Indonesian Journal of Electrical Engineering and Computer Science
Vol. 28, No. 1, October 2022, pp. 516~524
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v28.i1.pp516-524

Supervised learning using support vector machine applied to sentiment analysis of teacher performance satisfaction

Omar Chamorro-Atalaya[1], Dora Arce-Santillan[1], José Antonio Arévalo-Tuesta[2], Lilia Rodas-Camacho[2], Ronald Fernando Dávila-Laguna[3], Rufino Alejos-Ipanaque[4], Lilly Rocío Moreno-Chinchay[2]
[1]Faculty of Engineering and Management, Universidad Nacional Tecnológica de Lima Sur, Lima, Peru
[2]School of Economics, Universidad Nacional Federico Villarreal, Lima, Peru
[3]Faculty of Engineering, Universidad Cesar Vallejo, Lima, Peru
[4]Faculty of Administrative Sciences, Universidad Nacional del Callao, Lima, Peru

Article Info

Article history:

Received Feb 12, 2022
Revised Jul 30, 2022
Accepted Aug 5, 2022

Keywords:

Satisfaction
Sentiment analysis
Supervised learning
Suppor vector machine
Teacher performance

ABSTRACT

Satisfaction with teaching performance is an important measurement process in higher education institutions, for this reason, applying sentiment analysis to the opinions of university students through the support vector machine (SVM) Fine Gaussian supervised learning algorithm represents an important contribution to the academic literature. This article identifies the best classification algorithm according to performance parameters for predicting student satisfaction with teaching performance through sentiment analysis; the subsequent implementation of the research has the purpose of strengthening teaching practices, in addition to allowing continuous training of teaching for the benefit of student learning. This article has provided a compact predictive model, with literature review based on SVM and sentiment analysis techniques. Through the machine learning classification learner technique, it is identified that the SVM algorithm: Fine Gaussian SVM is the one with the best accuracy equal to 98.3%. Likewise, the performance metrics for the four classes of the model were identified, which have a sensitivity equal to 88.89%, a specificity of 98.04%, a precision of 99.21% and an accuracy of 98.85%.

Corresponding Author:

Omar Chamorro-Atalaya
Faculty of Engineering and Management, Universidad Nacional Tecnológica de Lima Sur
Lima, Peru
Email: ochamorro@untels.edu.pe

1. INTRODUCTION

The educational sector is incorporating information and communication technologies (ICT) in university activities in order to develop students' skills [1]. In particular, digital media, web applications and learning-teaching systems play a fundamental role in improving educational quality [2]-[4]. Educational institutions, in this search for opportunities to improve, have been identifying models to assess student satisfaction in line with trends in quality management and performance excellence [5], [6]. Christie *et al*. [7] point out, getting to know the dimension of student satisfaction with the institution they attend will allow identifying both positive and negative aspects, the latter being fundamental when determining strategies to improve education. Salas and Rueda [8] raises the importance of finding reliable ways to measure university student satisfaction, taking advantage of the explosive increase in the use of the Internet, because this would allow university institutions to know their reality and take corrective measures corresponding.

With the rapid growth of social networking applications, people use these platforms to express their opinions on everyday issues [9]. One of the platforms in which users can give an opinion is Twitter, through it the sentiment of the users is known [10], [11]. These opinions and comments can be extremely beneficial for organizations interested in knowing the public opinion about the services they offer. Ahmad *et al*. [12] this type of opinion can be obtained in another way through collection instruments such as questionnaires, which is undoubtedly a relatively arduous activity. Therefore, manually extracting an opinion from a large number of user comments is not feasible. Given this, one solution is to use an automatic method, used for the purpose of analyzing the polarization of users' sentiments.

Carrying out sentiment analysis is linked to the process that consists of representing opinions in terms of assessments, attitudes and emotions on a specific topic, generally sentiment analysis fulfills two tasks, firstly, to recognize the expressions of sentiment and define the orientation of the sentiment expressed by users [13], [14]. Barrett *et al*. [15] it is indicated that analyzing opinions is an activity linked to the natural language process (NLP) that allows identifying the opinions related to an object within a common context, the latter being a research technique that analyzes a determined sample of texts that are usually born in digital environments such as social networks [16]. To carry out sentiment analysis, methods can be applied with pre-established emotion dictionaries, also known as lexicons, or methods such as data mining and machine learning, which consists of building algorithms that learn to automatically classify large data sets [17]. Data mining allows us to identify patterns in large data sets, one of its characteristics is to be predictive, having the possibility of indicating what will happen, using statistics and probabilities of information that is hidden in stored data [18]. Currently, machine learning, which is a subfield of computer science and artificial intelligence, can be used to make predictions. Machine learning is a form of artificial intelligence that trains a virtual machine through data mining to automate data analysis processes, among other features [19].

