Indonesian Journal of Electrical Engineering and Computer Science
Vol. 27, No. 3, September 2022, pp. 1625~1634
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v27.i3.pp1625-1634

Supervised learning through k-nearest neighbor, used in the prediction of university teaching performance

Omar Chamorro-Atalaya[1], Nestor Alvarado-Bravo[2], Florcita Aldana-Trejo[3], Claudia Poma-Garcia[4], Carlos Aliaga-Valdez[5], Gutember Peralta-Eugenio[6], Abel Tasayco-Jala[7]
[1]Facultad de Ingeniería y Gestión, Universidad Nacional Tecnológica de Lima Sur, Lima, Perú
[2]Facultad de Ingeniería Química, Universidad Nacional del Callao, Lima, Perú
[3]Facultad de Ciencias Económicas, Universidad Nacional Federico Villarreal, Lima, Perú
[4]Facultad de Ingeniería Industrial, Universidad César Vallejo, Lima, Perú
[5]Facultad de Ciencias Administrativas, Universidad Nacional del Callao, Lima, Perú
[6]Facultad de Psicología, Universidad César Vallejo, Lima, Perú
[7]Facultad de Ciencias Empresariales, Universidad César Vallejo, Lima, Perú

Article Info

Article history:

Received Mar 23, 2022
Revised Jun 11, 2022
Accepted Jul 1, 2022

Keywords:

KNN
Prediction
Satisfaction
Supervised learning
Teaching performance

ABSTRACT

This study initially seeks to identify the most optimal supervised learning algorithm to be used in predicting the perception of teacher performance, and then to evaluate its performance indicators that validate its predictive capacity. For this, the MATLAB R2021a software is used; the experimental results determine that the supervised learning algorithm k-nearest neighbor weighted (weighted KNN) will be correct in 98.10% in predicting the perception of teaching performance, this has been validated by carrying out two evaluations through its performance indicators obtained in the confusion matrix and the receiver operating characteristic (ROC) curve, in the first evaluation an average sensitivity of 97.9%, a specificity of 99.1%, an accuracy of 98.8% and a precision of 96.7% are observed, thus validating the ability of the weighted KNN model to correctly predict the perception of teacher performance; while in the ROC curve, values of the area under the curve (AUC) equal to 0.99 and 1 are obtained, with this it is possible to validate the capacity that the model will have to distinguish between the 4 classes of the perception of the university teaching performance.

Corresponding Author:

Omar Chamorro-Atalaya
Facultad de Ingeniería y Gestión, Universidad Nacional Tecnológica de Lima Sur
Sector 3 Grupo 1A 03, Av. Central, Villa El Salvador 15834, Lima, Perú
Email: ochamorro@untels.edu.pe

1. INTRODUCTION

The so-called efficient teaching systems are oriented to the use of resources for the development of learning sessions, the same ones that generate large volumes of data, which is why their treatment is relevant today, since from the information obtained, it will be possible to improve decision-making that leads to generating a positive impact on university quality [1]-[3]. Since the appearance of artificial intelligence, data science, supervised learning, data mining and natural language processing, important progress has been made in the university educational field in the search for the improvement of educational quality [4]-[6]. In the framework of globalization, many organizations rely on data mining to monitor quality and performance indicators [7]. Thus, in the case of university academic organizations, the creation of alternative models for the evaluation of teaching performance constitutes a challenge for the authorities, since their adequate identification will allow them to contribute to the improvement of the activities involved in university management and, therefore, in improving educational quality [8]. An aspect to take into account is educational data mining, which is responsible for exploring indicators and parameters that come strictly from an academic environment, seeking to generate models of association in relation, under a prospective approach [9]-[11].

In the context of migration towards a teaching based on the use of technological tools, it is perceived in many cases that teachers have a poor use of technologies or virtual tools [12]-[14]; and it is that the teacher must have the capacity for adaptability and flexibility to the evolving environments of teaching that are being developed these days [15]-[17]; whose lack or lack influences the university educational quality, becoming a relevant and necessary aspect to permanently evaluate the teaching performance in various factors, with the purpose of implementing a quality education [18], [19]. In recent years, university educational institutions have carried out a poor systematization of teacher performance evaluation processes, so it is important to incorporate mechanisms that contribute to its assessment [20]; consistent with the evolution of data processing strategies [21], [22]; making these institutions more significant and even more effective [23].

There are today several methods based on machine learning that are relevant to analyze patterns linked to teaching performance, among which are the support vector machine (VSM), K-nearest neighbors (K-NN) or neural networks, showing effective results in the search to identify and understand influential factors in the improvement of academic quality [24], [25]. The K-NN method is one of the most frequently applied classification algorithms today in various research fields, being ideal for solving multi-category problems [26]. K-NN is a method in which, if the largest number of samples most adjacent to a sample belongs to a certain type, the sample called test or assay will belong to this [27], [28]. These classification models are evaluated through performance measures, comparing the predictions generated for a certain test or validation sample against the true classes of the same set [29].

