Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 239 (2024) 598–605

Procedia Computer Science

www.elsevier.com/locate/procedia

CENTERIS – International Conference on ENTERprise Information Systems / ProjMAN – International Conference on Project MANagement / HCist – International Conference on Health and Social Care Information Systems and Technologies 2023

Interpretable Success Prediction in a Computer Networks Curricular Unit Using Machine Learning

Catarina Félix de Oliveira[a,b], Sónia Rolland Sobral[a], Maria João Ferreira[a,c], Fernando Moreira[a,d,*]

[a]REMIT, Universidade Portucalense, Porto, Portugal
[b]LIADD-INESC TEC, Porto, Portugal
[c]ALGORITMI, Universidade do Minho, Guimarães, Portugal
[d]IEETA, Universiade de Aveiro, Aveiro, Portugal

Abstract

Today, higher education institutions are focused on understanding which factors are associated with the failure or success of students to, early on, be able to implement measures that can reduce the low performance of students and even dropout. The retention rate is positively and negatively influenced by factors belonging to several dimensions (personal, environmental, and institutional). We aim to use information from those dimensions to identify students enrolled in a Computer Networks course at risk of failing the subject. Besides, this needs to happen as early as possible, to be able to provide the students, for example, with extra support or resources to try to prevent that negative outcome. For predicting the grade level on the first test, the best accuracy obtained was 55%. However, most C-level grades were correctly classified, with 63% accuracy in predicting the students that are most at risk of failing, which is one of our main objectives. As for the prediction of the second test's grade level, the best accuracy obtained was 89% and concerned data regarding the students' interaction with the LMS together with students' grades history. All the C-level grades were correctly classified (100% accuracy) and so we were able to correctly predict every student at a high risk of failing. Using the procedure described in this paper, we are able to anticipate the students needing extra support, and provide them with different resources, to try to prevent their negative outcome.

Keywords: Learning Analytics; Educational Data Mining; Higher education; Dropout; Retention.

* Corresponding author.
E-mail address: fmoreira@upt.pt

1877-0509 © 2024 The Authors. Published by Elsevier B.V.

10.1016/j.procs.2024.06.212

1. Introduction

In most countries, education has been a priority for governments in the last decades, with particular importance for higher education because people with higher education have higher employment rates and are generally better paid [1]. Since the 1960s, there has been a significant increase in higher education enrolment [2], with an overall boost to 34% of the population. Based on the data provided in [3], estimates show that there will be a 56% increase in higher education enrolments by 2030. For example, approximately 47% of young adults have obtained an academic degree in the US, and over 80% of those who finish secondary education enroll in higher education [1,2]. While in Australia, 40% of young people have completed tertiary education [4], and in the UK, 42% of the active population in the labor market have tertiary education qualifications [5]. Finally, in China, more than 60% of students who finish secondary education attend higher education [6].

The figures presented above, particularly the growth in the number of students in higher education until 2030, lead higher education institutions (HEIs) to be very concerned with their prestige. This prestige is often reflected in international rankings and is based on a multitude of criteria, but the historical academic performance of students, that is, the failure or low performance of students is one of the essential metrics and, therefore, generates great concern for HEI decision-makers. This problem negatively affects both the student as an individual and the community. In this context, in [7], its shown that it is essential for HEIs to have a constant concern and to understand which factors are associated with the failure or success of students to, early on, be able to implement measures that can reduce the low performance of students and even its dropout [1].

In this context, student success can be defined as behaviors that aim to achieve academic goals, such as persistence in higher education and course completion [8]. For higher education accreditation agencies, one of the parameters observed with significant relevance, and which is included in the self-assessment guide of the courses, is the number of students who graduate and therefore is directly related to retention rates.

