Indonesian Journal of Electrical Engineering and Computer Science
Vol. 35, No. 2, August 2024, pp. 1219~1226
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v35.i2.pp1219-1226

Employing educational data mining techniques to predict programming students at-risk of dropping out

Niel Francis B. Casillano, Karen W. Cantilang
Department of Information Technology, College of Computer Studies (CCS), Eastern Samar State University, Borongan, Philippines

Article Info

Article history:

Received Feb 23, 2024
Revised Apr 9, 2024
Accepted Apr 13, 2024

Keywords:

Dropout
kNN
Logistic regression
Neural network
Predictive analytics

ABSTRACT

This research primary aimed at evaluating various predictive models in predicting programming students at risk of dropping out. It also aimed at identifying attributes that are significant in predicting students at risk of dropping. The educational data mining process (EDM) was utilized as the research framework. The study conducted a ten-fold cross-validation, revealing that the k-nearest neighbors (kNN) algorithm achieved the highest classification accuracy at 95.5%. The decision tree model followed closely with a 94.9% accuracy, logistic regression exhibited 94.4%, and the neural network model yielded a classification accuracy of 93.2%. Further analysis, including confusion matrices and receiver operating characteristic (ROC) curves, provided detailed insights into the models' performance. Notably, the decision tree algorithm excelled in identifying students who did not drop out, with a misclassification rate of 9 out of 30 for dropped students. Analysis also showed that students' assignments completed (AC), laboratory work (LW), and attendance (ATT) were the strongest predictors in identifying students at risk of dropping. Results of the study can be used by instructors to identify in advance student at risk of dropping and provide them with the necessary intervention to improve performance in programming.

Corresponding Author:

Niel Francis B. Casillano
Department of Information Technology, College of Computer Studies (CCS)
Eastern Samar State University
Borongan, Philippines
Email: nfcasillano@gmail.com

1. INTRODUCTION

Education is commonly seen as the cornerstone of a nation's prosperity and societal progress. A solid education system is vital for shaping the future of any society, paving the way for growth and advancement [1]. Consequently, monitoring of student performance is imperative to guarantee the cultivation of high-quality human resources. Extensive research in higher education focusing on student academic performance has been conducted to address persistent challenges such as academic underachievement, rising university dropout rates, and delays in graduation [2]. Poor academic performance raises issues such as lack of proper education and a scarcity of qualified human resources, both of which are detrimental to a country's development. This is why in every educational institution, monitoring of students' academic performance and achievement is crucial [3]. In higher educational institutions (HEIs) academic performance can be gauged through the student scores in various activities and assessments administered in every course. However, there is no general consensus as to how academic performance is best measured [4]. One major benefit of monitoring academic performance is the ability to identify top performing students and those who are at risk of failing and dropping out. Students who are at risk of failing, dropping out, or repeating subjects because of low performance have become the focus of concern of educational institutions [5]. Dropping out from the subject may ultimately lead to student attrition which does not only affect the student but also the teachers, the institution and the general public [6]. Therefore, predicting students at risk of academic drop-out is critical in improving student performance [7]. Phauk and Okazaki mentioned that it is imperative to develop an effective method of predicting academic performance for teachers to apply appropriate interventions to improve low performing students. One way to accomplish this task is through the use of educational data mining (EDM) techniques. Educational data mining is the application of typical data mining approaches to solve educational problems (yagci, [8]). Students' information, educational records, exam results, student engagement in class, and the frequency with which students ask questions are extracted and processed and used as inputs in the building of a model through EDM. Recently, EDM has proven to be an effective method for detecting hidden patterns in educational data, predicting academic progress, and predicting future outcomes to the learning/teaching environment [9]. A systematic review was conducted by [10] and highlighted the use of various machine learning techniques to identify at-risk students and predict dropout rates. Studies predominantly utilize datasets from university databases and online learning platforms, demonstrating the pivotal role of machine learning in improving student performance by predicting dropout risk. The paper of [11] examined nearly 70 papers to illustrate the diverse modern techniques extensively utilized for forecasting students' performance. The identified methodologies were artificial intelligence, primarily include machine learning, collaborative filtering, recommender systems, and artificial neural networks. Okereke *at el.* [12] mentioned that because of the multitude of predictive variables influencing student performance, it is essential to employ a feature selection mechanism, such as RapidMiner, to filter these variables. They utilized the decision tree, for training and testing purposes and noticed that the accuracy of predictions is contingent upon the datasets used for model training, while dissimilar datasets yield varying accuracy levels when subjected to the same algorithm.

