

Text mining and sentiment analysis of teacher performance satisfaction in the virtual learning environment

Omar Chamorro-Atalaya¹, Dora Arce-Santillan¹, José Antonio Arévalo-Tuesta²,
Lilia Rodas-Camacho², Genaro Sandoval-Nizama³, Rosa Valle-Chavez⁴, Yadit Rocca-Carvajal⁴

¹Faculty of Engineering and Management, Universidad Nacional Tecnológica de Lima Sur, Lima, Peru

²School of Economics, Universidad Nacional Federico Villarreal, Lima, Peru

³Department of Business Sciences, Accounting Professional School, Universidad Cesar Vallejo, Lima, Peru

⁴Department of International Business, Faculty of Business, Universidad Cesar Vallejo, Lima, Peru

Article Info

Article history:

Received Feb 7, 2022

Revised Jul 28, 2022

Accepted Aug 5, 2022

Keywords:

Sentiment analysis

Student satisfaction

Teacher performance

Text mining

Virtual learning

ABSTRACT

Although it is true that artificial intelligence and data science have become key tools that contribute to the improvement of many processes, identifying patterns and contributing to decision making, however, there are environments in which they are not yet being used relevantly and effectively. The objective of this study is to identify the relevant factors, based on the opinions expressed by the students through the social network Twitter regarding the perception of satisfaction with the teaching performance during the virtual learning environment. For which sentiment analysis and text mining are used under the Python programming language environment, through JupyterLab. As results, it was determined that a predominance of 57.27% of positive polarity, identifying that the relevant factors of student satisfaction with teaching performance, are related to the development of the teacher in the class sessions that contributes to the learning of the process control subject through the use of simulation tools such as simulink and tools linked to proportional integral derivative (PID) controllers; on the other hand, there is a percentage of negative polarity of 15.45% that belongs to the factors linked to the laboratory sessions in which graphic representation and block diagrams were used to explain the class session.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Omar Chamorro-Atalaya

Faculty of Engineering and Management, Universidad Nacional Tecnológica de Lima Sur

Sector 3 Grupo 1A 03, Av. Central, Villa EL Salvador 15834, Peru

Email: ochamorro@untels.edu.pe

1. INTRODUCTION

The quality of the service within the educational field, without a doubt, has become an indicator of good organizational practices at a global level [1], [2]. In the educational field, service quality has been related to a set of attributes of institutional management on which students evaluate the fulfillment of their expectations associated with their academic experience [3], [4]. The evaluation of academic quality is optimized when opinions and indicators of student satisfaction are incorporated, since the university student represents the main actor in the study of the quality of institutions [5], [6]. Likewise, in [7], [8] it is pointed out that university satisfaction is one of the fundamental objectives to determine the quality of virtual training in educational organizations; being a key factor the capacity and performance of the teachers in offering a response according to the expectations and needs of the students. Regarding teaching performance, in the virtual environment, in [9], [10] it is pointed out that the interactions between the teacher-student take place

through digital tools through information and communication technologies (ICT). Traditionally, the measurement of satisfaction is carried out through surveys. However, in the context of virtualization, it seeks to implement new methods and procedures, taking advantage of the increased use of the Internet and the production of data that it generates [11].

Today on the Internet millions of data are produced due to the massive use of social networks. All this large amount of data is attractive to different commercial, industrial and academic organizations [12], [13]. In addition to this, users actively participate on the internet leaving their own comments, opinions and reviews on all kinds of topics, due to this, researchers have been working for several decades on systems that allow large amounts of text to be analyzed automatically using techniques natural language processing (NLP) and data mining, among others [14], [15]. Indeed, one of the most important transformations today is the exponential growth of data that opens the possibility of new forms for the analysis of reality [16]. Thus, the value of information no longer lies in specific data, but in the way in which massive data is correlated to discover hidden patterns, trends and associations [17]. Until now, big data analysis has been used in different disciplines of scientific research whose purpose has been to solve complex problems, in the fields of environment, education, health, transportation, national security and biomedicine [18]. Data science is made up of three areas, big data, which is used to process data, data mining, whose purpose is to find patterns, and data visualization, whose purpose is to facilitate the understanding of information in a clear way and promote their socialization [19]. With regard to data mining, it is considered as the analysis of observed data sets, generally large volumes, with the aim of finding new relationships between variables, as well as the correct summary of said data sets in an understandable and useful [20].