Pioneering work on internal and external influences on a student's decision to remain or drop out of higher education studies (i.e., retention and dropout) was developed by Tinto [9] (Dropout-intentions Model). According to this author, there is a continuous but variable interaction between a series of educational, environmental, psychological, and social factors, which have been identified as determinants of academic performance and success in higher education studies [10]. Based on this seminal work, other authors [10] developed retention prediction models that allowed the development of measures to reduce attrition rates.

The use of Data Mining (DM) – analyzing large databases to generate new information – will serve, in the education area, to analyze student performance [7, 11]. DM methods applied to educational data are known as Educational Data Mining (EDM) [12]. With EDM, it is possible to develop methods to discover knowledge from data from the educational environment extracted from various domains (DM and machine learning, psychometrics and other areas of statistics, information visualization, and computational modelling). The EDM will make it possible to obtain essential/crucial information about the education process so that teachers can develop actions and measures to provide conditions to improve academic success. According to [13], there are three types of predictions in higher education: (i) predict the academic performance or GPA of students at the undergraduate level; (ii) predict the failure or withdrawal of students from a course; and (iii) predict student outcomes in specific courses.

We analyze data regarding students enrolled in a Computer Networks subject, to try to predict the students' performance in two tests. Our aim is to identify students at risk of failing the subject as soon as possible, to be able to provide them, for example, with extra support to try to prevent that outcome.

2. Background

Machine learning (ML) is the scientific field dealing with the ways in which machines learn from experience [14], is a subdivision of Artificial intelligence where algorithmic models are trained to perform specific tasks, recognizing, and learning patterns from the data they see, rather than through explicit computer programming by a human expert [15]. Machine learning techniques aim at analyzing the data (commonly represented in tabular form – datasets) to find meaningful patterns. ML problems can be divided into unsupervised and supervised learning problems, depending on the absence or presence of a dependent variable, respectively [16]. According with Moore [17] "a machine learns with respect to a particular task T, performance metric P, and type of experience E, if the system reliably improves its performance P at task T, following experience E. Depending on how we specify T, P, and E, the learning task might also be called by names such as data mining, autonomous discovery, database updating, programming by example, etc.”. Data mining and analytics may serve as a basic tool to promote the process from machine automation to information automation and then to knowledge automation [18], plays the key role in extracting interesting patterns (rules) from the data [19]. Data mining tasks can be classified into three main categories: prediction (the act of talking about the future), association (discovering interesting relationships among variables in big data), and clustering [20]. There are many examples of academic analytics: Enrolment Predictive Modelling at Baylor University, Predicting and Improving Student Retention at the University of Alabama, developing a Student Success Plan and Early Alert System at Sinclair Community College or Connecting Resource Utilization, Risk Level, and Outcomes at Northern Arizona University [21]. Educational data mining is the application of data mining and ML techniques to educational data [22] to early detect students at risk [23], analyzing data to find meaningful patterns. Surveys of educational data mining are presented in [22] and, more recently, in [24].

Some studies include sociodemographic data [25], psychometric data [26], statistical data, data related to course activity and course statistics [27], previous grades [28], and data concerning automatically evaluated programming exercises [29]. The aims stated for each study include improving the skills of each student in programming [30], predicting students’ performance [25] or program correctness [31], identify at-risk students, explore the effects of an instructional intervention for increasing student motivation [36], graphically model students’ progress [33], and also compare the effectiveness of different educational data mining techniques. Some unsupervised machine learning techniques [29], others use supervised ones as regression [26] and classification [25]. There are a lot of algorithms, like artificial neural networks / multilayer perceptron [28], decision trees [31], support vector machines [31], naive bayes [28], logistic [31] and linear [26] regression, J48 [25], Sequential minimal optimization [28], Markov models [29], decision rules, standard Voting, and object tracking procedures [33] and, finally, best-subset-regression and LASSO regression [31].