Supervised learning emerges from machine learning in which algorithms that work from labeled data are grouped, these algorithms use a data history to be trained with the purpose of predicting an output value [20]. These algorithms can be based on probabilistic models such as Naive Bayes (NB), logical models such as random forest (RF) or geometric models such as the support vector machine (SVM). The aforementioned algorithms have obtained the best results in sentiment analysis, a task that focuses on classifying tagged tweets into three classes: positive (P), negative (N) and neutral (NEU) of the language used in it [21], [22]. Within the group of algorithms with the highest performance, SVM is often applied as an automatic classification technique for polarity detection from textual data, it consists of identifying the hyperplane that best separates two or more classes of instances belonging to a data set [23], [24]. Instead of focusing on reducing the training error like other classification algorithms, SVM focuses on minimizing the generalization error by widening the margins between the separation hyperplane and the instances, with the purpose of minimizing the structural risk, proposed in the statistical theory of learning [25].

In this sense, applying sentiment analysis to the opinions of university students through the SVM Fine Gaussian supervised learning algorithm represents an important contribution to the academic literature. First, although this type of sentiment analysis can be found in several high-level database articles, research showing models in the field of education is scarce. In other words, this study can be used for future research that analyzes text sets to identify the best classification algorithms according to their performance parameters. Given the above, this article aims to identify the best classification algorithm according to performance parameters for predicting student satisfaction with teaching performance through sentiment analysis. Taking into account the above, this research is divided into five sections, including this introduction. In the Section 2, the literary review is theoretically detailed, in the Section 3 the methodology used for data collection and the data processing technique is described. The Section 4 shows the results and discusses them against other similar studies; finally, in the Section 5, the most relevant conclusions are presented and future research is presented.

2. LITERARY REVIEW

Emotion analysis, defined as an area of computational study of opinions, feelings and emotions expressed in texts, has been combined with machine learning, data mining and natural language processing techniques. In the area of education, it has been sought to apply the analysis of emotions in order to improve the teaching-learning process. There is research that demonstrates the advantages of using social networks to encourage the participation of university students to express themselves freely [26], [27].

Teacher evaluation is considered a resource to conduct the work of teachers according to the performance obtained and a source to measure their performance is the assessment by students, which is called a model based on the opinion of students [28]. In the work reported by Ortigosa *et al*. [29] they carry out an analysis of the global comments to the teachers, with which they conclude that it is a good indicator, since it reveals qualities of the teacher in his work. The continuous improvement in the teaching-learning process has focused its efforts to conclude the components that determine the teaching work carried out properly, instruments such as questionnaires for the evaluation of teaching by the students seek to have the opinion of the students, which have figured as one of the most examined and applied tools for this purpose [30].

In relation to relevant studies on this topic, we have in [31], the study of sentiment analysis of tweets that uses the machine learning approach with Naive Bayes and SVM supervised classification algorithms, which shows optimal results compared to traditional techniques. In the same line of research proposals, such as the one carried out in [32], have been carried out in a flexible way in terms of the application of sentiment classification techniques based on machine learning, confirming the importance of the domain to build accurate data extraction systems opinions, in addition to the influence of the size of the data set. Likewise, the research carried out in [33] focuses on the combination of two machine learning algorithms, SVM and decision tree rules, which highlight that the metrics of these algorithms are not only useful in reviewer classification, but also are free of undesirable biases that allows it to be considered optimal both in terms of utility and satisfaction. On the other hand, in the study of [34] the machine learning technique has been used for sentiment analysis, the comparative experiment revealed the superior precision of the method used in terms of extracting multiple review elements, in relation to other methods. Approaches such as the study of [35], where algorithms based on dictionaries are used to carry out the classification of sentiments, represent the importance of the use of these techniques. The study presented by [36] analyzes Twitter opinions using machine learning; the novelty of the proposed approach is that the publications acquire a weight for each comment.

3. METHOD

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. This period for acquiring comments or opinions is from week 9 to week 13 of the academic semester. The comments or opinions obtained were stored in a "csv" extension format, for later conditioning, which is done to obtain the polarization of the comments made by the twitter social network through sentiment analysis. The conditioning or pre-processing consists of 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. Next, the sentiment analysis of the tweets written in English is carried out; the result obtained will be the quantification of the sentiment contained in each comment or opinion written by the student. The polarization of comments or opinions are classified into positive polarity, neutral polarity and negative polarity. Likewise, the information collected was processed using the Classification Learner technique, to identify the best Machine Learning algorithm, through its performance parameters. Finally, once the algorithm for predicting teacher performance satisfaction with the best accuracy has been identified, the capabilities of the parameters. Figure 1 shows the acquisition, processing and identification diagram of the predictive model.