As indicated, data mining encompasses a series of modeling techniques aimed at different requirements of organizations, with respect to the educational sector, the use of these techniques can be seen in a series of works aimed at improving the quality of university education [21], [22]. Within data mining, text mining is found, which is the process of finding previously unknown implicit patterns that may be useful from a text repository [23]. Text mining is embedded in NLP techniques, which is a field of computer science, artificial intelligence and linguistics that studies the interactions between computers and human language, through the syntactic, semantic, pragmatic and morphological analysis; these rules, in combination with the information stored in computational dictionaries, define the patterns to be recognized through images, text and voice [24], [25]. Thanks to the expansion of technology, recent years have seen an exponential growth in subjective information available on the Internet. This phenomenon has given rise to interest in sentiment analysis (SA), a NLP task that identifies opinions related to an object [26]. SA, is an area of NLP, focuses on detecting, extracting, analyzing and classifying opinions, sentiments, evaluations and emotions published on different social networks (Facebook, Twitter, LinkedIn, among others) towards certain organizations or organizations people [27]. As indicated in [28] sentiment analysis seeks to extract an opinion and determine its polarity (positive, negative or neutral). Likewise, in [29], [30] it is pointed out that being the social network Twitter one of the most commonly used tools to determine the opinions of users about a wide variety of topics, this is not unrelated to the fact that, in the environment educational, sentiment analysis is used to evaluate student performance, teacher performance, academic desertion, among other relevant issues that aim to improve educational quality.

Given the above, this research aims to identify the most relevant factors, according to the satisfaction of university students with teaching performance during the virtual learning environment, through the text mining and sentiment analysis approach. In this sense, the research is organized as follows; Section 2 describes the literary review or state of the art on sentiment analysis and the assessment of polarity in the context of the use of social networks, section 3 describes the methodology and the proposed development, section 3 shows the results obtained, section 4 discusses the results, and finally section 5 presents the conclusions and future work.

2. LITERARY REVIEW

Understanding and processing emotions through sentiment analysis is an important and key factor in the approach to the emulation of the intelligence of the human being, as well as to identify the polarity related to a feeling, this way of approaching behavior analysis of people has aroused interest in the academic and scientific community [31]. With the great influx of information that is handled through social networks due to the high demand for active participation of people in these times, there has been an almost exponential increase in subjective data available on the internet, which has generated great interest. In establishing evaluations regarding the statements contained in these messages, in this way it is necessary and preponderant the need to use tools that detect, extract and structure said subjective information [32]. In a study on sentiment analysis applied to the processing of emotions and opinions of university students, it is concluded that sentiment analysis is a highly complex technique, both due to the implications of data

processing that involves NLP, cleaning and filtering, construction of classifiers, evaluation of results. However, it adds that it is possible to carry out sentiment analysis in different instruments for assessing university student satisfaction, allowing an efficient allocation of resources to programs or faculties within educational institutions [33]. Sentiment analysis through natural language processing is relevant since it allows to express evaluations about satisfaction linked to a product or service provided, being possible to obtain data to process from social networks [20]. However, performing sentiment analysis through any social network implies the use of techniques for data pre-processing, which in principle can be used the specific technique of Twitter or the traditional processing technique, in this regard, making an analysis on the performance of these pre-processing techniques, it was obtained that the results of the analysis in relation to precision, establish that both techniques have viable results, in the treatment of data, while in specific stages such as tokenization, labeling and elimination of empty words, in this sense, the specific techniques of Twitter show significant advantages [34].

3. METHOD

3.1. Research level and study unit

The research is of a non-experimental level, because no action is carried out on the students, which generates some type of alteration or modification of their comments made through the social network Twitter regarding satisfaction with the teaching performance. In this sense, the data is treated in its natural condition. In addition, the data will be collected weekly in each class session. The research work takes as a unit of analysis the comments or opinions expressed on the social network Twitter by the students enrolled in the course of automatic process control, of the professional school of mechanical and electrical engineering of the National Technological University of South Lima (UNTELS).

3.2. Description of the collected data

The sentiment analysis and text mining is developed under the Python programming language, through the JupyterLab environment, which will show the metrics of the Twitter account regarding the comments or opinions (tweets) made in each class session during the virtual learning environment. The acquisition of the data takes a period of 5 weeks, and is comprised from week 9 to week 13, with the purpose of obtaining a first result and then applying this model to the analysis of the 16 weeks of class that comprise a cycle academic. It should be noted that the total impressions obtained in the 5 weeks of data acquisition were 455, and the number of interactions or comments represented by "tweets" were 220. Figure 1 shows the code developed in Python.

```
df2=pd.read_csv("tweet_activity_metrics_process_control.csv")
df2.head()
```

Figure 1. Code developed in Python to obtain tweet metrics

As a result of the execution of the code shown, the information of the "tweet id" is obtained, which represents the identification code of the tweets generated by each class session. Another important aspect to highlight in the data collection is the evaluation of the "impressions" and "interactions" whose rate of impressions during week 9 to week 13, were 0.529 and 0.433, respectively; evidencing on average a 50% relationship between the visits that the account had and the participation in the comments; it is also evident that the participation of students with their comments on the Twitter account was 100%, since the number of students enrolled in the automatic process control subject was 25. Figure 2 shows the metrics of tweets generated during the data collection period.