To improve student progress, data from educational activities can be used to recognize patterns and make suggestions to improve student performance. Despite the concern to reduce the number of dropouts, it is also essential to understand which decisions need to be implemented at the strategic level concerning the policies, strategies, and actions that institutions carry out by curricular unit, by degree, or even as a whole. For this purpose, an initial search was carried out for academic papers published in journals and conferences in English that used Machine Learning for interpretable Success Prediction in a Computer Networks Curricular Unit. Using the query "(attrition OR retention OR dropout) AND ('higher education' OR tertiary OR university OR degree*) AND ('Computer Networking course' OR 'Networking course' OR 'Introduction Computer Networks course')" and limiting to the English language, it wasn't found any academic work that applied Machine Learning to Computer Networks Curricular Unit.

3. Methodology

The objective of this study is an interception of the three types of prediction indicated in [13], applied to the Computer Networks course of the Computer Engineering degree, to help teachers intervene early to improve student performance.

We analyze data regarding students enrolled in the first year of two courses: Informatics Engineering (IE) and Information Technology (IT). Both courses share some Curricular Units (CUs), as is the case of Algorithms and Programming (AP) during the first semester, and Computer Networks (CN) during the second. In the later, students need to perform two written tests. The tests are evaluated in the rage [0, 20], and an average grade above 10 mean that the student successfully completed the test. Students with grades below 7 fail the test and, therefore, the subject. Our aim is to use the data collected to predict the students’ performance in both tests. Ultimately, we wish to be able to anticipatedly identify students at risk of failing the subject, to provide them, for example, with extra support. We collected and integrated data from four different sources:

- Survey data (S): data collected from an online survey that the students were, voluntarily, asked to answer at the beginning of the school year;
- Interaction data (I): data based on the students’ interaction with the CN subject’s LMS: the proportion of available resources each student has interacted with;

- Grade data (G^{AP}, G^{CN}): data regarding the students' assessment in both CUs. In AP we consider values in the range [0,20]. In CN we consider two types of grades: first we try to predict the grades on the first test (G^{CN}_1), and, for that, we consider grade levels A (values above 10), B (values in the range [7,10[ ), and C (values below 7). When we try to predict the grade on the second test (G^{CN}_2), we consider the same levels for that grade. However, we also use G^{CN}_1 as feature and, in this case, it is a value in the range [0,20].
  With the collected data, we built eight different datasets, listed in Table 1.

Table 1. Datasets used for the study.

Dataset	Features	Sample size	Target
D1	I_1	62	
D2	$S + I_1$	22	G^{CN}_1
D3	$G^{AP} + I_1$	54	
D4	$S + G^{AP} + I_1$	22	
D5	$I_1 + I_2 + G^{CN}_1$	62	
D6	$S + I_1 + I_2 + G^{CN}_1$	22	G^{CN}_2
D7	$G^{AP} + I_1 + I2 + G^{CN}_1$	54	
D8	$S + G^{AP} + I_1 + I_2 + G^{CN}_1$	22	

The first dataset (D1) uses information of the students' interaction with the LMS before the first test (I1) to try to predict the grade on that test (G^{CN}_1). For the second dataset (D2), we added the data obtained in the surveys (S). The third dataset, besides the LMS interaction data, included the grades obtained in AP CU (GAP). The fourth dataset (D4) includes both. The remaining datasets were created with a similar process. However, instead of trying to predict the grade on CN's first test, they try to predict the grade on the second test (G^{CN}_2) and, for that, also include information of the students' interaction with the LMS between the first and second tests (I2).

A decision tree was fitted with Python's sklearn package for each dataset. The models were evaluated in terms of accuracy, precision, recall and F1-score, using 10-fold cross validation. The results obtained are presented on the next section.

4. Results

The Table 2 presents the accuracies the decision trees obtained for each dataset.

Table 2. Accuracies obtained for each dataset.