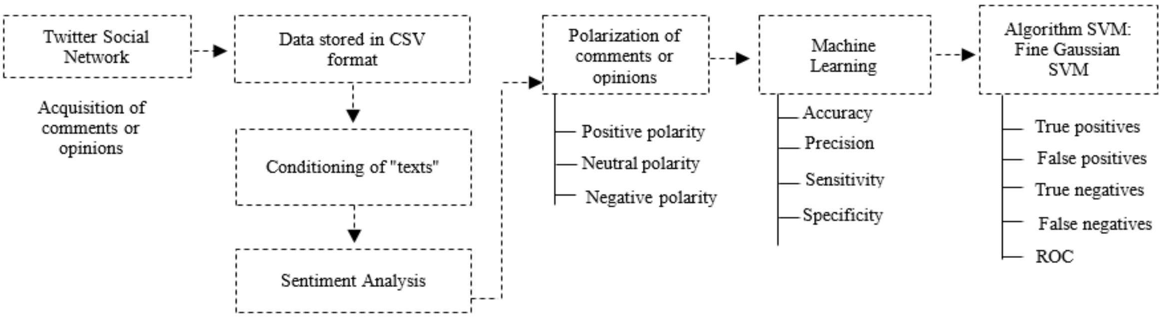


Figure 1. Diagram of acquisition, processing and identification of the predictive model

4. RESULTS AND DISCUSSION

As part of the results, the identification of the main parameter begins, which is the assessment of the accuracy with which the model has classified the instances in the training phase. Since it is a prediction model, we are interested in knowing if which algorithm performs predictions better. In this sense, Table 1 shows the results generated, which show that the SVM algorithm: Fine Gaussian SVM, is the one with the best accuracy of 98.3% for predicting student satisfaction with performance teacher through sentiment analysis. Being the accuracy the percentage of data that the model has classified correctly, it can be indicated that 98.3% of the number of positive predictions with this algorithm will be true.

Table 1. Determination of the classification algorithm

Algorithm	Accuracy
SVM: Fine Gaussian SVM	98.3%
SVM: Cubic SVM	93.1%
SVM: Quadratic SVM	93.1%
SVM: Linear SVM	83.6%

Selecting the classification algorithm according to the presented accuracy, the confusion matrix is shown, which compares the values that are predicted in the model with the real ones, through this analysis the sensitivity parameter will be identified, of the classes of the Fine Gaussian algorithm SVM, which is represented by the polarity of sentiments (positive polarity, negative polarity and neutral polarity) regarding student satisfaction with teaching performance, it is necessary to indicate that each column of the matrix represents the number of predictions for each class performed by the model, while each row reflects the actual values for each class. Similar to the development of the research using the SVM algorithm, the study by [37] indicates that they used the Twitter API to extract tweets. These were used to identify tweets as negative or positive; several algorithms were used to classify the tweets, obtaining the best results with the SVM classifier, which achieved an accuracy of 83% compared to the Naive Bayes (NB) which showed a classification accuracy of 82.7%. Likewise, Barhan and Shakhomirov [38] they proposed the supervised method to classify Twitter data. The results of this experiment showed that the SVM showed better performance than other algorithms with 88% accuracy. Similarly, Sadiq *et al*. [39] the experiment performed showed that the value stream mapping (VSM) algorithm performed better than the Naïve Bayes algorithm, achieving an accuracy of 81% and a retrieval accuracy of 74%.

Shown in Figure 2 are the false negative percentage (FNR), which represents the probability that the test will miss a true positive, and the true positive rate (TPR), which represents the probability that a positive result will be missed be actually positive. As can be seen, of the 3 polarities of sentiments towards teacher performance satisfaction, the neutral polarity (class 2) and positive polarity (class 1) show 100% sensitivity, this means that the Fine Gaussian SVM algorithm has a 100 % ability to correctly detect a true positive (TP), which are the correct predictions for both classes, from a false negative (FN). On the other hand, the negative polarity (class 1) shows a percentage of true positives of 66.7% and a percentage of false negatives of 33.3%, that is to say that the algorithm in 33.3% can show negative value predictions when the value should really be negative be positive.