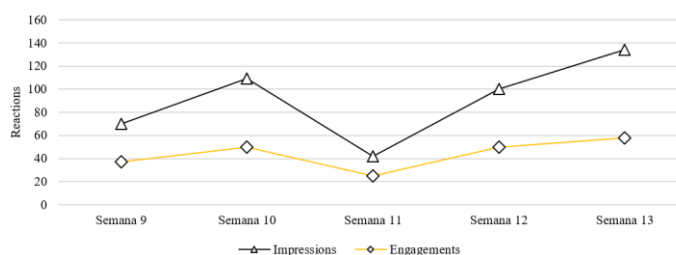


Figure 2. Metrics of generated tweets

3.3. Conditioning, sentiment polarization analysis and data extraction

Figure 3 shows the model for the identification of relevant factors of university student satisfaction with teaching performance during the virtual learning environment. As an initial part of the development, the acquisition of comments or opinions expressed in the Twitter social network is carried out, this stage consists of storing the data in comma separated values (CSV) format, conditioning the texts through the Python library and in this way performing the sentiment analysis using the inverse term frequency and document frequency (TF-IDF) approach, which has as its output the polarization of sentiments (positive, negative and neutral) and the extraction of data that will allow determining the relevant factors.

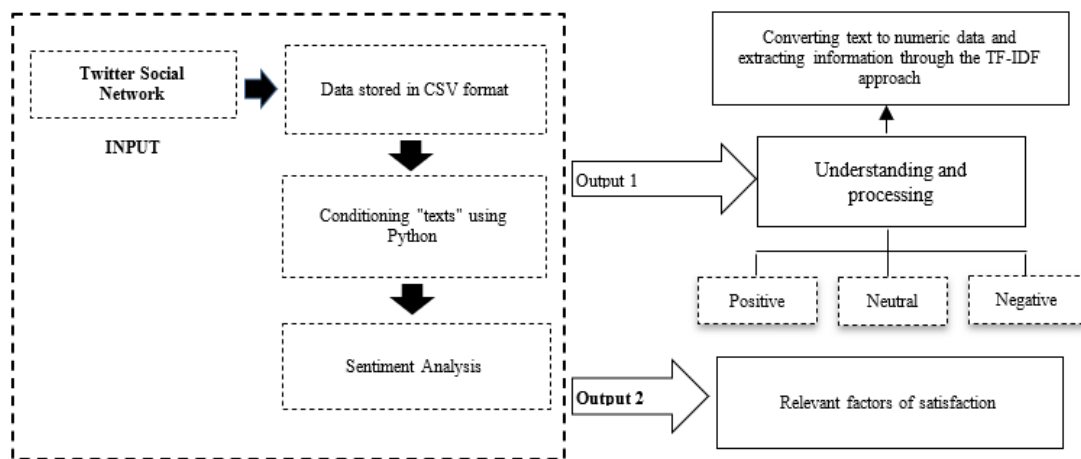


Figure 3. Model for the identification of relevant factors of student satisfaction

The model (Figure 3) shown specifies that from the data provided by the Twitter social network in file format with a "csv" extension, these will be conditioned to obtain the polarization of the comments made by the Twitter social network through the analysis of sentiments. The conditioning or pre-processing consists of eliminating repeated texts, empty spaces that are found in excess, between text and text, as well as converting all texts to lowercase. Next, the sentiment analysis of the tweets written in English is carried out, using Python libraries such as "nltk" and "vader"; the result obtained will be the quantification of the sentiment contained in each comment or opinion written by the student; the same ones that are classified in positive polarity (greater than 1), neutral polarity (equal to 0) and negative polarity (less than 1), representing the output (output 1) of the model used; the same one that will represent relevant information for decision making in the improvement of the virtual teaching-learning process of the process control subject.

In the proposed model (Figure 3) it is observed that the last stage is related to the extraction of texts, through the term frequency-inverse dense frequency (TF-IDF) approach; This indicator TF-IDF will allow selecting the texts that appear the most in each of the comments or opinions and at the same time that appear the least globally. Through this stage, the model allows identifying the most relevant factors, according to the satisfaction of university students with the teaching performance during the virtual learning environment. This second output (output 2) of the proposed model will also provide the value of TF-IDF, which represents the weight of a text "i" in a tweet "j". As indicated in [28] TF-IDF provides the best performance of Fscore approximately 77.8%~94.8%, compared to other approaches, so its use guarantees the functionality of your application in determining the most relevant aspects of the satisfaction.

