Quadratic vector support machine algorithm, applied to prediction of university student satisfaction

Omar Chamorro-Atalaya[1], Guillermo Morales-Romero[2], Yeferzon Meza-Chaupis[2], Elizabeth Auqui-Ramos[2], Jesús Ramos-Cruz[2], César León-Velarde[3], Irma Aybar-Bellido[4]
[1]Facultad de Ingeniería y Gestión, Universidad Nacional Tecnológica de Lima Sur, Lima, Perú
[2]Facultad de Ciencias, Universidad Nacional Enrique Guzmán y Valle, Lima, Perú
[3]Facultad de Humanidades, Universidad Tecnológica del Perú, Lima, Perú
[4]Facultad de Ciencias de la Educación y Humanidades, Universidad Nacional San Luis Gonzaga, Lima, Perú

Article Info

Article history:

Received Mar 19, 2022
Revised May 5, 2022
Accepted May 16, 2022

Keywords:

Algorithm
Prediction
Satisfaction
University students
Vector support machine

ABSTRACT

This study aims to identify the most optimal supervised learning algorithm to be applied to the prediction of satisfaction of university students. In this study, the IBM SPSS-25.0 software was used to test the reliability of the satisfaction questionnaire and the MATLAB R2021b software through the classification learner technique to determine the supervised learning algorithm. The experimental results determine a Cronbach's Alpha reliability of 0.979, in terms of the classification algorithm, it is validated that the quadratic vector support machine (SVM) has better performance metrics, being correct in 97.8% (accuracy) in the predictions of satisfaction of university students, with a recall (sensitivity) of 96.5% and an F1 score of 0.968. Likewise, when evaluating the classification model by means of the receiver operating characteristic curve (ROC) technique, it is identified that for the three expected classes of satisfaction the value of the area under the curve (AUC) is equal to 1, in such sense the predictive model through the SVM Quadratic algorithm, has a high capacity to distinguish between the 3 classes; i) dissatisfied, ii) satisfied and iii) very satisfied of satisfaction of university students.

Corresponding Author:

Omar Chamorro-Atalaya
Facultad de Ingeniería y Gestión, Universidad Nacional Tecnológica de Lima Sur
Lima, Perú
Email: ochamorro@untels.edu.pe

1. INTRODUCTION

Today there is a growing discussion about the areas in which artificial intelligence can be used, and although it is true that it is largely applied in industrial and technological environments, it is necessary to analyze its application in different scenarios of modern societies [1]. Artificial intelligence uses machine learning techniques to process large volumes of data, generated at every moment, by the simple fact of carrying out daily activities, as is the case in the university education sector [2]. In recent years, the analysis of data from university education has turned out to be quite limited, which has led to the lack of organizational policies that contribute to decision-making, improving educational quality [3], [4].

It is essential to safeguard the quality of educational processes, and a factor linked to this is university student satisfaction [5], [6]. Therefore, universities should be concerned about the educational quality they provide, evaluating indicators as part of a process of continuous improvement [7], [8]. However, it is important to bear in mind that in the face of the health emergency generated by COVID-19, academic activities have been moved to a virtual context, and rather than return to full attendance, today we are immersed in hybrid educational processes, in which quality and satisfaction need to be constantly measured [9], [10].

In this scenario, data science allows the modeling of the behavior of users involved in the educational process [11], classrooms in a virtual environment make continuous interaction of students with technological resources, as well as with social networks, generating a large amount of data [12], in which technology should contribute to accelerating the tracking, monitoring, storage and processing of the perception or assessment of their satisfaction [13], [14]. The search for indicators or factors linked to the prediction of university student satisfaction has given rise to various studies, however, traditional empirical methods have always been used [15]; However, today there is a tendency to use data mining and machine learning techniques that allow significant knowledge to be extracted, which can contribute to decision making [16]-[18]. Likewise, data mining presents relevant functions such as grouping, classification and identification of association patterns, defining through algorithms the prediction of certain variables such as student satisfaction, with a high degree of precision [19], [20].

There are several algorithms that help to determine the predictive model, among which we have K-nearest neighbor (K-NN), decision tree (DT), random forest (RF) and support vector machines (SVM) [21], [22]. In this regard, in [23] it is established that the SVM algorithm is a powerful method to build classifiers, thus allowing the prediction of one or more characteristics of vectors, which is called a hyperplane. Thus, SVM also responds to the statistical learning theory that consists of determining the best hyperplane that is equidistant from the closest one, to achieve a maximum margin on each side of the hyperplane, separating the members of the class from the variable under analysis [24]-[26]. In this sense, the purpose of this article is to show the performance metrics of the different algorithms possibly to be used in the prediction of university student satisfaction. Likewise, through a comparative analysis, the algorithm with the best performance will be determined.

2. METHOD

2.1. Population and sample

The population under analysis is made up of students in the condition of regular students and who are studying from the VII to the X cycle, during two consecutive academic semesters, of the 5 faculties of the National Technological University of Lima Sur, located in the district of Villa El Salvador in Peru; this population is made up of 869 students. With respect to the study sample, the survey was applied to the entire population, obtaining as a result that in the first semester there were 761 students surveyed and in the second semester 715 students, representing 87.57% and 82.28% respectively, thus achieving have a representative sample. However, in the case of determining the predictive model, the results of the survey of the two academic semesters were used, that is, of 1,476 students.