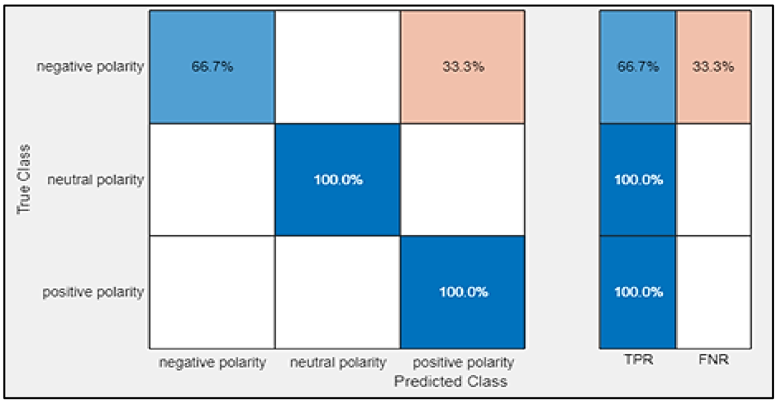


Figure 2. Rates of TPR and FNR in the Confusion Matrix

In Figure 3, the percentage of the positive predictive values (PPV) and the false discovery rate (FDR) are shown. As observed, the negative polarity (class 1) and neutral (class 2) show the highest precision value, in this case 100%, while the positive polarity (class 3) has a capacity of 97.6% and a percentage of false discoveries of 2.4%, that is to say, that only 97.6% of the sentiments really have positive polarities and only 2.4% of sentiments will be wrong in the prediction.

Figure 3. PPV and FDR rates in the Confusion Matrix

Below are receiver operating characteristic (ROC) curves that represent one method of establishing the accuracy of the algorithm. It should be noted that the closer the area under the curve (AUC) indicator is to one, the better the algorithm will perform. Figure 4 shows the ROC graph for class 1 (negative polarity), where it is evident that there is an accuracy of 98%. Another aspect to highlight is the discrimination threshold, whose values are 0.67 and 0 for TPR and TFP, respectively.

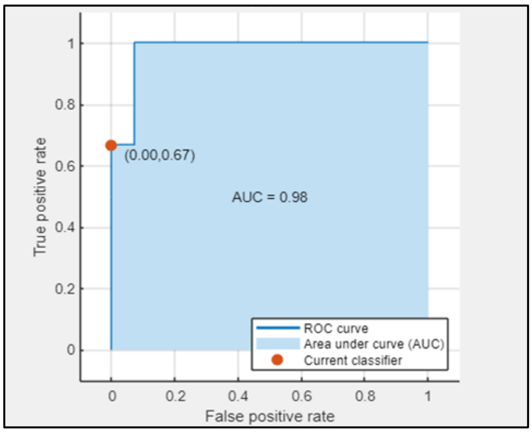


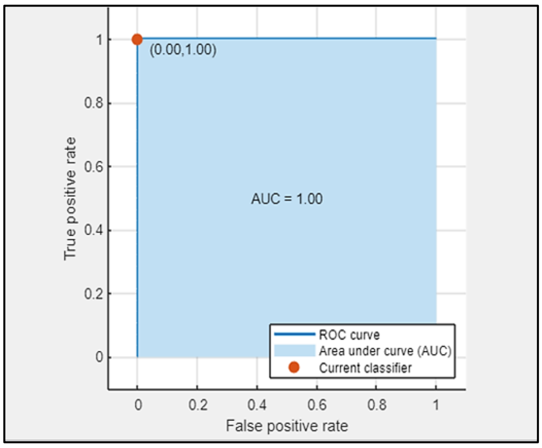
Figure 4. ROC curve for polarity negative

Continuing with the analysis of the performance of the algorithm in Figure 5, the relationship between sensitivity and specificity for classes 2 and 3 is shown. Figure 5(a) shows the ROC plot for classes 2 and 3, showing an accuracy of 100%. While the discrimination threshold is 1.00 for TPR and 0.00 for TFP. Figure 5(b) shows the ROC graph for class 3 (positive polarity), where it is evident that there is a precision of 100%. While the discrimination threshold is 1.00 for TPR and 0.06 for TFP.