3.4. Conditioning, sentiment polarization analysis and data extraction

To carry out this stage, the programming code is generated that will allow the file provided by Twitter to be loaded or read in the JupyterLab environment, called "tweets_process_control_29Dic21.csv". In addition, the "pandas" library will be imported into the program, with which it will be possible to read and write data, in different formats or extensions, in this investigation the file has a "csv" extension; Likewise, the "numpy" library will be used to create vectors or matrices in which the texts will be linked with their corresponding sentiment and polarity. Once this procedure has been carried out, the conditioning or pre-processing of the texts is carried out. Figure 4 shows the code used to carry out the conditioning of the texts contained in the comments on the Twitter account, using the Python libraries ("nltk", "sentiment.vader", "SentimentIntensityAnalyzer", "re"). This stage is responsible for eliminating repeated texts, individual

characters that are isolated, empty spaces that are too many, between text and text, as well as converting all texts to lowercase.

```

from nltk.sentiment.vader import SentimentIntensityAnalyzer

import nltk

import re

features = df.iloc[:,0].values

processed_features=[]

for sentence in range(0, len(features)):

    processed_feature=re.sub(r'W', ' ',str(features[sentence]))

    processed_feature=re.sub(r'^s+[a-zA-Z]s+', ' ',processed_feature)

    processed_feature=re.sub(r'^[a-zA-Z]s+', ' ',processed_feature)

    processed_feature=re.sub(r's+', ' ',processed_feature,flags=re.I)

    processed_feature=re.sub(r'^bs+', ' ',processed_feature)

    processed_feature=processed_feature.lower()

    processed_features.append(processed_feature)
    
```

Figure 4. Script for conditioning or pre-processing texts

Once the texts had been conditioned, and all of them normalized, we proceeded to determine the sentiment contained in each of thee tweets, written by the students of the process control subject. As evidenced in Table 1, the value of the sentiment of the first 5 tweets is shown, from which it was possible to assign the polarity of the sentiments. With this, the results of satisfaction will be obtained from the comparison of the number of comments with positive and negative polarity.

Table 1. Result of the sentiment contained in each tweet

Comment number	Comments on the satisfaction of the teaching performance	sentiment
1	It has been a good way to deepen the topic ...	0.4404
2	I would have liked the problems proposed in ...	0.2960
3	I found it interesting that the maximum overshoot ...	0.4019
4	The statements are the same as those learned...	0.0000
5	There are topics that were not clear in other...	-0.2924

Finally, the programming code is executed, which seeks to extract the most relevant texts, contained in the tweets, with positive and negative polarity, since these are closely related to the satisfaction and dissatisfaction of teacher performance during the virtual learning environment. The code makes use of libraries such as “stopwords” and “Brown” imported from the nltk.corpus library, the same ones that allow text processing. being the approach used for the extraction of TF-IDF texts; the “TfidfVectorizer” function is used; this function allows establishing the “max_features” parameter, which represents the number of “words” that appear most frequently and that must be extracted from all the tweets with positive and negative polarity generated by the students. This also specifies the “min_df” parameter, which establishes the minimum number of tweets in which the extracted words must be found; On the other hand, the “max_df” parameter is also specified, which represents the percentage of tweets in which the extracted texts must be found at most. In the investigation of [24] it is pointed out that it was possible to identify 93 general stop words, which helped in the determination of relevant factors of satisfaction. Furthermore, incorporating the stopword removal preprocessing step into the tamil text pooling, an improvement of 2.4 in F-score for TF-IDF was observed. Similarly, in [14] it is indicated that there is a significant improvement in Fscore when the data set is used after removing stopwords, so it is shown that the impact of removing stopwords is greater in TF-IDF than in TF-IDF that of fasttext. Likewise, in [28] it is pointed out that TF-IDF provides the best performance in determining the most relevant aspects of satisfaction. This was reflected in its results, where

the system made an F score with 0.939 using TF-IDF, then around 0.899 in the TF, due to this, it is indicated that the system will be fast and reliable.

4. RESULTS AND DISCUSSION

Figure 5 shows the result of the polarity of sentiments obtained through the comments or opinions of university students, carried out through Twitter, this analysis was carried out through Python. As shown in Figure 5, there is evidence of the predominance of positive sentiments in the comments or opinions expressed on the social network Twitter regarding satisfaction with teaching performance, finding a percentage distribution contained in the tweets of 57.27% of positive polarity, of 27.27% neutral polarity and 15.45% negative polarity.

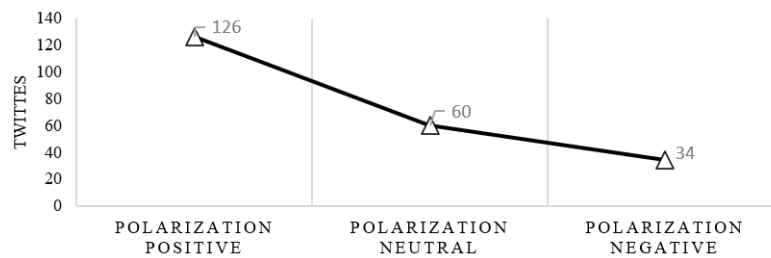


Figure 5. Distribution of the polarity of sentiments

Next, the extraction of predominant texts, generated under the TF-IDF approach, is considered, for which it is defined that "max_features" is equal to 10, "min_df" equal to 10 and "max_df" equal to 0.7. That is, that the 10 words are extracted most frequently, and that these appear in at least 10 tweets simultaneously and that they appear in a maximum of 70% of all tweets. In Figure 6, the results of the tweets with positive polarity are shown.

```
print(vectorizer.vocabulary_)
{'class': 1, 'interesting': 2, 'laboratory': 3, 'understand': 7, 'able': 0, 'us': 8, 'teacher': 6, 'use': 9, 'simulink': 5, 'pid': 4}

print(vectorizer.idf_)
[2.28883902 2.54835022 2.13061502 2.54835022 2.79966465 2.70869287
 2.89974811 2.89974811 3.13613689 3.13613689]
```

