Universidad Autónoma de Nuevo León

Facultad de Ingeniería Mecánica y Eléctrica

Subdirección Académica

Sentiment Analysis through a chatbot

por

Alexander Espronceda Gómez

como requisito parcial para obtener el grado de

ingeniería en tecnología de software

Enero 2022

Universidad Autónoma de Nuevo León

Facultad de Ingeniería Mecánica y Eléctrica

Subdirección Académica

Sentiment Analysis through a chatbot

por

Alexander Espronceda Gómez

como requisito parcial para obtener el grado de

ingeniería en tecnología de software

Enero 2022

Universidad Autónoma de Nuevo León

Facultad de Ingeniería Mecánica y Eléctrica

Subdirección Académica

Los miembros del Comité de Tesis recomendamos que la Tesis «Sentiment Analysis through a chatbot», realizada por el alumno Alexander Espronceda Gómez, con número de matrícula 1742000, sea aceptada para su defensa como requisito parcial para obtener el grado de Ingeniería en Tecnología de Software.

El Comité de Tesis

Dra. Satu Elisa Schaeffer
Asesor

Dra. Sara Elena Garza Villarreal
Revisor

Dr. Romeo Sánchez Nigenda
Revisor

Vo. Bo.

Dr. Fernando Banda Muñoz
Subdirección Académica

San Nicolás de los Garza, Nuevo León, enero 2022

Contents

List of Figures

List of Tables

Agradecimientos

———(WORK IN PROGRESS)———

Resumen

Alexander Espronceda Gómez.

Candidato para obtener el grado de Ingeniería en Tecnología de Software.

Universidad Autónoma de Nuevo León.

Facultad de Ingeniería Mecánica y Eléctrica.

Título del estudio: Sentiment Analysis through a chatbot.

Número de páginas: 19.

Objetivos y método de estudio: En esta tesis se propone generar un software conversacional que interprete el texto introducido por un usuario y determinar su estado de ánimo, y reaccione de acuerdo con éste por medio de frases predeterminadas.

El método de estudio utilizado hará un análisis comprensivo de las redes neuronales, así como también de la comprensión suficiente de algo tan voluble y a veces impredecible como lo es la mente humana.

Contribuciones y conlusiones: El algoritmo de entrenamiento utiliza un conjunto de datos específico para predecir lo más acertadamente posible qué está sintiendo una persona al momento de escribir alguna oración o frase.

Firma del asesor: ______________________________
Dra. Satu Elisa Schaeffer

Chapter 1

Introduction

Human beings are social beings, this is widely known. To survive, we must band together and communicate with each other, bonding in the process. This is thanks to a neural process called **empathy**, which is defined as a three-part process that happens in our brains (Elliott *et al.*, 2011). That roughly happens like this:

- Emotional simulation centered in the limbic system, which makes us mirror the emotional elements we're watching.
- Processing the perspective in the prefrontal and temporal cortex.
- Assessing the course of action to take, either showing compassion or doing something else. This is allegedly based in the obitofrontal cortex, as well as several other parts of the brain.

This is clearly what it's usually considered a human-only behavior, but there's been studies that indicate that apes, dogs and rodents have been observed to take action at the presence of distress signals, either from humans or other members of their own species (Preston y de Waal, 2002). If this is true, theoretically, a machine could be taught to process signals of distress and react accordingly using a learning algorithm.

1.1 Justification

At first, the objective was to create an algorithm that could serve as a makeshift therapy chatbot that people could use when they were confused about their own feelings, but as time has passed, a lot of things have happened in my life regarding people with close-to-none empathy. This project could prove especially useful towards people who have trouble discerning when to console someone or having an idea of how other people or even themselves feel, such as the case of people with Asperger's Syndrome or other forms of high-functioning autism. To this end, the decision was made to work on this project.

1.2 Hypothesis

Empathy consists in a pattern of neurochemical reactions triggered by different situations. Machine learning could learn to identify these patterns. The hypotesis of this thesis is that machine learning could help people with a vague sense of empathy or self-knowledge to discern what they're feeling.

1.3 Objectives

In this section, the objectives proposed for this paper are established.

1.3.1 General Objectives

The objective of this project is to determine how the person that writes the input text is feeling according to the words in it. This could be achieved thanks to the technology present in machine learning algorithms and an extensive amount of datasets.