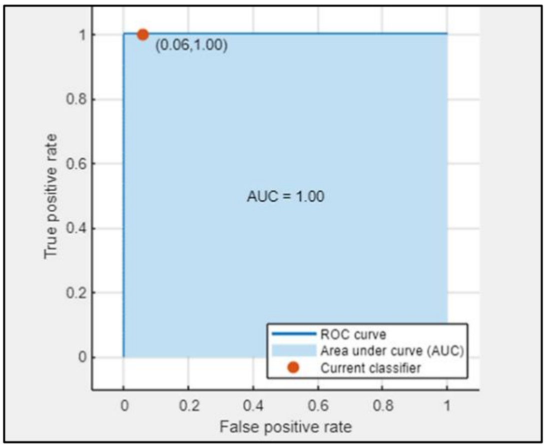
Through what is obtained in Table 2, the performance metrics of the Fine Gaussian SVM algorithm are shown, for the four classes (1: negative polarity, 2: neutral polarity and 3: positive polarity), in general it is possible to visualize optimal parameters of performance with a Precision of 99.21%, a Sensitivity of 88.89%, a Specificity of 98.04% and an Accuracy of 98.85%. The results obtained show an accuracy of the Fine Gaussian SVM algorithm of 98.85%, with which it can be pointed out that the predictive model presented will show optimal performance, supported by what was obtained in [10], where it is pointed out that the results related to the accuracy of the classifier correspond to 72%. These results can help predict interests or future trends with greater confidence, as was verified in a test where tweets were classified according to the positive or negative polarity of their sentiment.

In relation to other works that have contributed similar topics, the SVM algorithm used in [21] indicates that the support vector machine algorithm is the best used to correctly classify the sentiment of the tweets, whether positive or negative, for such this reason, the experiments were performed using a large training dataset and the algorithm achieved a high accuracy of around 87%. Regarding the use of machine learning and text mining in the area of education, in [40] it is pointed out that the Ensemble Bagged Trees classification algorithm shows an accuracy of 81.3%, for the 4 classes (levels of satisfaction) of the predictive model satisfaction of teaching performance in the virtual environment. In the same way, in [41] a supervised learning model is carried out for the predictive system of personal and social attitudes of university students of professional engineering careers, through the logistic regression kernel algorithm, an accuracy of 91.96%

is obtained, a precision of 79.09%, a Sensitivity of 75.66% and a Specificity of 92.09%. Similarly, Salas and Rueda [8] it is pointed out that the decision tree technique allows identifying 8 predictive models on the interaction and communication of students in the social network Facebook during the teaching and learning process. Likewise, as indicated in [37], 3 predictive models were made through the decision tree technique, which allowed the students of the Basic Applied Statistics subject to be more motivated and satisfied to use the application of data science during the teaching-learning process. Similarly, in Atalaya *et al*. [22], in order to improve educational quality, the k-nearest neighbors (K-NN) algorithm was used in the predictive analysis of the quality of the university administrative service in the virtual environment, identifying that the algorithm's metrics have an accuracy of 92.77 %, a sensitivity of 86.62% and a specificity of 94.7%, with a total accuracy of 85.5%. Finally, another investigation that is important to compare our findings is that of [42], where the students' academic studies performances were analyzed and predicted using three data mining techniques: decision tree, multilayer of perception and Naïve Bayes. Being this last algorithm that showed a prediction accuracy of 86%, thereby helping teachers to detect those students who are expected to obtain a low grade.



(a)



(b)

Figure 5. ROC curve for (a) neutral polarity and (b) positive polarity

Table 2. Fine Gaussian SVM algorithm performance parameters

	Sensitivity	Specificity	Accuracy	Precision
Negative polarity	66.67%	100.00%	98.28%	100.00%
Neutral polarity	100.00%	100.00%	100.00%	100.00%
Positive polarity	100.00%	94.12%	98.28%	97.62%
Total	88.89%	98.04%	98.85%	99.21%

5. CONCLUSION

In this age of technology and digitization, Twitter becomes a rich source for sentiment analysis and text mining. The objective of this article is to identify the best classification algorithm according to performance parameters for the prediction of student satisfaction with teaching performance through sentiment analysis. This article has provided a compact predictive model, with literature review based on SVM and sentiment analysis techniques. Through the automatic learning classification technique, it is identified that the SVM algorithm: Fine Gaussian SVM, is the one that presents a better accuracy of 98.3%, thus validating that SVM is one of the techniques of most used classification for polarity detection from textual data. Likewise, the performance metrics of the Fine Gaussian SVM algorithm were identified for the four classes (1: negative polarity, 2: neutral polarity and 3: positive polarity), which have a precision of 99.21%, a sensitivity of 88.89%, a specificity of 98.04% and an accuracy of 98.85%, for the prediction of student satisfaction with teaching performance. It is recommended to implement this proposed algorithm, for the identification of factors that affect university satisfaction with the quality of the educational service in order to take the pertinent corrective action.