Figure 6. Extraction of words with positive polarity

As shown in Figure 6, the terms with the highest IDF turned out to be "simulink" and "pid", with a value of 3.1361, this result allows establishing that the relevant factors that contribute significantly to student satisfaction during the learning environment virtual, are those related to teaching performance in class sessions that allow understanding of laboratory sessions through the use of simulation tools such as simulink and PID (proportional-integrator-derivative) controllers. The results are similar to those obtained in [30] where 76.6% of the tweets are marked as positive; this shows that the students, who actively use Twitter, have a good perception and experience in the use of online learning applications, likewise, the large number of positive sentiments towards the virtual course is related to the study material used, which is easy to understand. In [27] the study of perception based on sentiment analysis, on the development of academic activities by students during the confinement period, a positive polarity of 33.24% and a percentage of contribution of the negative polarity of 17.73 %, which is equivalent to a ratio of 2 to 1, allowing the conclusion that Systems Engineering students, despite the adaptation process, have a positive perception of the academic processes and look optimistically at the possibility of improving the practices developed in learning virtual.

In order to establish the same level of performance and precision in the word extraction process, Figure 7 shows the results of the negative polarity tweets. regarding Figure 7, the terms with the highest IDF

turned out to be “diagrams” and “block”, with a value of 2.1574; These extracted texts allow us to establish that the relevant factors that generated student dissatisfaction during the virtual learning environment were linked to the laboratory sessions in which graphic representation and block diagrams were used to explain the class session; evidently it was not possible for the students to understand in a significant way the handling of these tools. regarding the identification of satisfaction and dissatisfaction factors related to teaching performance through the acquisition of comments or opinions expressed on the social network Twitter, in [22] it is pointed out that discussion forums have a great impact on the student's experience, because discussion forums are a means for students to clearly express their thoughts, opinions, and struggles regarding learning content, likewise, textual interactions by students enhance sentiment and cognitive engagement toward the course. As indicated in [24], the predictive model through sentiment analysis allowed students of the applied basic statistics subject to be motivated and satisfied to use the application of data science during the virtual teaching-learning process.

```
print(vectorizer.vocabulary_)
{'laboratory': 4, 'graphs': 3, 'able': 0, 'block': 1, 'diagrams': 2}

print (vectorizer.idf_)
[1.84729786 1.9903987 2.15745279 2.15745279 2.15745279]
```

Figure 7. Extraction of words with positive negative

The results obtained show that it is possible to obtain a reference point on student satisfaction using text mining techniques and sentiment analysis, in the same way as indicated in [11] it is possible to perform sentiment analysis in different assessment instruments of student satisfaction allowing an efficient allocation of resources to programs or faculties within educational institutions. for its part, in [23] where the EmotiBlog was evaluated through sentiment analysis, focusing on the automatic detection and classification of sentences with subjective information, the validity of EmotiBlog was demonstrated, suggesting for this reason to apply the developed model.

5. CONCLUSION

This study shares the ideas of various authors on the use of text mining and sentiment analysis in the educational field to improve teaching-learning conditions, since it allowed the construction of a predictive model on the satisfaction of university students with teaching performance during the virtual learning environment using the Python programming language through the JupyterLab environment. The results of the study show regarding the polarity that there is a predominance of 57.27% of positive sentiments in the comments or opinions expressed on the social network Twitter, identifying that the relevant factors of student satisfaction during the virtual learning environment are related with the teaching performance in the class sessions that allow to achieve the understanding of the sessions through the use of the laboratory through simulation tools such as simulink and PID controllers; on the other hand, there is a percentage of negative polarity of 15.45% that belongs to the factors linked to the laboratory sessions in which graphic representation and block diagrams were used to explain the class session.

For future work, it is recommended to expand the line of research, through a predictive study and implement sentiment analysis and text mining techniques in the satisfaction analysis of the dimensions that make up university quality, in order to make timely decisions that allow improving the teaching-learning process during the semester studied.

ACKNOWLEDGEMENTS

Thanks and appreciation to the “Universidad Nacional Tecnológica de Lima Sur”.

REFERENCES




- [1] J. Grau-Valldosera, J. Minguillón, and A. Blasco-Moreno, “Returning after taking a break in online distance higher education: from intention to effective re-enrollment,” *Interactive Learning Environments*, vol. 27, no. 3, pp. 307–323, Apr. 2019, doi: 10.1080/10494820.2018.1470986.
- [2] O. F. C. Atalaya, T. N. D. Leyva, D. Y. A. Santillan, J. I. C. Bedriñana, D. M. B. Pichilingue, and A. Roman-Gonzalez, “Satisfaction of the Graduate for the continuous improvement of educational quality in UNTELS,” *Advances in Science, Technology and Engineering Systems Journal*, vol. 4, no. 5, pp. 151–157, 2019, doi: 10.25046/aj040520.

