirigida a: asesorar a los estudiantes sobre la planificación general de sus estudios y a facilitar y fomentar el establecimiento de grupo de trabajo y de apoyo entre iguales. Favorecer esa planificación implica que los servicios de orientación de las universidades puedan detectar a tiempo si los estudiantes tienen una formación previa y competencial suficiente en la materia, si son capaces de hacer una adecuada gestión del tiempo frente a las dificultades, sus motivaciones, así como si sus expectativas son realistas respecto a su capacidad para superar las dificultades. Respecto a la tutoría entre iguales, esta se convierte en “una potente herramienta para ayudar a las universidades a alcanzar sus objetivos estratégicos, en especial, la mejora de la retención y el éxito de sus estudiantes de nuevo ingreso. Los beneficios para la integración del alumnado nuevo, para el desarrollo de competencias del estudiantado veterano y para la consolidación de una cultura universitaria de participación” (Casado -Muñoz, Lezcano- Barbero y Colomer -Feliu, 2015, p.172).

Apoyo mediado por el docente

En relación a las causas de abandono relacionadas con el docente, entre otras, García Areito (2019) menciona el escaso seguimiento y supervisión de los docentes, así como diseños instructivos con materiales y recursos educativos inadecuados. La incorporación de videos e imágenes que facilitan la comprensión de los materiales docentes, son parte de los recursos educativos que se incorporan en las asignaturas para promover el éxito académico y la no deserción (Rojas y Castro, 2016, p. 124.). “De acuerdo a las nuevas formas de aprendizaje, el profesor debe desarrollar ciertas competencias digitales que le permitan interactuar correctamente con las formas de aprendizaje del estudiante, por lo que se demanda de este, no solo la creación de contenidos tradicionales, sino contenidos que se puedan adaptar a la secuencia, el ritmo, la navegación, la pedagogía y la presentación del alumno (Olvera-Castillo, y García-Almeida, 2019, sin página). El análisis de los estilos de aprendizaje facilita el desarrollo de actividades motivacionales (Ferrero y Oloriz, 2015). Estas estrategias fueron utilizadas en el diseño de sesiones presenciales en pequeño grupo (6 estudiantes). “En estos encuentros solo se revisaron las actividades llevadas a cabo por cada estudiante y se los orientó para avanzar en aquellos casos que no estaban interpretando adecuadamente las consignas. Se trató de que la intervención fuese solo de carácter orientativo al trabajo con los materiales, evitando realizar explicaciones de temas o resolución de casos de manera grupal” (p.6). Escobar (2015) recomienda el uso de una enseñanza basada en problemas, basada “fundamentalmente en el proceso investigativo en el que el alumno es el protagonista y es quien debe consolidar soluciones a problemas emanados de la interacción social, siendo el maestro una guía constante del proceso” (p. 2). “Todos los estudios sobre aprendizajes basados en problemas (ABP) y las prácticas docentes centradas en el alumnado mejoran la calidad de los aprendizajes y la integración y compromiso institucional, pues los aprendizajes adquiridos van más allá del conocimiento, incluyendo estrategias cognitivas más elevadas de aplicación y análisis. Además, los estudiantes se motivan tratando de solucionar los problemas, lográndose con ello una alta participación e implicación en el estudio, y una mejor preparación para resolver problemas del mundo laboral y social (Cabrera, 2015, p. 45).