1.3.2 Specific Objectives

- Generating an algorithm capable of detecting key words related to mood in text with TensorFlow.
- Predicting successfully the mood according to the input given.
- Giving feedback on the input, reinforcing it if it's positive or giving empathetic words if it's negative

1.4 Metodology

The tools that are used in this paper are mostly Python-based, such as TensorFlow, a neural network framework. This, combined with natural language processing tools and several filtering techniques will be used to achieve – or at least approach as close as possible to – the expected results.

1.5 Structure

The content in this thesis is divided in several chapters, each one of them talking about different information about either the topics that are relevant to the scope of this project or the general process that has happened to reach the goal.
In the second chapter, topics like relevant concepts are discussed and expanded upon, also related content and similar projects are looked upon and compared to this project.
In the third chapter, a general approach to the project's process is described, with some screenshots of the relevant information.

Chapter 2

Background

Technology in the past decades has been advancing exponentially. So much, in fact, that we can relegate data analysis to them for better accuracy and reliability than what a human can possibly achieve. This is what it's called Machine Learning (sometimes referred only as ML) There's a variety of scenarios where it comes in handy, such as pattern recognition, which relates extensively to most of this project's work.

In this chapter, some key concepts will be explained for easier comprehension of this thesis and the project itself as a whole.

2.1 Machine Learning

Machine learning can be described, broadly and figuratively speaking, as a black box where some data is inserted as an input and numbers come out of it as an output (Zhang, 2020). Some more advanced models of ML allow some internal parameters inside this figurative black box to be able to be tampered with, so that some characteristics of the input data can have effect on the output, these parameters are called *weights* (Ayodele, 2010). Most ML algorithms have two stages: training and validation:

- Training processes the inputs and makes educated guesses, and in case of guessing incorrectly, depending on the obtained result, the weights are changed accordingly.

- Validation is as simple as it sounds, some input is fed to the algorithm and information needs to be compared to the real results to test the accuracy percentage.

One of these models that is one of the most used nowadays is the one called *Neural Network*.

2.1.1 Neural Network

A neural network works by using *neurons*, they utilize layers that individually weigh the input given to them from the initial text or, if this has been processed already, from another neuron (Ayodele, 2010). Likewise, similar to how biological brains work, these algorithms can only predict reliably if given enough data to train and validate their outputs with.

2.2 Sentiment Analysis

Sentiment Analysis (or Opinion Mining, as it is also known) as a tool for data analysis is arguably a recent happening. The term was coined in 2003 and has evolved ever since (Kumar y Teeja, 2012). This type of data analysis has a lot of potential usages that have yet to be implemented in the daily life.

2.2.1 Concept

The specific execution of the algorithm varies depending on the intended purpose, but the concept and process that is used is generally the same:

- The sentence to analyze is broken down to its component parts, this process is called *tokenization*, and the resulting products are called, fittingly, *tokens*.
- Every token is then tagged, making it part of an internal dictionary or *lexicon*
- A score is assigned to every token depending on the used dataset.

The end score could be left as-is or can be reintroduced to the algorithm for a multi-layered approach depending on its focus. (Appel *et al.*, 2015)

2.2.2 Tokenizing

Tokenizing is the process that happens while making tokens, the way it works is very straightforward: every word in the lexicon that a machine can read is assigned a number for easier reading. Let's take the following example:

This is an example text

We can tell there are 6 words in the example phrase. So the tokenizing process would make the example look in the following way:

1, 2, 3, 4, 5, 6

Where 1 corresponds to the word "This", 2 corresponds to "is", 3 to "an" and so on.

The interesting part about this process would happen if we used another example phrase, like the following:

$\boxed{\text{This is another example}}$

If we did the tokenization process, it would be processed in this way:

$\boxed{1, 2, 7, 4}$

Since the internal lexicon already knows some of the words in this second example, it reuses their token, adding new ones (in this example, "another" is 7) if needed.

This is fairly useful for a machine learning algorithm, since it won't have to compare such massive amount of characters in a string each time, and it would only need to evaluate integers. Whether it's frequency or comparison.