Meanwhile, one of the most dreaded subjects in any computing related program is Programming. Programming as a core computing course is both intimidating and daunting. A number of students who took it find the course uninteresting [13]. A paper published in 2019 concluded that the average success rate of students in introductory programming worldwide ranges from 67%-72% [14]. Eastern Samar State University-College of Computer Studies has been offering BS Computer Science and BS Information Technology since 2004. Computer programming is one of its core fundamental courses. Student performance in introductory and advanced computer programming averages a little over 60%. A large number of students show poor performance during first half of programming classes that makes students at risk for failure and ends up dropping out. This data has been validated in recent study of Co and Casillano which mentioned that programming subjects are one of the biggest predictors of students' on-time graduation [15].

To answer this pressing issue of student performance in computer programming and to help predict and identify students at risk of dropping, the researcher decided to investigate and aim to produce an effective predictive model using educational data mining techniques. This study primarily aimed to develop a predictive model that will identify students at-risk of dropping in computer programming using educational data mining techniques. Specifically, this study aims to i) predict programming students at risk of dropping out using the following educational data mining techniques: decision tree, k-nearest neighbors, logistic regression, and neural networks, ii) determine the features that best affect the determination of programming students at risk of dropping, and iii) evaluate the performance of all predictive models applied.

2. METHODS

2.1. Research design

In this study, the educational data mining process [16] was employed. This methodology offers a systematic approach illustrated in Figure 1, guiding the creation of a robust model. This systematic approach delves beyond surface-level analysis, revealing intricate patterns and trends. Leveraging sophisticated mathematical computations and algorithms, data mining efficiently dissects data, enabling the prediction of future events. The principles and techniques of data mining are versatile, finding applicability across diverse sectors, including education. The process involves the sequential execution of the following steps: i) data collection, ii) initial data preparation, ii) statistical analysis, iv) data preprocessing, v) implementation of data mining, and vi) evaluation of results.

Figure 1. Educational data mining process [15]

2.2. Data collection, preparation and pre-processing

The information for this study was retrieved from the archives of the college of computer studies, focusing on the grade records of students who pursued BS information technology and BS computer science and completed programming 1 and programming 2 courses during the years 2017, 2018, 2019, and 2020. The grades obtained from programming 1 and 2 were systematically collected and converted into a digital format, specifically a spreadsheet file, to facilitate subsequent analysis and processing. In its original state, the gathered data, commonly referred to as raw data, was typically unsuitable for meaningful analysis and modeling. Datasets resulting from the integration of information from multiple sources might have exhibited issues such as missing data, inconsistencies, errors, miscoding, and duplications. Therefore, preliminary processing of the raw data was imperative to address and rectify these potential issues. Following data collection and preparation, data is cleaned to remove missing data, data noise, and inconsistency, ensuring that the quality of prediction is not affected. It should be highlighted, however, that models such as random forests and decision trees can handle missing data [17].

2.3. Data description

The process of attribute selection for the predictive model was guided by a thorough analysis, as explained in the research by [18]. This methodology was carefully aligned with the dataset compiled by instructors from Eastern Samar State University at the conclusion of the semester. These aggregated datasets were then fed into Educational Data Mining (EDM) models, adhering to the conventions delineated in Table 1, ensuring the precision and reliability of the predictive analytics.