Referencias

- Alamán, X., Carro, R. M., Cobos, R., Gómez, J., Jurado, F., Ruano, P. M., ... & Torrado, J. C. (2019). GHIA: Modelado de Estudiantes, Analítica de Aprendizaje, Atención a la Diversidad y e-Learning. *Informática Educativa Comunicaciones*, 30(30).
- Burgos, C., Campanario, M. L., Peña, D. d. I., Lara, J. A., Lizcano, D., y Martínez, M. A. (2018). Datamining for modeling students' performance: A tutoring action plan to prevent academic dropout. *Computers and Electrical Engineering*, 66, 541-556. doi: 10.1016/j.compeleceng.2017.03.005
- Cabrera, L. (2015). Efectos del proceso de Bolonia en la reducción del abandono de estudios universitarios: datos para la reflexión y propuestas de mejora. *Revista Fuentes*, 16, 39-62.
- Cabrera, L., Bethencourt, J.T., Alvarez, P. y González, M. (2006). El problema del abandono de los estudios universitarios. *Revista Electrónica de Investigación y Evaluación Educativa*. RELIEVE, 12.
- Casado -Muñoz, R., Lezcano- Barbero, F. & Colomer -Feliu, J. (2015). Diez pasos clave en el desarrollo de un programa de mentoría universitaria para estudiantes de nuevo ingreso. *Revista Electrónica Educare* 19 (2), 155-179. doi: <http://dx.doi.org/10.15359/ree.19-2.10>
- Casaravilla, A., Del Campo, J. M., García, A., & Torralba, M. R. (2012) Análisis Del Abandono En Estudios De Ingeniería. Comunicación presentada en la II Conferencia Latinoamericana sobre el Abandono en la Enseñanza Superior, Portoalegre, Brasil.

Cobos, R., Jurado, F., & Blázquez-Herranz, A. (2019). A Content Analysis System That Supports Sentiment Analysis for Subjectivity and Polarity Detection in Online Courses. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, 14(4), 177-187.

Escobar, J. (2016). El uso del aprendizaje basado en problemas como instrumento para la disminución de la deserción académica: el caso del proyecto integrador en la Universidad la Gran Colombia. *Congresos CLABES*.

Ferrero, E. L y Oloriz, M. G. (2015). Aplicación de estrategias motivacionales para mejorar la enseñanza de matemática introductoria en la educación superior. *Congresos CLABES*.

García Aretio, L. (2019) El problema del abandono en estudios a distancia. Respuestas desde el Diálogo Didáctico Mediado RIED. *Revista Iberoamericana de Educación a Distancia*, 22 (1), <https://doi.org/10.5944/ried.22.1.22433>

González, M. C., Álvarez, P. R., Cabrera, L. y Bethencourt, J. T. (2007). El abandono de los estudios universitarios: factores determinantes y medidas preventivas. *Revista española de pedagogía*, 65(236), 71-85.

Olvera-Castillo, N.L. y García-Almeida, C. Y. (2019). El rol del docente en el aprendizaje adaptativo. *Revista digital FILHA*, 21

Rojas, R. y Castro, E. (2016). Estrategias de acompañamiento para promover la permanencia de la población estudiantil en la UNED de Costa Rica. En *Claves innovadoras para la prevención del abandono en instituciones de educación abierta y a distancia* (pp. 105-146). Republica dominicana : UAPa

Torrado, M. y Figuera, P. (2019). Estudio longitudinal del proceso de abandono y reingreso de estudiantes de Ciencias Sociales. El caso de Administración y Dirección de Empresas. *Educación*, 55(2), 401-417.

Bibliografía:

Álvarez- Pérez. P. Lopez Aguilar, D. (2017). Estudios sobre deserción académica y medidas orientadoras de prevención en la universidad de la laguna (España) *Revista Paradigma*, 38 (1), 48 – 71.

Esteban, M., Bernardo, A. B., Tuero, E., Cerezo, R. y Núñez, J. C. (2016). El contexto sí importa: identificación de relaciones entre el abandono de titulación y variables contextuales. *European Journal of Education and Psychology*, 9(2), 79-88.

Rodríguez, E. J., Rabazo, A. E. y Naranjo, D. (2015). Evidencia empírica de la adquisición de la competencia de resolución de problemas. *Perfiles educativos*, 37(147), 50-66.

3. Metodología empleada en el proyecto

A continuación se presenta el Plan de Trabajo Detallado que se presentó en la propuesta, con comentarios sobre su ejecución.

FASE 1 (F1). Recopilación de datos del curso académico 2016-17.

Tarea 1.1 Recopilación de datos de actividad de usuarios de cursos alojados en la plataforma Moodle.

Solicitud de los archivos de registros web de cursos Moodle.

Tarea 1.2. Recopilación de datos institucionales.

Solicitud al Vicerrectorado de Estudiantes de datos institucionales.