2.2. Level and research design

The level of research is predictive, this is due to the fact that it seeks to determine the algorithm that presents the best performance to predict university student satisfaction from the 39 indicators that are part of the questionnaire, which are also called predictors. The research design consists of a non-experimental type, this is due to the fact that no action is exerted on the population prior to the application of the survey, which alters their perception. Figure 1 shows the representation of the process used to obtain the predictive algorithm of the variable under study.

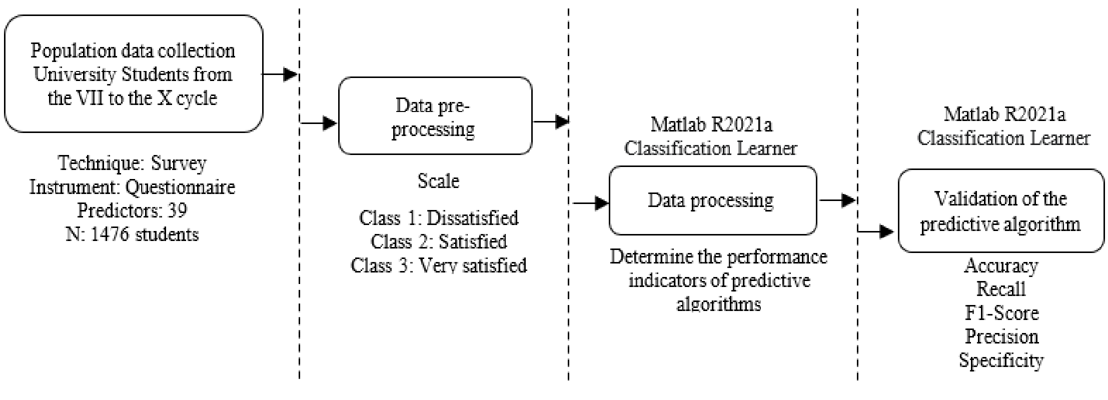

Figure 1. Process used to obtain the predictive algorithm

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

The technique used in this research was the survey, which was carried out online, through the University's website; this survey used as an instrument a questionnaire composed of 6 dimensions and with a total of 39 questions. For the coding of the answers, the Likert scale was used, ranging from 1 (dissatisfied), 2 (somewhat satisfied), 3 (satisfied) and 4 (very satisfied). Table 1 shows the data collection instrument, as well as the 39 questions and the validity of each component through Cronbach's Alpha, contained using the SPSS V25 software.

It should be noted that the output variable (target) is the student satisfaction of each student, and the predictive elements are the 39 indicators; Thus, for the purposes of defining the classes that the predictive algorithm will assign to the variable under study, the Scale was used, using the 30% and 70% percentiles in such a way that the target classes are redefined in 3 classes, these being "Class 1: dissatisfied", "Class 2: Satisfied" and "Class 3: Very satisfied". Table 2 shows the statistics obtained from the SPSS V.25 software, to obtain the new classification of the levels of the output variable. Table 3 shows the Scale obtained, in which the classes for the variable under study are specified.

Table 1. 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
I1	To work as a team	0.978	0.979
I2	To solve problems and cases of the specialty	0.978	
I3	To act with autonomy and initiative	0.978	
I4	To compare own ideas with others	0.978	
I5	To speak in public with appropriate language	0.978	
I6	To have a positive attitude towards change and innovation	0.978	
I7	Assume a self-education (self-learning and continuing education)	0.978	
I8	To master practical professional skills	0.978	
I9	To work under pressure	0.978	
I10	To have investigative skills	0.978	
I11	Respect schedules; does not miss virtual class without notice	0.978	
I12	Their mastery of the subjects of the courses they develop	0.978	
I13	Your teaching methodology	0.978	
I14	His firmness so that students respect the university rules	0.978	
I15	The quality of virtual classes, practices and other types of academic activity	0.978	
I16	The treatment of students during the virtual class	0.978	
I17	Impartiality in grading students	0.978	
I18	Oral and written expression that you have in the virtual class	0.978	
I19	Your identification with the institution	0.978	
I20	Your professional suitability	0.978	
I21	Availability of books of your specialty on the university website	0.978	
I22	Bibliographic information search system, on the university website	0.978	
I23	Availability of virtual library	0.978	
I24	The efficiency of the work of administrative staff via remote	0.978	
I25	Quality of attention of the administrative staff via remote	0.978	
I26	The treatment of the student is cordial and timely	0.978	
I27	The information provided to the student is pertinent	0.978	
I28	Psychopedagogical services	0.978	
I29	Quality of care in the health unit for students	0.978	
I30	Welfare university	0.978	
I31	Registrations and license plates	0.978	
I32	Assume studies with responsibility, seriousness and dedication	0.978	
I33	With the pride of belonging to the university	0.978	
I34	With the commitment to leave the name of the university high	0.978	
I35	With the respect you show for the authorities, teachers and administrative staff	0.978	
I36	With the respect that you treat your colleagues	0.978	
I37	With the treatment you receive from your colleagues	0.978	
I38	With your interest in being better every day	0.978	
I39	With your commitment to the surrounding society	0.978	