Table 1. Data description

Attributes (course)	Possible values	Description
SG	Poor/good/average	Semestral grade
STS	Poor/good/average	Subject test score
AC	Yes/no	Assignment completed
ATT	Poor/good/average	Attendance
LW	Yes/no	Class lab work
DR	Yes/no	Student drop-out status

2.4. Predictive model

Decision tree model: the decision tree model is a fundamental algorithm that categorizes data into nodes based on class purity. It serves as a precursor to random forest. The tree in orange is a specialized program capable of handling both discrete and continuous information [19]. Tree parameters:

- Induce binary tree: constructs a binary tree with two child nodes.
- Min. number of instances in leaves: ensures the algorithm avoids splits with fewer instances than the specified threshold.
- Do not split subsets smaller than: restricts the algorithm from splitting nodes with instances below a specified number.
- Limit the maximal tree depth: controls the depth of the classification tree.
- Stop when majority reaches [%]: halts node splitting after reaching a specified majority threshold.

k-nearest neighbors (k-NN) model: the k-NN model, a supervised learning technique, is employed for both regression and classification. It predicts the class for test data by computing distances between the test data and all training points. The algorithm selects the K number of points most similar to the test data, assessing the likelihood of test data belonging to each of the 'K' training data classes. In regression, the value is the average of the chosen 'K' training points [20].

Logistic regression: logistic regression models the probability of a discrete outcome given an input variable. Commonly used for binary outcomes, logistic regression is applicable to situations with more than two discrete outcomes through multinomial logistic regression. It is a valuable tool in classification tasks, aiding in determining if a new sample fits best into a category, particularly in aspects of cybersecurity such as attack detection [21]. Artificial neural networks (ANNs): artificial neural networks, computational networks inspired by biology, are employed in this study, focusing on multilayer perceptrons (MLPs) with backpropagation learning methods. MLPs, featuring input, hidden, and output layers, are commonly used for a wide range of issues in supervised ANNs [22].

2.5. Model evaluation

The assessment of predictive models will be executed using the test and score widget within orange. This widget serves a dual purpose. Initially, it generates a table containing various performance metrics for classifiers, such as classification accuracy and area under the curve. Additionally, it produces evaluation data that can be employed by other widgets, such as receiver operating characteristic (ROC) analysis and confusion matrix, to scrutinize classifier performance.

2.6. Machine learning tool used

All aspects of data preparation, predictive modeling, and model evaluation were executed through the use of Orange, a robust data mining software. Orange stands out as a free and open-source platform, equipped with a comprehensive suite of tools for machine learning and data visualization. Leveraging Orange's capabilities facilitated the creation of visually intuitive workflows for effective data analysis, enhancing the study's analytical depth and efficiency [23].

2.7. Ethical consideration

Adherence to ethical standards is paramount when conducting research, especially studies involving personal information from individuals. To safeguard privacy, the names of students were anonymized through code transformation (e.g., stud1, stud2, stud3, stud4...) before their respective grades underwent preprocessing. Furthermore, only adjectival ratings (such as poor/good/very good) and categorical data (Yes/No) were utilized instead of numeric grades. The researchers' sole objective in undertaking this study is to develop a predictive model for identifying students at risk of dropping out using educational data mining techniques.

3. RESULTS AND DISCUSSION

A total of 177 instances of student data from 2017 to 2020 were analyzed. Each student data were preprocessed to contain its corresponding features/attributes specifically, semestral grade (SG), subject test score (STS), assignments completed (AC), attendance (ATT), class lab work (LW), and dropout status (DR) (see Figure 2). The student data was then fed to four predictive models, specifically, decision tree model, kNN, logistic regression and neural networks, similar machine learning techniques were used by [24]. Similarly, examined and evaluated 30 chosen articles and unveiled five primary prediction techniques [25]: artificial neural networks (ANNs), decision trees, support vector machines (SVM), k-nearest neighbors (KNN), and naïve Bayes. The models were then subjected to the Test an score widget and ROC Analysis to identify its accuracy and performance (see Figure 3).