Esta Fase se completó en su totalidad. El apoyo del vicerrector de Tecnologías de la Información de la UCM y de la decana de la Facultad de Comercio y Turismo (FCyT) nos ha posibilitado tener acceso institucional a todas las asignaturas de todos los estudios de Grado que se imparten en esta Facultad. Dado el carácter de estas titulaciones los datos analizados se corresponden con una gran variedad de áreas de conocimiento: asignaturas impartidas por un total de 34 departamentos diferentes de la UCM. Por tanto, los resultados obtenidos pueden ser extensibles a otros campos de conocimiento.

FASE 2 (F2). Análisis estadístico de los datos y desarrollo del modelo de alerta

Tarea 2.1. Depuración de los datos: detección y tratamiento de datos perdidos y atípicos

Tarea 2.2. Creación de un modelo de alerta a través de la aplicación de modelos estadísticos.

Tarea 2.3. Entrenamiento del modelo de alerta.

Tarea 2.4. Validación y fiabilidad del modelo de alerta.

Esta Fase se completó en su totalidad. Parte de los resultados fueron detallados en la ponencia "Portabilidad de modelos predictivos del rendimiento académico: análisis del "efecto profesor"", dentro de la III Jornada de Big-Data en Educación: Definiendo factores de riesgo, celebrada el pasado día 26 de noviembre de 2018 en la Facultad de Comercio y Turismo de la Universidad Complutense de Madrid.

FASE 3 (F3). Diseño del sistema recomendaciones.

Tarea 3.1. Creación de un sistema de recomendaciones personalizado para mejora del rendimiento. Se elaborará un plan de ruta que recoja prácticas que faciliten el aprendizaje comprensivo y recomendaciones para que las partes interesadas (profesor y alumno) puedan tomar decisiones con respecto a los alumnos en riesgo de abandono.

Esta Fase se ha completado parcialmente. En base al análisis efectuado se está elaborando un plan de ruta que recoge prácticas que faciliten el aprendizaje comprensivo en entornos similares a la realidad del futuro desempeño profesional y recomendaciones que posibiliten la reducción del número de abandonos.

FASE 4 (F4). Informe final y difusión de resultados

Tarea 4.1. Elaboración del informe final de resultados y propuesta de mejoras.

Tarea 4.2. Plan de difusión de resultados.

Tarea 4.3. Elaboración de la memoria final del proyecto.

Esta Fase se ha completado. Tal como ya se ha señalado, el pasado día 26 de noviembre de 2018 se organizó una Jornada de difusión en la Facultad de Comercio y Turismo de la Universidad Complutense de Madrid. La jornada se desarrolló de 10:00 a 14:00 y constó de cuatro charlas: las primeras tres charlas fueron impartidas por miembros del equipo del proyecto y se basaron en la exposición de los resultados obtenidos a lo largo de la ejecución del proyecto. La última charla fue impartida por un ponente invitado (el profesor Javier Vidal, de la Universidad de León, especialista en Orientación y Tutoría en la Universidad) y llevó por título "Algunas reflexiones sobre el abandono universitario". Esta ponencia nos pareció de especial interés dado el perfil del ponente: Javier Vidal es Doctor en Filosofía por la Universidad de Salamanca y Profesor Titular de la Universidad de León en el área de Métodos de Investigación y Diagnóstico en Educación. En la Universidad de León ha sido Director del Programa Institucional de Calidad (1996-2004) y Vicerrector de Planificación y Evaluación (2004-2006). Ha colaborado con la ANECA (Agencia Nacional de Evaluación de la Calidad y Acreditación) en las áreas de Evaluación Institucional y Acreditación, así como del

Observatorio de Inserción Laboral. Desde junio de 2006 hasta abril de 2008 ha sido Director General de Universidades del Ministerio de Educación y Ciencia, y desde abril de 2008 a abril de 2009, Director General de Evaluación y Ordenación del Sistema Educativo del Ministerio de Educación, Política Social y Deporte. Como investigador, sus trabajos se encuentran en el ámbito de estudio de la educación superior, especialmente evaluación de la investigación, la evaluación institucional y las relaciones entre formación y empleo. Ha colaborado con la Comisión Europea como consultor en varias ocasiones.