		Survey data			
		No		Yes	
		G^{CN}_1	G^{CN}_2	G^{CN}_1	G^{CN}_2
AP	No	50%	85%	36%	64%
	Yes	54%	89%	55%	55%

The accuracy obtained for dataset D1 (trying to predict the grade on the first test with only information regarding the students' interaction with the LMS before that) was 50%. When adding the survey information (D2), the accuracy decreased to 36%. This seems to suggest that the survey data does not help to explain the grades. If, besides interaction data, we consider the grades obtained in AP (D3), the accuracy obtained is 54%. This value is increased to 55% when we also include the survey data (D4). The higher accuracy obtained in predicting G^{CN}_1 was obtained when considering, besides interaction data, both the survey data and the data regarding the grades previously obtained in AP (D4).

For dataset D5 (trying to predict the grade on the second test with only information regarding the students' interaction with the LMS), the accuracy obtained was 85%. When adding the survey data (D6), the accuracy decreased to 64%. Once again, this seems to suggest that the survey data does not help to explain the grades. When adding the AP grades to the interaction data (D7), the accuracy obtained was 89%. Finally, when all the information is considered, the accuracy decreases to 55%. The higher accuracy obtained in predicting G^{CN}_2 was obtained when considering the interaction data together with the data regarding the grades obtained in AP.

With these results we can state that, considering previously obtained grades (weather in different CUs as in the same) is advantageous for predicting the students' performance. Another observation that can be made is that the survey data helps predict the students' performance in the beginning but, after that there are other factors (e.g., LMS interaction, grade obtained on the first test) that are more accurate in explaining the grade obtained in the second test.

The decision tree obtained for the best performing model for the prediction of the first test is shown in Figure 1

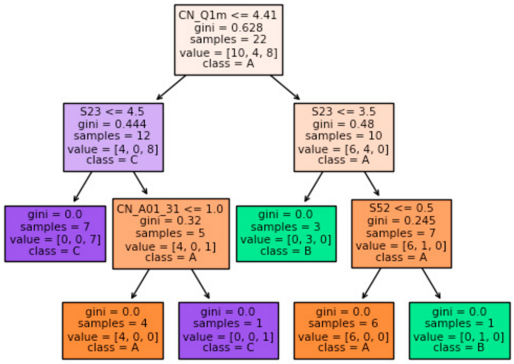

Fig. 1. Decision tree obtained for the best performing model for the prediction of the first test.

This decision process of this tree is explained next. If feature CN_Q1m (1st quiz: higher grade for each student) 4.41, it will look at feature S23 (Satisfaction with colleagues). If this is below than 4.5, then the result on the first test is a C-level grade. Otherwise, it will look at feature CN_A01_31 (interacting with an LMS activity), the result is an A-level grade. Otherwise, the result is a C-grade level. Back to the root of the tree, if the value in CN_Q1m is higher than 4.4, it will look at feature S23. If this is below 3.5 the result is a B-level grade, otherwise it will look at feature S52 (Other hobbies). If this value is below 0.5 (the student does not have other hobbies), the grade level is A, otherwise it is B.

The confusion matrix for this model is presented in Figure 2.

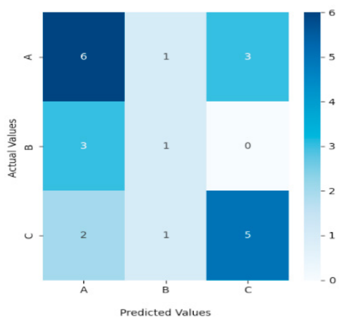

Fig. 2. Confusion matrix obtained for the best performing model for the prediction of the first test.

The confusion matrix in the figure shows that, for A-level grades, the majority (60%) was correctly classified, while 1 (10%) was misclassified as B and 3 (30%) as C. For B-level grades, only one (25%) was correctly classified, while 3 (75%) were misclassified as A. Regarding the C-level grades, the majority (63%) were correctly classified, while 2 (25%) were misclassified as A and 1 (13%) as B. With this, we can see that the model has a 63% accuracy in predicting the students thar are most at risk of failing, which is one of our main objectives.

