

Analysis of the Strengths and Weaknesses in Teaching Methodologies through Students' Commentaries using Sentiment Analysis

Roberto Ramírez Monroy
A01366943
Tecnológico de Monterrey

Alitzel Macías Infante
A01373166
Tecnológico de Monterrey

I. INTRODUCTION

The field of educational research seeks to make improvements in education in aspects such as educational planning, teaching and learning, curriculum development, and understanding of students needs, amongst others [1]. Furthermore, a key factor in the success of students learning is the teacher's knowledge of the subject, his/her ability to implement new strategies, to design appropriate assessment tools, and to address different learning styles [2]. Therefore, the analysis of feedback from students about professor's methods of teaching are highly important since they can give insights in weaknesses and strengths of specific subjects.

In this work the dataset ECOAS will be analyzed using machine learning algorithms. ECOAS is a dataset from Tecnológico de Monterrey, in which students evaluate their teachers from each subject at the end of the course. The evaluation consists of a numerical mark from 1 to 10 in specific aspects of the way of teaching, and also commentaries about the professor. In particular, we will not be working with the numerical mark, but only with the comments from the dataset.

The stakeholders of the present analysis are Tecnológico de Monterrey, other educational institutions, mainly at undergraduate level, and researchers of the field of educational investigation, and from educational innovation.

In related previous work, in [3], Ceballos et al. use the student survey ECOA to compare the teaching performance of only teachers vs research teachers. The results they found were that both groups perform similarly; however, they did found differences depending on attributes such as the professor's age, gender, and research level.

On the other hand, in [4] Kandhro et al. used sentiment analysis to analyze students' sentiments from teacher's evaluation through a Long Short-Term memory Model (LSTM). Although in [4] they did not use the obtained model to find correlations between the sentiment of the students' evaluation of their teachers and aspects such as the professors' methodology of teaching, this model is useful for this future kind of analysis. In the present work, for example, pre-processed comments by sentiment analysis will be used to analyze weaknesses and strengths in different categories of teachers' methodologies.

Given the data that we have at our disposal, what we are

trying to achieve is to know what competences in teachers are the ones that students care more about, in the positive and negative senses. And if this competences depend on more factors, as the campus of the school, or the subject in question, including the department and the group of the teacher.

II. DATA AND METHODS

A. Datasource Description

The data that was given to us, by the professor, consist on near 70,000 records of the student ECOAS, where we have data like the campus of the student and teacher, the department of the teacher, and the most important fields for this analysis; one comment from the student to the teacher, and some pre-calculated document sentiment score (from -1 to 1) where a negative value means that the comment has a negative connotation towards the teacher, and a positive value a positive connotation.

Another pre-classified columns we have are the category and subcategory of the comment, where we have 3 different values for categories (knowledge transfer, general capability and emotional intelligence) and a total of 10 subcategories. Letting us know a better classification of the comments of the students.

B. Variables

The independent variables of the dataset that are going to be used are:

- *Campus*: Campus of the institution where the particular subject and teacher were evaluated.
- *Division*: School where the subject belongs (e.g. School of Science and Engineering)
- *Department*: Specific department the subject being evaluated belongs to (e.g. Computer Science, Civil Engineering)
- *Question*: The question the comment analyzed is responding to. The two different questions of the dataset are *Comments to the professor*: and *Why wouldn't I recommend the professor?*.
- *Subject*: The subject where the comments belong to.
- *Document Sentiment*: A numerical variable ranging from -1 to 1 that was obtained from sentiment analysis of

the comments and describes how positive or negative a comment is.

- *Sentiment*: A variable that has 3 possible categories for a comment: whether it is neutral, positive or negative.

On the other hand, the dependent variable used in the analysis is:

- *Subcategory*: The category the comment belongs to. It is related to aspect of the professor's way of teaching that the student is commenting on.

C. Data Preparation

Running a quick initial analysis of all the data, we found that some of the columns would not be useful for the purpose we are setting in this document. Like the campuses coordinates, the campus size, and another column with almost no data, called EntitiesList.

Also, we are not going to take into deep consideration the rows without significant data in the comment, because not all students leave a real/complete comment, so we have some registers with just dots, spaces, etc. And these comments are not classified into any category nor subcategory.

Even if we have more positive than negative reviews (42982 vs 10173) and 4672 neutral comments, we don't think that a data balancing procedure should be made at the beginning, but we may test with a reduced dataset, where we would reduce the number of the positive rows.

One feature engineering we think something that could be useful is creating a row in the dataset, which will have the length of the comment. In order to calculate if this length could be correlated to the sentiment score of the comment or another features of the dataset.

D. Descriptive Analysis

In our data, we have 6 numeric features, describing different values, where the ones we are more interested about are the document sentiment and the word count of the comment, because the other values have a qualitative behaviour, being ID's of the subject or the group of the teacher that is being evaluated. We can see the correlation matrix of these values in fig 1

Also, we analysed the distribution of the sentiment scores (fig. 2) and the word count (fig. 3)

And after cleaning the data, keeping only the rows with a classification on the category of the comment, we can see the distribution of all the comment sentiments in fig. 4. Where we can observe that there are more positive evaluations than negative ones, and the neutral ones are the ones that have less registers.

Another analysis tool we used, are the word clouds, just to get a better look on the words that are more used in positive and negative evaluations. First we used the *nltk* library to get a list of stop words on Spanish, and then we computed the word clouds for the positive and negative datasets. We can see the results of the positive data in fig. 5 and the negative one in fig. 6.