In this sense, the objective of the article is to determine the performance indicators of the supervised learning method through K-NN, applied to the prediction of teaching performance. So also, describe its indicators such as accuracy, specificity, precision and sensitivity. The research is carried out in order to generate a prediction model that contributes to the improvement of teaching performance, the contribution of this type of study is based on making the early forecast of student satisfaction, which is linked to their academic performance, whose variable is of the utmost importance for the higher institution, in this way it will allow them to make decisions for improvement in a timely manner during the development of the academic cycle and not when it has already finished, taking into consideration the perception of the students in real time.

2. METHOD

2.1. Technique and instrument for data collection and validation of results

The technique used for data collection is the survey, and the instrument is made up of a questionnaire that includes 20 questions, divided into six factors for evaluating teacher performance. The data under study were collected during the academic years 2018, 2019 and 2020, resulting in a total of 963 evaluations, for each indicator of the data collection instrument. Table 1 shows the distribution of teachers who were evaluated in each academic semester.

In relation to the reliability of the data collection instrument, the content of the instrument is validated through [30], in said investigation, it is argued, the choice of the indicators, that, in general terms, their selection was made by the academic organizing committee of the higher institution. Table 2 specifies the factors that make up the data collection instrument, as well as its corresponding Cronbach's Alpha, in order to demonstrate the reliability of the data collected.

Table 1. Distribution of professors evaluated by academic semester

Academic semester	2018-I	2018-II	2019-I	2019-II	2020-I	2020-II
Number of teachers evaluated	157	160	164	165	161	156

Table 2. Validation of data collected through Cronbach's Alpha

Coding	Indicators of the data collection instrument	Cronbach's alpha if the indicator is excepted	Cronbach's alpha in general
IDD1	Planning skill	0.993	0.994
IDD2	Ability of didactic strategies	0.992	
IDD3	Communication skills	0.992	
IDD4	Ability to manage learning sessions	0.994	
IDD5	Ability to interact positively with students	0.992	
IDD6	Global evaluation of teaching performance	0.994	

2.2. Research design

The research design is of a non-experimental type, this is due to the fact that the variable under study is not influenced, but on the contrary, the analysis is carried out on its natural state. Figure 1 shows the stages that will lead to obtaining the predictive model of teacher performance through the supervised learning technique. In the first place, the data collection stage is highlighted, the same as having identified the predictive elements and the "target" or output variable (perception of teaching performance), we proceeded to use the MATLAB R2021a software with the purpose of identifying the performance indicators of the method to be used as part of the supervised learning, in order to obtain the predictive classification algorithm.

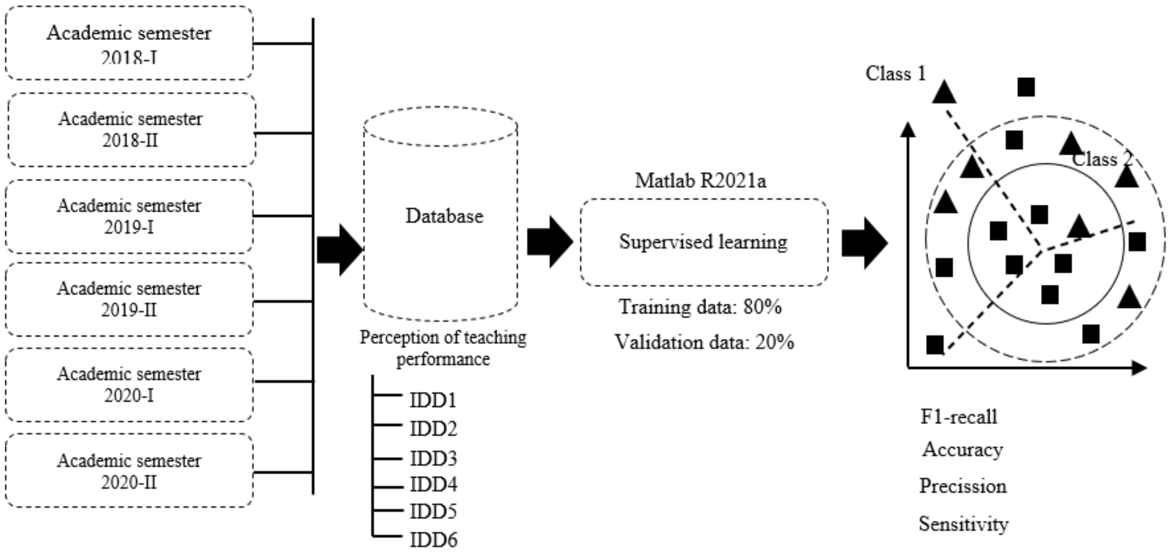

Figure 1. Stages of data processing through supervised learning

3. LITERARY REVIEW

The evaluation of the teaching performance corresponds to issuing a value judgment regarding the fulfillment of the responsibilities in the teaching-learning process, in order to obtain valid, objective and reliable information, and, finally, to determine the achievements reached by the students, as well as the development of work areas. Therefore, it is crucial to emphasize that this evaluation constitutes the fundamental strategy to improve educational quality. Without efficient teachers, education cannot be optimized or transformed [4]. In general, it can be said that it is necessary to evaluate teacher performance so that education, both virtual and face-to-face, reaches its potential. Therefore, predicting student dropout, teaching performance and academic performance in distance learning on virtual platforms is one of the main concerns of educational data mining (EDM) or educational data mining [31].

