

Article

LENNA (Learning Emotions Neural Network Assisted): An Empathic Chatbot Designed to Study the Simulation of Emotions in a Bot and Their Analysis in a Conversation

Rafael Lahoz-Beltra ^{1,2,*} and Claudia Corona López ¹
¹ Department of Biodiversity, Ecology and Evolution (Biomathematics), Faculty of Biological Sciences, Complutense University of Madrid, 28040 Madrid, Spain; ccoronal@ucm.es

² Modeling, Data Analysis and Computational Tools for Biology Research Group, Complutense University of Madrid, 28040 Madrid, Spain

* Correspondence: lahozraf@ucm.es



Citation: Lahoz-Beltra, R.; López, C.C. LENNA (Learning Emotions Neural Network Assisted): An Empathic Chatbot Designed to Study the Simulation of Emotions in a Bot and Their Analysis in a Conversation. *Computers* **2021**, *10*, 170. <https://doi.org/10.3390/computers10120170>

Academic Editors:
Katia Lida Kermanidis,
Manolis Maragoudakis,
Phivos Mylonas
and Fernando Bobillo

Received: 10 October 2021

Accepted: 10 December 2021

Published: 13 December 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/>).

Abstract: Currently, most chatbots are unable to detect the emotional state of the interlocutor and respond according to the interlocutor's emotional state. Over the last few years, there has been growing interest in empathic chatbots. In other disciplines aside from artificial intelligence, e.g., in medicine, there is growing interest in the study and simulation of human emotions. However, there is a fundamental issue that is not commonly addressed, and it is the design of protocols for quantitatively evaluating an empathic chatbot by utilizing the analysis of the conversation between the bot and an interlocutor. This study is motivated by the aforementioned scenarios and by the lack of methods for assessing the performance of an empathic bot; thus, a chatbot with the ability to recognize the emotions of its interlocutor is needed. The main novelty of this study is the protocol with which it is possible to analyze the conversations between a chatbot and an interlocutor, regardless of whether the latter is a person or another chatbot. For this purpose, we have designed a minimally viable prototype of an empathic chatbot, named LENNA, for evaluating the usefulness of the proposed protocol. The proposed approach uses Shannon entropy to measure the changes in the emotional state experienced by the chatbot during a conversation, applying sentiment analysis techniques to the analysis of the conversation. Once the simulation experiments were performed, the conversations were analyzed by applying multivariate statistical methods and Fourier analysis. We show the usefulness of the proposed methodology for evaluating the emotional state of LENNA during conversations, which could be useful in the evaluation of other empathic chatbots.

Keywords: empathic chatbot; sentiment simulation; sentiment analysis; entropy of emotional state changes; chatbot performance

1. Introduction

The design of a chatbot and, thus, software agent with which a human can hold a conversation using natural language is a proposal that dates back to the 1960s. During the decade of the sixties, Weizenbaum [1] designed ELIZA, a chatbot that emulated a Rogerian psychotherapist, i.e., imitating psychologist Carl Rogers [2], during an interview with a patient. ELIZA followed a passive model, since the conversation with a human was kept active by means of a strategy consisting of reflecting the patient's statements or answers in the questions that ELIZA asks over and over again. Years later, [3] proposed PARRY, a chatbot that simulated a paranoid schizophrenic subject by implementing a behavioral or mental model based on concepts and pre-established judgments of these concepts. Compared to ELIZA, chatbot PARRY followed an inverse strategy and maintained the conversation by following a paranoid pattern that was the result of displaying its fears and anxieties during the conversation.

In the 1960s and 1970s, both ELIZA and PARRY designs were inspired by the Turing test and used a common strategy: For each statement made by the interlocutor, the chatbot

identifies keywords and retrieves the answer from a rule base or response table. If the interlocutor's sentence does not include some key words, then the chatbot ultimately resorts to predefined answers in order to keep the conversation going. However, PARRY, unlike ELIZA, was the first chatbot exhibiting emotions to its interlocutor [4]. Currently, one of the goals of artificial intelligence is to design chatbots with the ability to detect the emotions of a human interlocutor and to exhibit or simulate emotional behavior.

The first theory of emotions was proposed by Darwin in 1872 [5] by assuming that there were discrete and small numbers of basic emotions, which would later result in a biological conception of emotions. In the biological realm, emotions are "affective states" based on physiological responses to stimuli. Currently, the design of emotionally aware chatbots (EAC) is based on the replacement of the rule base by machine learning or affective computing [6] techniques and on a computational definition of emotions [7]. For example, under this paradigm, an emotion can be defined as a complex construct that could be represented as a two-dimensional Russell model [8]. In this case, the chatbot architecture included an emotion classifier detecting the emotion expressed in the interlocutor's sentence. Thereafter, the chatbot generates a response that is most appropriate depending on the emotion expressed by the interlocutor, e.g., happiness, fear, etc.

The use of chatbots has been increasing in recent times and is widely used in the commercial sector in areas such as customer service or shopping assistance, as well as those dedicated to multiple purposes, e.g., SIRI or Alexa. Although most chatbots are designed in order to manage interactions of a commercial nature, for example, through a website [9], research on chatbots applied to medicine has experienced a major explosion. In special circumstances, a conversation between a chatbot and a patient allows, in some clinical situations, better diagnosis and treatment, improved patient monitoring and improved data collection from him/her, offering better and more personalized care. For example, in psychology [10,11] and psychiatry [12], there are studies that suggest the use of chatbots when performing psychological evaluations and for the treatment of mental health. However, despite these applications of bots in mental health in medicine, there is a lack of chatbots designed to study complex clinical problems. One of the problems we are concerned with is related to the effect of a subject's mental state on the physiological mechanisms that determine their state of health. A subject's mental state, i.e., the mood, can be affected by the environment or by its interaction with other individuals, for instance, through a conversation between the subject and other individuals.

The main novelty of this study has been the protocol with which it is possible to analyze the conversations between a chatbot and an interlocutor, whether the latter is a person or another chatbot. For this purpose, we have designed a minimally viable prototype of an empathic chatbot, named LENNA, with which we have been able to evaluate the usefulness of the proposed protocol. The protocol first obtains the Shannon entropy of the emotional states (stress, excitement, depression and healthy) that LENNA experiences during a conversation, the number of words that reflect each of the emotions (anger, fear, anticipation, trust, surprise, sadness, joy and disgust) expressed in the course of a conversation and the values of the univariate statistics (minimum value, first quartile, median, mean, third quartile and maximum value) that are obtained by applying sentiment analysis techniques to the conversation. Secondly, based on the data obtained from a conversation, multivariate statistical analysis methods were applied, and the conversation was represented by means of Fourier analysis, observing different patterns that reflect emotional aspects of the conversation held with LENNA.

In this paper, we designed an empathic chatbot, which we have named LENNA (Learning Emotions Neural Network Assisted), for which its architecture serves to simulate and study how the emotional state of an interlocutor can affect our emotional state, the latter subject being simulated with a chatbot. This kind of chatbot could be useful in the design of a virtual patient for the purpose of studying diseases that are due partly to a certain affective or emotional state. From a physiological point of view, there is a relationship between an altered mood or emotional state with disorders in the production of

hormones and the induction of an inflammatory process. Currently, according to different studies, we know that an inflammatory process can ultimately cause different pathologies, such as cancerous tumors. More specifically, it has been observed how states of depression or low mood are able to trigger the hormonal system, producing cortisol, cytokines, etc., and, ultimately, resulting in tissue inflammation, which is one of the pathways by which the development of cancer can eventually arise [13,14]. In addition to the case of cancer, other pathologies are the result of mood alterations affecting hormones, such is the case of arterial hypertension and its relationship with obesity, since the former may be the cause of the latter [15]. In a previous study [16], we took the first steps in the simulation with a chatbot of a virtual patient, assuming that this patient has an altered mood, suffering from stress or depression. Mood was revealed by the words used during a conversation, both in terms of content and frequency of use, depending on the frequency of “chatbot hormone levels”. Given a stressful stimulus, we then simulated the kinetics of hormones involved as well as how hormone levels triggered chronic inflammation, which may promote the development of colon cancer.

This study is motivated by the aforementioned scenarios, and the main goal is a protocol designed to analyze conversations with an empathetic chatbot, i.e., a chatbot with the ability to recognize emotions of its interlocutor from the vocabulary used, classifying the words in different emotional categories. It is intended that in the future such a model can be used to simulate the relationship between different altered mood states and the hormonal disorders produced by these states. A model of this kind is useful for modeling pathologies associated with changes in hormonal levels and for studying these pathologies without the need for real patients, which may open up new avenues of research in psychology and psychiatry, the design of empathic chatbots in artificial intelligence, etc. In this study, we designed an “empathic chatbot”; that is, a program that is able to understand the emotional state of the person conversing with the chatbot and that can respond accordingly is designed, which makes interactions more natural and similar to having conversations between a program and a human being [17]. There are already applications such as SERMO or EMMA, which are able to recognize words and provide appropriate responses to the person’s feelings. Thus, they suggest activities or exercises that can be beneficial for the user according to their mood [7]. This type of chatbot recognizes and processes different emotional states by means of so-called sentiment analysis, also known as opinion mining. This consists, in short, of analyzing language to extract and classify subjective information contained in words. In this manner, a chatbot detects the feelings, opinions, attitudes, etc., of the person with whom the chatbot is having a conversation.

In this paper, LENNA was designed as follows. In the chatbot, the extraction of words to detect the emotions was carried out using classifiers, specifically Bayesian classifiers [18,19] in combination with artificial neural networks [20,21]. The artificial neural network was included in order to connect, in the future, the chatbot with an artificial endocrine system [22] through the neural network. Thus, physiological mechanisms have not been included in the present simulation but will be developed in a future study. As mentioned above, the main contribution of this study is the methodology we have used for the analysis of conversations between LENNA and an interlocutor, regardless of whether the latter is a person or another bot. A novel approach has been the use of Shannon entropy for measuring the changes in the emotional state experienced by the chatbot during a conversation, including also the analysis of the sentiment to the conversation. Once simulation experiments were performed, the conversations were analyzed as mentioned above, applying multivariate statistical methods and Fourier analysis to the obtained data.