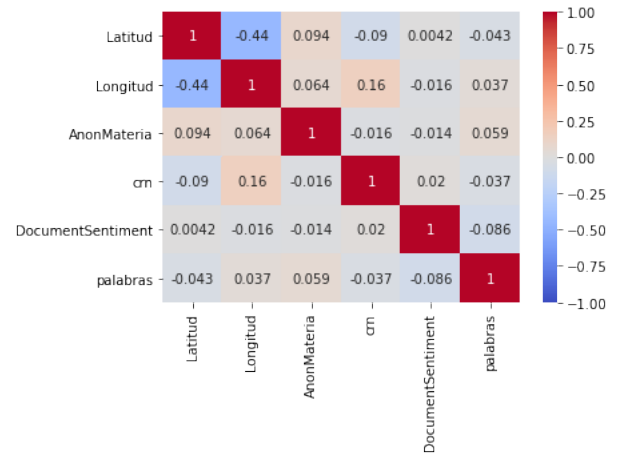


Fig. 1. Correlation Matrix of numeric columns

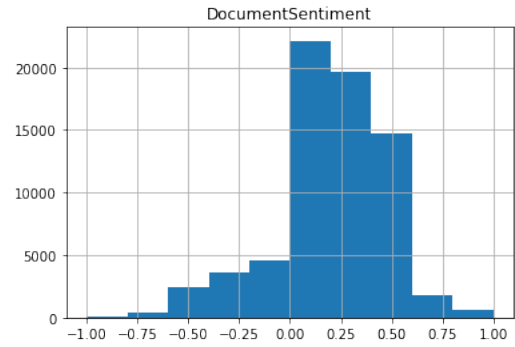


Fig. 2. Sentiment score column distribution

In this project the type of learning we are going to use, will be supervised learning, because we already have some labels of the *categories* of the comments, so we could train some models with a part of the data with those labels, and then test in other part of the data, also having the labels to compare and rate the models.

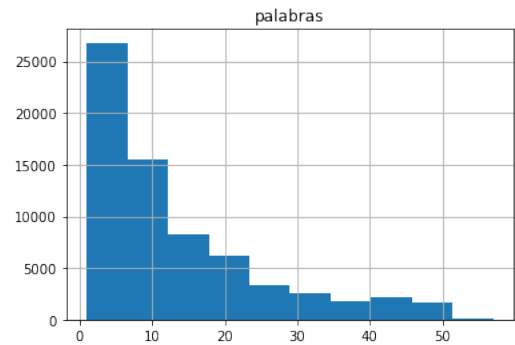


Fig. 3. Word count column distribution

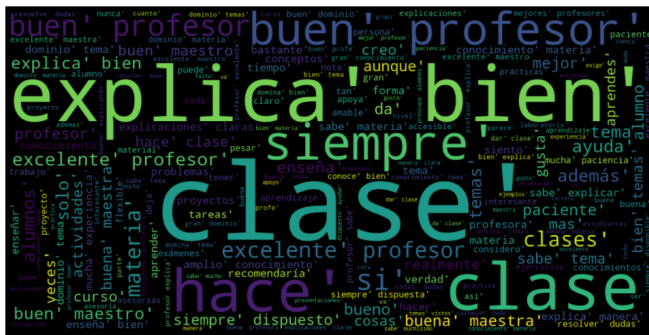
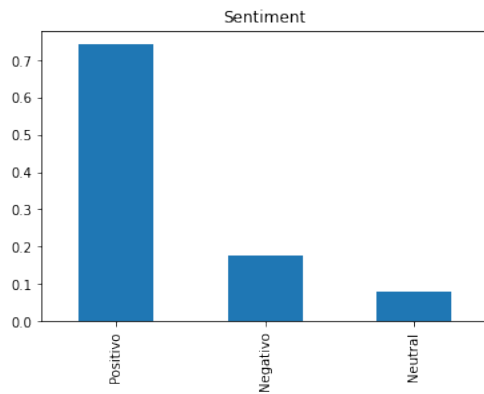


Fig. 5. Positive evaluations word cloud

E. ML techniques to apply

The first proposal of technique to apply in order to answer the question we are targeting, is to use an Artificial Neural Network, where the input would be some information of the subject, the sentiment of the evaluation and more data we could get or create, to predict the *category/subcategory* of the evaluation, so then we could know the opportunity areas of improvement or the strengths of the teacher, helping them improve the way they teach.

Another technique we may use is SVM (Support Vector Machines) for the same classification purpose, where we could also use Principal Components to reduce the input dimension.



Fig. 6. Negative evaluations word cloud

The last technique we may use to analyze the data is a clustering algorithm. This is in order to have variety in the methods use to analyze the data. Clustering algorithms can help us to detect patterns in the independent variables such as *Department* and *Sentiment*, that may help classify the category the comments belong to. In particular, we may be using K-means since in this algorithm we can select the number of clusters we want to work on.

REFERENCES

- [1] R. Kapur, "Significance of Research in Education", 2018, Retrieved from: https://www.researchgate.net/publication/323833808_Significance_of_Research_in_Education.
- [2] T. Stafford, The Importance of Educational Research In the Teaching of History, *Canadian Social Studies* vol. 40, n. 1, 2006. Retrieved from: <https://files.eric.ed.gov/fulltext/EJ1073947.pdf>
- [3] H. Ceballos, M. Chavez, and F. Cantu, A Data Analytics Approach to Contrast the Performance of Teaching (only) vs Research Professors, *International Journal on Interactive Design and Manufacturing*, Unpublished.
- [4] I. Kandhro, S. Wasi, K. Kumar, M. Rind, and M. Ameen, Sentiment Analysis of Students' Comment by using Long-Short Term Model, *Indian Journal of Science and Technology*, vol 12(8) pp. 1-16, 2019, DOI:10.17485/ijst/2019/v12i8/141741.