Yuliansyah *et al*. [32] it is pointed out that in various academic investigations interactions with decision trees, naive Bayes and k-nearest neighbors (k-NN) were analyzed, observing that these algorithms have a greater predictive power of performance. On the other hand, a systematic review published by Sokkhey and Okazaki [25] cites numerous articles dated between 2009 and 2016 that use data mining to predict not only academic desertion but, in general, the satisfaction of students with the factors of the teaching process, using system logs. Similar achievements were made by An *et al*. [26] in their study on early prediction of performance in virtual environments using machine learning and learning management system (LMS) logs. Studies such as the ones mentioned above show that performance could be predicted by applying machine learning algorithms to the recording of the different interactions with the virtual platform.

4. RESULTS AND DISCUSSION

Initially, the supervised learning algorithm that will be used in predicting the perception of university teaching performance will be identified, for which the classification learner technique is used through the MatLab environment. The first indicator to determine the most optimal algorithm is the accuracy validation, this indicator shows us how close the result is to a true prediction, that is, to correct positive predictions. Table 3 shows the most optimal algorithms to be used in predicting the perception of university teaching performance.

Table 3 shows that the most optimal algorithm to be used in predicting the perception of university teaching performance is the weighted k-nearest neighbor supervised learning algorithm (weighted KNN), because, with this algorithm, the model predictive will be 98.10% correct in university teaching performance. Once the most optimal algorithm has been initially identified, its performance is visualized through the performance indicators (sensitivity, specificity, accuracy and precision), this analysis is carried out for each of its teaching performance classes (class 1: regular, class 2: good, class 3: deficient, and class 4: very good) and in general, through the confusion matrix tool. The confusion matrix is represented by four quadrants, true positive (TP), false positive (FP), true negative (TN) and false negative (FN), these quadrants are very important, because they are used in the evaluation of performance indicators. Table 4 shows the results of the quadrants of the confusion matrix.

Table 4 shows us in the true positive (TP) quadrant, 280 predictions in class 1, 449 in class 2, 103 in class 3 and 112 in class 4, this means that the model will make that number of predictions for each class, correctly in the positive class; while in the true negative (TN) quadrant, 665 predictions are shown in class 1, 504 in class 2, 850 in class 3 and 849 in class 4, this means that the model will make that number of predictions for each class, correctly in the negative class; On the other hand, in the false positive (FP) quadrant, 8 predictions are shown in class 1, 3 predictions in class 2, 7 predictions in class 3 and no predictions in class 4, this means that the model will perform that number of predictions for each class incorrectly in the positive class. Finally, in the false negative (FN) quadrant, 9 predictions are shown in class 1, 6 predictions in class 2, 2 predictions in class 3 and 1 prediction in class 4, this means that the model will perform that number of predictions for each class incorrectly in the negative class. In the following Table 5, the results of the sensitivity and specificity indicators are shown, through these indicators it will be possible to observe in percentage terms the successes and errors of the supervised learning algorithm when used in the prediction of the perception of teaching performance academic.

Table 3. Accuracy validation results

	Accuracy (Validation)
Weighted KNN	98.10%
Trilayered Neural Network	97.80%
Medium KNN	97.70%
Narrow Neural Network	97.70%
Fine Gaussian SVM	97.60%
Fine KNN	97.60%
Medium Neural Network	97.40%
Wide Neural Network	97.40%
Logistic Regression Kernel	97.30%

Table 4. Results of the quadrants of the confusion matrix

		Positive Classes		Negative Classes	
		TP	FP	TN	FN
Class 1	Regular	280	8	665	9
Class 2	Good	449	3	504	6
Class 3	Deficient	103	7	850	2
Class 4	Very Good	112	0	849	1

Table 5. Sensitivity and specificity matrix for class

Class 1	Regular	96.9%	0.70%	2.40%	0%
Class 2	Good	1.3%	98.7%	0%	0%
Class 3	Deficient	1.9%	0%	98.1%	0%
Class 4	Very Good	0%	0.90%	0%	99.1%
	TPR	96.9%	98.7%	98.1%	99.1%
	FNR	3.10%	1.30%	1.90%	0.90%

The results in Table 5 show a sensitivity that is represented by the true positive rate (TPR) of 96.9% in class 1, 98.7% in class 2, 98.1% in class 3 and 99.1% in class 4, this means that the predictive model has the ability to discriminate the positive classes from the negative ones, in the percentages described according to each class. While the specificity indicator, which is represented by the false negative rate (FNR), shows an error rate of 3.10% in class 1, 1.30% in class 2, 1.90% in class 3 and 0.90% in class 4, this represents the probability that the model is confounded when predicting a negative class when it is a positive class. In the following Table 6, the results of the precision indicator are shown, through this indicator it will be possible to know the capacity of the weighted KNN supervised learning algorithm to detect only relevant data.