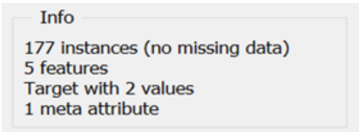

Figure 2. Dataset information

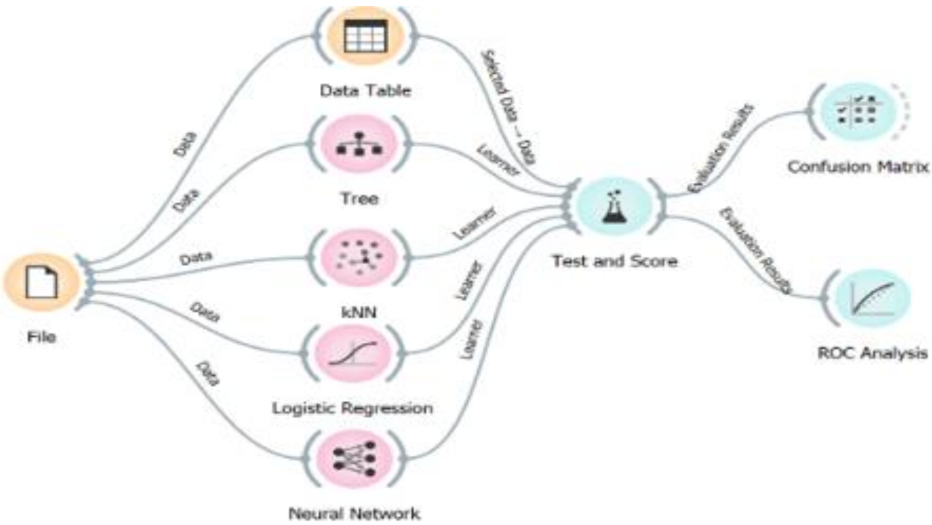

Figure 3. Orange framework

3.1. Test and score results

After undergoing a ten (10) fold cross validation, the following results were obtained: It can be gleaned in the evaluation results (Figure 4) that the model with the highest classification accuracy is kNN with a classification accuracy of 95.5%. It was followed by the decision tree model (94.9%), logistical regression (94.4) and finally by neural network (93.2%). Several researches have already been conducted that showed kNN as a good algorithm in predicting student performance. The study of [26] that predicted university-level academic performance through machine learning mechanisms, concluded that KNN is the model that best predicts academic performance for each of the semesters, followed by decision trees. Nugroho *et al.* [27] concluded that experiments conducted to identify student academic performance using the KNN Algorithm yielded clear and accurate results. Nouri *et al.* [28] used several machine learning models to predict student performance and kNN turned out to be the best performing model among all models used. Mulyani *et al.* [29] experimented on using an ensemble learning model to interfere students at risk of dropping. They claim that combining kNN, random forest, and logistic regression improves prediction performance, with an average improvement of 7.74% in the harmonic mean of accuracy and recall (F1-score) over previous work.

The performance of each model in accurately classifying student outcomes was further evaluated by examining confusion matrices. In Figure 5, representing the kNN model, it is evident that out of the 30 data entry identified as "dropped," 7 were mistakenly classified as "not dropped" or "no." The decision tree algorithm (Figure 6) exhibited the poorest performance in identifying dropouts, misclassifying 9 out of 30 instances tagged as "dropped" as "not dropped" or "no." On the other hand, the logistic regression model (Figure 7) misclassified 2 as "dropped" and 8 i as "not dropped," while the neural network model (Figure 8) misclassified 4 as "dropped" and 8 as "not dropped.". It is also worth noting that the decision tree model achieved the most accurate classification of students who did not drop from the course/program.

