Online-Journals.org

iJEP International Journal of Engineering Pedagogy

iJEP | eISSN: 2192-4880 | Vol. 16 No. 1 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijep.v16i1.56637

PAPER

Prediction of Secondary-School Student Cognitive Performance in Informatics

Dimitrios A. Varsos, **Nikolaos C. Zygouris**, **Kostas M. Kolomvatsos**, **Antonios N. Dadaliaris**, **Georgios Dimitriou** (✉)

University of Thessaly, Lamia, Greece

dimitriu@uth.gr

ABSTRACT

Machine learning techniques for the prediction of performance in the learning process are primarily studied at the post-secondary level of education. Data sets are usually large at that level, resulting in predictive models that have a high accuracy. In contrast, limited research has been conducted at the secondary school level, mostly due to the typically small data sizes and the unique educational challenges of that level. In the present study we aimed to address such issues by implementing a model to predict the final performance of lower secondary school students in a course on Informatics, using a teaching scenario that combined the flipped classroom, a variation of the jigsaw technique and educational robotics. Given the student grades from the first third of the year, as well as demographics data, we used an IBM data analytics tool to create and test several predictive models. The CHAID algorithm achieved the highest accuracy (82.14%) and AUC value, outperforming others like Quest, C5, BN, and RF. The tool also pinpointed the educational activity that was the most significant predictor of final performance, indicating its strong instructional value. Despite the relatively small dataset, the results suggest that careful parameter selection can yield models that predict learner performance with a high accuracy and assist educators in continuously improving their teaching for better student outcomes.

KEYWORDS

predictive modeling, machine learning, classification algorithms, flipped classroom, jigsaw technique, educational robotics (ER), STEM epistemology

1 INTRODUCTION

The task of designing, composing and evaluating the teaching process is of uttermost importance in education. An instructor can choose for this task one or more from a wide variety of tools, i.e., methods, techniques, technologies and applications, many of which come from scientific fields other than education. The choice of, the appropriate combination of, and the way such tools are used by the instructor of a course, are important in achieving the purpose and specific targets of the course, in

Varsos, D. A., Zygouris, N. C., Kolomvatsos, K. M., Dadaliaris, A. N., Dimitriou, G. (2026). Prediction of Secondary-School Student Cognitive Performance in Informatics. *International Journal of Engineering Pedagogy (iJEP)*, 16(1), pp. 66–86. https://doi.org/10.3991/ijep.v16i1.56637

Article submitted 2025-05-15. Revision uploaded 2025-12-18. Final acceptance 2026-01-02.

continuously improving the course, and in identifying those factors that have the greatest impact on the student success rate.

Learning computer programming is a complex process that requires high-level cognitive skills, such as the application of logical rules and problem-solving [1], [2]. The level of difficulty in learning computer programming for novice student programmers of all ages increases when the traditional learning method is applied, which involves using a simple text editor for writing and compiling programs in a high-level programming language [3]. An alternative method is the use of educational robotics (ER) in combination with block-based programming environments. This approach has the advantage of utilizing an easier code development environment, as it eliminates all syntax errors, and commands are transferred with simple drag-and-drop actions [4]. Additionally, the outcome of the program is more understandable to students as they see a robot performing specific tasks [5]. Furthermore, ER promotes the four essential skills of the 21st century, which are collaboration, creativity, communication, and critical thinking [6], while simultaneously providing a fun and engaging learning environment [7].

The flipped classroom is an innovative method of active learning [8] with a student-centered focus [9]. It aims to reverse the way in which the content of a lesson is taught and applied. In the traditional form of a lesson, the content is usually presented live through in-class meetings and is then applied in the form of homework [10]. In contrast, in a flipped classroom, students study the new educational material at home before the lesson, and then apply the knowledge they have gained during the live in-class meetings [11], [12], [13], [14]. The educational material focuses on activities that encourage active student participation, is distributed online, and can include videos, presentations, slideshows, e-books, websites, handouts, etc. [15], [16]. The distribution of educational content, so that students can access it online, can be done through an asynchronous e-learning platform [17], [18], [19]. In this way, students learn independently at their own pace at home [10], while there is more available time in face-to-face lessons in the classroom for exercises, projects, and discussions [14].