The results in Table 6 show an accuracy that is represented by the positive predicted values (PPV) of 97.2% in class 1, 99.3% in class 2, 93.6% in class 3 and 100% in class 4, This means that the predictive model has the ability to predict the positive classes, in the percentages described according to each class. While the error rate, represented by the false discovery rate (FDR), shows 2.80% in class 1, 0.70% in class 2, 6.40% in class 3 and 0.0% in class 4. By validating optimal values of accuracy, sensitivity, specificity and precision in the 4 classes of the supervised learning model, it can be stated that the weighted KNN algorithm perfectly handles the prediction of the perception of university teaching performance. Having performed the specific analysis (4 classes), it is shown in the following Table 7, the average results of the performance indicators (accuracy, sensitivity, specificity and precision).

Table 6. Precision matrix for class

Class 1	Regular	97.2%	0.40%	6.40%	0%
Class 2	Good	2.10%	99.3%	0%	0%
Class 3	Deficient	0.70%	0%	93.6%	0%
Class 4	Regular	0%	0.20%	0%	100.0%
	PPV	97.2%	99.3%	93.6%	100.0%
	FDR	2.80%	0.70%	6.40%	0%

Table 7. General indicators of the weighted KNN algorithm

		Sensitivity	Specificity	Accuracy	Precision
Class 1	Regular	96.9%	98.8%	98.2%	97.2%
Class 2	Good	98.7%	99.4%	99.1%	99.3%
Class 3	Deficient	98.1%	99.2%	99.1%	93.6%
Class 4	Very Good	99.1%	100.0%	99.9%	100.0%
Total		97.9%	99.1%	98.8%	96.7%

The results in Table 7 show the general values of the performance indicators of the weighted KNN algorithm, where an average sensitivity of 97.9%, a specificity of 99.1%, an accuracy of 98.8% and a precision of 96.7% are observed. These percentages validate that the weighted KNN supervised learning algorithm has the ability to correctly identify the perception of university teaching performance among its 4 evaluation classes, correctly identifying positive and negative classes. Regarding the identification of the weighted k-nearest neighbor algorithm (weighted KNN), which indicates through the accuracy indicator that the predictive model will be 98.8% correct in the perception of university teaching performance, this result is similar and can be indicated which has better values than the one carried out in [31], where it is pointed out that the most efficient model to predict dropout in e-learning courses was with the K-NN algorithm, obtaining an accuracy of 94%.

Likewise, it is similar to the study carried out in [32] where it is indicated that the K-nearest neighbor algorithm presents the best prediction results of users in virtual education environments with an accuracy of 91%. Lee *et al.* [33] where the authors point out that an accuracy greater than 90, it is possible to optimally predict the graduation of the students. This statement is supported by [34], in this study the authors state that in supervised learning the precision of the algorithm depends on the accuracy, for this reason it is important that this indicator provides us with optimal values. In relation to the evaluation of the performance indicators through the confusion matrix where an average sensitivity of 97.9%, a specificity of 99.1% and precision of 96.7% are observed, which validate the capacity of the weighted KNN algorithm to correctly predict the perception of university teaching performance. Qu *et al*. [34] it is pointed out that an accuracy of 97.2% and a sensitivity of 96.5% validate the capacity of the optimal performance of the classifier algorithm. Similarly, Lee *et al*. [33], accuracy, precision and sensitivity values of 87.44%, 52.84% and 50.68%, respectively, were obtained, which will be applied in the prediction of students who graduate on time. As indicated in [35] a high percentage of sensitivity is important because it reflects the ability of the supervised learning model to predict positive classes and a high percentage of accuracy reflects the ability to distinguish only relevant data.

Next, the performance of the classification model is evaluated by means of the receiver operating characteristic curve (ROC) technique, taking into account that the closer the value of the area under the curve (AUC) is to 1, it can be indicated that the predictive model through the weighted KNN algorithm will have optimal performance when used in predicting the perception of university teaching performance. In Figure 2, the ROC curve for class 1 is shown. Figure 2 shows a true positive rate (TPR) of 97% and the probability of a false prediction of the regular performance of the university professor, represented by the false negative rate (FNR) of 1%, with an optimal value of the area under the 0.99 curve.

Figure 3 shows the ROC curve for class 2. As evidenced in the figure, the TPR is 99%. While the probability of a false prediction of good university teaching performance is represented by the FNR of 1%, with an optimal value of the low area of the curve of 1. Figure 4 shows the ROC curve for class 3. As can be seen, the TPR is 98%. While the probability of a false prediction of poor university teaching performance, represented by the FNR of 1%, with an optimal value of the area under the curve of 1.

Finally, Figure 5 shows the ROC curve for class 4. As evidenced in the figure, the TPR is 99%. While the probability of a false prediction of low university teaching performance, represented by the FNR is 0.0%, with an optimal value of the area under the curve of 1. In this way, the optimal performance of the supervised learning model through the weighted KNN algorithm to be used in predicting the perception of university teaching performance.