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

N (sample)		Minimum	Maximum	Percentiles	
Valid	1476	39.00	156.00	30%	104.00
Lost	0			70%	117.00

Table 3. Scale for the student satisfaction variable

Interval	Class
39 to 104	Dissatisfied
105 to 117	Satisfied
118 to 156	Very satisfied

3. RESULTS AND DISCUSSION

Through the Matlab software, the classification learner technique is used, with which the validation of the supervised learning algorithm that will be applied to the prediction of the satisfaction of university students is carried out. This type of validation measures the quality of the data grouping, that is, it refers to how close the result of a measurement is to the true value. Figure 2 shows the most optimal algorithms to be applied on the data.

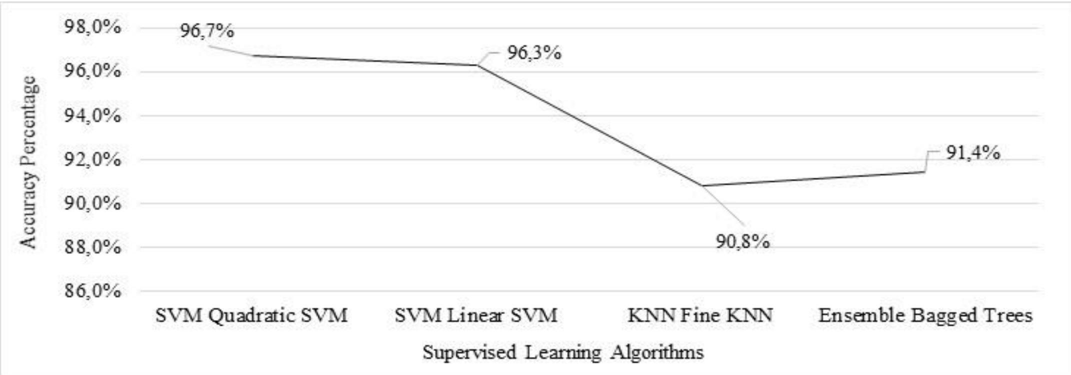

Figure 2. Algorithm accuracy validation

As shown in Figure 2, the most optimal algorithm regarding the internal validation of accuracy is the quadratic vector support machine (SVM Quadratic) algorithm, which has an accuracy of 96.7%, that is, applying the SVM Quadratic algorithm, the model was 96.7% correct in predicting the satisfaction of university students. Regarding the identification of the quadratic vector support machine algorithm as the one that presents the most optimal accuracy and its choice for the prediction of satisfaction of university students, in [27], the frequent use of machine learning techniques for mining is highlighted of educational data, including the decision to use the support vector machine (SVM) algorithm, therefore, it can be indicated that the use of machine learning algorithms can be a great alternative to solve educational research problems. Although the results provide us with an optimal percentage of accuracy of the SVM Quadratic algorithm, it is important to evaluate the performance of the other metrics of this trained model (recall, specificity, accuracy, precision, recall, F1-score), according to the classes foreseen for the predictive model (class 1: dissatisfied, class 2: satisfied and class 3: very satisfied) and in a general way. Thus, in order to further deepen this analysis and validate the algorithm to be used with greater support, the results of the confusion matrix are presented, which, through its external validation metrics, will allow us to observe the percentage of successes and errors of the algorithms classified as the most optimal when going through the learning process on the data.

Initially, the recall or Sensitivity metric will be analyzed, which is the percentage of positive cases detected and is represented by the true positive rate (TPR), while the Specificity metric is the percentage of negative cases detected, in this case it is represented by the rate of false negatives (FNR). Note that true positives (TP) are data points classified as true by the model that are actually positive (meaning they are correct), and false negatives (FN) are data points that the model identifies as true negatives that are actually positive (meaning they are wrong). Table 4 shows the results of the sensitivity and specificity metrics, which reflect the ability of our algorithm to discriminate positive cases from negative ones.

Table 4. Comparative recall and especificidad results for each class

	SVM Quadratic SVM	SVM Linear SVM	KNN Fine KNN	Ensemble Bagged Trees
Class 1	96.9%	95.1%	87.9%	91.3%
Class 2	98.1%	98.7%	93.6%	91.8%
Class 3	94.4%	93.6%	89.5%	90.8%

The results of Table 4 show that the SVM quadratic algorithm presents better recall and specificity values compared to the other algorithms, highlighting that, of its 3 classes, class 2 (satisfaction level: satisfied) shows the highest percentage of sensitivity, this means that the predictive model has the ability to discriminate positive cases from negative ones, in 98.1%; In other words, the algorithm has a very low error rate of 1.9%, in which 0.5% can be mistaken in predicting that a student is dissatisfied when in fact he or she is satisfied, and 1.4% can be mistaken in predicting that a student is very satisfied when in reality he is only satisfied. Next, the precision metric will be analyzed, which is the percentage of positive predictions detected, that is, in the capacity of a classification algorithm to detect only relevant data, with respect to this metric we must take into account that the lower the dispersion, the greater the precision of the algorithm. The results are shown in Table 5.