2. Methods

2.1. LENNA Architecture

The empathic bot LENNA was modeled according to Figure 1, and modeling is based on the hybridization of a chatbot—with natural language processing (NLP) features—with Bayesian classifiers. The chatbot model corresponds to the ELIZA type [1], which originally

simulated a conversation between a psychoanalyst, the program, a patient and the user. The chatbot was implemented in Python language by adapting and modifying the eliza.py program developed by [23]. The result has been a chatbot that simulates different emotional states as a conversation takes place, depending on the phrases said by the interlocutor. According to Figure 1, any conversation between an interlocutor and LENNA begins in a neutral or initial emotional state or state 0, writing LENNA as the text shown below.

Therapist

Talk to the program by typing in plain English, using normal upper- and lower-case letters and punctuation. Enter “quit” when done.

Hello. My name is LENNA. How are you feeling today?

>

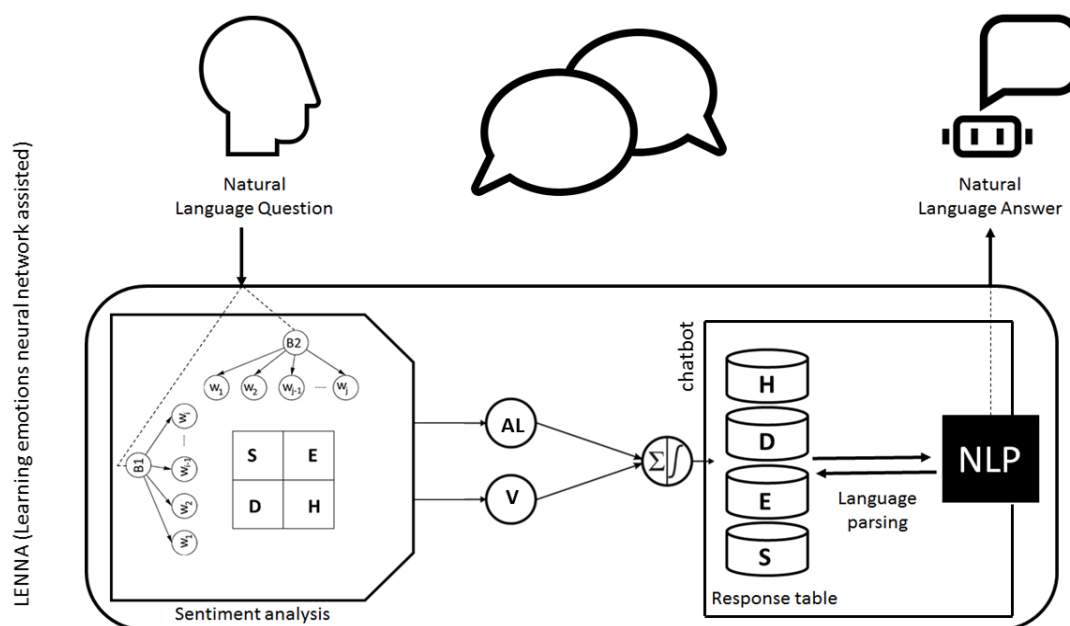


Figure 1. General organization of LENNA depicting the response tables for the emotional states (H = healthy, D = depressed, E = excitation and S = stress), natural language processing (NLP) module, perceptron neural network with the inputs AL (arousal level) and V (valence) and the sentiment analysis module and Bayesian classifiers B1 and B2.

Once the interlocutor writes a response, the text is analyzed by means of sentiment techniques, i.e., via Bayesian classifiers, in order to identify and extract information about LENNA’s interlocutor attitude or intention and, therefore, his/her emotional state. The model assumes that the “emotional state” of the chatbot is the result of the secretion of hormones and neurotransmitters, manifesting the final physiological reaction in its language. In the present study, we simulated emotions using artificial intelligence techniques (AI), simulating the emotional state of a chatbot. Obviously, in a human being, emotions are the result of complex mechanisms not simulated in the present study involving specific areas of the brain, particularly the limbic system [24].

Although studies by Ekman [25] on facial expression suggest six basic emotions [26], i.e., anger, disgust, fear, happiness, sadness and surprise, in the present study we reduced the states space of the interlocutor and, therefore, of LENNA to four “mood” states. These

states are not necessarily emotional states, and they consist of the following states: stress (state 1), excitation (state 2), depression (state 3) and a fourth state that we have called healthy (state 4). These four states correspond to the four sectors resulting from the values of two variables (Figure 1): arousal level AL, i.e., the degree of excitement and calm, and valence V, i.e., a measure of the degree of pleasant and unpleasant (i.e., positive and negative) feelings [27]. Consequently, once the conversation started (state 0), the emotional state of the LENNA bot changes to one of the four states shown in Table 1.

Table 1. LENNA’s Emotional States.

	Unpleasant 0	Pleasant 1
High arousal 0	Stress 00 (state 1)	Excitation 01 (state 2)
Low arousal 1	Depression 10 (state 3)	Healthy 11 (state 4)

Sentiment analysis of the interlocutor’s sentences is performed by means of two possible architectures. In the first design, two naïve Bayesian classifiers were designed, denoted as Bayes 1 and Bayes 2 (Figure 2), and this design is named name LENNA 1. In the second approach—naming this bot model as LENNA 2—we retain Bayes 1 and substitute Bayes 2 by a polarity test (Figure 3).

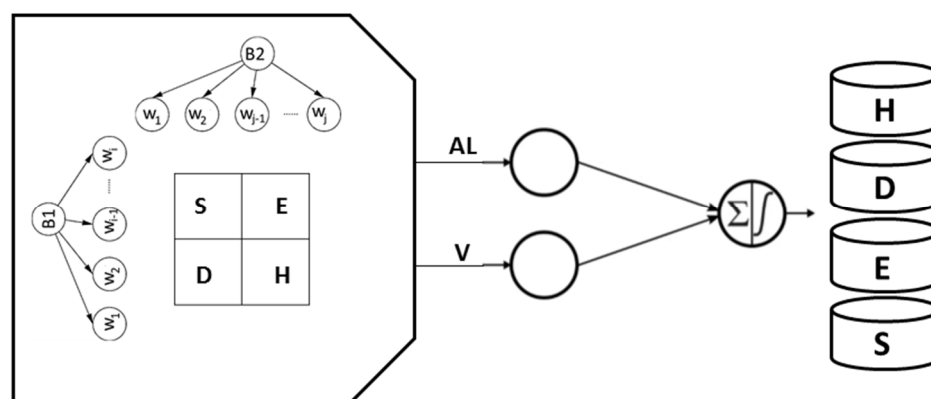


Figure 2. LENNA 1 default architecture.

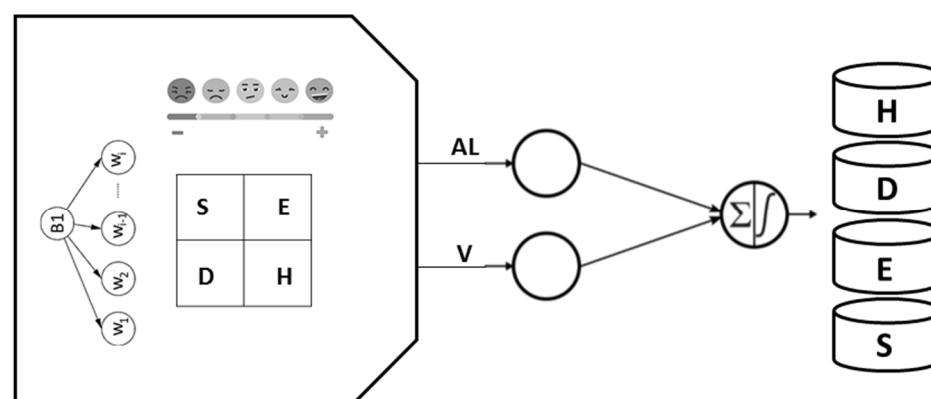


Figure 3. LENNA 2 alternative architecture.

Once the emotional states of the bot were defined, a set of words were selected, reflecting the emotional state of an interlocutor. When such words are detected by LENNA, it enables the chatbot to recognize the interlocutor’s emotion. Table 2 shows the words recognized by LENNA from the interlocutor’s text.

Table 2. Interlocutor vocabulary recognized by LENNA.

Stress	Excitation	Depression	Healthy
Alarmed, tense, angry, afraid, annoyed, distressed, frustrated, fed up, resentful, mad, irritated, fear, anxiety, agitated, furious and bitter.	Aroused, astonished, excited, delighted, happy, surprised, determined, awe, amusement, joyful, optimistic, enthusiastic, loving, pleased, charmed and grateful.	Miserable, sad, gloomy, depressed, bored, droopy, tired, worried, taken back, dull, anxious, guilty, lonely, disappointed, indifferent, fatigued, desperate and troubled.	Pleased, glad, serene, content, at ease, satisfied, relaxed, calm, hopeful, powerful, empowered, sure, dynamic, ambitious, confident, peaceful and comforted.

Let t be the interlocutor's text in response to the statement made by LENNA and $C = \{c_1, c_2, \dots, c_n\}$ is a given number of classes. If we assume two classes in the Bayesian classifier, c_1 and c_2 , then from the vocabulary of Table 2, it will be possible to define two subsets or bag of words, c_1 and c_2 . Once the interlocutor writes a text t in response to LENNA, the bot counts the number of times that each word had appeared in the text, classifying text t in class c_1 or c_2 . Applying Bayes' rule and eliminating the denominator, we will have that for a text t and class c :

$$p(c|t) = p(t|c) \cdot p(c)$$

where $p(c)$ is the probability of a class or bag of words c . Therefore, given a sentence or text of the interlocutor with the words w_1, w_2, w_3, \dots , we will rewrite the probability $p(t|c)$ as $p(w_1, w_2, w_3, \dots, w_i | c)$. The Bayesian model states that given a word w_i extracted from text t , the probability that the class is c_1 given w_i is described as follows.

$$p(c_1/w_i) = \frac{p(w_i/c_1) \cdot p(c_1)}{p(w_i)}$$

For c_2 , the probability is described as follows.

$$p(c_2/w_i) = \frac{p(w_i/c_2) \cdot p(c_2)}{p(w_i)}$$

The probability of a class $p(c_i)$ is calculated as follows:

$$p(c_i) = \frac{n(c_i)}{N}$$

where $n(c_i)$ the number of words in class c_i , and N the total number of words in the training set. Given a class c_i , the probability that a word w_i belongs to that class is given by the following:

$$p(w_i/c_i) = \frac{n_{w_i}(c_i)}{n(c_i)}$$

where $n_{w_i}(c_i)$ is the number of words w_i in the class c_i .