To enhance the assessment of model accuracy, a ROC curve was employed. This graphical representation illustrates the true positive rate against the false positive rate of the models. The outcomes of the ROC analysis for the model are depicted in Figures 9 and 10. In these figures, both true positive rates for "Yes" and "No" results are depicted by orange, purple, pink, and green lines, all consistently surpassing the 0.5 threshold marked by a red dotted line. This observation indicates that the model exhibits more true positive results compared to false positives, suggesting a commendable performance of the model.

Evaluation Results

Model	AUC	CA	F1	Precision	Recall
kNN	0.894	0.955	0.953	0.955	0.955
Tree	0.804	0.949	0.945	0.952	0.949
Neural Network	0.946	0.932	0.930	0.930	0.932
Logistic Regression	0.928	0.944	0.941	0.942	0.944

Figure 4. Evaluation results

Predicted

Actual	No	Yes	Σ
No	146	1	147
Yes	7	23	30
Σ	153	24	177

Figure 5. Confusion matrix for kNN

Predicted

Actual	No	Yes	Σ
No	147	0	147
Yes	9	21	30
Σ	156	21	177

Figure 6. Confusion matrix for decision tree

Predicted

Actual	No	Yes	Σ
No	145	2	147
Yes	8	22	30
Σ	153	24	177

Figure 7. Confusion matrix for logistic regression

Predicted

Actual	No	Yes	Σ
No	143	4	147
Yes	8	22	30
Σ	151	26	177

Figure 8. Confusion matrix for neural network

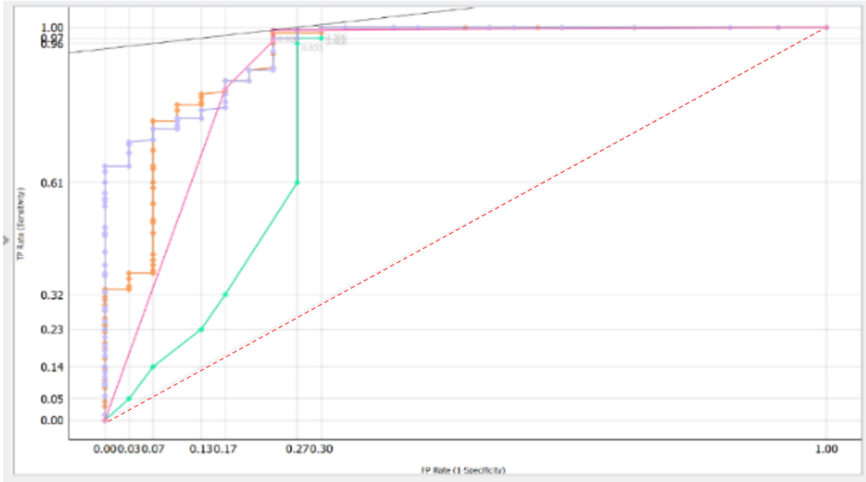

Figure 9. ROC analysis for “No” (Did not Drop)

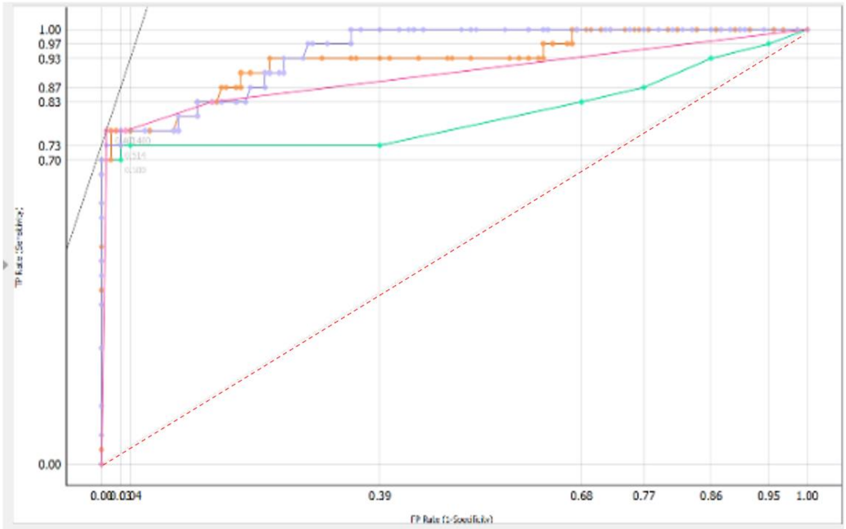

Figure 10. ROC analysis for “Yes” (Dropped)