Table 5. Comparative precision results for each class

	SVM quadratic SVM	SVM linear SVM	KNN fine KNN	Ensemble bagged trees
Class 1	99.3%	99.5%	97.3%	96.7%
Class 2	94.6%	93.0%	86.3%	88.6%
Class 3	97.6%	98.4%	92.1%	90.4%

The results of Table 5 show that the SVM quadratic algorithm presents better accuracy values compared to the other machine learning algorithms, highlighting that, of its 3 classes, class 1 (satisfaction level: dissatisfied) shows the highest percentage high positive predicted values (PPV), that is, 99.3% of the values actually have positive polarities, while the error rate, represented by the false discovery rate (FDR), is 0.7%. Having obtained the most optimal recall and precision metrics in the SVM quadratic algorithm so far, it can be pointed out that the chosen machine learning model perfectly handles the prediction in the 3 foreseen classes. Given this, we proceed to determine the accuracy values for the three classes under analysis. Table 6 shows the percentages of correct positive predictions.

Table 6. Comparative accuracy results for each class

	SVM Quadratic SVM	SVM Linear SVM	KNN Fine KNN	Ensemble Bagged Trees
Class 1	98.8%	98.4%	95.6%	96.4%
Class 2	96.7%	96.3%	90.8%	91.4%
Class 3	97.9%	97.9%	95.2%	95.0%

The results of Table 6 show that the SVM quadratic algorithm presents better accuracy metrics in its three expected classes, highlighting that, of its 3 classes, class 1 (level of satisfaction: dissatisfied) shows the highest percentage of accuracy, that is, 98.8% of the time the SVM quadratic algorithm was correct in predicting the dissatisfaction of university students. Obtaining the values of the metrics according to each class and comparing the most optimal algorithms, Figure 3 shows the general percentages of all the performance metrics analyzed for the trained model (recall, specificity, accuracy, precision, recall, and F1-score).

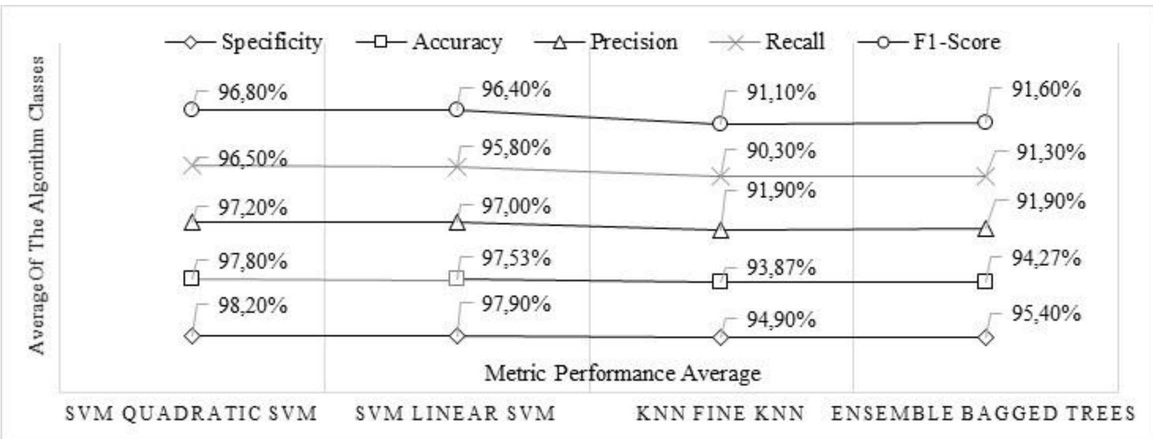

Figure 3. General performance metric

In general, it can be shown that the SVM quadratic algorithm presents the most optimal metrics for the predictive model of the satisfaction of university students, regarding specificity, the SVM quadratic algorithm will have 98.2% the ability to identify the satisfaction of the students between its 3 levels; regarding the accuracy, it can be indicated that the SVM quadratic algorithm will have a 97.8% capacity to determine correct positive predictions; Regarding accuracy, the SVM quadratic algorithm will have a 97.2% ability to identify positive predictions regarding student satisfaction levels; Regarding sensitivity (recall), the SVM quadratic algorithm will have 96.5% the ability to correctly detect the level of satisfaction among students. Likewise, once the sensitivity (recall) and precision metrics have been obtained, the F1 score is determined, which is calculated as the weighted average of both metrics, where F1 reaches its best performance when the score is 1, as shown in Figure 2, the F1 score is equal to 0.968 (96.8%), with these results the optimal performance of the predictive model is supported through the SVM quadratic algorithm.