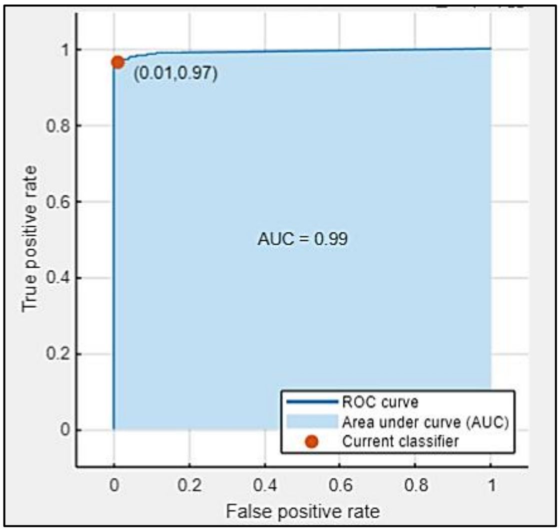

Figure 2. ROC curve for class 1 (regular)

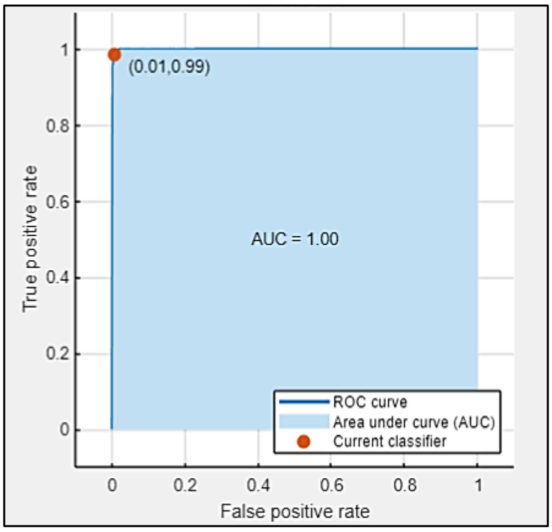

Figure 3. ROC curve for class 2 (good)

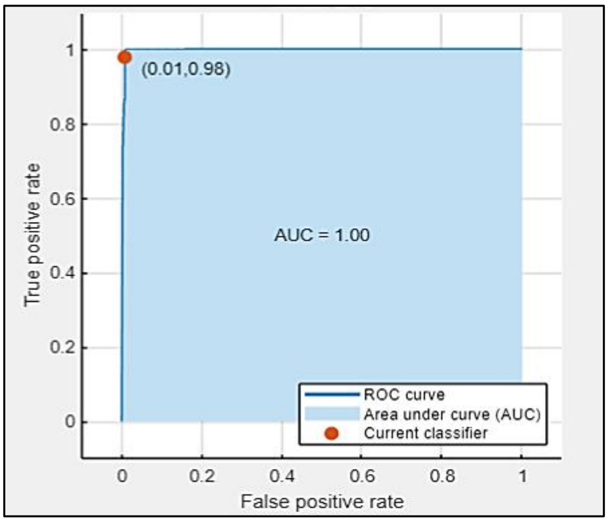

Figure 4. ROC curve for class 3 (deficient)

Analyzing the precision of the algorithm interpreting the AUC/ROC indicator, we can say that the results achieved are good enough to support the relevance of the use of the predictive model in the perception of teacher performance. In relation to the results of the Receiver Operating Characteristic Curve (ROC) where values of the area under the curve (AUC) equal to 0.99 and 1 are obtained, and a True Positive Rate (TPR) equal to 98% and 99%, with which the capacity that the model will have to distinguish between the 4 classes of the perception of university teaching performance can be validated, our study is supported by the one carried out in [36] where an area under the ROC curve of 0.805 and an area under the ROC curve of 0.805 were obtained accuracy of 75.42% in the prediction of academic risks in engineering careers, stating that the results achieved are good enough to support the relevance of the use of models in prediction; Likewise, in [37], AUC values equal to 0.785 and 0.833 were obtained, with a TPR of 88%, affirming an optimal prediction of students with low satisfaction in the virtual learning environment. As indicated in [36] the results of the classification and forecasting process when using the KNN algorithm, represents benefits in the ease of interpretation and comparability of the results, contextualizing this result in the university environment, for a teacher it is important to know the different ways in which a student relates to the educational process and much better if they can take actions in a timely manner to facilitate student learning and improve their satisfaction, re-conceptualizing the role of the teacher in terms of managing virtual environments Learning.

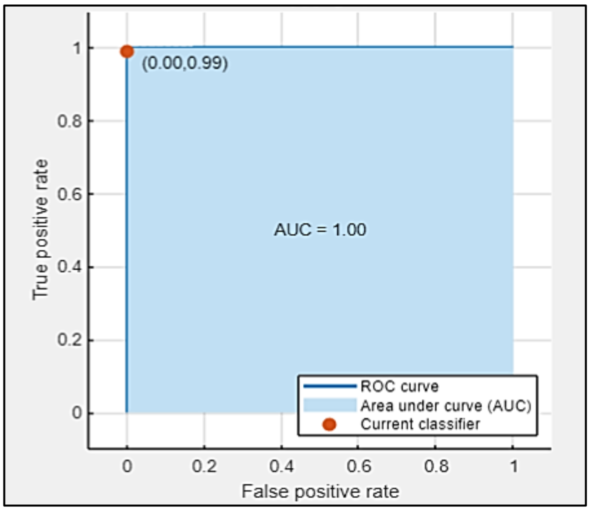

Figure 5. ROC curve for class 4 (very good)