In order to promote a favorable learning environment in the classroom, the application of collaborative learning techniques plays an important role, where students interact with their peers to acquire knowledge and develop essential skills [20], [21]. One of the many collaborative teaching techniques is the jigsaw technique. It is a useful technique that allows a group of students to cover multiple topics simultaneously within a set timeframe [22]. The most important difference between the jigsaw technique and other collaborative learning methods is that each member of the group is assigned a portion of the educational task, allowing each member to contribute equally and meaningfully to the group [23].

When education partially moves online, as in the case of the flipped classroom model, monitoring the student progress in mastering the lesson becomes a challenge for educators. Various methodologies have been employed over time to support educators in identifying students in need and intervening early if necessary [24]. A prominent example of such a methodology is student success predictive models [25].

Predictive modeling in general is a process in data science that is used to infer outcomes on uncertain future events. It involves creation of predictive models (PMs) with machine learning algorithms, which analyze current and historical data sets, in order to detect patterns, observe trends, and so to be able to predict future behavior [26], [27]. More specifically, starting with sets of collected data, predictive modeling creates mathematical models to predict future values or classes of attributes that are the prediction targets. This means that predictive modeling relies on the hypothesis that a set of known data can be used for the prediction of the values or the classes of a new data set concerning the specific attributes being observed [28].

Predictive modeling constitutes a basic methodology of research in education, focusing on the prediction of student performance. In particular, it can be utilized for the evaluation of performance, for the improvement of the educational process and the guidance on the student cognitive steps. Further, for providing feedback and adapting teaching directives based on student learning behaviors, for the evaluation of teaching material and processes. Finally, for the detection of abnormal learning behaviors and problems in order to achieve a deeper understanding of educational circumstances [29], [30].

In total, predictive modeling in teaching and learning can be used in areas such as educational theories analytics [31], analysis of pedagogical strategies [32], analysis of programming codes in Informatics courses [33], cooperational learning and prediction of team project group assessment [34], prediction of student performance, and in particular of students in danger to fail for in-time intervention [35], [36], [37], course support and improvement, e.g. regarding the course context and the sequence of course activities [38], [39], improvements in foreign language learning [40].

Literature reviews in predictive modeling in education generally concentrate on large data sets gathered from online learning, post-secondary education and face-to-face teaching along with computer-based instruction. There are few studies related to secondary school students, even fewer focusing on Informatics courses centered around learning programming [41], [42].

It is well understood that large data sets give better prediction accuracy than small data sets. However, it is typical in secondary education to have small class sizes, with teachers covering several classes of a total of not more than a few hundred students per teacher [43]. In such educational settings, the distinct differences between the teaching methods that different teachers follow, as well as the differences in school policies and capabilities, can have a profound effect with significant variations on the student overall learning process [44], [45]. This means that any attempt to produce large data sets for the creation of PMs may result in low performance prediction accuracy, when such sets consist of combined data collected from different educators, or even different schools [46]. However, past experience has shown that PMs are invaluable in education, as they provide tailor-made insights that can be challenging for educators to pinpoint and respond to during the typical flow of teaching [47]. For this reason, it becomes a necessity to focus research on small classes and thus small prediction data sizes.

Another important aspect that advocates the development of PMs for small educational data sets is the trend to use blended learning in secondary education [48]. In face-to-face teaching, and for courses that are covered in several weekly hours, many cognitive patterns may become obvious to the teacher [49]. However, this is not the case when distant learning methods, such as the flipped classroom, are involved, especially for courses like Informatics that are covered in only one or two hours per week. In such cases, cognitive patterns are less obvious, e.g. the nuanced connections between how a student engages in class discussions and their practical achievements in distant class tasks like programming assignments [50]. Here PMs have the knack for identifying such patterns [38].