In relation to the results obtained, which indicate that the SVM quadratic algorithm presents an Accuracy of 97.8%, which will have the capacity to determine correct positive predictions of the satisfaction of university students, this is similar to the study carried out in [21] where it is pointed out that the experimental results show that the accuracy of the classification of the training set through the support vector machine algorithm has an accuracy of 99.58%, which provides great reliability when applied to the satisfaction of students with the online course platform. Similarly, in [28] where a predictive model is proposed through the SVM algorithm, an accuracy of 92.18% is validated, in this way the classification accuracy is supported, obtaining a more effective machine learning algorithm to predict customer satisfaction university students. This can be answered in the study carried out by [21], where it is stated that in machine learning and data mining the accuracy of the algorithm depends on the accuracy of the data classification, for this reason it is important that this metric provides us with optimal values. In addition, the study carried out in [19] supports the statements made, because it is pointed out that, by applying the predictive model, the graduation of students has been successfully predicted, this statement is made possible by obtaining an accuracy of 90%.

Likewise, regarding the results obtained from the precision, sensitivity (recall) and F1 score metrics, with values of 97.2%, 96.5% and 96.8% respectively, it can be indicated that the classification model of the proposed algorithm has good robustness and is equal to as good as the model carried out in [21], where values of the metrics accuracy, sensitivity (recall) and F1 score of 97.16%, 96.45%, 96.80%, respectively, are obtained. In turn, our results show greater robustness than those obtained in [19] in which it was possible to obtain values of accuracy, precision, recall and F1 score of 87.44%, 52.84%, 50.68% and 51.73% respectively, to be applied in predicting students graduating on time. This affirmation can be sustained in the study of [29] where it is pointed out that these metrics reflect a high capacity of the classification model, because, the greater the recall, the greater the capacity of the model to recognize positive instances, the greater the Whatever the accuracy, the capacity of the model to distinguish instances will be reflected, finding the F1 score as the combination of the two metrics, in this sense, the higher the F1 score, the more solid the classification model will be. Validated the use of the SVM Quadratic algorithm in the predictive model of the satisfaction of university students, the classification model is evaluated by means of the receiver operating characteristic curve (ROC) technique, which allows us to visualize the balance between the rate of true positives (TPR) and the false negative rate (FNR). Figure 4 shows the receiver operating characteristic (ROC) curve for class 1, representing dissatisfied college students.

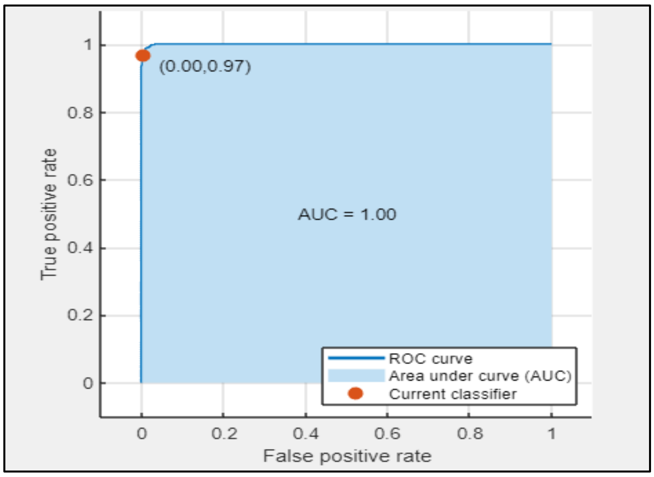

Figure 4. ROC curve for class dissatisfied

Regarding the receiver operating characteristic curve (ROC), it should be taken into account that the closer the value of the area under the curve (AUC) is to 1, this will represent a more optimal performance of the predictive model through the SVM quadratic algorithm. As shown in Figure 4, the percentage of the rate of true positives (sensitivity) is 97% and the probability of a false prediction (specificity) in the class of dissatisfied university students is 0%, this is supported by a value of the area under the curve of 1. Similarly, Figure 5 shows the receiver operating characteristic (ROC) curve for class 2, representing satisfied college students. As can be seen in Figure 5, the general performance of the classification model has an area under the curve value of 1, with the percentage of the true positive rate being 98% and the probability of a false prediction in the class of students satisfied university students of 4%. Finally, in Figure 6, the receiver operating characteristic (ROC) curve is shown for class 3, representing highly satisfied college students.