- [3] E. Duque and Y. Gómez, "Conceptual evolution of service quality perception measurement models: a view from higher education," *Sum of Business*, vol. 5, no. 12, pp. 180-191, Jul. 2014, doi: [https://doi.org/10.1016/S2215-910X\(14\)70040-0](https://doi.org/10.1016/S2215-910X(14)70040-0)
- [4] T. Chen, L. Peng, X. Yin, J. Rong, J. Yang, and G. Cong, "Analysis of user satisfaction with online education platforms in china during the COVID-19 Pandemic," *Healthcare*, vol. 8, no. 3, p. 200, Jul. 2020, doi: [10.3390/healthcare8030200](https://doi.org/10.3390/healthcare8030200).
- [5] M. González-Peiteado, M. Pino-Juste, and M. Penado-Abilleira, "Study of the satisfaction perceived by UNED students with their life university," *RIED. Revista Iberoamericana de Educación a Distancia*, vol. 20, no. 1, p. 243, Jan. 2017, doi: [10.5944/ried.20.1.16377](https://doi.org/10.5944/ried.20.1.16377).
- [6] D. A. B. Osma, C. G. C. Estupiñan, and I. A. P. Arteaga, "University students who work: subjectivity, construction of meaning and dissatisfaction," *Revista Virtual Universidad Católica del Norte*, no. 58, pp. 96-115, Oct. 2019, doi: [10.35575/rvucn.n58a4](https://doi.org/10.35575/rvucn.n58a4).
- [7] N. E. L. Maldonado, S. R. R. López, I. S. R. Rodríguez, and M. A. C. García, "Digital tools in times of covid-19: perception of higher education teachers in Mexico," *RIDE Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, vol. 12, no. 23, Dec. 2021, doi: [10.23913/ride.v12i23.1108](https://doi.org/10.23913/ride.v12i23.1108).
- [8] D. Vlachopoulos and A. Makri, "Quality teaching in online higher education: the perspectives of 250 online tutors on technology and pedagogy," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 06, p. 40, Mar. 2021, doi: [10.3991/ijet.v16i06.20173](https://doi.org/10.3991/ijet.v16i06.20173).
- [9] N. H. N. Aziz, H. Haron, and A. F. Harun, "ICT-supported for participatory engagement within E-learning community," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 20, no. 1, p. 492, Oct. 2020, doi: [10.11591/ijeecs.v20.i1.pp492-499](https://doi.org/10.11591/ijeecs.v20.i1.pp492-499).
- [10] M. del R. C. Alcántar, C. I. Torres, and M. L. M. Gómez, "Perception of students regarding the use of ICT and learning English," *Apertura*, vol. 10, no. 2, pp. 50-63, Oct. 2018, doi: [10.32870/Ap.v10n2.1391](https://doi.org/10.32870/Ap.v10n2.1391).
- [11] N. Z. Dina, R. T. Yunardi, A. A. Firdaus, and N. Juniarta, "Measuring user satisfaction of educational service applications using text mining and multicriteria decision-making approach," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 17, p. 76, Sep. 2021, doi: [10.3991/ijet.v16i17.22939](https://doi.org/10.3991/ijet.v16i17.22939).
- [12] K. Ravi and V. Ravi, "A survey on opinion mining and sentiment analysis: tasks, approaches and applications," *Knowledge-Based Systems*, vol. 89, pp. 14-46, Nov. 2015, doi: [10.1016/j.knosys.2015.06.015](https://doi.org/10.1016/j.knosys.2015.06.015).
- [13] N. A. Mansor, N. Abdullah, and H. A. Rahman, "Towards electronic learning features in education 4.0 environment: literature study," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 19, no. 1, p. 442, Jul. 2020, doi: [10.11591/ijeecs.v19.i1.pp442-450](https://doi.org/10.11591/ijeecs.v19.i1.pp442-450).
- [14] M. S. F. Fayaza and F. F. Farhath, "Towards stopwords identification in tamil text clustering," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 12, 2021, doi: [10.14569/IJACSA.2021.0121267](https://doi.org/10.14569/IJACSA.2021.0121267).
- [15] E. R. Gómez-Torres, R. Jaimes, O. Hidalgo, and S. Luján-Mora, "Influence of social networks on the analysis of sentiment applied to the political situation in Ecuador," *Enfoque UTE*, vol. 9, no. 1, pp. 67-78, Mar. 2018, doi: [10.29019/enfoqueute.v9n1.235](https://doi.org/10.29019/enfoqueute.v9n1.235).
- [16] N. S. A. Rahman, L. Handayani, M. S. Othman, W. M. Al-Rahmi, S. Kasim, and T. Sutikno, "Social media for collaborative learning," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 1, p. 1070, Feb. 2020, doi: [10.11591/ijece.v10i1.pp1070-1078](https://doi.org/10.11591/ijece.v10i1.pp1070-1078).