The capability of PMs to forecast future student performance based on current and historical data [51], is crucial not only for immediate educational adjustments [52] but also for long-term curriculum development [53]. It allows educational institutions to be proactive rather than reactive, adapting to the evolving student educational needs. For informatics education, in particular, predictive modeling in the teaching process can help us move away from the traditional teaching model towards a more customized and supportive learning experience [54].

The main purpose of this study is to find a PM using machine learning algorithms that will assist lower secondary school students in achieving a more comprehensive understanding of the subject of Informatics. In fact, the research presented herein is part of a wider study on the development of new methods for the teaching of computer programming, as well as for providing basic elements of the STEM (Sciences, Technology, Engineering, Mathematics) epistemology, to secondary school students.

In developing an educational setting, a teaching scenario (TS) was designed for learning programming, as part of the Informatics curriculum taught in lower secondary schools. The TS is based on three interconnected elements: a) ER, b) the application of the flipped classroom approach using an asynchronous e-learning platform, and c) the application of a variation of the jigsaw technique as a collaborative teaching method in face-to-face lessons. The idea behind this approach of teaching computer programming is to make students enjoy the class more, in order to increase their motivation for learning and improve their performance [55].

This study hypothesizes that a predictive model can be developed to determine the highest percentage of success for secondary school students in an Informatics course. Additionally, it explores the possibility that instructors can receive ongoing feedback during the academic year, which can be used to enhance the course content and teaching methods in real time. These hypotheses assume that the integration of predictive analytics into the educational process will not only forecast student outcomes with a high degree of accuracy but also provide actionable insights that facilitate adaptive teaching strategies tailored to student needs and course dynamics.

2 MATERIALS AND METHODS

2.1 Participants

The number of students of the Music High School of Lamia in the school year 2022–2023 that participated in this research was 140 (68 males, 72 females), aged 12–15 years old (mean 13.62, SD 1.06). Of these, 48 students attended 1st, 46 students attended 2nd and 46 attended 3rd grade.

None of the 140 students had learning disabilities, a history of major medical conditions, psychiatric disorders, developmental disorders, or significant visual or auditory impairments, as confirmed by their school medical records. All human data used in this research were collected in compliance with the Helsinki Declaration, and the permission of the Ethics and Deontology Committee of the University of Thessaly (28-7-22/1). Finally, students also provided signed letters of consent from their guardians for the participation in the research.

2.2 Implementation

The course organization was based on the two separate but interrelated components of a flipped classroom model: In-class and distant education. Remotely, students learn, read and comprehend the theoretical part of computer programming, using appropriate teaching material, and cooperate and communicate with each other on an e-learning platform. In live in-class meetings, students utilize in practice what they learnt theoretically through teamwork.

Structure of distance learning course. The course teaching material was distributed through the e-learning platform. Since students were assumed to have no prior knowledge of programming and robotics, the material was the same for all three grades.

The course was divided in three units and each unit had three lessons. Each lesson in the platform contained a quiz exercise and a homework assignment. Students had first to read and try to understand the theoretical part of the lesson, and then continue with the quiz and the homework. The quiz was comprised of five questions, of either correct-wrong or matching type. It was being automatically graded, at the 0 to 20 scale, and each question taking 4 points. After the quiz exercise was completed, the students could go ahead with the assignment, a small compulsory personal homework of a text format, which had to be submitted by students and corrected and graded by the instructor at the 0 to 20 scale.

Structure of face-to-face learning course. In live in-class meetings, students utilize in practice through team work what they learnt theoretically. To organize students in teams and enable the collaboration learning we use a variant of the jigsaw technique. The jigsaw technique can be applied by the teacher in 10 steps [56]. Our variation of the jigsaw technique compared to the original is the following:

Steps 1, 2 are the same. In step 3, first we break the problem down into sub-problems and assign roles based on the problem-solving process. Then, in step 4, students are investigating the part of their problem. In step 5 students collaborate and exchange information. We do not apply steps 6, 7, and 8, but all the members of the team at the same time collaborate, learn from each other and focus on the part of their work, seeing all the stages of solving the problem at the same time. Step 9 is the same with the jigsaw method. Step 10 differs, and specifically at the end of the lesson session, the members of each team meet for a few minutes and make an evaluation of their work up to that point, to find out what did not go well in the whole process of working together, and to improve it the next time.