4. Recursos humanos

La mayoría del equipo lo formaron profesores de las Facultades de Comercio y Turismo y de Ciencias Económicas y Empresariales. Además, se contó con la participación de 2 estudiantes al considerarse que la inclusión de su visión mejorará este estudio en prácticamente todas sus fases. La totalidad de los profesores tuvieron sus asignaturas dadas de alta en el Campus Virtual de la UCM.

Este problema presentaba todas las características asociadas a lo que en los últimos años ha dado en denominarse “Big Data” (Mayer-Schönberger y Cukier, 2013). Estas técnicas tienen su origen en métodos estadísticos aplicados en bases de datos con alta dimensión. En ese sentido, la pertenencia de la mayoría de los profesores a la Unidad Docente de Estadística e Investigación Operativa (Métodos de Decisión) resultó totalmente apropiada.

Además, varios de los profesores participantes en el proyecto tenían experiencia específica en el tratamiento de datos multivariantes, tanto a nivel docente (impartición de la asignatura “Técnicas Estadísticas Multivariantes en Gestión Comercial”, de tercer curso del grado en Comercio) como investigador.

El equipo docente lo completaron tres profesores más: uno de la Universidad Politécnica de Madrid, otro de la Universidad Internacional de La Rioja (con amplia experiencia en el uso de plataformas virtuales, dado el carácter no presencial de las enseñanzas que imparten) y una profesora de la Universidad de Alcalá, especialista en las áreas de didáctica de las matemáticas y educación adaptativa. En definitiva, se ha tratado de reunir especialistas tanto en educación como en el tratamiento cuantitativo y cualitativo de datos.

Dada la complejidad técnica del problema planteado, también se incorporó al equipo un miembro del PAS de la UCM. Se trata de un técnico especialista de los servicios informáticos, con experiencia previa de varios años como programador, desarrollador de software educativo y técnico de soporte interno para Microsoft Ibérica SRL; actualmente es T.E. II de aulas y servicios informáticos de la UCM.

Por último, aunque no fue formalmente un miembro del equipo, ha trabajado con nosotros durante el desarrollo del proyecto un colaborador en tareas informáticas y estadísticas. Su nombre es Carlos Enrique Huane Estrada, y él fue el principal responsable de la construcción de la base de datos relacional y de la implementación en Python de los diferentes métodos estadísticos utilizados. Parte de los resultados obtenidos pueden encontrarse en su Trabajo Fin de Máster “Creación de una base de datos relacional y primeros pasos en un proyecto Data Mining”, presentado en el Máster de Ingeniería Matemática de la UCM.

5. Desarrollo de las actividades

Las fases de trabajo siguieron los objetivos expuestos anteriormente, y que consideramos se describen de manera suficiente tanto en el punto 2 de objetivos conseguidos como en el punto 3 de metodología.

6. Anexos

Anexo 1. Póster III Jornada Big Data en Educación



111 Jornada de INNOVación

Big DATA en EDUCación

Definiendo factores de riesgo

Organiza:

Proyecto* de Innovación "Big Data en Educación 111: recomendaciones personalizadas de actuación ante estudiantes en riesgo de abandono"

*Universidad Complutense de Madrid (UCM)

*Universidad de Alcalá (UAH)

26 DE NOVIEMBRE DE 2018

Carlos Enrique Huane Estrada (UCM)
Evento de Innovación e Implementación de Big Data en Educación

Juan Luis Peñabza Figueroa (UCM)
Curso de Big Data en Educación
Rendimiento Académico de los Estudiantes de la Facultad de Comercio y Turismo

Francisco Javier Vidal García (UCM)
Algunas reflexiones sobre el abandono universitario

José Luis Arroyo (UCM)
Ponencia de los Profesores de la Facultad de Comercio y Turismo

10:00 a 10:10 - Inauguración

10:15 a 10:40 - Carlos Enrique Huane Estrada

10:45 a 11:10 - Juan Luis Peñabza Figueroa

11:15 a 11:40 - Pausa

11:45 a 12:10 - José Luis Arroyo

12:10 a 13:00 - Francisco Javier Vidal García

13:00 a 13:30 - Mesa redonda