5. CONCLUSION

We can affirm that the supervised learning model through the weighted KNN algorithm will fit properly and will be able to make a good prediction on the test data set in terms of general accuracy, as well as the accuracy of the 4 performance perception classes university teacher (class 1: fair, class 2: good, class 3: deficient, and class 4: very good). The prediction of the weighted KNN algorithm has an accuracy of 98.8% and a precision of 96.7%, thus validating the ability of the model to correctly predict and distinguish between the 4 classes of the perception of university teaching performance, likewise its implementation will allow obtaining information in real time for decision-making of university academic management, since the University under analysis continues to use traditional mechanisms such as surveys that are carried out at the end of the cycle, however, this algorithm will allow decisions to be made in a timely manner and not when the academic year ends. As future work, it is recommended to expand the line of research, by determining models of the other professional careers of the higher institutions, which must in turn cover the 10 academic cycles, this analysis is relevant, because the satisfaction of the student university, is directly related to the performance and professional skills acquired during their stay at the university; likewise, it is suggested to extend the line of research in other careers and for the 10 cycles, of the same, also carrying out; In the same way, a model can be identified to detect students in a situation of academic risk in real time, which will give the opportunity to carry out educational interventions in a timely manner that will reduce the problem of poor academic performance.

ACKNOWLEDGEMENTS

Thanks and appreciation to the “Universidad Nacional Tecnológica de Lima Sur” and the School of Mechanical and Electrical Engineering.

REFERENCES

[1] S. Jia and Y. Pang, “Teaching quality evaluation and scheme prediction model based on improved decision tree algorithm,” *International Journal of Emerging Technologies in Learning*, vol. 13, no. 10, pp. 146–157, Oct. 2018, doi: 10.3991/ijet.v13i10.9460.

[2] E. B. Canon, J. E. Q. Salamanca, and E. B. Canon, “Predictive model of the learning progress of uniminuto students applying machine learning techniques,” *Revista Conrado*, vol. 17, no. 83, pp. 305–310, 2021.