2.3 Framework for analyzing the data

The framework for our analysis follows a comprehensive machine learning pipeline. This includes data collection, attribute selection, model training and validation using cross-validation, evaluation through multiple performance metrics, and interpretation via predictor importance analysis. The goal is to develop a predictive model that can be applied during the academic year to identify students who may need additional support and help educators make informed, data-driven decisions.

We first collect measurable data related to student academic performance, as well as demographic data. Once we have gathered the necessary data, we prepare and define attributes appropriately. Academic data provide attributes that are treated as ordinal variables due to their discrete scoring system, while demographic data provide attributes that are handled according to their data type. We then split the dataset into a training and a test set to evaluate model performance. We also apply 10-fold cross-validation on the training set to reduce the risk of overfitting and improve the model's reliability.

For model development, we use the IBM SPSS Modeler Version 18.4 software package that allows for automated modeling and testing of several classification algorithms and ranking them based on their overall accuracy. The algorithms tested are C5, Logistic Regression (LR), Bayesian Network (BN), LSVM, Random Trees, XGBoost Linear, XGBoost Tree, CHAID, QUEST, C&R Tree, and Neural Net. The algorithms that exhibit the best accuracy are validated against the test set to ensure they generalize well and are not overfitted. To assess the performance of the models, standard classification metrics are used, including overall accuracy, precision, recall, F1-score, and the area under the ROC curve (AUC).

Using the above analysis framework, we determine which algorithm performs best in predicting final student outcomes. Additionally, by analyzing predictor importance, we identify which input attributes have the greatest influence on the model's predictions.

3 RESULTS

3.1 Data collection

To construct a PM, it is necessary to have measurable attributes of the research subject, with a definite output and a data set that collects results from iterative events. In the area of education, such events are courses that are repeated each year, where data can concern course performance (e.g., enrollment, test grades, grade averages, final grade and success level), as well as interactions between individual students and a teaching platform (e.g., browsing behavior, quizzes and interactive exercises). Also, they may include data from cooperating students (e.g., text communication), administrative data (e.g., school, school district, teachers), demographic data (e.g., gender, age, nationality) and data related to student psychology (e.g. motivations, emotional situations) [57], [58].

In our case, the data we had access to were the grades from the answered quiz exercises, the grades from the homework assignments, the final grade for the course, and demographic data, such as the class grade, age, and gender of the students.

3.2 Attribute selection

According to Moreno-Marcos et al. [59], the appropriate selection of attributes (or features) can often be more important than the prediction algorithms themselves. To select the attributes – predictors of our model, we need to find out which factors of the subject are influencing the prediction.

In online learning environments, data from activities related to blended learning, such as those in a flipped classroom, can be used to forecast student performance. Additionally, attributes associated with assignments were found to be potentially significant predictors [57]. Clickstream data can be acceptable predictors when exercises are not available, but they do not add prediction power if attributes related to exercises are used [60]. Previous grade achievements of students and their demographics were identified as the most important factors that can be used to predict academic success in post-secondary education [61].

According to Xu et al., student performance can be forecasted early on, but the predictions tend to be less stable at such a course stage [57]. In the middle stage of the course lessons, however, performance can be more reliably predicted, allowing teachers to adjust their teaching strategies promptly. This makes it possible to intervene and provide personalized support, tailoring instruction to the individual needs of each student.

In the current research, we explore the possibility of creating a PM for the student final performance, using only the grades from the first unit of the quiz exercises and homework assignments, as well as demographic data, such as the class grade, age and gender of the students. The reason we are using only the grades from the first three lessons corresponding to the first unit, is that it is closer to the middle stage of the lessons and not too early [57]. Furthermore, we want the model outcomes to be useful during the same period, that is, with only early grades available, in order not only to predict the final grade of a student, but also to allow the teacher to assess the student's mastery of the subject. Additionally, the teacher should be able to receive information during the school year about which factors might affect the prediction of student performance and be able to improve it as teaching progresses. Finally, we want to create a PM to be continuously applicable in future school years, thus using data solely from within the same school year.