By applying the previous model, we designed two naive Bayesian classifiers that make it possible to simulate an emotional state in LENNA. Thus, based on a word w_i from the interlocutor's text t , LENNA selects one of the four possible emotional states from Table 1. The first classifier or Bayes 1 classifies a word w_i written in the interlocutor's text t according to its arousal level AL (Figure 1) into one of two possible categories: low arousal (bag of words 1) or high arousal (bag of words 0). The training set of Bayesian network Bayes 1 was provided by the following set of words.

("alarmed", "neg"), ("tense", "neg"), ("angry", "neg"), ("afraid", "neg"), ("annoyed", "neg"), ("distressed", "neg"), ("frustrated", "neg"), ("fear", "neg"), ("anxiety", "neg"), ("ag-

itated", "neg"), ("furious", "neg"), ("bitter", "neg"), ("irritated", "neg"), ("mad", "neg"), ("resentful", "neg"), ("fed up", "neg"), ("aroused", "neg"), ("astonished", "neg"), ("excited", "neg"), ("delighted", "neg"), ("happy", "neg"), ("surprised", "neg"), ("determined", "neg"), ("awe", "neg"), ("amusement", "neg"), ("joyful", "neg"), ("optimistic", "neg"), ("enthusiastic", "neg"), ("loving", "neg"), ("pleased", "neg"), ("charmed", "neg"), ("grateful", "neg"), ("miserable", "pos"), ("sad", "pos"), ("gloomy", "pos"), ("depressed", "pos"), ("bored", "pos"), ("droopy", "pos"), ("tired", "pos"), ("worried", "pos"), ("taken back", "pos"), ("shocked", "pos"), ("dull", "pos"), ("anxious", "pos"), ("guilty", "pos"), ("lonely", "pos"), ("disappointed", "pos"), ("indifferent", "pos"), ("fatigued", "pos"), ("desperate", "pos"), ("troubled", "pos"), ("pleased", "pos"), ("glad", "pos"), ("serene", "pos"), ("content", "pos"), ("at ease", "pos"), ("satisfied", "pos"), ("relaxed", "pos"), ("calm", "pos"), ("confident", "pos"), ("hopeful", "pos"), ("peaceful", "pos"), ("comforted", "pos"), ("powerful", "pos"), ("empowered", "pos"), ("sure", "pos"), ("dynamic", "pos"), ("ambitious", "pos"). In the training set, "neg" labels class 0 and "pos" labels class 1.

Likewise, we designed a second classifier or Bayes 2, which can be used to classify a word w_i from the interlocutor according to its valence V (Figure 1) in one of two possible categories: unpleasant (0) or pleasant (1). In the Bayes 2 classifier, the training set comprised the following list of words.

("alarmed", "neg"), ("tense", "neg"), ("angry", "neg"), ("afraid", "neg"), ("annoyed", "neg"), ("distressed", "neg"), ("frustrated", "neg"), ("fear", "neg"), ("anxiety", "neg"), ("agitated", "neg"), ("furious", "neg"), ("bitter", "neg"), ("irritated", "neg"), ("mad", "neg"), ("resentful", "neg"), ("fed up", "neg"), ("aroused", "pos"), ("astonished", "pos"), ("excited", "pos"), ("delighted", "pos"), ("happy", "pos"), ("surprised", "pos"), ("determined", "pos"), ("awe", "pos"), ("amusement", "pos"), ("joyful", "pos"), ("optimistic", "pos"), ("enthusiastic", "pos"), ("loving", "pos"), ("pleased", "pos"), ("charmed", "pos"), ("grateful", "pos"), ("miserable", "neg"), ("sad", "neg"), ("gloomy", "neg"), ("depressed", "neg"), ("bored", "neg"), ("droopy", "neg"), ("tired", "neg"), ("worried", "neg"), ("taken back", "neg"), ("shocked", "neg"), ("dull", "neg"), ("anxious", "neg"), ("guilty", "neg"), ("lonely", "neg"), ("disappointed", "neg"), ("indifferent", "neg"), ("fatigued", "neg"), ("desperate", "neg"), ("troubled", "neg"), ("pleased", "pos"), ("glad", "pos"), ("serene", "pos"), ("content", "pos"), ("at ease", "pos"), ("satisfied", "pos"), ("relaxed", "pos"), ("calm", "pos"), ("confident", "pos"), ("hopeful", "pos"), ("peaceful", "pos"), ("comforted", "pos"), ("powerful", "pos"), ("empowered", "pos"), ("sure", "pos"), ("dynamic", "pos"), ("ambitious", "pos"). In this training set, "neg" refers to class 0 and "pos" to class 1.

Once a word w_i from the interlocutor has been recognized and depending on the class in which Bayes 1 and Bayes 2 (or polarity test) classify w_i , respectively, we obtained the LENNA's emotional state (Table 3). Such a state depends on output $AL = \{0, 1\}$ from Bayes 1 and output $V = \{0, 1\}$ from Bayes 2 (or polarity test).

Table 3. Classification of LENNA's emotional states from two naive Bayesian networks.

Bayes 1	Bayes 2/Polarity Test	
	Stress 00 (state 1)	Excitation 01 (state 2)
	Depression 10 (state 3)	Healthy 11 (state 4)

After sentiment analysis of the interlocutor's text, AL and V outputs are sent to an artificial neural network (Figure 4). As mentioned above, the artificial neural network acts as a connection port, i.e., in a future project, it opens up the possibility to connect the empathic chatbot model to other models, e.g., physiological models related to emotions. This allows us to simulate the secretion of hormones associated with an emotional state or the simulation of other physiological models. Since the modeling of the artificial gland has not been included in the present study, the neural network performs only in order

to establish an association between the outputs of the Bayesian classifiers and a given emotional state.

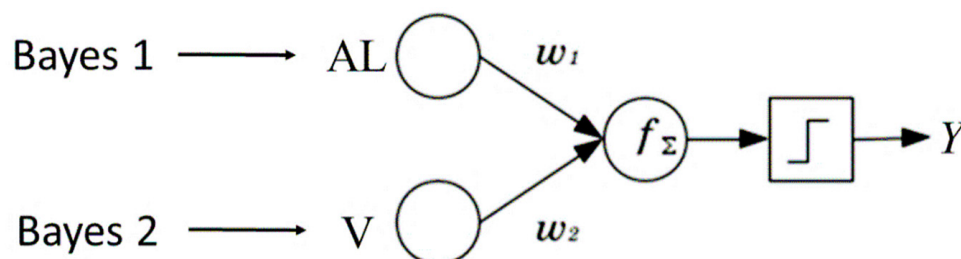


Figure 4. LENNA bot perceptron neural network (for explanation, see text).

The outputs AL and V of Bayesian classifiers are the inputs of a perceptron neural network, d being the desired output $y(t)$ (Figure 4). The perceptron was trained (Table 4) in order to associate the inputs (AL and V) with the output (d). Training was conducted according to the perceptron learning rule [28] with a bias equal to 0.3 and a learning rate r equal to 0.05:

Table 4. Perceptron training matrix.

AV	V	d
0	0	0.00
0	1	0.50
1	0	0.75
1	1	1.00

1. Initialize (randomly) the weights associated with connections w_1 and w_2 (Figure 4).
2. For each one of the inputs, AL and V, of the neural network, we calculate the output $y(t)$ of the network. Thus, once we have calculated the *net* value of the following:

$$net = AV \cdot w_1 + V \cdot w_2 + bias$$

We obtain output $y(t)$, which is given by the following sigmoid activation function.

$$y(t) = \frac{1}{1 + \exp(-net)}$$

3. In cases where y is different from d , which is the neuron with an error, the weights of the connections should be modified according to the following learning rule.

$$w_i(t+1) = w_i(t) + r(d_i - y_i)x_i, x_i = \{AV, V\}$$

In order to modulate the neural network with an artificial hormonal system in the future, we have proceeded to define function $z(t)$. Currently, such a function is established in LENNA with a relationship between each output value $y(t)$ of the neural network and one of the four possible emotional states.

$$z(t) = \begin{cases} 0.00, S \\ 0.50, E \\ 0.75, D \\ 1.00, H \end{cases}$$

The model described above corresponds to the organization depicted in Figure 1. The software implementing LENNA was written in Python 3.8 language modifying, as

stated above, the ELIZA chatbot developed by [23], using the TextBlob v0.16.0 library [29] for the implementation of the Bayesian classifiers and the polarity test. Note that by adopting the organization of Figure 1 as a prototype, two versions were designed, LENNA 1 (Figures 1 and 2) and LENNA 2 (Figure 3), for which their only difference was the procedure that implements valence V calculation. The architecture of LENNA 1 is the one described above, i.e., the sentiment analysis or opinion mining using two naive Bayesian networks. On the other hand, in LENNA 2, the second Bayesian classifier (Bayes 2) was replaced by a polarity analysis classifying a word w_i extracted from the interlocutor's text t as positive (1) or negative (0).

2.2. Simulation Experiments

The simulation experiments consisted of carrying on a conversation between a speaker and LENNA until the conversation is ended by means of the reserved word "quit". In order to evaluate the performance of the proposed chatbot, the following simulation experiments were conducted:

- A human interlocutor knowing the LENNA vocabulary converses with LENNA. The interlocutor in this experiment is labeled as HKV (human knows vocabulary).
- A human interlocutor has a conversation with LENNA, but the person does not know the vocabulary with which LENNA has been trained. The interlocutor in this experiment is labeled as HIV (human ignores vocabulary).
- A bot, i.e., an artificial interlocutor, converses with LENNA. Obviously, the bot ignores LENNA vocabulary. The following bots were selected from a chatbot repository [30]: alicia, AMA, bible, brain_bot, Einstein, eliza, parry, plumber, robo_woman and rosie, and they exhibit different "personalities". For example, eliza and alicia are the classic Weizenbaum bots that imitate the style of a "rhonegerian psychotherapist", while Einstein and plumber have a specialized conversation. Others such as parry simulate a subject suffering from paranoid schizophrenia or bible who always responds with quotes from Genesis.