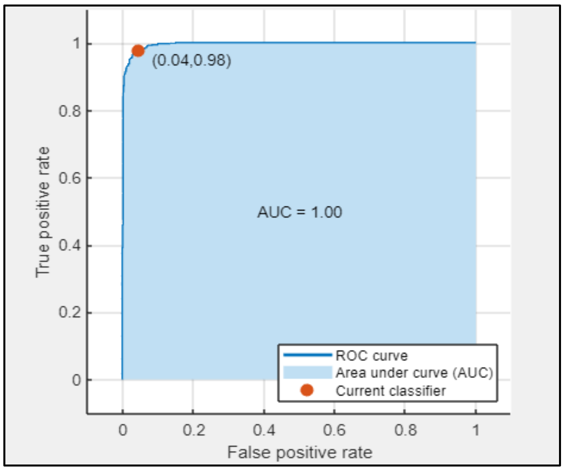

Figure 5. ROC curve for class satisfied

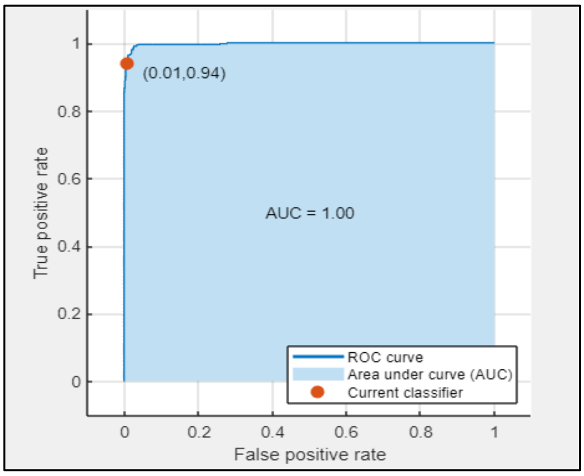

Figure 6. ROC curve for class very satisfied

As seen in Figure 6, the percentage of the rate of true positives is 94% and the percentage of the rate of false negatives or error rate in the class of very satisfied university students is 1%, this is supported by a value of the area under the curve of 1, at the end of the evaluation of the predictive model through the SVM Quadratic algorithm is validated and supports its optimal performance when applied to the satisfaction of university students. Regarding the results obtained from an area under the curve (AUC) value of 1, the present study shows better metrics obtained than the study carried out in [30], in which an area under the curve value equal to 0.785 was obtained and 0.833, with an accuracy of 88.11% in identifying the group of students with low satisfaction with online education. In addition, in relation to the satisfaction of university students, the cited study reveals that the two greatest predictors are the component of self-efficacyexpectations regarding the online modality and the social dimension of the perception of the online experience. Regarding recommendations in [1], it is suggested to provide courses, learning activities, methods to improve the learning experience and maximize the satisfaction of university students.

4. CONCLUSION

The findings of this study validate the performance of the metrics of the quadratic support vector machine (SVM quadratic) algorithm when applied to the prediction of satisfaction of university students, being correct in 97.8% (Accuracy) in the predictions, with a recall (sensitivity) of 96.5% and an F1 score of 0.968. Likewise, when evaluating the classification model by means of the receiver operating characteristic curve (ROC) technique, it is identified that for the three expected classes of satisfaction the value of the area under the curve (AUC) is equal to 1, in such sense the predictive model through the SVM quadratic algorithm, has a high capacity to distinguish between the 3 classes: i) dissatisfied, ii) satisfied and iii) very satisfied of satisfaction of university students.

ACKNOWLEDGEMENTS

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

REFERENCES

[1] I. M. Ho, K. Y. Cheong, and A. Weldon, "Predicting student satisfaction of emergency remote learning in higher education remote during COVID-19 using machine learning techniques," *Plos One*, vol. 16, no. 4, pp. 1-27, Apr. 2021, doi: 10.1371/journal.pone.0249423.

[2] E. J De la Hoz, A and T. J. Fontalvo, "Machine learning methodology for the classification and prediction of users in virtual environments of education," *Technological information*, vol. 30, no. 1, pp. 247-254, Feb. 2019, doi: 10.4067/S0718-07642019000100247.

[3] D. D. Figaredo, J. Reich, and J. A. R. Valiente, "Learning analytics and data-driven education: An expanding field," *RIED*, vol. 23, no. 2, pp. 33-39, Jul. 2020, doi: 10.5944/ried.23.2.27105.

[4] E. A. Franco, R. E.L. Martínez, and V. H. M. Domínguez, "Predictive models of academic risk in computing careers with educational data mining," *RED*, vol. 21, no. 66, pp. 1-36, Apr. 2021, doi: 10.6018/red.463561.

[5] P. L. Sanz, Y. M. Bustos, Y. M. Acuña, K. N. Pinto, S. Z. Saldías, and A. Rodríguez-Fernández, "Satisfaction with higher education as an indicator of quality: the case of a Chilean public university," *Cuban Journal of Higher Education*, vol. 40, no. 3, pp. 1-16, Jul. 2021.

[6] N. D. Lescano, H. I. M. Huamán, D. M. D. De Angulo, and A. V. Muñoz, "Level of satisfaction in quality training at the University level," *Journal of Interdisciplinary Studies in Social Sciences*, vol. 23, no. 2, pp. 247-266, May. 2021, doi: 10.36390/telos232.04.

[7] N. M. Almusharraf and S. H. Khahro, "Studens' satisfaction with online learning experiences during the COVID-19 Pandemic," *International Journal of Emerging Technologies in Learning*, vol. 15, no. 21, pp. 246-267, Nov. 2021, doi: 10.3991/ijet.v15i21.15647.

[8] H. F. Mella, M. M. Navarro, and M. J. R. Riquelme, "Analysis of satisfaction of students of the Faculty of Engineering of Universidad de Talca," *Chilean Journal of Engineering*, vol. 18, no. 3, pp. 350-363, Dec. 2010, doi: 10.4067/S0718-33052010000300009.