Next, in Table 3, we present the classification report for this model, that presents the model evaluation in terms of precision, recall, F1-score, and support.

Table 3. Classification report obtained for the best performing model for the prediction of the first test.

	precision	recall	f1-score	support
A	0.55	0.60	0.57	10
B	0.33	0.25	0.29	4
C	0.62	0.62	0.62	8
accuracy			0.55	22
macro avg	0.50	0.49	0.49	22
weighted avg	0.54	0.55	0.54	22

For the second test, the decision tree obtained for the best performing model is depicted in Figure 3.

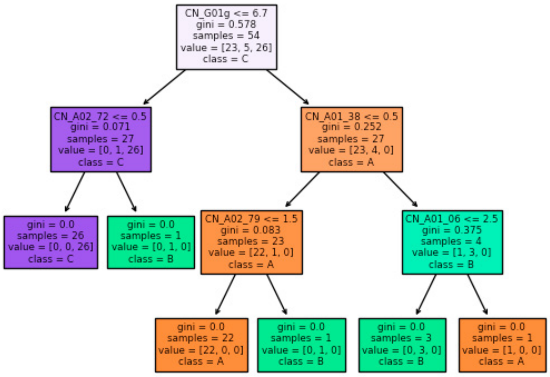

Fig. 3. Decision tree obtained for the best performing model for the prediction of the second test.

This decision tree starts by looking at feature CN_G01g (grade on the first test). If it is lower than 6.7, it will look at feature CN_A02_72 (interacting with an LMS activity, between the first and the second test). If this value is below 0.5 (the student did not interact with this activity) the result is a C-level grade, otherwise a B-level grade. Going back to the tree root, if the grade on the first test is above 6.7, it will look at feature CN_A01_38 (interacting with an LMS activity, before the first test). If its value is below 0.5 (the student did not interact with that activity), the tree will look at CN_A02_79 (interacting with an LMS activity, between the first and the second test). If its value is below 1.5 (the student interacted with the activity once or never), the result will be an A, otherwise a B. Going back to CN_A01_38, if its value is higher than 0.5, the tree will look at feature CN_A01_06 (interacting with an LMS activity, before the first test). If this value is lower than 2.5 (the student interacted with this activity twice or less times), the result will be a B-level grade, otherwise an A-level grade.

The confusion matrix for this model is presented in Figure 4.

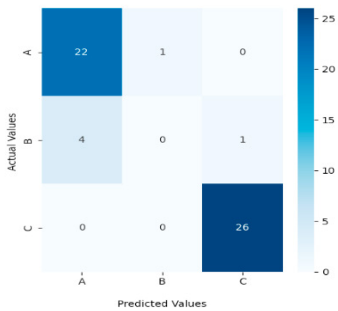

Fig. 4. Confusion matrix obtained for the best performing model for the prediction of the second test.

In the figure we can see that there are few misclassifications. For A-level grades only one value was misclassified as a B, while 96% were correctly classified. For B-level grades, 4 values are misclassified as A and one as C, with no correct predictions. For the C-level grades, all the values were correctly classified. The model used for the first test grade prediction had an accuracy of 63% in predicting students at high risk of failing. Now, for the second test, this accuracy is increased to 100%, which means that it could correctly predict every student at high risk of failing. This allows us to anticipate the students in need of extra support, trying to prevent their negative outcome.

Next, in Table 4, we present the classification report for this model, that presents the same evaluation metrics presented before for the prediction of the first test grade.

Table 4. Classification report obtained for the best performing model for the prediction of the second test.

	precision	recall	f1-score	support
A	0.85	0.96	0.90	23
B	0.00	0.00	0.00	5
C	0.96	1.00	0.98	26
accuracy			0.89	54
macro avg	0.60	0.65	0.63	54
weighted avg	0.82	0.89	0.85	54

5. Conclusions

The digital transformation has brought a profound change in the higher education system and created various opportunities and facilities for today's students, such as by providing study materials through an LMS and active teaching methods. However, there are several barriers and obstacles; namely, the increase in the number of students attending higher education is accompanied by high dropout rates. In this context, a study was conducted where data on students registered in a Computer Network course for a Computer Engineering degree was used to predict students' performance in two tests. This study is justified since research in the application of ML to Computer Networks Curricular Unit were not found.