The above experiments were conducted in two batches of trials: one with LENNA 1 and the other with LENNA 2. Therefore, by applying the above protocol, 60 possible conversations were conducted, 10 talks of each kind of conversation with a total of 6 possible conversations: bot-LENNA 1, bot-LENNA 2, HKV-LENNA 1, HIV-LENNA 1, HKV-LENNA 2 and HIV-LENNA 2.

2.3. Statistical Analysis

In this paper, we present a novel procedure for evaluating the emotional state transitions exhibited by a chatbot during the course of a conversation using LENNA as an empathic bot. The changes in emotional state were analyzed from the recorded texts of conversations held between an interlocutor, either human or bot, with LENNA. The approach is based on the combination of machine learning techniques, particularly sentiment analysis, with multivariate statistical analysis methods. Statistical analyses, both univariate and multivariate, were conducted with STATGRAPHICS Centurion 18 Version 18.1.12 statistical package.

In each experiment, the conversation was recorded in addition to a sequence of emotional states that LENNA showed throughout the conversation (stress, excitement, depression or healthy). The sequence of emotional states was analyzed by calculating Shannon entropy $H(e)$ with a web tool [31] in order to analyze the randomness of the sequence of emotional states through which LENNA passed during the conversation:

$$H(e) = -\sum_e p(e) \cdot \log p(e)$$

where $p(e)$ is the probability of an emotional state $e = 1, 2, 3, 4$ for stress, excitement, depression and healthy, respectively. Entropy allows us to measure the predictability of an emotional state, and the more predictable the emotional state of the chatbot, the closer

the entropy is to 0; moreover, the more unpredictable its emotional state is, the closer the entropy is to a value of 1.

We first analyzed entropy by applying univariate statistical methods, comparing with an ANOVA test the mean value of entropy in the six classes of possible conversations between LENNA and an interlocutor. Secondly, we applied sentiment analysis or text mining techniques to the texts of the recorded conversations. Sentiment analysis allows us to conduct quantitative analysis of the text by extracting subjective information from an analysis of the polarity, i.e., the positive or negative connotation of the language used, inferring from the emotional state of the text the “mental state” or mood of the conversation between an interlocutor and LENNA. By applying text mining, the conversations expressed in the texts were analyzed with the Syuzhet 1.0.6 package [32] under RStudio Version 1.1.419. The procedure was conducted according to the following experimental protocol.

The texts of the conversations between an interlocutor and LENNA were collected and normalized, i.e., cleaned. Following this, they were tokenized or fragmented into smaller strings or sentences. Sentiment was evaluated with NRC Emotion Lexicon (NRC Word-Emotion Association Lexicon, Version 0.92) [33]. This lexicon comprised list of English words and their associations with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy and disgust) and two sentiments (negative and positive). The result of this analysis was a sentiment vector in which its length is the number of sentences in the text, wherein the values represent the evaluation of the sentiment in one sentence of the text. For each sentence, valence was also obtained by calculating its value as the difference between the number of positive and negative words as well as the number of words associated with the above emotions and sentiments. The percentage of words associated with an emotion and sentiment in the text as a whole was also calculated. From the values of the sentiment vector, we examined how the emotions are distributed throughout the text. To this end, several univariate statistics were obtained, i.e., minimum value, first quartile, median, mean, third quartile and maximum value, with which an overall assessment of each conversation was obtained. Next, in order to elucidate how the conversation narrative is structured and, thus, how the frequency of words expressing positive or negative feelings changes over the course of the conversation, we obtained a trajectory graph. A trajectory graph is a representation of narrative time versus emotional valence. With this plot, it was possible to analyze which sentence had a positive or negative emotion expressed, if the conversation evolved on neutrality, if a certain narrative had a happy ending, etc. Finally, in order to eliminate the extreme values of sentiment, we obtained a Fourier transform by turning the trajectory graph into another equivalent graph that is independent of the length of the conversation [34].

Finally, in the third place, we obtained a data matrix for each conversation, and the rows are the values of the sentiment vector (minimum value, first quartile, median, mean, third quartile and maximum value), entropy and the total number of words associated with each emotion (anger, fear, anticipation, trust, surprise, sadness, joy and disgust). Likewise, in the matrix, the sentiments (negative and positive) were represented in columns. From the data matrix, we performed a principal component analysis and a discriminant analysis, which included a corresponding biplot and a graph of the discriminant functions with the classification of the conversations, respectively.

3. Results

From the 60 conversations held with LENNA and the calculation of entropy in the sequence of emotional states in each conversation, we obtained, in the Shapiro–Wilk test, a p -value equal to 0.37986, concluding that entropy fits to a normal distribution $N(0.12, 0.05)$ (Figure 5). Likewise, the evaluation of homoscedasticity or equality of variances among the experimental groups was evaluated with a Levene’s test, confirming this equality by obtaining a p -value of 0.7262.

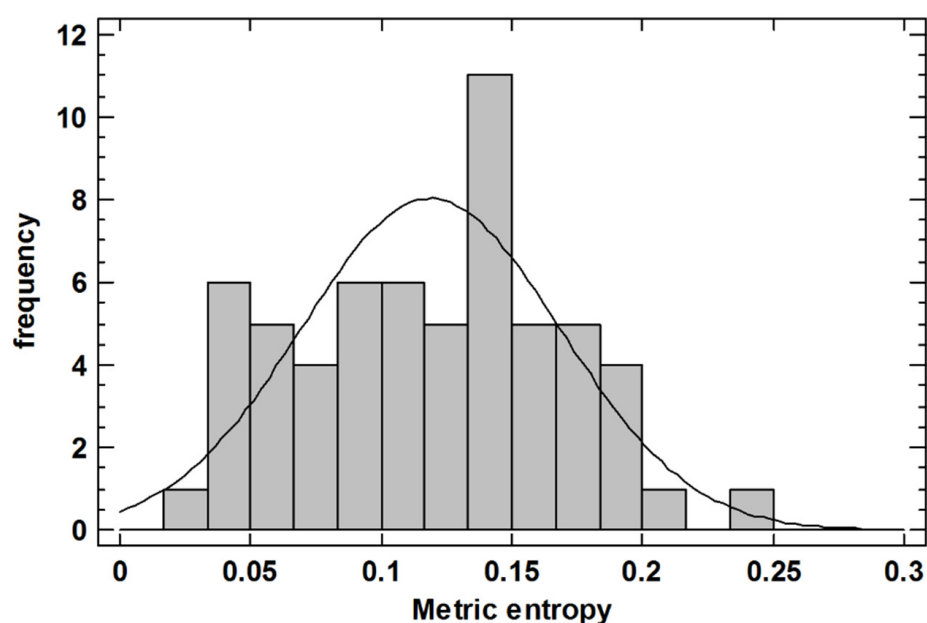


Figure 5. Frequency histogram of the entropy of the sequence of emotional states in a conversation with LENNA.

Table 5 shows the mean and standard deviation in each of the six possible types of conversations between LENNA and an interlocutor. In order to evaluate whether or not there are significant differences between the experimental groups, we performed an analysis of variance or ANOVA obtaining a p -value equal to 0.0001. Although the ANOVA assumptions are met, we conducted a non-parametric Kruskal–Wallis test of median comparison by obtaining a p -value equal to 0.000186507. Therefore, we can conclude that there are significant differences among the mean entropies or medians of two or more types of conversations. According to Fisher’s Least Significant Difference (LSD) test and Multiple Range Tests as well as the box-and-whisker plot (Figure 6), we concluded that there are significant differences between the conversations bot-LENNA1 and bot-LENNA2. Significant differences were also observed between the conversations bot-LENNA1 and those in which a person knows LENNA’s vocabulary (HKV-LENNA1 and HKV-LENNA2). There are also significant differences between HIV-LENNA1 with bot-LENNA2 and HIV-LENNA2 conversations. Finally, the differences are also significant between conversations HIV-LENNA1 and HKV-LENNA2 as well as between the latter conversations (HKV-LENNA2) and HIV-LENNA2.

Table 5. Statistical summary of the six types of conversations between LENNA and an interlocutor.

Group	n	Average	Standard Deviation
1	10	0.078221	0.051652
2	10	0.134791	0.0489143
3	10	0.134751	0.0426167
4	10	0.088339	0.0403524
5	10	0.166452	0.0260079
6	10	0.110264	0.0312346
Total	60	0.118803	0.0496933

Conversation (1) bot-LENNA 1, (2) bot-LENNA 2, (3) HKV-LENNA 1, (4) HIV-LENNA 1, (5) HKV-LENNA 2 and (6) HIV-LENNA 2.

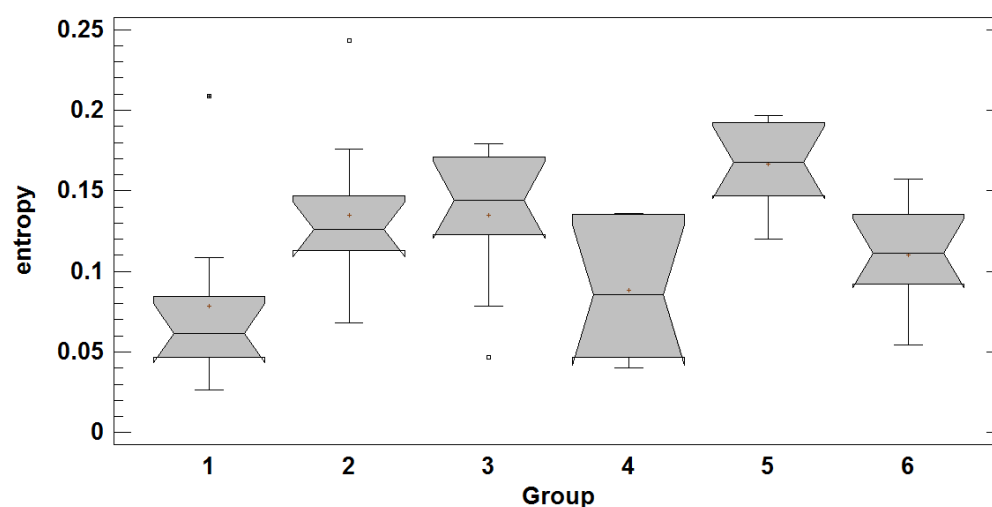


Figure 6. Box-and-whisker plot for Metric entropy in each experiment (Group) showing median entropy (notch) and its mean (cross). Conversation (1) bot-LENNA 1, (2) bot-LENNA 2, (3) HKV-LENNA 1, (4) HIV-LENNA 1, (5) HKV-LENNA 2 and (6) HIV-LENNA 2.