The rank widget in Orange data mining was used to identify the attributes or features that best determines the prediction of whether a student will drop a course or program. The widget shows the gain ratio which is a ratio of the attribute's intrinsic information to the information gain, which lowers the bias towards multivalued features that arises in information gain. This method is used to determine which attributes are most relevant. As can be viewed in Figure 11, the top 3 attributes that are relevant in classifying students at risk of dropping out are AC, laboratory work (LW), and ATT.

		#	Gain ratio
1	C AC	2	0.350
2	C LW	2	0.262
3	C ATT	3	0.156
4	C SG	3	0.121
5	C STS	3	0.067

Figure 11. Gain ratio for student attributes

This is consistent with the study of [30] which mentions that academic performance is one of the mostly influential factors that influence dropping out of students. The result is also in line with the findings of [31] which mentions that low overall score in laboratory activities indicates that the student is prone to fail or drop out. This data reveals that attendance to programming classes and completion of a student's activity both assignments and laboratory works are imperative in identifying students who will continue with a program or will drop out of it. Programming instructors must ensure that students are able to complete and perform well in programming activities to ensure his engagement in the class thereby improving attendance and completion of laboratory activities and assignments.

4. CONCLUSION

In this study, the researcher explored various predictive models to identify programming students at risk of dropping out using educational data mining techniques. Through rigorous evaluation, the kNN algorithm emerged as the most accurate, achieving a classification accuracy of 95.5%, closely followed by the decision tree model at 94.9%. Logistic regression and neural networks also demonstrated respectable accuracy rates of 94.4% and 93.2%, respectively. Moreover, analysis revealed the significance of attributes such as assignments completed AC, LW, and ATT as predictors of student dropout risk. These findings offer valuable insights into effective predictive modeling for identifying at-risk students in programming courses, enabling educators to proactively intervene and support students, thereby enhancing overall retention rates and academic success. However, it is important to acknowledge the limitations of our study. Conducted within a specific institution and focused solely on programming courses, the generalizability of the findings to other academic disciplines or institutions may be limited. Additionally, while significant predictors of dropout risk were identified, other unexplored factors may also influence student retention. Future research should consider expanding the scope to encompass additional variables and longitudinal data, thereby enhancing predictive accuracy and providing a more comprehensive understanding of student retention dy.

REFERENCES

[1] HR Management DTi, "Education policy development and its importance," Retrieved from LinkedIn. Accessed: Oct. 19, 2023. [Online]. Available: https://www.linkedin.com/pulse/education-policy-development-its-importance-designedtoinspire-g4gwf/

[2] A. Namoun and A. Alshanqiti, "Predicting student performance using data mining and learning analytics techniques: A systematic literature review," *Applied Sciences (Switzerland)*, vol. 11, no. 1, pp. 1–28, Dec. 2021, doi: 10.3390/app11010237.

[3] R. Vilanova *et al.*, "Data-driven tool for monitoring of students performance," *IFAC-PapersOnLine*, vol. 52, no. 9, pp. 190–195, 2019, doi: 10.1016/j.ifacol.2019.08.188.

[4] U. Akpur, "Critical, Reflective, Creative Thinking and Their Reflections on Academic Achievement," *Thinking Skills and Creativity*, vol. 37, p. 100683, September 2020, doi: 10.1016/j.tsc.2020.100683.

[5] M. G. Gallego, A. P. Perez de los Cobos, and J. C. G. Gallego, "Identifying students at risk to academic dropout in higher education," *Education Sciences*, vol. 11, no. 8, p. 427, Aug. 2021, doi: 10.3390/educsci11080427.