[9] G. X. V. Paredes, K. G. C. Málaga, K. P. Chipa, A. J. F. Calsina, and J. D. S. Condori, "Student satisfaction regarding virtual education during Covid-19," *Scientiarvm Postgraduate Journal*, vol. 6, no. 1, pp. 13-17, Jan. 2020, doi: 10.26696/sci.epg.0107.

[10] N. Aziz *et al.*, "ICT-supported for participatory engagement within E-learning community," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 20, no. 1, pp. 492-499, Oct. 2020, doi: 10.11591/ijeecs.v20.i1.pp492-499.

[11] R. A. S. Rueda, R. D. S. Rueda, and J. A. S. Silis, "Analysis of students' perception of the use of information and communication technologies in the educational process trough data science," *Electronic Journal Quality in Higher Education*, vol. 10, no. 1, pp. 96-124, May. 2019, doi: 10.22458/caes.v10i1.2463.

[12] D. Kucak, V. Juricic, and G. Dambié, "Machine learning in Education - a Survey of Current Research Trends," *29th DAAAM International Symposium on Intelligent manufacturing and Automation*, Vienna, Austria, doi: 10.2507/29th.daaam.proceedings.059.

[13] B. Bahati and E. Mukama, "Measuring learner satisfaction with formative e-assessment strategies," *International Journal of Emerging Technologies in Learning*, vol. 14, no. 7, pp. 61-79, Apr. 2019, doi: 10.3991/ijet.v14i07.9120.

[14] S. N. Azizan, A. S. H. lee, G. Crosling, G. Atherton, B.V. Arulanandam, C. E. Lee, and R. B. A. Rahim, "Online learning and COVID-19 in higher Education: The value of IT Models in Assessing Students' Satisfaction," *International Journal of Emerging Technologies in Learning*, vol. 17, no. 3, pp. 245-278, Feb. 2022, doi: 10.3991/ijet.v17i03.24871.

[15] J. Z. Ramirez, "Prediction factors of student satisfaction in online courses," *Ibero-American Journal of Distance Education*, vol. 19, no. 2, pp. 217-235, Jul. 2016, doi: 10.5944/ried.19.2.15112.

[16] A. R. Páez, N. D. G. Ramírez, and D. S. Guzmán, "Constrution and implementation of a model to predict the academic performance of university studens using the Naïve Bayes Algorithm," *Dialogues on Education*, vol. 10, no. 19, pp. 1-18, May. 2019, doi: 10.32870/dse.v0i19.509.

[17] A. Siddique, A. Jan, F. Majeed, A. I. Qahmash, N. N. Quadri, and M. O. A. Wahab, "Predicting academic performance using an efficient model based on fusion of classifiers," *Applied Sciences*, vol. 11, no. 24, pp.1-19, Dec. 2021, doi: 10.3390/app112411845.

[18] F. Afrin, M. S. Rahaman, and M. Hamilton, "Minning student respondes to infer student satifaction predictors," *Seventh International Conference on Learning and Teaching in Computing and Engineering (LaTiCE'20)*, Jun. 2020, doi: 10.48550/arXiv.2006.07860.

[19] 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 Bussiness Intelligence*, vol. 7, no. 1, pp. 67-73, Apr. 2021, doi: 10.20473/jisebi.7.1.67-73.
[20] O. M. H. Bedoya, M. L. Trujillo, and C. E. M. Echeverry, "Predictive model for the identification of factors associated with the search time of the first job in university graduates," *Virtual Journal of the Catholic University of the North*, vol. 58, pp. 3-18, Jul. 2019, doi: 10.35575/rvucn.n58a6.
[21] L. C. Yuan *et al.*, "Study on the university students' satisfaction of the wisdom tree massive open online course platform based on parameter optimization intelligent algorithm," *Conference Collection IMETI2020*, vol. 104, Dec. 2021, pp. 1–22, doi: 10.1177/00368504211054256.
[22] N. M. G. Purnamasari, M. A. Fauzi, Indriati, and L. S. Dewi, "Cyberbulling identification in twitter using support vector machine and information gain based feature selection," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 18, no. 3, pp. 1994-1500, Jun. 2020, doi: 10.11591/ijeecs.v18.i3.pp1494-1500.
[23] S. Huang, N. Cai. P. P. Pacheco, S. Narrandes, Y. Wang, and W. Xu, "Applications of support vector machine (SVM) learning in cancer genomics," *Cancer Genomics & Proteomics*, vol. 15, no. 1, pp. 41-51, Jan. 2018, doi: 10.21873/cgp.20063.
[24] C. Matovelle, P. Buestan, and E. T. Carrasco, "Use of artificial intelligence with learning machines in the construction of predictive expert systems applied to the educational field in engineering," *Educational research challenges during the COVID19 pandemic*, pp. 529-542, 2021.
[25] A. Kesumawati and D. T. Utari, "Predicting patterns of student graduation rates using Naïve Bayes Classifier and support vector machine," *AIP Conference Proceedings 2018*, Oct. 2018, doi: 10.1063/1.5062769.
[26] N. M. Ali, N. A. A. Aziz, and R. Besar, "Comparison of microarray breast cancer classification using support vector machine and logistic regression with LASSO and boruta feature selection," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 20, no. 2, pp. 712-719, Nov. 2020, doi: 10.11591/ijeecs.v20.i2.pp712-719.
[27] A. Abu, M. Al-Emran, and K. Shaalan, "Factors affecting students' performance in higher education: a systematic review of predictive data mining techniques," *Technology, Knowledge and Learning*, vol. 24, pp. 567–598, Apr. 2019, doi: 10.1007/s10758-019-09408-7.
[28] C. Yu, J. Wu, and A. Liu, "Predicting learning outcomes with MOOC clickstreams," *Education Science*, vol. 9, no. 2, pp. 1–15, May. 2019, doi: 10.3390/educsci9020104.
[29] 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, pp. 1–16, Dec. 2019, doi: 10.3390/e21121216.
[30] M. F. Musso and E. C. Cascallar, "Prediction of satisfaction with virtual education in university students: use of machine learning," *AACC Meetings, XVIII National Meeting and VII International Meeting of the AACC*, 2021.