As referred to in the previous section for the study, a collection of datasets was defined, with the first one using information from the students' interaction with the LMS before the first test to predict the score on that test. For the second dataset, the data obtained from the surveys conducted for the students was added. The third dataset, in addition to the interaction data with the LMS, included the grades obtained, and the fourth dataset included both. The remaining datasets were created with a similar process. However, instead of trying to predict the score on the first test, it was attempted to predict the score on the second test, and to do this, information on students' interaction with the LMS between the first and second tests was also included. With this information, it was possible to predict which students need extra support and provide them with different resources to avoid a negative result. Although the study results have been very positive, the same needs a larger sample to give more sustainability to the results obtained.

As future work, we will extend the study to other CU of the same course and of different courses of a higher education institution. In addition, we intend to use other classification algorithms besides decision trees.

Acknowledgments

This work was supported by the FCT – Fundação para a Ciência e a Tecnologia, I.P. [Project UIDB/05105/2020]

References

[1] Ma, Pender, and Welch. (2016) “Education Pays: The Benefits of Higher Education for Individuals and Society.” https://files.eric.ed.gov/fulltext/ED572548.pdf

[2] Roser,, and Ortiz-Ospina. (2013) “Tertiary Education.” https://ourworldindata.org/tertiary-education

[3] Choudaha, and Van Rest. (2018) “Envisioning pathways to 2030: Megatrends shaping the future of global higher education and international student mobility.” https://www.studyportals.com/wp-content/uploads/2018/01/Report-Envisioning-Pathways-to-2030-Studyportals-2018.pdf

[4] Norton, Cherastidtham, and Mackey. (2018) “Mapping Australian higher education 2018.” https://grattan.edu.au/wp-content/uploads/2018/09/907-Mapping-Australian-highereducation-2018.pdf

[5] Office of National Statistics. (2017) “Graduates in the UK labour market: 2017.” https://dera.ioe.ac.uk/30577/1/Graduates%20in%20the%20UK%20labour%20market%202017.pdf

[6] China Power Team. (2020) “How does education in China compare with other countries?” https://chinapower.csis.org/education-in-china/

[7] Oliveira, Sobral, Ferreira, and Moreira. (2021) "How Does Learning Analytics Contribute to Prevent Students' Dropout in Higher Education: A Systematic Literature Review." *Big Data Cogn. Comput.*, **5**, 64.

[8] Crisp, Taggart, and Nora. (2015) "Undergraduate Latina/o Students: A Systematic Review of Research Identifying Factors Contributing to Academic Success Outcomes." *Review of Educational Research* **85** (**2**): 249–274.

[9] Tinto. (1975) "Dropout from higher education: A theoretical synthesis of recent research." *Review of Educational Research* **45** (**1**): 89–125.

[10] Cornelius, Wood, and Lai. (2016) "Implementation and evaluation of a formal academic-peer-mentoring programme in higher education." *Active Learning in Higher Education* **17** (**3**): 193–205.

[11] Baradwaj, and Pal. (2011) "Mining educational data to analyze students' performance." *International Journal of Advanced Computer Science and Applications* **2** (**6**): 63-69.

[12] Baker, and Yacef. (2009) "The state of educational data mining in 2009: A review and future visions." *J. of Edu. Data Mining* **1** (**1**): 3-17.

[13] Alturki, Hulpus, and Stuckenschmidt. (2020) "Predicting academic outcomes: A survey from 2007 till 2018." *Technology, Knowledge and Learning*, 1-33.