2.3 Similar Applications

The algorithm proposed on this paper is, of course, not the only sentiment analysis application by a long stretch. There are many applications that already apply this kind of Machine Learning for several purposes, like Movie Review algorithms detecting sentiment from IMDB (Pang *et al.*, 2002), or Koko, which uses the OpenAI API which is a counseling app for distressed teenagers (Morris *et al.*, 2018).
It's important to mention GPT-3 as a whole as well, which, to date, it's one of the most impressive AI algorithm to be developed, the downsides being that it's still in beta phase, it's super resource-heavy, and its access is reserved to businesses through a fee, very expensive to use for the general public, especially students as myself. That's why in this paper, TensorFlow is used, which is free to use, doesn't need a lot of resources to work and it's portable once it's trained.

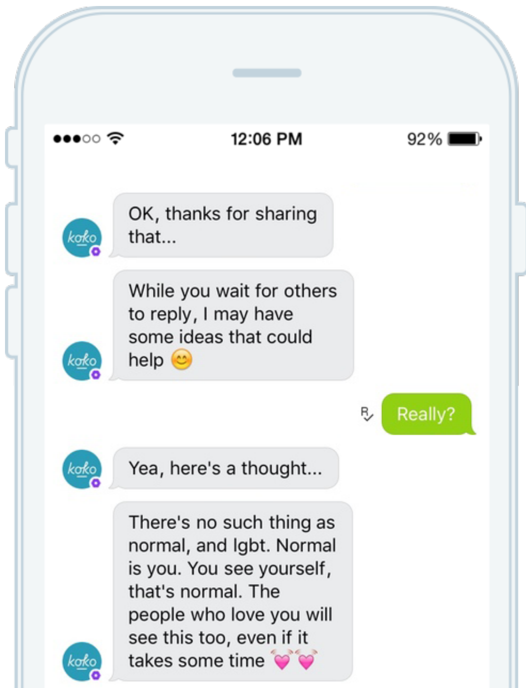

Figure 2.1: Screenshot of the Koko app, obtained from `https://www.koko.ai/` on 04/21/2021

Chapter 3

Project Development

3.1 Datasets

There are several datasets on the internet, but none of them have the amount of sheer volume and actually useful data that is required for this task. The closest available was used, however, and it brought relatively acceptable levels of accuracy (Calefato *et al.*, 2019). This dataset, paired with NLTK processing, stopwords and truncating words and verbs commonly used in the English language, was able to pinpoint if the input had a positive, neutral or a negative feeling about 40% of the time, approximately. This is not really a good number for such a small amount of labels, but it's an improvement nonetheless. Previous versions with different approaches, combination of layers and datasets had less than 20% of accuracy.

3.1.1 Pre-filtering

Since the dataset that was chosen was imported straight from Twitter with little to no filtering, some cleanup had to be done to ensure peak performance. The first problem was the punctuation marks, which were easy to filter out. The issues came after this with the so-called stopwords, which are words that don't really contribute to the overall meaning of the text. Luckily, NLTK[1] has its own repository of these words, so it was implemented. There was also an issue where verbs in different tenses were evaluated very differently, so a stemmer was implemented, which truncated words to its most basic features (aptly named stems) and prevented the loss to keep rising that much between epochs.

[1]Natural Language Toolkit, tool used specifically for these case scenarios. `https://www.nltk.org/`

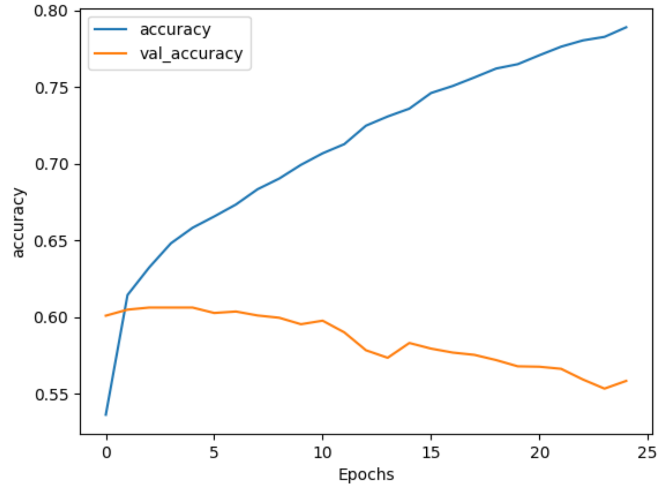

Figure 3.1: Accuracy Graph of the Algorithm Training on May 2020, with no NLTK stemming