As explained above, the input attributes that were used for the PM were (i) the student grades in the first three quiz exercises, (ii) the student grades in the first three homework assignments and (iii) demographic data about the class grade, age, and gender of the students i.e., a total of nine input attributes. The single output attribute was the student final grade. The grades range between 0 and 20, but due to the grading methodology used, the grade of each individual quiz exercise and homework assignment did not take continuous values in the [0, 20] interval, but rather discrete values, e.g., 0, 4, 8, 12, 16, 20, or 0, 5, 10, 15, 20. This was important for the way attributes were declared at the PM creation. Tables 1 and 2 give the frequencies of student grades in the three quiz exercises and the three homework assignments, using the aforementioned discrete grade values.

Table 1. Frequencies of student grades regarding the three quiz exercises

Exercise/Grade		0	4	8	12	16	20	Total
Exercise1	f(%)	0(0)	5(3,57)	14(10,00)	26(18,57)	42(30,00)	53(37,86)	140(100)
Exercise2	f(%)	0(0)	13(9,29)	18(12,86)	22(15,71)	38(27,14)	49(35,00)	140(100)
Exercise3	f(%)	0(0)	7(5,00)	17(12,14)	38(27,14)	44(31,43)	34(24,29)	140(100)

Table 2. Frequencies of student grades regarding the three homework assignments

Assignment/Grade		0	5	10	15	20	Total
Assignment1	f(%)	0(0)	0(0)	44(31.43)	59(41.14)	37(26.43)	140(100)
Assignment2	f(%)	0(0)	0(0)	32(22.86)	46(32.86)	62(44.29)	140(100)
Assignment3	f(%)	0(0)	0(0)	82(58,57)	0(0)	58(41.43)	140(100)

In this study, the quiz and homework grade attributes were declared as ordinal (non-binary), whereas the final grade was declared as flag. We coded the final grade with the value 1 (High) for students with a final grade higher than 16, and the value 0 (Low) for students with a final grade lower than or equal to 16. We set a threshold of 16 because we were interested in creating a PM to predict which students would have a final grade exactly or below 16, by the time we know, through quiz exercises and homework assignments, the first third of their yearly performance, and we thus have obtained the additional information we seek regarding the percentage of mastery of the Informatics course up to that time.

3.3 Dataset option

Since the dataset consisted of 140 entries, we did not divide it into three separate sets (training, validation and test), but rather into two sets, more specifically an 80% training and a 20% test set. When the dataset is of small or medium size, it is common to perform the validation phase on the training set only, using multiple cross-validation [62]. We opted to use the 10-fold multiple cross-validation on the training set in order to avoid PM overfitting, as used in similar research [63], [64].

3.4 Model creation

To create our PM, we used the software package IBM SPSS Modeler Version 18.4. The steps we followed are shown in Table 3.

Table 3. Steps in the predictive modeling process using IBM SPSS modeler

Step	Process	Description
1	Data Input	Data including grades and demographic details are entered.
2	Type Declaration	Attribute types are specified for the data entered, with roles defined for each attribute.
3	Auto Classifier	The system automatically classifies data using various machine learning algorithms.
4	Model Ranking	Models are ranked based on their overall accuracy.
5	Cross-Validation	Cross-validation is performed with a specified number of folds (e.g., ten) to ensure model reliability.
6	Analysis	Data is analyzed to select the best performing algorithms.
7	Results	The final ranked list of algorithms, termed "SPSS golden nuggets", is displayed.

The attribute roles are either Input or Target. In particular, we defined as Input all attributes related with the demographic data (Gender, Age, Class grade), the student grades in the first three quiz exercises (Exercise1, 2, and 3) and the first three homework assignments (Assignment1, 2, and 3) and as Target the final grade attribute.