ACKNOWLEDGEMENTS
Thanks and appreciation to the "Universidad Nacional Tecnológica de Lima Sur".

REFERENCES
[1] P. Ramkissoon, L. J. Belle, and T. Bhurosy, "Perceptions and experiences of students on the use of interactive online learning technologies in Mauritius," *International Journal of Evaluation and Research in Education*, vol. 9, no. 4, pp. 833–839, Dec. 2020, doi: 10.11591/ijere.v9i4.20692.
[2] A. R. Bernaola, M. Aldude, J. Estrada, V. Señas, and L. A. Arenas, "Analysis of the use of technological tools in university higher education using the soft systems methodology," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 7, 2020, doi: 10.14569/IJACSA.2020.0110754.
[3] E. A. Lagunas, J. R. L. Cuevas, and E. P. Palencia, "Perception of students towards the quality of private universities in Monterrey," *Revista Iberoamericana de Educacion Superior*, vol. 6, no. 17, pp. 58–76, Sep. 2015, doi: 10.1016/j.rides.2015.10.003.
[4] 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, pp. 492–499, Oct. 2020, doi: 10.11591/ijeecs.v20.i1.pp492-499.
[5] R. A. Majid and J. C. Hasim, "The effectiveness of frog VLE implementation: Students' perspective," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 14, no. 1, pp. 381–387, Apr. 2019, doi: 10.11591/ijeecs.v14.i1.pp381-387.
[6] N. D. Lescano, H. I. M. Huamán, D. M. D. de Angulo, and A. V. Muñoz, "Satisfaction level in training quality at university level," *Telos Revista de Estudios Interdisciplinarios en Ciencias Sociales*, vol. 23, no. 2, pp. 247–266, May 2021, doi: 10.36390/telos232.04.
[7] B. A. Christie, K. K. Miller, R. Cooke, and J. G. White, "Environmental sustainability in higher education: What do academics think?," *Environmental Education Research*, vol. 21, no. 5, pp. 655–686, Jul. 2015, doi: 10.1080/13504622.2013.879697.
[8] R. S. Rueda and R. S. Rueda, "Impacto de la red social Facebook en el proceso educativo superior de las matemáticas considerando la ciencia de datos," *Nóesis. Revista de Ciencias Sociales y Humanidades*, vol. 28, no. 55–1, pp. 23–42, May 2019, doi: 10.20983/noesis.2019.3.2.
[9] A. Alsaeedi and M. Zubair, "A study on sentiment analysis techniques of Twitter data," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 2, 2019, doi: 10.14569/IJACSA.2019.0100248.
[10] P. S. Holgado, M. M. M. Acera, and D. B. Herrero, "From data-driven to data-feeling: sentiment analysis in real-time of messages in Spanish about Scientific communication using machine learning techniques," *Anuario Electrónico de Estudios en Comunicación Social "Disertaciones,"* vol. 13, no. 1, Jan. 2020, doi: 10.12804/revistas.urosario.edu.co/disertaciones/a.7691.
[11] K. Mrhar, L. Benhiba, S. Bourekkache, and M. Abik, "A Bayesian CNN-LSTM Model for sentiment analysis in massive open online courses MOOCs," *International Journal of Emerging Technologies in Learning*, vol. 16, no. 23, pp. 216–232, Dec. 2021, doi: 10.3991/ijet.v16i23.24457.
[12] M. Ahmad, S. Aftab, M. S. Bashir, N. Hameed, I. Ali, and Z. Nawaz, "SVM optimization for sentiment analysis," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 4, pp. 393–398, 2018, doi: 10.14569/IJACSA.2018.090455.
[13] C. A. A. Pastrana and C. F. O. Andrade, "Mandatory social isolation: a sentiment analysis using machine learning," *Suma de Negocios*, vol. 12, no. 26, pp. 1–13, Dec. 2021, doi: 10.14349/sumneg/2021.v12.n26.a1.
[14] W. Al-Ghaith, "Developing lexicon-based algorithms and sentiment lexicon for sentiment analysis of saudi dialect tweets," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 11, pp. 83–88, 2019, doi: 10.14569/IJACSA.2019.0101112.
[15] N. Barrett, J. H. W. Jahnke, and V. Thai, "Engineering natural language processing solutions for structured information from clinical text: Extracting sentinel events from palliative care consult letters," *Studies in Health Technology and Informatics*, vol. 192, no. 1–2, pp. 594–598, 2013, doi: 10.3233/978-1-61499-289-9-594.
[16] Q. Tul *et al.*, "Sentiment analysis using deep learning techniques: a review," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 6, 2017, doi: 10.14569/ijacsa.2017.080657.
[17] A. Muslim, A. B. Mutiara, R. Refianti, C. M. Karyati, and G. Setiawan, "Comparison of accuracy between long short-term memory-deep learning and multinomial logistic regression-machine learning in sentiment analysis on twitter," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 2, pp. 747–754, 2020, doi: 10.14569/ijacsa.2020.0110294.
[18] E. Teso, M. Olmedilla, M. R. M. Torres, and S. L. Toral, "Application of text mining techniques to the analysis of discourse in eWOM communications from a gender perspective," *Technological Forecasting and Social Change*, vol. 129, pp. 131–142, Apr. 2018, doi: 10.1016/j.techfore.2017.12.018.
[19] D. B. Fernández, D. Gil, and S. L. Mora, "Application of machine learning in predicting performance for computer engineering students: A case study," *Sustainability (Switzerland)*, vol. 11, no. 10, p. 2833, May 2019, doi: 10.3390/su11102833.
[20] E. J. D. L. Hoz, E. J. D. La Hoz, and T. J. Fontalvo, "Methodology of machine learning for the classification and prediction of users in virtual education environments," *Informacion Tecnologica*, vol. 30, no. 1, pp. 247–254, 2019, doi: 10.4067/S0718-07642019000100247.
[21] S. Al-Otaibi *et al.*, "Customer satisfaction measurement using sentiment analysis," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 2, pp. 106–117, 2018, doi: 10.14569/IJACSA.2018.090216.
[22] O. F. C. Atalaya *et al.*, "K-NN supervised learning algorithm in the predictive analysis of the quality of the university administrative service in the virtual environment," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 25, no. 1, pp. 521–528, Jan. 2022, doi: 10.11591/ijeecs.v25.i1.pp521-528.
[23] A. Ben-Hur and J. Weston, "A user's guide to support vector machines.," in *Methods in molecular biology (Clifton, N.J.)*, vol. 609, 2010, pp. 223–239.
[24] E. G. Nihad, E.-N. El Mokhtar, Z. Abdelhamid, and A. A. Mohammed, "Hybrid approach of the fuzzy C-Means and the K-Nearest neighbors methods during the retrieve phase of dynamic case based reasoning for personalized Follow-up of learners in real time," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 6, p. 4939, Dec. 2019, doi: 10.11591/ijece.v9i6.pp4939-4950.
[25] X. Zhu, S. Kiritchenko, and S. M. Mohammad, "NRC-Canada-2014: Recent Improvements in the Sentiment Analysis of Tweets," in *8th International Workshop on Semantic Evaluation, SemEval 2014 - co-located with the 25th International Conference on Computational Linguistics, COLING 2014, Proceedings*, 2014, pp. 443–447, doi: 10.3115/v1/s14-2077.
[26] M. G. G. Lizarraga, M. T. B. Traver, and M. B. Y. Diaz, "Cyberactivism: A new form of participation for University students," *Comunicar*, vol. 24, no. 46, pp. 47–54, Jan. 2016, doi: 10.3916/C46-2016-05.