- [17] J. Cabero and A. Martínez, "ICT and the initial training of teachers. Models and digital skills," *Teachers, Journal of Curriculum and Teacher Training*, vol. 23, no. 3, pp. 247-268, Sept. 2020, doi: <https://doi.org/10.30827/profesorado.v23i3.9421>
- [18] J. Serrano-Guerrero, J. A. Olivás, F. P. Romero, and E. Herrera-Viedma, "Sentiment analysis: a review and comparative analysis of web services," *Information Sciences*, vol. 311, pp. 18-38, Aug. 2015, doi: [10.1016/j.ins.2015.03.040](https://doi.org/10.1016/j.ins.2015.03.040).
- [19] B. Agarwal and N. Mittal, "Machine learning approach for sentiment analysis," in *Prominent feature extraction for sentiment analysis*, 2016, pp. 21-45, doi: [10.1007/978-3-319-25343-5_3](https://doi.org/10.1007/978-3-319-25343-5_3).
- [20] E. Haddi, X. Liu, and Y. Shi, "The role of text pre-processing in sentiment analysis," *Procedia Computer Science*, vol. 17, pp. 26-32, 2013, doi: [10.1016/j.procs.2013.05.005](https://doi.org/10.1016/j.procs.2013.05.005).
- [21] F. F. Balahadia, M. C. G. Fernando, and I. C. Juanatas, "Teacher's performance evaluation tool using opinion mining with sentiment analysis," in *2016 IEEE Region 10 Symposium (TENSYMP)*, 2016, pp. 95-98, doi: [10.1109/TENCONSpring.2016.7519384](https://doi.org/10.1109/TENCONSpring.2016.7519384).
- [22] A. Chanaa and N. El Faddouli, "E-learning text sentiment classification using hierarchical attention network (HAN)," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 13, p. 157, Jul. 2021, doi: [10.3991/ijet.v16i13.22579](https://doi.org/10.3991/ijet.v16i13.22579).
- [23] J. Fernández, E. Boldrini, J. M. Gómez, and P. Martínez-Barco, "Sentiment analysis and opinion mining: The EmotiBlog corpus," *Procesamiento de Lenguaje Natural*, vol. 47, 2011.
- [24] H. Jamshidifarsani, S. Garbaya, T. Lim, P. Blazevic, and J. M. Ritchie, "Technology-based reading intervention programs for elementary grades: An analytical review," *Computers & Education*, vol. 128, pp. 427-451, Jan. 2019, doi: [10.1016/j.compedu.2018.10.003](https://doi.org/10.1016/j.compedu.2018.10.003).
- [25] L. Hao and L. Hao, "Automatic identification of stop words in chinese text classification," in *2008 International Conference on Computer Science and Software Engineering*, 2008, pp. 718-722, doi: [10.1109/CSSE.2008.829](https://doi.org/10.1109/CSSE.2008.829).
- [26] O. Y. Adwan, M. Al-Tawil, A. Huneiti, R. Shahin, A. A. Zayed, and R. Al-Dibsi, "Twitter sentiment analysis approaches: a survey," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 15, p. 79, Aug. 2020, doi: [10.3991/ijet.v15i15.14467](https://doi.org/10.3991/ijet.v15i15.14467).
- [27] G. E. Chanchi, M. A. Ospina, and M. E. Ospino, "Sentiment analysis of the perception of systems engineering students from the University of Cartagena on the academic activities developed during confinement due to COVID-19," *Espacios*, vol. 41, no. 42, Nov. 2020, doi: [10.48082/espacios-a20v41n42p21](https://doi.org/10.48082/espacios-a20v41n42p21).
- [28] E. Rslan, M. H. Khafagy, K. Munir, and R. M. Badry, "English semantic similarity based on map reduce classification for agricultural complaints," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 12, 2021, doi: [10.14569/IJACSA.2021.0121231](https://doi.org/10.14569/IJACSA.2021.0121231).
- [29] J. Watts, "Assessing an online student orientation: impacts on retention, satisfaction, and student learning," *Technical Communication Quarterly*, vol. 28, no. 3, pp. 254-270, Jul. 2019, doi: [10.1080/10572252.2019.1607905](https://doi.org/10.1080/10572252.2019.1607905).
- [30] S. F. Persada, A. Oktavianto, B. Miraja, R. Nadlifatin, P. F. Belgiawan, and A. A. P. P. Redi, "Public perceptions of online learning in developing countries: a study using the ELK stack for sentiment analysis on Twitter," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 09, p. 94, May 2020, doi: [10.3991/ijet.v15i09.11579](https://doi.org/10.3991/ijet.v15i09.11579).
- [31] E. Cambria, "Affective computing and sentiment analysis," *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 102-107, Mar. 2016, doi: [10.1109/MIS.2016.31](https://doi.org/10.1109/MIS.2016.31).