Table 4 shows the list of the classification algorithms that exhibited the best overall classification accuracy [51] based on the training set. From the algorithms shown, CHAID, QUEST, and C5 are Decision Tree (DT) algorithms, some of the most popular in machine learning [65]. According to Table 4, the CHAID classifier exhibited an overall accuracy of 86.607% and AUC equal to 0.908. The QUEST and C5 classifiers had the same overall accuracy value but yielded a lower AUC value equal to 0.782. The BN classifier exhibited an overall accuracy of 90.179% and AUC equal to 0.954. In contrast, the Random Forest (RF) classifier exhibited the highest overall classification accuracy equal to 100%, as well as the highest AUC value of 1.0. We concluded that the best classifier based on the training set was the RF classifier.

Table 4. Algorithm ranking with the auto classifier option

Algorithm	Overall Accuracy	Area Under Curve (AUC)
CHAID	86.607	0.908
QUEST	85.714	0.782
C5	85.714	0.782
Bayesian Network	90.179	0.954
Random Forest	100	1.0

3.5 Evaluation of the model on the test set

After having created the PM to predict final student grade and validated it by evaluating its performance using a 10-fold cross-validation on the training set only, we moved on to evaluate the PM and its five best classification algorithms on the test set. To this end, we introduced into the application the remaining 20% of the data entries that comprise the test set. We then obtained an analysis of its performance based on this set.

Figure 1 shows the overall classification accuracy of the five algorithms CHAID, QUEST, C5, BN, and RF, based on the training set, against the respective overall accuracy of the algorithms, based on the test set.

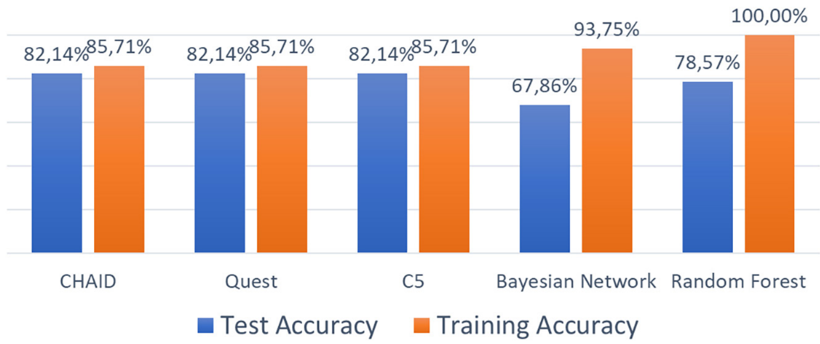

Fig. 1. Training and testing accuracy for all classification algorithms

From Figure 1, we can observe that while the RF algorithm had 100% accuracy on the training set, it did not show the same good accuracy based on the test set. In fact, it exhibited an accuracy of 78.57%, which was lower than the accuracy of the three DT algorithms CHAID, Quest, and C5. This may be due to the smaller size of the test set compared to the training set. That is, the RF algorithm requires a large amount of data to perform good accuracy in classifying data. Indeed, according to Herrera et al. [50], the RF algorithm demonstrates better accuracy when classifying large sets of data.

Next the analysis focused on sensitivity and specificity of the five classification algorithms, based on the precision, the recall, and the F1-score metrics [39], with the corresponding values given in Table 5. More specifically, the CHAID classifier exhibits a Precision of 0.895, a Recall of 0.850 and an F1-score of 0.872. The same values have the three metrics for the QUEST and C5 classifiers. The RF algorithm exhibits lower respective values than CHAID, QUEST, and C5, i.e., a precision of 0.889, a recall of 0.800 and an F1-score of 0.842. In contrast, the BN algorithm exhibits the lowest corresponding values than all other algorithms, i.e., a precision of 0.789, a recall of 0.750, and an F1-score of 0.769.