[6] M. C. Pascoe, S. E. Hetrick, and A. G. Parker, "The impact of stress on students in secondary school and higher education," *International Journal of Adolescence and Youth*, vol. 25, no. 1, pp. 104–112, Dec. 2020, doi: 10.1080/02673843.2019.1596823.

[7] J. Niyogisubizo, L. Liao, E. Nziyumva, E. Murwanashyaka, and P. C. Nshimyumukiza, "Predicting student's dropout in university classes using two-layer ensemble machine learning approach: A novel stacked generalization," *Computers and Education: Artificial Intelligence*, vol. 3, p. 100066, 2022, doi: 10.1016/j.caeai.2022.100066.

[8] G. Mahajan and B. Saini, "Educational data mining: a state-of-the-art survey on tools and techniques used in EDM," *International Journal of Computer Applications & Information Technology*, vol. 12, no. 1, p. 310, 2020, [Online]. Available: www.ijcait.com

[9] E. Fernandes, M. Holanda, M. Victorino, V. Borges, R. Carvalho, and G. Van Erven, "Educational data mining: predictive analysis of academic performance of public school students in the capital of Brazil," *Journal of Business Research*, vol. 94, pp. 335–343, Jan. 2019, doi: 10.1016/j.jbusres.2018.02.012.

[10] B. Albreiki, N. Zaki, and H. Alashwal, "A systematic literature review of student' performance prediction using machine learning techniques," *Education Sciences*, vol. 11, no. 9, p. 552, Sep. 2021, doi: 10.3390/educsci11090552.

[11] J. L. Rastrollo-Guerrero, J. A. Gómez-Pulido, and A. Durán-Domínguez, "Analyzing and predicting students' performance by means of machine learning: A review," *Applied Sciences (Switzerland)*, vol. 10, no. 3, p. 1042, Feb. 2020, doi: 10.3390/app10031042.

[12] G. Okereke, C. H. Mamah, E. C. Ukekwe, and H. C, Nwagwu, "A Machine Learning Based Framework for Predicting Student's Academic Performance," *Physical Science & Biophysics Journal*, vol. 4, no. 2, 2020, doi: 10.23880/psbj-16000145.

[13] K. Sunday, P. Ocheja, S. Hussain, S. S. Oyelere, O. S. Balogun, and F. J. Agbo, "Analyzing student performance in programming education using classification techniques," *International Journal of Emerging Technologies in Learning*, vol. 15, no. 2, pp. 127–144, Jan. 2020, doi: 10.3991/ijet.v15i02.11527.

[14] Simon *et al.*, "Pass rates in introductory programming and in other STEM disciplines," in *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, New York, NY, USA: ACM, Dec. 2019, pp. 53–71. doi: 10.1145/3344429.3372502.

[15] J. Co and N. F. Casillano, "Predicting on-time graduation based on student performance in core introductory computing courses using decision tree algorithm," *Jurnal Pendidikan Progresif*, vol. 11, no. 3, pp. 650–658, 2021, doi: 10.23960/jpp.v11.i3.202116.

[16] A. C. C.-R. de Sousa, C. A. B. de Oliveira, and J. L. C. M. Borges, "Using academic performance to predict college students dropout: A case study," *Educação e Pesquisa*, vol. 44, no. 0, Oct. 2018, doi: 10.1590/s1678-4634201844180590.

[17] E. Alyahyan and D. Düştegör, "Predicting academic success in higher education: literature review and best practices," *International Journal of Educational Technology in Higher Education*, vol. 17, no. 1, p. 3, Dec. 2020, doi: 10.1186/s41239-020-0177-7.