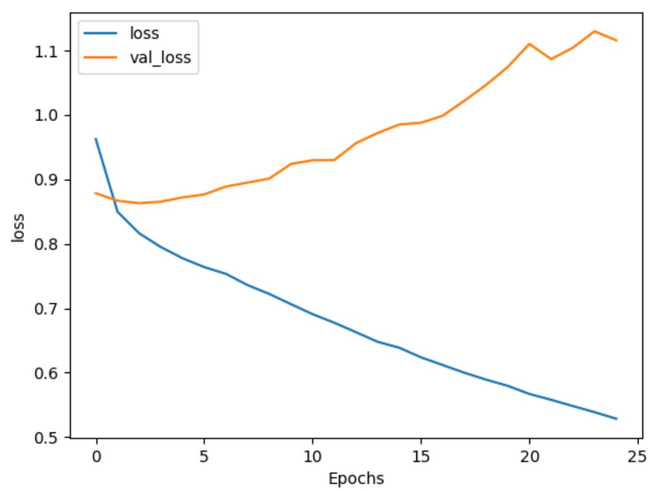

Figure 3.2: Loss Graph of the Algorithm Training on May 2020, with no NLTK stemming

3.1.2 Filtering

The dataset itself has several different sentiment labels to analyze, the ones being considered in the scope of this paper are:

- Sadness
- Neutral
- Happiness
- Fun
- Worry
- Boredom

But since they're not evenly distributed, leaving them as-is led to very inaccurate results, so a generalistic approach was opted for, classifying the end results in "Good", "Neutral" and "Bad" depending on the overall wellness percieved from the input. This final filter works only with the training data, and works as follows:

- Sadness and Worry are in the "Bad" category.
- Neutral and Boredom are in the "Neutral" category.
- Happiness and Fun are in the "Good" category

Using a more complicated classification process would take an even amount of data in every category. Which, at the time of writing, no dataset readily available has.

3.2 Algorithm Used

A bidirectional LSTM algorithm was used with a softmax activation end layer. After much, much testing *rmsprop* was chosen as the optimizer because of its slightly better results overall. The internal lexicon is limited to 5000 items, and the maximum length of any given phrase after filtering is 30 characters. The training consists in 25 epochs on 75% of the dataset on a random arbitrary order, using the remaining 25% for validation instead.