[27] N. Altrabsheh, M. Cocea, and S. Fallahkhair, "Learning sentiment from students' feedback for real-time interventions in classrooms," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8779 LNAI, 2014, pp. 40–49.
[28] A. Gewerc, L. Montero, and M. Lama, "Collaboration and social networking in higher education," *Comunicar*, vol. 21, no. 42, pp. 55–62, Jan. 2014, doi: 10.3916/C42-2014-05.
[29] A. Ortigosa, J. M. Martín, and R. M. Carro, "Sentiment analysis in Facebook and its application to e-learning," *Computers in Human Behavior*, vol. 31, no. 1, pp. 527–541, Feb. 2014, doi: 10.1016/j.chb.2013.05.024.
[30] A. Martínez, M. Sánchez, and J. Martínez, "Student opinion questionnaires on teaching performance: An institutional strategy for the evaluation of medical teaching," *Electronic journal of educational research*, vol. 12, no. 1, pp. 1–18, 2010.
[31] F. L. Cruz, J. A. Troyano, F. Enríquez, F. J. Ortega, and C. G. Vallejo, "'Long autonomy or long delay?' the importance of domain in opinion mining," *Expert Systems with Applications*, vol. 40, no. 8, pp. 3174–3184, Jun. 2013, doi: 10.1016/j.eswa.2012.12.031.
[32] H. J. Min and J. C. Park, "Identifying helpful reviews based on customer's mentions about experiences," *Expert Systems with Applications*, vol. 39, no. 15, pp. 11830–11838, Nov. 2012, doi: 10.1016/j.eswa.2012.01.116.
[33] L. Chen, L. Qi, and F. Wang, "Comparison of feature-level learning methods for mining online consumer reviews," *Expert Systems with Applications*, vol. 39, no. 10, pp. 9588–9601, Aug. 2012, doi: 10.1016/j.eswa.2012.02.158.
[34] M. Eirinaki, S. Pisal, and J. Singh, "Feature-based opinion mining and ranking," *Journal of Computer and System Sciences*, vol. 78, no. 4, pp. 1175–1184, Jul. 2012, doi: 10.1016/j.jcss.2011.10.007.
[35] E. Kontopoulos, C. Berberidis, T. Dergiades, and N. Bassiliades, "Ontology-based sentiment analysis of twitter posts," *Expert Systems with Applications*, vol. 40, no. 10, pp. 4065–4074, Aug. 2013, doi: 10.1016/j.eswa.2013.01.001.
[36] S. Mohamed and A. Ezzati, "A data mining process using classification techniques for employability prediction," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 14, no. 2, pp. 1025–1029, May 2019, doi: 10.11591/ijeecs.v14.i2.pp1025-1029.
[37] M. Anjaria and R. M. R. Guddeti, "Influence factor based opinion mining of Twitter data using supervised learning," in *2014 6th International Conference on Communication Systems and Networks, COMSNETS 2014*, Jan. 2014, pp. 1–8, doi: 10.1109/COMSNETS.2014.6734907.
[38] A. Barhan and A. Shakhomirov, "Methods for sentiment analysis of Twitter messages," in *Proceedings of the 12th conference of fruct association*, 2012, pp. 215–222, doi: https://doaj.org/article/2ced321baf3243b1bebc74166e04932f.
[39] S. Hussain, N. A. Dahan, F. M. Ba-Alwib, and N. Ribata, "Educational data mining and analysis of students' academic performance using WEKA," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 9, no. 2, pp. 447–459, Feb. 2018, doi: 10.11591/ijeecs.v9.i2.pp447-459.
[40] K. H. Susheelamma and K. M. Ravikumar, "Student risk identification learning model using machine learning approach," *International Journal of Electrical and Computer Engineering*, vol. 9, no. 5, pp. 3872–3879, Oct. 2019, doi: 10.11591/ijece.v9i5.pp3872-3879.
[41] H. Jamshidifarsani, S. Garbaya, T. Lim, P. Blazevic, and J. M. Ritchie, "Technology-based reading intervention programs for elementary grades: An analytical review," *Computers and Education*, vol. 128, pp. 427–451, Jan. 2019, doi: 10.1016/j.compedu.2018.10.003.
[42] A. Mueen, B. Zafar, and U. Manzoor, "Modeling and Predicting students' academic performance using data mining techniques," *International Journal of Modern Education and Computer Science*, vol. 8, no. 11, pp. 36–42, Nov. 2016, doi: 10.5815/ijmecs.2016.11.05.