BIOGRAPHIES OF AUTHORS

Omar Chamorro-Atalaya is an electronic engineer, graduated from the National University of Callao, with a Master's degree in Systems Engineering and a doctoral student at the Faculty of Administrative Sciences at UNAC. Researcher recognized by CONCYTEC. Research professor at the 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.

Guillermo Morales-Romero is Doctor in Educational Sciences, Master in Systems Engineering, Lawyer, Graduate in Mathematics and Computer Science, Master in Public Management, graduated from the systems engineering program, Professor at the Enrique Guzmán y Valle National University of Education. 24 years in University Teaching in the programs of Engineering, Computer Science, Law, Education, Statistics and others, Professor at the Postgraduate School of the UNE. Author of scientific articles in journals indexed in databases: Scopus, Wos. Computer Auditor and computer security. He can be contacted at email: gmorales@une.edu.pe.

Yerferzon Meza-Chaupis Professor Associated with Exclusive Dedication 40 Hours of the National University of Education, Enrique Guzmán y Valle. With 36 years of professional experience in Regular Basic Education, as well as managerial staff, former director and deputy director and former coordinator of the Educational Network V of the district of independence, UGEL 02 MINEDU. Academic Degree of Master in Educational Administration at the Cesar Vallejo University, Title of Second specialty in School Management with Pedagogical Leadership at the Cayetano Heredia Peruvian University. Doctoral studies in Public Management and Governance at Cesar Vallejo University, Author of scientific articles in journals indexed in databases: Scopus, Latindex and Others. He can be contacted at email: ymeza@une.edu.pe.

Elizabeth Auqui-Ramos Magister in the mention of University Teaching, with completed doctorate studies in the mention of Educational Sciences, studies in Industrial Engineering. Proficiency in English at an intermediate level, Italian and Quechua at a basic level, graduates in didactics of mathematics and pedagogical management and innovation for educational quality and competitiveness, with 11 years of experience in teaching at higher and university levels. She can be contacted at email: eauqui@une.edu.pe.

Jesús Ramos-Cruz Doctor of Education, Master in Administration and Planning of Higher Education, Bachelor of Educational Sciences in the specialty of Mathematics and Public Accountant, professor hired at the National University of Education Enrique Guzmán y Valle in the Academic Department of Mathematics and Informatics and at the Private University "San Juan Bautista" of Ica in the Professional School of Human Medicine, Former Postgraduate Professor at the National University of "Huancavelica" and the National University "San Luis Gonzaga" of Ica, Former Director of the Institute of Public Technological Higher Education "Pisco"-Ica. Appointed professional with 25 years of experience in Non-University Higher Education. He can be contacted at email: jesus.ramos@une.edu.pe.

César León-Velarde Doctor in Education, Master in Education: Measurement, Evaluation and Accreditation of Educational Quality, Master in Educational Management, Graduate and Bachelor in Education, Bachelor in Sacred Theology. Professor of Philosophy and CDCSS. With more than 10 years of experience as a teacher in private and national universities in Peru. He currently works at the Technological University of Peru (UTP). Recognized by a private university as a teacher of academic excellence and as one of the best teachers in that university. He interesting in e-learning education in which I have been able to specialize over the years and obtain pleasing results and recognition. He can be contacted at email: c19593@utp.edu.pe.

Irma Aybar-Bellido Master in Administration and Planning of Higher Education, Doctor of Education, teacher in the Auxiliary category of the National University "San Luis Gonzaga" of Ica, faculty of educational sciences and humanities, full-time teacher of the "private" University San Juan Bautista” de Ica, from the Professional School of Human Medicine and Civil Engineering, hired professor at the “Enrique Guzmán y Valle” La Cantuta University in the Faculty of Sciences, postgraduate professor at the “Enrique Guzmán y Valle” National University EXPRO itinerant modality, co-author of scientific articles in scientific journals, with 29 years of teaching experience in Regular Basic Education and former Education specialist of the AGP of the UGEL Huaytará-DREH-MINEDU. She can be contacted at email: irma.aybar@unica.edu.pe.