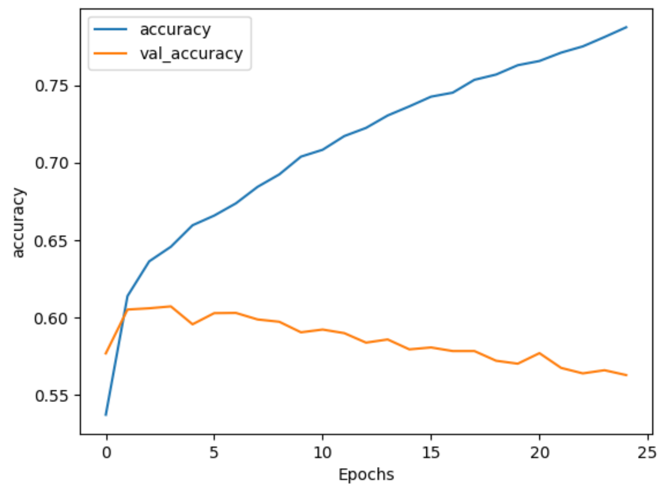

Figure 3.3: Accuracy Graph of the Algorithm Training on May 2021

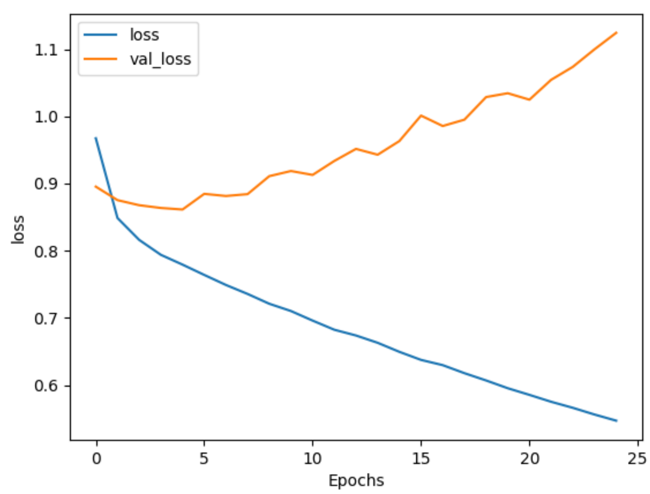

Figure 3.4: Loss Graph of the Algorithm Training on May 2021

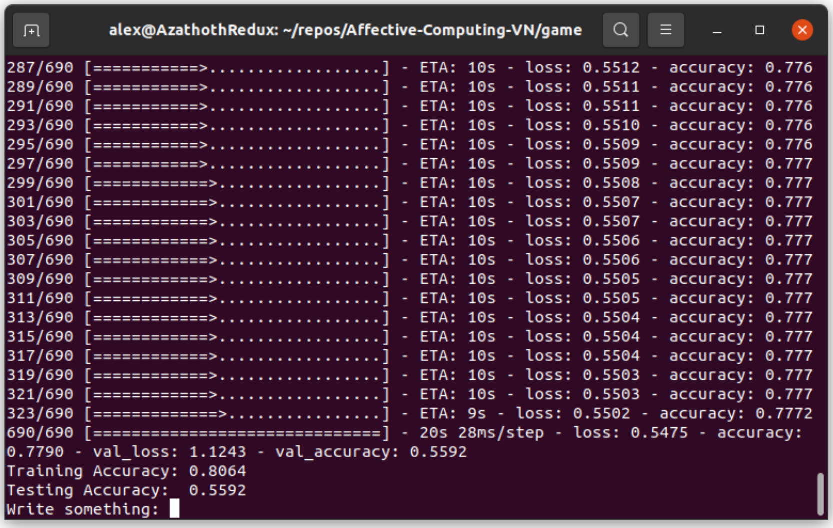

Figure 3.5: Debugging of the Trained Model

3.3 Interface

Originally, *Ren'py*[2] was the chosen framework for this project's interface to work with, but – unfortunately for the proposed usage – it only works with Python 2.7, which makes it incompatible with TensorFlow 2.0. Making a bridge between Python 2 and 3 would inevitably generate more issues that would take more time to solve, so it was scrapped in favor of the *pygame* library.

[2] An open-source Python framework focused mostly in the development of visual novels and other videogame formats. `https://www.renpy.org/`