[14] I. Kavakiotis, Tsave, Salifoglou and Vlahavas. (2017) "Machine Learning and Data Mining Methods in Diabetes Research." *Computational and Structural Biotechnology Journal*, (**15**)**:** 104-117.

[15] Du, Emam, and Gniadecki. (2020) "Review of Machine Learning in Predicting Dermatological Outcomes." *Frontiers in Medicine* **7**.

[16] Felix, and Sobral (2020) "Predicting students' performance using survey data." *IEEE Global Engineering Education Conference*, EDUCON.

[17] Mitchell. (2006) "The Discipline of Machine Learning." *Machine Learning Department, CMU.*

[18] Ge, Song, Ding, and Huang. (2017) "Data Mining and Analytics in the Process Industry: The Role of Machine Learning." *IEEE Access* (**5**): 20590-20616.

[19] Tsai, Chiang, and Yang. (2014) "Data mining for internet of things: A survey." *IEEE Communications Surveys and Tutorials* **16** (**1**): 77-97

[20] Huang, Cheng, Huang, Lin, Wang, Chuang, and Wu. (2021) "Establishing a Dynamic Capital Structure Model for Company Sustainability Performance Using Data Mining Techniques." *Sustainability* **13**.

[21] Campbell, DeBlois, and Oblinger. (2007) "Academic Analytics: A New Tool for a New Era." *EDUCAUSE Review* **42** (**4**), 2007.

[22] Fernandes, Holanda, Victorino, Borges, Carvalho, and VanErven (2019) "Educational data mining: Predictive analysis of academic performance of public school students in the capital of brazil." *Journal of Business Research* (**94**): 335–343.

[23] Aldowah, Al-Samarraie, and Fauzy (2019). "Educational data mining and learning analytics for 21st century higher education: A review and synthesis." *Telematics and Informatics* (**37**): 13-49.

[24] Bakhshinategh, Zaiane, ElAtia, and Ipperciel. (2018) "Educational data mining applications and tasks: A survey of the last 10 years." *Education and Information Technologies* **23** (**1**): 537–553.

[25] Ayub, and Karnalim. (2017) "Predicting Outcomes in Introductory Programming using J48 Classification." *World Transactions on Engineering and Technology Education* **15** (**2**): 132-136.

[26] Alvarez, Wise, Altermatt, and Aranguiz. (2019) "Predicting academic results in a modular computer programming course." *LALA 2019 - Proceedings of the 2nd Latin American Conference on Learning Analytics*, Valdivia, Chile.

[27] Piech, Sahami, Koller, Cooper, and Blikstein. (2012) "Modeling how students learn to program.". *IGCSE'12 - Proceedings of the 23rd ACM Technical Symposium on Computer Science Education*.

[28] Sivasakthi. (2017) "Classification and prediction-based data mining algorithms to predict students' introductory programming performance." *International Conference on Inventive Computing and Informatics (ICICI)*.

[29] Gil, Fernández-Alemán, Trujillo, García-Mateos, Luján-Mora, and Toval. (2018) "The Effect of Green Software: A Study of Impact Factors on the Correctness of Software." *Sustainability* **10** (**10**).

[30] Figueiredo, and García-Peñalvo. (2019) "Teaching and learning strategies of programming for university courses." *TEEM'19: Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality*, 2019.

[31] Costa, Fonseca, Santana, Araújo, and Rego. (2017) "Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses." *Computers in Human Behavior* (**73**): 247-256.

[32] Burn (2006) "Why Should I Care? Student Motivation In An Introductory Programming Course." *Annual Conference & Exposition, Chicago*.

[33] Gilal, Jaafar, Omar, Basri, Aziz, Khand, and Hasan. (2017) "Suitable Personality Traits for Learning Programming Subjects: A Rough-Fuzzy Model." *International Journal of Advanced Computer Science and Applications (IJACSA)* **8** (**8**), 2017.