[3] A. Muniasamy and A. Alasiry, "Deep learning: The impact on future eLearning," *International Journal of Emerging Technologies in Learning*, vol. 15, no. 1, pp. 188–199, Jan. 2020, doi: 10.3991/IJET.V15I01.11435.
[4] W. Liu, "An improved back-propagation neural network for the prediction of college students' English performance," *International Journal of Emerging Technologies in Learning*, vol. 14, no. 16, pp. 130–142, Aug. 2019, doi: 10.3991/ijet.v14i16.11187.
[5] S. Wang, D. Liu, N. Wang, and Y. Yuant, "Design and implementation of an online python teaching case library for the training of application oriented talents," *International Journal of Emerging Technologies in Learning*, vol. 15, no. 21, pp. 217–230, Nov. 2020, doi: 10.3991/IJET.V15I21.18191.
[6] Y. Fan, Y. Liu, H. Chen, and J. Ma, "Data mining-based design and implementation of college physical education performance management and analysis system," *International Journal of Emerging Technologies in Learning*, vol. 14, no. 6, pp. 87–97, Mar. 2019, doi: 10.3991/ijet.v14i06.10159.
[7] A. Abdelhadi, S. Zainudin, and N. S. Sani, "A regression model to predict key performance indicators in higher education enrollments," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 1, pp. 454–460, 2022, doi: 10.14569/IJACSA.2022.0130156.
[8] F. F. Hérnandez, M. S. Mendiola, and A. M. González, "Prediction model of the academic performance of the students of the basic cycle of the medicine career from the evaluation of the teaching performance," *Revista Mexicana de Investigación Educativa*, vol. 21, no. 70, pp. 975–991, 2016.
[9] E. A. Gamie, M. S. A. El-Seoud, and M. A. Salama, "Comparative analysis for boosting classifiers in the context of higher education," *International Journal of Emerging Technologies in Learning*, vol. 15, no. 10, pp. 16–26, 2020, doi: 10.3991/ijet.v15i10.13663.
[10] N. Priyasadie and S. M. Isa, "Educational Data mining in predicting student final grades on standardized indonesia data pokok pendidikan data set," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 12, pp. 212–216, 2021, doi: 10.14569/IJACSA.2021.0121227.
[11] A. R. Páez and D. S. Guzmán, "Diseño de un modelo para automatizar la predicción del rendimiento académico en estudiantes del IPN/Design of a model to automate the prediction of academic performance in students of IPN," *RIDE Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, vol. 8, no. 16, pp. 246–266, Feb. 2018, doi: 10.23913/ride.v8i16.340.
[12] I. M. Gómez-Trigueros, "Digital teaching competence and space competence with TPACK in social sciences," *International Journal of Emerging Technologies in Learning*, vol. 15, no. 19, pp. 37–52, Oct. 2020, doi: 10.3991/ijet.v15i19.14923.
[13] N. A. Atan, Z. Tasir, M. F. Ali, M. S. Rosli, and M. N. H. M. Said, "Students' performance and perceptions towards authentic learning environment based on visual thinking continuum," *International Journal of Emerging Technologies in Learning*, vol. 16, no. 23, pp. 56–73, Dec. 2021, doi: 10.3991/ijet.v16i23.27455.
[14] P. J. A. L. Rose and T. A. C. Mary, "An early intervention technique for at-risk prediction of higher education students in cloud-based virtual learning environment using classification algorithms during COVID-19," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 1, pp. 612–621, 2022, doi: 10.14569/IJACSA.2022.0130174.
[15] L. F. M. Siu, J. L. Rodriguez, R. E. R. Salazar, "Soft skills and teacher performance at the higher level of education," *Propósitos y Representaciones*, vol. 9, no. 1, p. 1, 2021, [Online]. Available: http://www.scielo.org.pe/pdf/pyr/v9n1/2310-4635-pyr-9-01-e1038.pdf.
[16] G. Picón, G. Gonzáles, and J. Paredes, "Performance and teacher training in digital skills in non-face-to-face classes during the COVID-19 pandemic," *Revista Científica Internacional*, vol. Vol. VIII, Número 1, no. ISSN 2311-7559, 2021.
[17] W. Wuryaningsih, D. H. Susilastuti, M. Darwin, and A. C. Pierewan, "Effects of web-based learning and F2F learning on teachers achievement in teacher training program in Indonesia,", *International Journal of Emerging Technologies in Learning (IJET)* vol. 14, no. 21, p. 123, Nov. 2019, doi: 10.3991/ijet.v14i21.10736.
[18] H. Wang and J. Cui, "Evaluation of Teaching effect in higher educational institutions and identification of its influencing factors," *International Journal of Emerging Technologies in Learning*, vol. 16, no. 17, pp. 226–239, Sep. 2021, doi: 10.3991/ijet.v16i17.24901.
[19] E. A. R. Barrantes, "Análisis cuantitativo de la evaluación del desempeño docente en la Universidad de Costa Rica," *Actualidades Investigativas en Educación*, vol. 21, no. 1, pp. 1–27, Jan. 2021, doi: 10.15517/aie.v21i1.42425.
[20] M. Wang, Z. Zhang, and M. Xu, "Postgraduate training performance evaluation applying weighting and comprehensive fuzzy evaluation methods," *International Journal of Emerging Technologies in Learning*, vol. 16, no. 1, pp. 124–133, Jan. 2021, doi: 10.3991/IJET.V16I01.18231.
[21] E. G. Suarez and R. M. Toro, "Evaluación del desempeño docente: Preparación para el aprendizaje de los estudiantes en el Marco de Buen Desempeño Docente," *Propósitos y Representaciones*, vol. 6, no. 2, p. 407, Aug. 2018, doi: 10.20511/pyr2018.v6n2.236.
[22] N. Kanya, A. B. Fathoni, and Z. Ramdani, "Factors affecting teacher performance," *International Journal of Evaluation and Research in Education (IJERE)*, vol. 10, no. 4, p. 1462, Dec. 2021, doi: 10.11591/ijere.v10i4.21693.
[23] C. J. C. Aquino, B. T. Afalla, and F. L. Fabelico, "Managing educational institutions: School heads' leadership practices and teachers' performance," *International Journal of Evaluation and Research in Education*, vol. 10, no. 4, pp. 1325–1338, Dec. 2021, doi: 10.11591/IJERE.V10I4.21518.
[24] S. C. Suh, B. N. A. Upadhyaya, and N. V. A. Nadig, "Analyzing personality traits and external factors for stem education awareness using machine learning," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 5, pp. 1–4, 2019, doi: 10.14569/ijacsa.2019.0100501.
[25] P. Sokkhey and T. Okazaki, "Study on dominant factor for academic performance prediction using feature selection methods," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 8, pp. 492–502, 2020, doi: 10.14569/IJACSA.2020.0110862.
[26] Y. An, M. Xu, and C. Shen, "Classification method of teaching resources based on improved KNN algorithm," *International Journal of Emerging Technologies in Learning (IJET)*, vol. 14, no. 04, p. 73, Feb. 2019, doi: 10.3991/ijet.v14i04.10131.
[27] Y. Lou, "Storage and allocation of english teaching resources based on k-nearest neighbor algorithm," *International Journal of Emerging Technologies in Learning*, vol. 14, no. 17, pp. 102–113, Sep. 2019, doi: 10.3991/ijet.v14i17.11188.
[28] A. Ashari, I. Paryudi, and A. Min, "Performance comparison between naïve bayes, decision tree and k-nearest neighbor in searching alternative design in an energy simulation tool," *International Journal of Advanced Computer Science and Applications*, vol. 4, no. 11, 2013, doi: 10.14569/ijacsa.2013.041105.
[29] K. Calva, M. Flores, H. Porras, and A. C. Martínez, "Academic performance prediction model for the propedeutic course of the Escuela Pólitecnica Nacional and the implementation of an automated supervised learning model," *Latin-American journal of Computing (LAJC)*, vol. 8, no. 2, pp. 60–71, 2021.