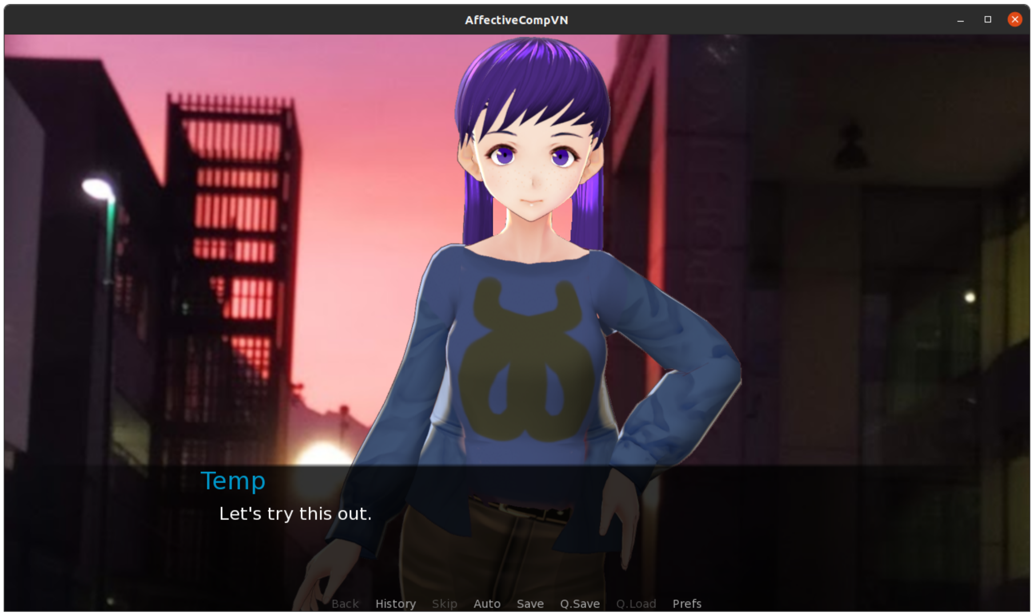

Figure 3.6: First version of the interface using Ren'py

Figure 3.7: Reacting positively to text in the "Good" category

The current interface is a hybrid between a Pygame screen, where the assistant appears to react to the input, and the console, where a person can input text to be analyzed.

3.3.1 Assistant

As for the character that's being used, it's also gone through some changes. Originally the idea was to make a low-poly character render to work with, but since 3D modeling-from-scratch skills exceed the scope of this paper, an alternative software was selected instead. Namely *VRoid*.

The main purpose for this assistant is to make people feel like it's her that they're talking to and not to some faceless machine, while also making it easier to the eyes. A more realistic, less animated style could have been used, but a friendly, less prone to uncanny valley approach to the design was opted for with this in mind.

Figure 3.8: First attempt at 3D modeling an assistant.

Figure 3.9: Assistant Ver. 2, now using VRoid.

Figure 3.10: Assistant Ver. 3, the current design.

Bibliography

Appel, O., F. Chiclana y J. Carter (2015), «Main concepts, state of the art and future research questions in sentiment analysis», *Acta Polytechnica Hungarica*, **12**(3), págs. 89–91.

Ayodele, T. O. (2010), «Introduction to Machine Learning», .

Calefato, F., F. Lanubile y N. Novielli (2019), «Experimental datasets for sentiment analysis and emotion mining - Emotion Mining Toolkit (EMTk)», URL `https://doi.org/10.5281/zenodo.2575509`.

Elliott, R., A. C. Bohart, J. C. Watson y L. S. Greenberg (2011), «Empathy.», *Psychotherapy*, **48**, págs. 1–2.

Kumar, A. y M. S. Teeja (2012), «Sentiment analysis: A perspective on its past, present and future», *International Journal of Intelligent Systems and Applications*, **4**(10), págs. 2–4.

Morris, R. R., K. Kouddous, R. Kshirsagar y S. M. Schueller (2018), «Towards an Artificially Empathic Conversational Agent for Mental Health Applications: System Design and User Perceptions», *J Med Internet Res*, **20**(6), pág. e10 148, URL `http://www.jmir.org/2018/6/e10148/`.

Pang, B., L. Lee y S. Vaithyanathan (2002), «Thumbs up? Sentiment classification using machine learning techniques», *arXiv preprint cs/0205070*.

Preston, S. D. y F. de Waal (2002), «The communication of emotions and the possibility of empathy in animals.», , págs. 2–3.

ZHANG, X.-D. (2020), «Machine learning», en *A Matrix Algebra Approach to Artificial Intelligence*, Springer, págs. 223–440.

Resumen autobiográfico

Alexander Espronceda Gómez

Candidato para obtener el grado de
Ingeniería en Tecnología de Software

Universidad Autónoma de Nuevo León
Facultad de Ingeniería Mecánica y Eléctrica

Tesis:
Sentiment Analysis through a chatbot

Nací el 17 de Noviembre de 1998 en Monterrey, Nuevo León, el mayor de los hijos de José Artemio Espronceda Estrada y Yadhira Lizet Gómez García.
Soy el primer hijo de la generación en mi familia, por lo que nunca sentí pertenecer, ya que mis tíos eran mucho más grandes que yo y mis primos mucho más pequeños. Por ello, siempre me encontraba pensando maneras de comunicarme con todos ellos "en su idioma" y lo lograba con relativo éxito. Pero a la persona que nunca pude entender fue a mi madre. Así que la mayoría de la inspiración de este proyecto se lo atribuyo a ella.
Me apasiona mucho el área de Análisis de Datos y Aprendizaje Máquina (Machine Learning), así como áreas como el Diseño de Videojuegos y la Psicología, por lo que este proyecto es la culminación entre mis pasiones más grandes para concluir la carrera de Ingeniería de Tecnología de Software.