From the above statistical analyses, we conclude the following. In the case of a human interlocutor familiar with the vocabulary of LENNA (HKV), the sequence of emotional states that LENNA experiences during the course of the conversation exhibits greater randomness (Figure 6, groups 3 and 5) than in conversations where the interlocutor ignores (HIV) vocabulary (Figure 6, group 4 and 6). Thus, the fact that an interlocutor knows LENNA's vocabulary allows the speaker to provoke a change in the emotional state of the chatbot during a conversation. In the conversations held between LENNA and the other bot, conversations between a bot and LENNA 2 (Figure 6, group 2) exhibiting a higher average entropy can be observed; thus, greater randomness and uncertainty in the sequence of emotional states of a conversation than in conversations between a bot and LENNA 1 were also observed (Figure 6, group 1). Consequently, in LENNA 2, the fact that Bayes 2 was replaced by the polarity test provided by TextBlob library increases LENNA's sensitivity to the vocabulary of the human interlocutor. Thus, LENNA 2 architecture increases the number of emotional state changes in the chatbot. Interestingly, there is no significant differences in average entropy between conversations held between LENNA 1 and a bot (Figure 6, group 1) and the conversations held between LENNA 1 and a person ignoring the LENNA vocabulary (HIV, Figure 6, group 4). In both, we obtained the minimum entropy value.

In all simulation experiments, the conversations between a bot and LENNA 1 ended in a situation of locally trapping one or the other interlocutor repeating the same phrases, resulting in a deadlock of the conversation. According to Table 6, in all the conversations and independently of the bot, the depressed emotional state (3) predominates. One possible explanation is as follows. Since the LENNA 1 interlocutor is a bot instead of a person, the sentences expressed by the bot have a low or no arousal level AL, a fact that is detected by the naive Bayesian classifier (Bayes 1). However, in LENNA 2, a higher sensitivity of LENNA to bot vocabulary results in a greater number of emotional state changes and, therefore, to higher entropy (Table 7). It is interesting to note how, in LENNA 2, due to the low value of the arousal level AL, there is an alternation between periods of healthy emotional state (4) and periods of depressed state (3). Likewise, according to Tables 6 and 7 bots with similar architecture, e.g., alice and eliza, they present very similar entropy values. Other bots programmed to have specialized conversations on a topic, e.g., plumber speaks about plumbing, show high entropy values. In contrast, conversations with parry, the bot that simulates a paranoid schizophrenic patient, exhibited the minimum value of entropy and, consequently, the maximum predictability in conversations with LENNA. A curious result in some conversations between a bot and LENNA 2 was the sentence in which the

bot expressed its status “as a bot”, i.e., the recognition of being a simulation or having been programmed by a third party.

Table 6. Sequence of emotional states in a conversation between a bot and LENNA 1.

Bot	Emotional States	Entropy (Mean Entropy 0.07822)
alice	0333334333333	0.05948
AMA	033333	0.10834
bible	0333333	0.08452
brain_bot	0333333333	0.04690
Einstein	0333333343333	0.0594
eliza	0333333344433333	0.06390
parry	03333333333	0.03995
plumber	034333	0.20860
robo_woman	0333333	0.08452
rosie	03333333333333	0.02652

Table 7. Sequence of emotional states in a conversation between a bot and LENNA 2.

Bot	Emotional States	Entropy (Mean Entropy 0.13479)
alice	0343434333	0.12955
AMA	03333	0.14439
bible	04343334443	0.12260
brain_bot	03444443344	0.11279
Einstein	04343434	0.17570
eliza	03434344443	0.12020
parry	03333333	0.06795
plumber	044334	0.24319
robo_woman	0333333	0.08452
rosie	01433343334	0.14702

Figure 7 shows the biplot for the first two principal components with the first three principal components explaining 82.3045% of the variability of the original data. According to the figure, the correlation between the value of positive emotions (joy, trust, surprise and anticipation) and the value of positive sentiment (pos) is observed, as well as between negative emotions (fear, anger disgust and sadness) and the value of negative sentiment (neg). Likewise, the correlation between entropy (entropy) and the median value (Me) extracted from the sentiment vector can be observed. In the first component (PC1), all weights are similar and positive for the eight emotions, i.e., joy, trust, surprise, anticipation, fear, anger disgust and sadness, and the two sentiments (pos and neg). Moreover, PC1 is weakly influenced by entropy (entropy) and the median (Me) of the sentiment vector. Thus, PC1 summarizes the result of sentiment analysis. In contrast, in the second component (PC2), although entropy has no effect on the component value and the weights are similar, the sign of the emotions' weights is positive and negative for positive and negative sentiments, respectively. Finally, the third principal component (PC3) barely picks up the influence of emotions, mainly reflecting the influence of entropy.

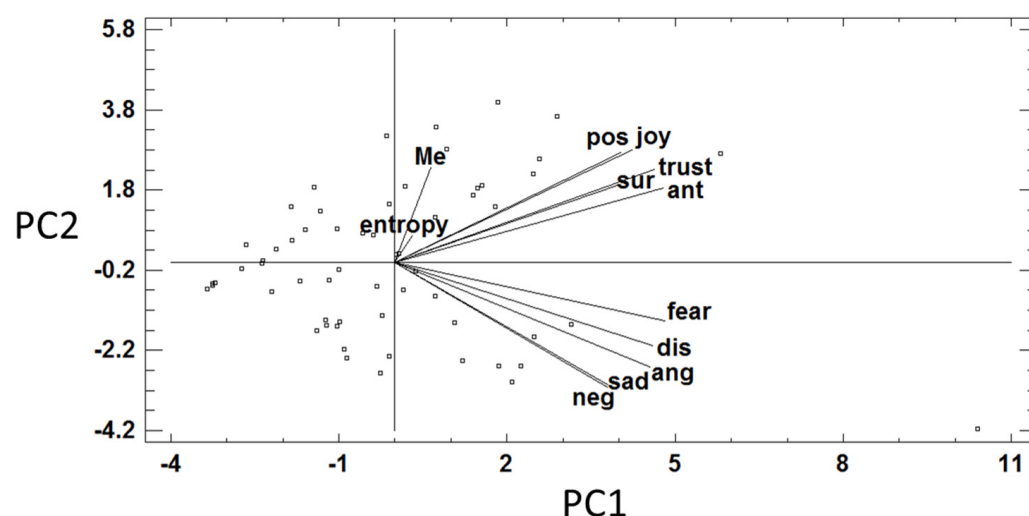


Figure 7. Biplot showing 60 conversations as a function of the first (PC1) and second (PC2) principal components. The figure shows emotions (ang = anger, fear = fear, ant = anticipation, trust = trust, sur = surprise, sad = sadness, joy = joy and dis = disgust) and two sentiments (neg = negative and pos = positive).

Figure 8 shows 60 conversations according to F1 and F2 classification functions, with two differentiated groups: G1 group is composed of conversations between a bot and LENNA and G2 group is made up of the conversations between a person and LENNA as long as the person knows LENNA vocabulary (HKV). It is interesting to note that between these two clusters of conversations, G1 and G2, the figure shows scattered conversations between a person and LENNA when a person ignores the LENNA vocabulary (HIV). Discriminant analysis conducted with all conversations shows how both groups differed significantly on the type of conversation, since the discriminant function is statistically significant (p -value of 0.0000). In discriminant analysis, 71.67% of the conversations were correctly classified. This is a very important result of the present study.

In the conversations with LENNA, the Fourier analysis (Figures 9 and 10) of the graph representing emotional valence, i.e., the difference between the number of words expressing positive emotions and those reflecting negative emotions, with respect to the narrative time allowed us to conclude the following. We concluded that in general there is a trend to form certain patterns during the conversation. In a conversation in which positive sentiment predominates and, therefore, the number of words expressing positive emotions is greater than the number of words expressing negative emotions, the Fourier plot usually shows a pattern formed by two positively connected peaks. For example, Figure 9 shows, in a frame, positive emotional valence, i.e., the time interval in a conversation in which the chatbot, LENNA in the experiment, “felt” joy and, therefore, attractive emotions. Thus, the bot was “feeling good” because the number of words reflecting positive emotions was higher than the number of words expressing negative emotions. In some cases, one of the peaks extended over time while in other cases, between the two peaks, a smaller one with a negative sign is observed (Figure 9a). This pattern appears in conversions with LENNA, either with a person or with another chatbot. Likewise, when a conversation is dominated by negative sentiments, a negative sign peak is often observed at the beginning of the conversation (Figure 10). That is, in this case, the chatbot experienced more negative emotions than positive ones during a certain time interval in the course of the conversation with another interlocutor. It is important to mention that conversations of this kind were only observed when the conversation was between LENNA and a human interlocutor.