- [32] Y. M. Sanzón, D. Vilariño, M. J. Somodevilla, C. Zepeda, and M. Tovar, "Models to detect the polarity of messages in social networks," *Research in Computing Science*, vol. 99, no. 1, pp. 29–42, Dec. 2015, doi: 10.13053/rcs-99-1-3.
- [33] Y. S. Montes, M. G. Prettel and M. T. Boutureira, "Emotions in university life: analysis of the relationship between emotional self-knowledge and emotional self-regulation in adolescents and university students," *Psicogente*, Vol. 21, no. 40, pp. 422-439, Jul-Dec. 2018, doi: <https://doi.org/10.17081/psico.21.40.3081>
- [34] D. Ramachandran and R. Parvathi, "Analysis of Twitter specific preprocessing technique for tweets," *Procedia Computer Science*, vol. 165, pp. 245–251, 2019, doi: 10.1016/j.procs.2020.01.083.

BIOGRAPHIES OF AUTHORS






Omar Chamorro-Atalaya    is an electronic engineer, graduated from the National University of Callao (UNAC), with a Master's degree in Systems Engineering and a doctoral student at the Faculty of Administrative Sciences at UNAC. Researcher recognized by CONCYTEC (National Council of Science, Technology and Technological Innovation – Peru). Research professor at the National Technological University of South Lima (UNTELS), he teaches courses on automatic process control and industrial automation, and design of control panels and electrical control. He can be contacted at email: ochamorro@untels.edu.pe.






Dora Arce-Santillan    is an electrical mechanical engineer, with a Master's Degree in Electrical Engineering with a mention in Engineering Project Management at the National University of Callao (UNAC). She is the author of scientific articles indexed in Scopus and WoS. She reviewer of scientific articles in international journals indexed to Scopus. She can be contacted at email: yvonne.asantillan@gmail.com.






José Antonio Arévalo-Tuesta    is a professional in economics, with master's studies in administration and educational management, doctorate in economics and education, with experience in business management, strategic planning of public and private institutions, design, planning, execution and evaluation of strategic private and social projects. He works in the university system, making a university teaching career at the Federico Villarreal National University (UNFV), ascending up to the category of main professor. Certified researcher by renacyt - sinacyt - concytec code: P0079433 occupying different positions, among them, director of the economic research institute, president of the board of directors and director of the pre-university center, dean of the faculty of economic sciences period (2007-2017), period (2015 - 2019), in the field of public administration, he worked in the work and social promotion sector, as general director of planning, and deputy executive director of the general directorate of administration. Also, as a member of the board of directors of the company mercado mayorista s.a. (emmsa). Director of the public technological institute manuel seoane corrales. Professor at master's level in strategic planning, investment projects and programs, public economics - research methodology, epistemology, territorial development in different institutions. Currently member of the board of directors of the national strategic planning center, ceplan and head of the postgraduate unit of the UNFV faculty of economics. He can be contacted at email: jarevalotu@unfv.edu.pe.






Lilia Rodas-Camacho    is a Professional in Law and Political Science. Titled and Collegiate; Master's degree in University Teaching and educational management, with solid experience in Legal Consulting and in the legal practice of the professional career. With more than 10 years of experience as a Professional Lawyer and University Professor. She works in the university system, completing his university teaching career at the Federico Villarreal National University. with personal skills in the exercise of oral litigation of law. Ease in the performance and development of university teaching and flexibility to adapt to change and make decisions under pressure. She can be contacted at email: lrodas@unfv.edu.pe.






Genaro Edwin Sandoval Nizama    is Doctor of Administration from the Accounting career. He is a graduate of the Inca Garcilazo de la Vega University. He has more than 10 years of professional experience. He currently works at the Cesar Vallejo University, in the Department of Business Sciences and the Accounting professional school. He can be contacted at email: gsandoval255@hotmail.com



Rosa Valle-Chavez    is Doctor in Education, Master in University Teaching and Bachelor of Industrial Relations graduated from the Universidad Particular San Martín de Porres. I am a responsible, proactive person, with initiative, ability to work in a team, multidisciplinary and committed to research work. I have 8 years of work experience in Teaching at Institutes and Universities. I currently work at the Private University César Vallejo Sede Callao and Piura. She can be contacted at email: rvallec@ucvvirtual.edu.pe



Yadit Rocca-Carbajal    is Master of Business Administration (MBA). Degree in International Business with work experience in private and public companies. My academic, human and work training has focused on the development and implementation of useful proposals in the business field. I am a professional committed to research, leadership and teamwork. Specialized in business intelligence, digital marketing and business training for micro and small entrepreneurs. She currently coordinator of the Professional School of International Business at the Cesar Vallejo University - Callao. She can be contacted at email: yrocca@ucv.edu.pe