[18] A. Aleryani, W. Wang, and B. De La Iglesia, "Dealing with missing data and uncertainty in the context of data mining," 2018, pp. 289–301. doi: 10.1007/978-3-319-92639-1_24.
[19] M. Kantardzic, *Data mining: concepts, models, methods, and algorithms*. Wiley Interscience, 2003.
[20] A. Christopher "K-Nearest neighbor." Retrieved March 25, 2022, from https://medium.com/swlh/k-nearest-neighbor-ca2593d7a3c4
[21] T. W. Edgar and D. O. Manz, "Exploratory study. research methods for cyber security," pp. 95–130, 2017.
[22] Y. S. Park and S. Lek, "Artificial neural networks: multilayer perceptron for ecological modeling," in *Developments in Environmental Modelling*, vol. 28, 2016, pp. 123–140. doi: 10.1016/B978-0-444-63623-2.00007-4.
[23] Orange data mining, "Orange visual programming," *Tree*, 2015, [Online]. Available: https://orange3.readthedocs.io/projects/orange-visual-programming/en/latest/widgets/model/tree.html
[24] A. S. Hashim, W. A. Awadh, and A. K. Hamoud, "Student performance prediction model based on supervised machine learning algorithms," *IOP Conference Series: Materials Science and Engineering*, vol. 928, no. 3, p. 032019, Nov. 2020, doi: 10.1088/1757-899X/928/3/032019.
[25] Y. Baashar, G. Alkawsi, N. Ali, H. Alhussian, and H. T. Bahbouh, "Predicting student's performance using machine learning methods: a systematic literature review," in *2021 International Conference on Computer & Information Sciences (ICCOINS)*, IEEE, Jul. 2021, pp. 357–362. doi: 10.1109/ICCOINS49721.2021.9497185.
[26] L. E. Contreas-Bravo, N. Nieves-Pimiento, and K. G. Guerrero, "Prediction of university-level academic performance through machine learning mechanisms and supervised methods," *Ingenieria (Colombia)*, vol. 28, no. 1, p. e19514, Nov. 2023, doi: 10.14483/23448393.19514.
[27] A. Nugroho, O. R. Riady, A. Calvin, and D. Suhartono, "Identification of student academic performance using the KNN algorithm," *Engineering, MAthematics and Computer Science (EMACS) Journal*, vol. 2, no. 3, pp. 115–122, Sep. 2020, doi: 10.21512/emacsjournal.v2i3.6537.
[28] J. Nouri, M. Saqr, and U. Fors, "Predicting performance of students in a flipped classroom using machine learning: towards automated data-driven formative feedback," *ICSIT 2019 - 10th International Conference on Society and Information Technologies, Proceedings*, pp. 79–82, 2019.
[29] E. Mulyani, I. Hidayah, and S. Fauziati, "Dropout prediction optimization through SMOTE and ensemble learning," in *2019 2nd International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2019*, IEEE, Dec. 2019, pp. 516–521. doi: 10.1109/ISRITI48646.2019.9034673.
[30] Nurmalitasari, Z. A. Long, and M. F. Mohd Noor, "Factors influencing dropout students in higher education," *Education Research International*, vol. 2023, pp. 1–13, Feb. 2023, doi: 10.1155/2023/7704142.
[31] M. M. Tamada, R. Giusti, and J. F. de M. Netto, "Predicting students at risk of dropout in technical course using LMS logs," *Electronics (Switzerland)*, vol. 11, no. 3, p. 468, Feb. 2022, doi: 10.3390/electronics11030468.

BIOGRAPHIES OF AUTHORS

Niel Francis B. Casillano is an Associate Professor at the College of Computer Studies at Eastern Samar State University Philippines. He has a BSc in Information Technology and Master of Arts in Teaching-Computer Science. He can be contacted at email: nfcasillano@gmail.com.

Karen W. Cantilang is an Assistant Professor at the College of Computer Studies at Eastern Samar State University Philippines and is presently the Program Head of the Information Technology Department. She holds a B.Sc. degree in Information Technology and a Master in Information Technology. She can be contacted at email: yenessu@gmail.com.