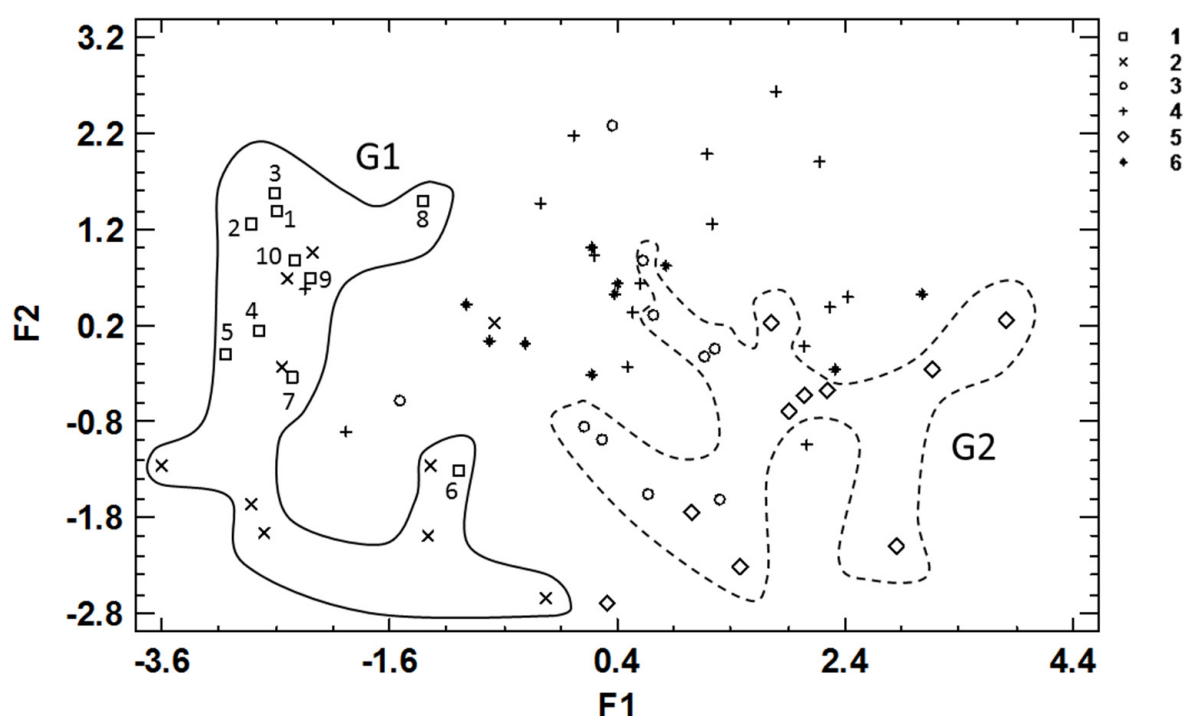


Figure 8. Discriminant analysis (F1 and F2 are classification functions) of the conversations. Group G1 is composed of conversations between a bot (1 = alice, 2 = parry, 3 = rosie, 4 = Einstein, 5 = bible, 6 = plumber, 7 = eliza, 8 = brain_bot, 9 = AMA and 10 = robo_woman) and LENNA. Group G2 is composed of conversations between a person and LENNA in the case where the person knows LENNA's vocabulary (HKV). Conversations between a person ignoring LENNA's vocabulary (HIV) and LENNA correspond to the scattered points between two clusters. In the figure, the classes of points refer to the following conversations: 1 = bot-LENNA 1, 2 = bot-LENNA 2, 3 = HKV-LENNA 1, 4 = HIV-LENNA 1, 5 = HKV-LENNA 2 and 6 = HIV-LENNA 2.

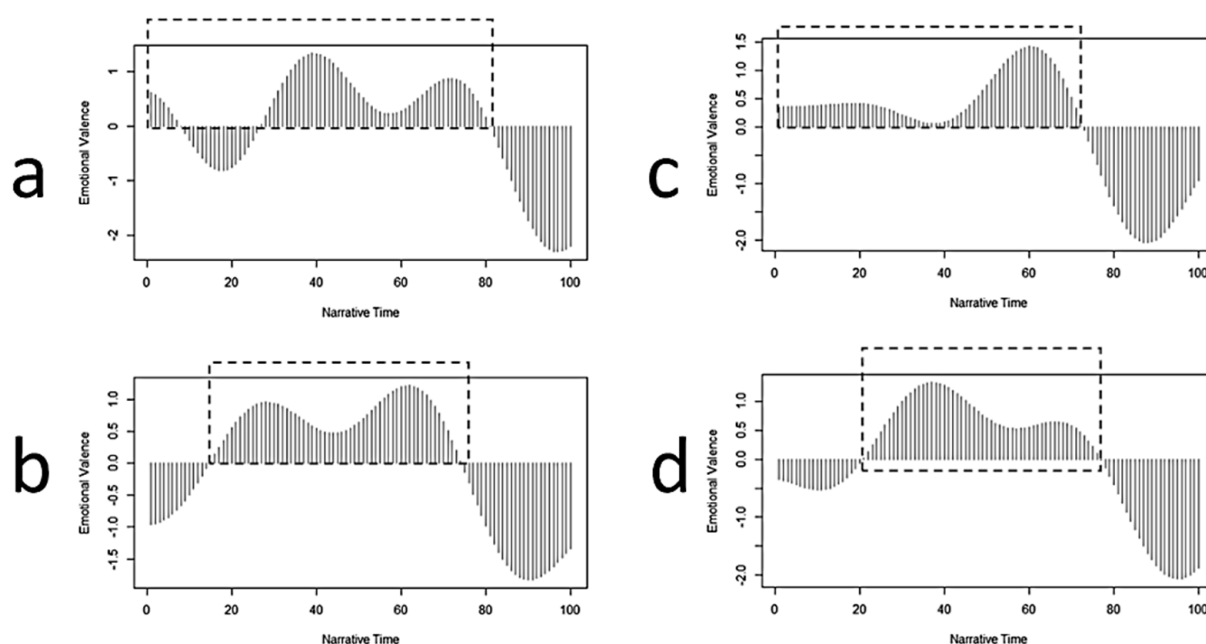


Figure 9. Fourier transform of the emotional valence with respect to the narrative time of the conversations: (a) alice-LENNA 2, (b) eliza-LENNA 2, (c) parry-LENNA 2 and (d) person HIV-LENNA 1 (for explanation, see text).

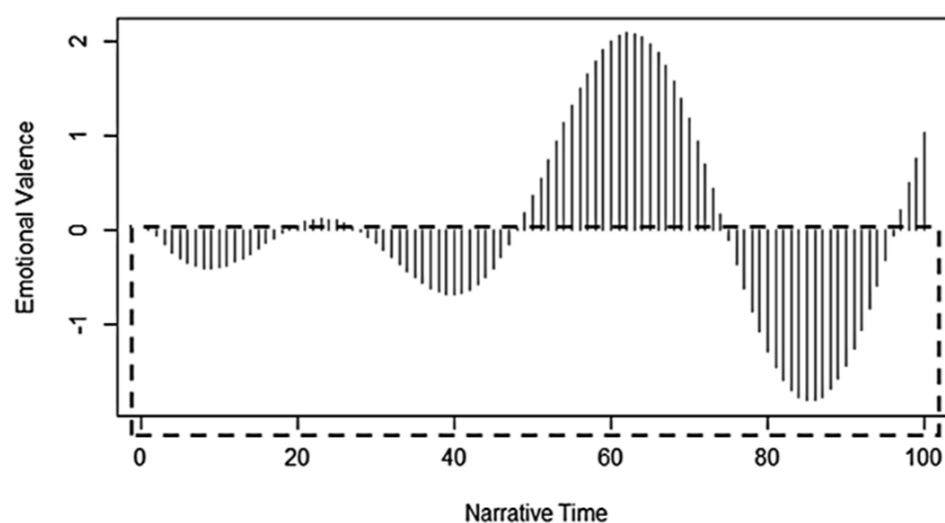


Figure 10. Fourier transform of the emotional valence with respect to the narrative time of the conversation between a person HIV and LENNA 1 (for explanation, see text).

4. Conclusions

In this study, we have shown the usefulness of a protocol oriented to evaluate the evolution and execution of an empathic chatbot during a conversation. For this purpose, we have designed an elementary model of chatbot that we have called LENNA, simulating conversations with a person or with other bots. The protocol obtains Shannon entropy and different data from the sentiment analysis method applied to the conversations. By applying multivariate statistical analysis, a correlation between entropy and the median valence (sentiment) in a conversation was obtained, obtaining two classification functions F1 and F2 that allowed discriminating the nature of the interlocutor in a conversation held between LENNA and the speaker, regardless of whether the interlocutor was a bot, a person or a person knowing or not knowing the vocabulary used by LENNA. Finally, Fourier analysis allowed us to find patterns throughout the narrative time period, which again depended on the type of interlocutor, i.e., whether it was a person or a bot. In our opinion, our methodology is not only useful for evaluating chatbots but also in other fields of artificial intelligence where conversation analysis is required.

5. Discussion

In this study, we have designed a first version of an empathic chatbot planning to simulate a virtual patient in the future suffering from pathological states or diseases caused in part by an altered emotional state. However, the main novelty of this paper is the use of entropy to evaluate emotional state changes during a conversations as well as the general protocol we followed for the analysis of the conversations between LENNA and an interlocutor that is either a person or a bot. In addition, the chatbot has been implemented by means of an architecture that we have called LENNA, and its main feature is the hybridization between different AI techniques. In particular, the simulation of empathy in the chatbot results in a combination of sentiment analysis techniques, i.e., Bayesian classifiers, with a perceptron neural network. Once LENNA is trained, the system allows the chatbot to select and classify words of a conversation, triggering its emotional state. The conducted experiments lead us to conclude that LENNA adequately simulates the transit by using different emotional states, according to the sentences used by the interlocutor when communicating with the chatbot. However, although both versions of the chatbot meet the desired goals, there are differences between LENNA 1 and LENNA 2. We observed how the use of two Bayesian classifiers in LENNA 1 provides less “sensitivity” to the interlocutor’s sentences with respect to LENNA 2. Thus, LENNA 2 is more sensitive than LENNA 1 to the interlocutor’s statements due to the replacement of one of the Bayesian classifiers by a polarity value given by a Python library.

As we mentioned above, a novel issue of the study is the use of entropy as a measure of the changes in the emotional state in a chatbot and, therefore, of its predictability (Tables 6–9). Based on entropy, it was observed that there is greater randomness and, therefore, unpredictability in the case of conversations between LENNA and a person who knows the vocabulary used in the training of the chatbot. Furthermore, the obtained results suggest that LENNA 2 is more emotionally empathetic than LENNA 1. For the above reasons and since the sensitivity of the chatbot depends on its architecture, in the future we propose to study the use of Bayesian classifiers with a training set larger than the 68 words that we implemented in the current version. In addition, in a future version of LENNA, it would be useful to include new emotional states that are not simulated in the current version in order to improve the interaction between the interlocutor and the chatbot, rendering it more natural, among other features for improving the current model.

Table 8. Sequence of emotional states in a conversation between a person and LENNA 1.