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 an 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 an 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.

Ronald Fernando Dávila-Laguna is an Industrial Engineer by profession, Doctor of Administration from the Cesar Vallejo University. He is a magister with a mention in "Social development management" from the "Hermilio Valdizan" National University of Huánuco. He has diplomas in Public and Private Investment Projects, Direction and Management of Organizational Development, methodology of scientific research. He is a speaker at national and international conferences. He has work experience as head of area, production and marketing in the public and private sectors. He currently works as a teacher at the Cesar Vallejo University, the Scientific University of the South, and the Catholic University Sedes Sapientiae. He can be contacted at email: rdavilala@ucvvirtual.edu.pe.

Rufino Alejos-Ipanaque is a Doctor of Administration; Master in Education and management. Bachelor of Business Administration. Collegiate, graduated with a master's degree in strategic business administration; diploma in research and university teaching; administrative audit; municipal regional public management; experience in administrative management in public and private, non-governmental companies; experience in establishing, interacting and maintaining a good level of management and interpersonal relationships with clients, university teaching experience. He can be contacted at email: ralejosi@unac.edu.pe.

Lilly Rocío Moreno Chinchay is a Doctor in Administration. Master in National Development Planning and Bachelor of Economics. University Professor at the Federico Villarreal National University, in the Faculty of Economic Sciences. She can be contacted at email: lmoreno@unfv.edu.pe.