[30] O. C. Atalaya, G. M. Romero, N. T. Loli, B. C. Salas, S. G. Mendoza, and C. L. Velarde, "Evaluation of teaching performance in the virtual teaching-learning environment, from the perspective of the students of the professional school of mechanical engineering," *International Journal of Emerging Technologies in Learning*, vol. 16, no. 15, pp. 244–252, Aug. 2021, doi: 10.3991/ijet.v16i15.23091.

[31] I. Urteaga, L. Siri, and G. Garófalo, "Predicción temprana de deserción mediante aprendizaje automático en cursos profesionales en línea," *RIED. Revista Iberoamericana de Educación a Distancia*, vol. 23, no. 2, p. 147, Jul. 2020, doi: 10.5944/ried.23.2.26356.

[32] H. Yuliansyah, R. A. P. Imaniati, A. Wirasto, and M. Wibowo, "Predicting students graduate on time using C4.5 algorithm," *Journal of Information Systems Engineering and Business Intelligence*, vol. 7, no. 1, p. 67, Apr. 2021, doi: 10.20473/jisebi.7.1.67-73.

[33] C. Y. Lee *et al.*, "Study on the university students' satisfaction of the wisdom tree massive open online course platform based on parameter optimization intelligent algorithm," *Science Progress*, vol. 104, no. 3_suppl, p. 003685042110542, Jul. 2021, doi: 10.1177/00368504211054256.

[34] S. Qu, K. Li, B. Wu, X. Zhang, and K. Zhu, "Predicting student performance and deficiency in mastering knowledge points in MOOCs using multi-task learning," *Entropy*, vol. 21, no. 12, p. 1216, Dec. 2019, doi: 10.3390/e21121216.

[35] O. M. B. Herrera, M. L. Trujillo, and C. E. M. Echeverry, "Modelo predictivo para la identificación de factores socioculturales asociados al tiempo de búsqueda del primer empleo en egresados universitarios," *Revista Virtual Universidad Católica del Norte*, no. 58, pp. 3–18, Oct. 2019, doi: 10.35575/rvucn.n58a6.

[36] E. J. D. L. Hoz, E. J. D. L. 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, Feb. 2019, doi: 10.4067/S0718-07642019000100247.

[37] E. A. Franco, R. E. L. Martínez, and V. H. M. Domínguez, "Modelos predictivos de riesgo académico en carreras de computación con minería de datos educativos," *Revista de Educación a Distancia (RED)*, vol. 21, no. 66, Apr. 2021, doi: 10.6018/red.463561.

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.

Nestor Alvarado-Bravo is a research professor at the National University of Callao in Peru, graduated in Psychology; He teach at the university from 1998 to the present, at the Faculty of Chemical Engineering. He is a Master in Clinical Psychology and a Doctor in Psychology at the university level she carry out research on issues of Psychology, University Education and Management. He is a thesis advisor. He carry out consulting and external training for small and micro companies. He can be contacted at email: nmalvaradob@unac.edu.pe.

Florcita Aldana-Trejo is a research professor at the Federico Villarreal National University of Peru. She has a degree in economics. She teaches at the university from 1995 to the present, at the Faculty of Economic Sciences. She is a Master in Research and University Teaching and a Doctor in Administration. At the university level. She carries out research on economics, university education and management. She is a thesis advisor and out consulting and external training for small and micro companies. She can be contacted at email: faldana@unfv.edu.pe.

Claudia Poma-Garcia is a Doctor of Education, Cesar Vallejo University (UCV) in Lima-Peru. Master in Public Management Cesar Vallejo University. Economist Graduated from the Federico Villarreal National University (UNFV). College of Economists of Lima. She has studied a Diploma in Foreign Trade Operations, Safety and Health at Work. She currently works as a University Professor at the Cesar Vallejo University, and at the National University of Callao. She can be contacted at email: cpomagar@ucvvirtual.edu.pe.

Carlos Aliaga-Valdez is a Doctor of Administration. Master in Strategic Business Management. MBA from the Université du Québec à Montréal. Graduated from the Master's Degree in Financial Sciences from the Federico Villarreal University. Specialization in Finance at the School of Business Administration (ESAN). Degree in Administrative Sciences from the National University of San Marcos; chartered professional. In his professional activity, he has worked as an official in the areas of Personnel, Logistics, Banking Operations and Finance in banking and industrial companies. He has participated in the development of business strengthening and improvement projects financed by the Inter-American Development Bank. He can be contacted at email: craliagav@unac.edu.pe.

Gutember Perarlta-Eugenio is a Psychologist with a master's degree in psychological intervention from the César Vallejo University. Research professor who teaches curricular experiences linked to scientific research at the César Vallejo University and the Catholic University of Trujillo. External advisor in the preparation and development of research projects in social sciences. He can be contacted at email: gperaltae@ucvvirtual.edu.pe.

Abel Tasayco-Jala is a senior lecturer at the Cesar Vallejo University, Peru in the subjects of research, thesis and topics related to administration. As well as making scientific publications on human resources, knowledge management and educational issues. He can be contacted at email: atasaycoj@ucvvirtual.edu.pe.