HKV (Mean Entropy 0.13475)		HIV (Mean Entropy 0.08833)	
Emotional States	Entropy	Emotional States	Entropy
01111333134	0.15243	0333331333	0.09219
03333333334	0.07871	0434333043433333333	0.06262
0232233322	0.13610	03333333333	0.03995
0433342434	0.17219	03143133333	0.13556
0333333333	0.04690	03333333333	0.03995
03131431313	0.15243	0333333333	0.04690
04344231333	0.17925	03333333433	0.07871
0231333333	0.13568	0334233333	0.13568
012133413343	0.17122	0333233343	0.13568
03223322323	0.12260	03431333333	0.11615

Table 9. Sequence of emotional states in a conversation between a person and LENNA 2.

HKV (Mean Entropy 0.16645)		HLV (Mean Entropy 0.11026)	
Emotional States	Entropy	Emotional States	Entropy
013231333111	0.13750	03333433343	0.09962
03444434433	0.12020	03334331313	0.13556
02123142222	0.16978	0333333334	0.09219
0333313431	0.15710	0333033333430303344444333	0.05437
04432411334	0.19255	03333333343	0.07871
02423333422	0.16573	0343133343	0.15710
01113211331	0.14702	044443444	0.10960
03332343421	0.18549	02433343334	0.14702
01144312242	0.19660	0333343343	0.11568
02434114334	0.19255	03343434333	0.11279

An interesting fact to note is that in conversations between LENNA and other bots, the resulting conversations have lower entropy than compared with those held with persons in which the conversations are much more predictable and, therefore, foreseeable. This not only means that LENNA expresses fewer emotional states; fewer changes takes place from one state to another. We have also observed how conversations between two chatbots experience a phenomenon of local trapping, i.e., they “enter into a loop” or end abruptly. This fact could be used in the future to design a CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) that gives a computer the ability to distinguish whether the conversation it is having is with a human being or another chatbot.

It is important to note that the use of entropy aims to predict an emotional state by ultimately measuring the fluctuation of a series of transient emotional states that reflect

how an individual, in this case, a chatbot, “feels on a psychic level” at a particular instant in time. Suppose that a bot can be in one of three emotional states A, B and C. Let us perform two experiments in which we calculate the entropy in the sequence of emotional states ABC through which a bot has passed in a conversation, LENNA in our case, and then we calculate it in the sequence of states CBA recorded in another conversation with the bot. Effectively, there are no differences in Shannon entropy values in one and another conversation in which the entropy value in the example equals 0.52832. That is to say that, in both cases, the sequence of emotional states through which the bot has passed in the course of one conversation and the other exhibited the same degree of randomness and, therefore, predictability. Another important issue is the property regarding the “source” that emits one of the three symbols, i.e., the emotional state of the bot is such that one emotional state is independent of the previous one. Thus, the future emotional state of the bot at time $t+1$ must be independent of its emotional state at t . Obviously, being an empathic bot, the future emotional state of the bot depends on the sentence said by its interlocutor. Consequently, certain words stated by the interlocutor trigger a transition to another emotional state that differs from the current one (or not). Therefore, how would we justify that Shannon entropy is still useful for measuring emotional state changes when it is possible that the sequence of emotional states is not random? Suppose we had recorded the above mentioned sequences, i.e., ABC and CBA, in two different conversations between the bot and a speaker. Obviously, the entropy value of the sequence of emotional states would be similar in the two conversations. We assume that instead of considering the symbols representing individual emotional states as in the first scenario, we calculate the entropy of a block or set of concatenated symbols, i.e., ABC and CBA. Since, in the example, we consider three emotional states and, therefore, a set of three symbols $\{A, B, C\}$, the joint entropy of a block [35] in the example will be $H(u, v, w)$:

$$H(u, v, w) = - \sum_u \sum_v \sum_w p(u, v, w) \log(p(u, v, w))$$

for which its value can be calculated if we know the joint density function $f(u, v, w)$, i.e., $p(u, v, w)$. If we also take into account that entropy is additive and if the sequence of emotional states fulfills the condition of independent events, then $H(u) + H(v) + H(w)$ will be equal to $H(u, v, w)$. Now, what if the sequences of emotional states were not random? If the sequences were not random or $f(u, v, w)$ is unknown, then $H(u) + H(v) + H(w)$ will be at most $H(u, v, w)$ concluding in such a case that entropy must be lower or equal to the one we have calculated assuming that the sequence fulfills the independence condition.

In summary, if, in the experiments performed with LENNA, the emotional states recorded in a sequence were not independent, then the actual entropy values would be lower with their maximum value being the one calculated and shown in Tables 5–9. Indeed, for the calculation of entropy in a sequence of dependent emotional states the probabilities could be calculated on the basis of a Markov chain. However, in case where the emotional states through which LENNA passes are not independent, the application of Shannon’s expression would still be valid. That is, according to the above reasoning, the calculated value of entropy would be an upper bound.

From a methodological point of view, another interesting issue of the present study is the application of sentiment analysis techniques to a conversation held between an interlocutor and a chatbot. For instance, in the biplot obtained from principal component analysis, the correlation between the entropy and the median extracted from the sentiment vector suggests how the analysis of emotional states including entropy results in similar conclusions as the analysis based on the median extracted from the sentiment vector. Moreover, the possibility of classifying a conversation by using discriminant analysis, i.e., whether it has been between two chatbots or between a person and a chatbot, opens a very interesting path for the application of sentiment analysis techniques in the analysis and assessment of chatbot performance.

In this study, we have used two Naïve Bayes Classifiers in the NPL module to obtain LENNA's emotional state for the following reasons. The use of Naïve Bayes Classifiers has been chosen among other machine learning techniques, e.g., SVM, because it allows us to classify properly the responses of a user when categorizing a text using fewer hardware resources (CPU and memory). It is also an efficient technique in terms of execution time, and it performs well with small training sets [36]. Moreover, Bayesian Classifiers are common in modeling scenarios where humans coexist with chatbots, which makes it an appropriate choice when designing a minimally viable chatbot prototype. For instance, at present, malware and spam are distributed by chatbots by using different media such as chat networks, online games, etc. A Naïve Bayes Classifier in combination with other classifiers, e.g., with an entropy classifier, is a useful tool for distinguishing between a human user and a chatbot that represents a threat [37]. Another example is in medicine by successfully identifying the symptoms of respective diseases [38,39] and in bots designed to conduct job interviews [40], etc. However, machine learning techniques other than Naïve Bayes Classifiers have been successfully used. For example, in the field of medicine, "medical bots" have been designed for disease diagnosis using Support Vector Machines [41].

In the present study, instead of using a multi-class Bayesian Classifier, we have used two binary Bayesian Classifiers from the TextBlob library. In the current version of the chatbot, once a word in a sentence of the interlocutor is detected, the word is then classified in one of the four classes of emotions (Table 2). The classification is the result of the intersection of the outputs of the two Bayesian networks: stress (00), excitement (01), depression (10) and healthy (11). In the future, when we modify LENNA's architecture with virtual glands including a hormonal system, we will replace the two binary Bayesian Classifiers by a Bayesian multi-class classifier, e.g., using the Scikit-Learn Python library.

In other fields different from AI, the designed empathic chatbot enables, in the future, the possibility of modeling and simulating diseases for which its origin could be related to the emotional state of a subject. As already mentioned throughout the paper one possible approach would be the connection of LENNA's neural network with an artificial endocrine system [22]. Such a system would allow the application of a chatbot for the clinical study of patients suffering from post-traumatic stress disorder and in people suffering from severe depressive disorders [42] where the hypothalamic–pituitary–adrenal axis promotes the secretion of hormones that ultimately encourages the release of stress-related hormones. The inclusion of a perceptron model that receives the output of the Naive Bayes Classifiers seems to be an unnecessary addition to LENNA architecture. However, this is not the case since, in this study, we present an empathic bot that is minimally viable and can be extended in the near future. Our intention is to extend and complete the architecture that supports the chatbot by providing a perceptron with a bio-inspired emotion model. The model relies on the ability of neurons to communicate with hormonal glands, resulting in more complex and realistic activity patterns than those of an artificial neural network. Therefore, our future strategy will be to endow LENNA with virtual glands and a feedback mechanism between the neural network and the hormonal system in the style of the EMANN model [43]. In fact, the bio-inspired simulation of emotions improves the performance of artificial neural networks. Although in our case the purpose is different, there are currently useful applications of artificial endocrine systems, i.e., artificial neural networks modulated by hormones. For instance, applications where the magnitude of the actuators must be controlled, such as autonomous robots able to sail similarly to a ship [44] or intelligent systems designed to predict overpressure in the atmosphere [45]. In summary, empathic chatbots are still a research topic for which its development can be expanded. For example, one option is to add the recognition of facial microexpressions by mapping them in the interlocutor [46]. Obviously, these are possibilities that open up a wide range of potential ideas to be addressed in the future. However, to date, LENNA has been designed with elementary architecture that defines the chatbot as a minimally viable prototype. For this reason, the emotional plane or pleasure-arousal plane of the chatbot

has been chosen in a similar manner to [47], i.e., reducing the possible emotions to the four quadrants of the plane (stress, excitement, depression and calm).

The design of empathic chatbots that express emotions is an open field in which advances will arise from merging concepts and techniques coming from different areas of knowledge such as neurobiology [48] and AI [49]. This approach results in more realistic models of emotion, where the emotional behavior of a bot is closer to a human being. In addition, it will be necessary to design protocols by combining different techniques, e.g., sentiment analysis and multivariate statistical analysis with data mining, among other emerging techniques, enabling an adequate analysis of the emotions experienced by a chatbot during a conversation.

Author Contributions: C.C.L. has collaborated in the Introduction and carried out simulation experiments that were used in her Master of Healthcare Biology 2020–2021, Complutense University of Madrid. R.L.-B. devised the general problem, the model and wrote the Python routines. He has also supervised the work of the second author and wrote this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

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

References

1. Weizenbaum, J. ELIZA—A computer program for the study of natural language communication between man and machine. *Commun. ACM* **1966**, *9*, 36–45. [\[CrossRef\]](#)
2. Rogers, C. A theory of therapy, personality and interpersonal relationships as developed in the client-centered framework. In *Psychology: A Study of a Science. Volume 3: Formulations of the Person and the Social Context*; Koch, S., Ed.; McGraw Hill: New York, NY, USA, 2010.
3. Colby, K.M.; Hilf, F.D.; Weber, S.; Kraemer, H. Turing-like indistinguishability tests for the validation of a computer simulation of paranoid processes. *Artif. Intell.* **1972**, *3*, 199–221. [\[CrossRef\]](#)
4. Pamungkas, E.W. Emotionally-aware chatbots: A survey. *arXiv* **2019**, arXiv:1906.09774.
5. Darwin, C. *The Expression of the Emotions in Man and Animals*; John Murray: London, UK, 1872.
6. Picard, R.W. Affective computing for future agents. In *Cooperative Information Agents IV—The Future of Information Agents in Cyberspace. CIA 2000*; Klusch, M., Kerschberg, L., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2000; Volume 1860. [\[CrossRef\]](#)
7. Ghandeharioun, A.; McDuff, D.; Czerwinski, M.; Rowan, K. EMMA: An emotion-aware wellbeing chatbot. In Proceedings of the 8th International Conference on Affective Computing and Intelligent Interaction (ACII), Cambridge, UK, 3–6 September 2019; pp. 1–7.
8. Posner, J.; Russell, J.A.; Peterson, B.S. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev. Psychopathol.* **2005**, *17*, 715–734. [\[CrossRef\]](#)
9. Gupta, S.; Borkar, D.; De Mello, C.; Patil, S. An e-commerce website based chatbot. *Int. J. Comput. Sci. Inf. Technol.* **2015**, *6*, 1483–1485.
10. Ho, A.; Hancock, J.; Miner, A.S. Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *J. Commun.* **2018**, *68*, 712–733. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Romero, M.; Casadevante, C.; Montoro, H. How to create a psychologist-chatbot. *Psychol. Pap.* **2020**, *41*, 27–34.
12. Denecke, K.; May, R.; Deng, Y. Towards emotion-sensitive conversational user interfaces in healthcare applications. *Stud. Health Technol. Inform.* **2019**, *264*, 1164–1168.
13. Zunszain, P.A.; Hepgul, N.; Pariante, C.M. Inflammation and depression. *Curr. Top. Behav. Neurosci.* **2013**, *14*, 135–151.
14. Murata, M. Inflammation and cancer. *Environ. Health Prev. Med.* **2018**, *23*, 50–58. [\[CrossRef\]](#)
15. Raman, J.; Smith, E.; Hay, P. The clinical obesity maintenance model: An integration of psychological constructs including mood, emotional regulation, disordered overeating, habitual cluster behaviours, health literacy and cognitive function. *J. Obes.* **2013**, *2013*, 240128. [\[CrossRef\]](#)
16. Lahoz-Beltra, R.; Rodriguez, R.J. Modeling a cancerous tumor development in a virtual patient suffering from a depressed state of mind: Simulation of somatic evolution with a customized genetic algorithm. *Biosystems* **2020**, *198*, 104261. [\[CrossRef\]](#)

17. Spring, T.; Casas, J.; Daher, K.; Mugellini, E.; Abou Khaled, O. Empathic response generation in chatbots. In Proceedings of the 4th Swiss Text Analytics Conference (SwissText 2019), Winterthur, Switzerland, 18–19 June 2019.
18. Pang, B.; Lee, L. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.* **2008**, *2*, 1–135. [\[CrossRef\]](#)
19. Wilson, T.; Wiebe, J.; Hoffmann, P. Recognizing contextual polarity in phrase-level sentiment analysis. In Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT '05), Vancouver, BC, Canada, 6–8 October 2005; Association for Computational Linguistics: Stroudsburg, PA, USA, 2005; pp. 347–354. [\[CrossRef\]](#)
20. Dayhoff, J.; Deleo, J. Artificial neural networks: Opening the black box. *Cancer* **2001**, *91*, 1615–1635. [\[CrossRef\]](#)
21. Tzirakis, P.; Trigeorgis, G.; Nicolaou, M.A.; Schuller, B.W.; Zafeiriou, S. End-to-end multimodal emotion recognition using deep neural networks. *IEEE J. Sel. Top. Signal. Process.* **2017**, *11*, 1301–1309. [\[CrossRef\]](#)
22. Xu, Q.-Z.; Wang, L. Recent advances in the artificial endocrine system. *J. Zhejiang Univ.-Sci. Comput. Electron.* **2011**, *12*, 171–183. [\[CrossRef\]](#)
23. Strout, J.; Epler, J. Eliza.py, ELIZA in Python. 2017. Available online: <https://github.com/jezhiggins/eliza.py> (accessed on 8 October 2021).
24. MacLean, P.D. Some psychiatric implications of physiological studies on frontotemporal portion of limbic system (visceral brain). *Electroencephalogr. Clin. Neurophysiol.* **1952**, *4*, 407–418. [\[CrossRef\]](#)
25. Ekman, P.; Sorenson, E.R.; Friesen, W.V. Pancultural elements in facial displays of emotions. *Science* **1969**, *164*, 86–88. [\[CrossRef\]](#)
26. Shiotani, M.N. Ekman's theory of basic emotions. In *The Sage Encyclopedia of Theory in Psychology*; Harold, L.M., Ed.; Sage Publications: Thousand Oaks, CA, USA, 2016; pp. 248–250.
27. Yu, L.-C.; Lee, L.-H.; Hao, S.; Wang, J.; He, Y.; Hu, J.; Lai, K.R.; Zhang, X. Building Chinese affective resources in valence-arousal dimensions. In Proceedings of the NAACL-HLT 2016, San Diego, CA, USA, 12–17 June 2016; pp. 540–545.
28. Lahoz-Beltra, R. *Bioinformática: Simulación, Vida Artificial e Inteligencia Artificial*; Ediciones Díaz de Santos: Madrid, Spain, 2004.
29. Loria, S. TextBlob Documentation. Release v0.16.0. 2020. Available online: <https://textblob.readthedocs.io/en/dev/> (accessed on 8 October 2021).
30. BOT LIBRE. The Open Source Chatbot and Artificial Intelligence Platform. Available online: <https://www.botlibre.com/> (accessed on 8 October 2021).
31. Kozłowski, L. Shannon Entropy Calculator. 2021. Available online: <https://www.shannonentropy.netmark.pl/> (accessed on 8 October 2021).
32. Jockers, M. Syuzhet. Release 1.0.6. 2020. Available online: <https://cran.r-project.org/web/packages/syuzhet/syuzhet.pdf> (accessed on 8 October 2021).
33. Mohammad, S.M.; Turney, P.D. NRC Emotion Lexicon. Release 0.92. Available online: <https://saifmohammad.com/WebPages/AccessResource.htm> (accessed on 8 October 2021).
34. Jockers, M.L. Revealing Sentiment and Plot Arcs with the Syuzhet Package. 2015. Available online: <https://www.matthewjockers.net/2015/02/02/syuzhet/> (accessed on 8 October 2021).
35. Schürmann, T.; Grassberger, P. Entropy estimation of symbol sequences. *Chaos* **1996**, *6*, 414–427. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Helmi Setyawan, M.Y.; Awangga, R.M.; Efendi, S. Comparison of multinomial naive Bayes algorithm and logistic regression for intent classification in chatbot. In Proceedings of the 2018 International Conference on Applied Engineering (ICAE), Batam, Indonesia, 3–4 October 2018; pp. 1–5. [\[CrossRef\]](#)
37. Smys, S.; Haoxiang, W. Naïve Bayes and entropy based analysis and classification of humans and chat bots. *J. ISMAC* **2021**, *3*, 40–49.
38. Anapuma, C.V. Chatbot disease prediction and treatment recommendation using machine learning. *High Technol. Lett.* **2021**, *27*, 354–358.
39. Zygałło, A.; Kozłowski, M.; Janicki, A. Text-Based emotion recognition in English and Polish for therapeutic chatbot. *Appl. Sci.* **2021**, *11*, 10146. [\[CrossRef\]](#)
40. Sarosa, M.; Junus, M.; Hoesny, M.; Sari, Z.; Fatnuriyah, M. Classification technique of interviewer-bot result using naïve Bayes and phrase reinforcement algorithms. *Int. J. Emerg. Technol. Learn. IJET* **2018**, *13*, 33–47. [\[CrossRef\]](#)
41. Tamizharasi, B.; Livingston, J.; Rajkumar, S. Building a medical chatbot using support vector machine learning algorithm. *J. Phys. Conf. Ser.* **2021**, *1716*, 012059. [\[CrossRef\]](#)
42. Kasckow, J.W.; Baker, D.; Geraciotti, T.D., Jr. Corticotropin-releasing hormone in depression and post-traumatic stress disorder. *Peptides* **2001**, *22*, 845–851. [\[CrossRef\]](#)
43. Thenius, R.; Zahadat, P.; Schmickl, T. EMANN—A model of emotions in an artificial neural network. In Proceedings of the ECAL 2013: The Twelfth European Conference on Artificial Life, Sicily, Italy, 2–6 September 2013; pp. 830–837. [\[CrossRef\]](#)
44. Sauzé, C.; Neal, M. Artificial endocrine controller for power management in robotic systems. *IEEE Trans. Neural Netw. Learn. Syst.* **2013**, *24*, 1973–1985. [\[CrossRef\]](#)
45. Temeng, V.; Yevenyo Ziggah, Y.; Arthur, C. A novel artificial intelligent model for predicting air overpressure using brain inspired emotional neural network. *Int. J. Min. Sci. Technol.* **2020**, *30*, 683–689. [\[CrossRef\]](#)
46. Xu, F.; Zhang, J.; Wang, J.Z. Microexpression identification and categorization using a facial dynamics map. *IEEE Trans. Affect. Comput.* **2017**, *8*, 254–267. [\[CrossRef\]](#)
47. Yan, F.; Iliyasu, A.M.; Jiao, S.; Yang, H. Quantum Structure for Modelling Emotion Space of Robots. *Appl. Sci.* **2019**, *9*, 3351. [\[CrossRef\]](#)

-
48. Arbib, M.A.; Fellous, J.-M. Emotions: From brain to robot. *TREND Cogn. Sci.* **2004**, *8*, 554–561. [[CrossRef](#)]
 49. Samani, H.A.; Saadatian, E. A multidisciplinary artificial intelligence model of an affective robot. *Int. J. Adv. Robot. Syst.* **2012**, *9*, 6. [[CrossRef](#)]