Table 5. Precision, recall, and F1-score for all five classifiers

Classifier	Precision	Recall	F1-Score
CHAID	0.895	0.850	0.872
QUEST	0.895	0.850	0.872
C5	0.895	0.850	0.872
Bayesian Network	0.789	0.750	0.769
Random Forest	0.889	0.800	0.842

From the above Precision values, it follows that the CHAID, QUEST and C5 algorithms exhibit greater sensitivity and specificity compared to the BN and RF algorithms. This means they can more accurately find how many of the students that actually received a score greater than 16 were correctly predicted (sensitivity) and how many of the students predicted to receive a score greater than 16 actually received a score greater than 16 (specificity). This explains the reason why the CHAID, QUEST and C5 classifiers result in a higher total accuracy than the BN and the RF algorithms.

The ROC curves for the AUC values of the five classification algorithms are depicted in Figure 2. We can see that the CHAID classifier has an AUC value of 0.863, the QUEST and the C5 classifiers have an AUC value of 0.8, the RF classifier exhibits an AUC value of 0.797 and the BN classifier has an AUC value of 0.713. We can thus note that the CHAID classifier has the best AUC value, i.e., a value that is closer to 1, than all other three algorithms. According to [66], the higher the value of the AUC metric for a classifier, the better are the classification results. Therefore, among all classification algorithms tested, we can conclude that the CHAID classifier exhibits the best total prediction accuracy.

Despite the early evaluation of the algorithms on the training set, which suggested that the RF classifier seems to excel in the classification accuracy in the prediction of student grades, the evaluation on the test set proved that the PM created, based on that algorithm, was not as good in classification as suggested, since it exhibited lower values than CHAID, QUEST and C5 in all evaluation metrics, i.e., Overall accuracy, Precision, Recall, F1-score and AUC. It is very likely that the model created on the RF classification algorithm exhibited overfitting on the training data set, despite the use of the 10-fold multiple cross-validation, with the result of not being able to perform a high-quality classification on the unknown data of the test set. Finally, the difference in the rest four algorithms was made by the higher value in the AUC metric of the CHAID classifier, as compared with the AUC of QUEST, C5, and BN.

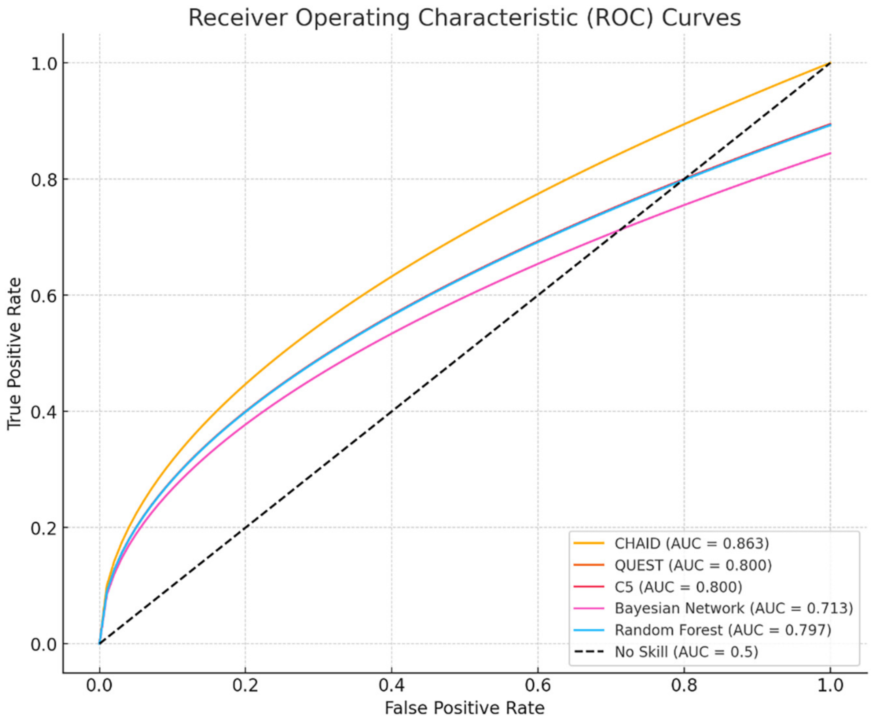

Fig. 2. ROC curves with the respective AUC values for the five classifiers tested

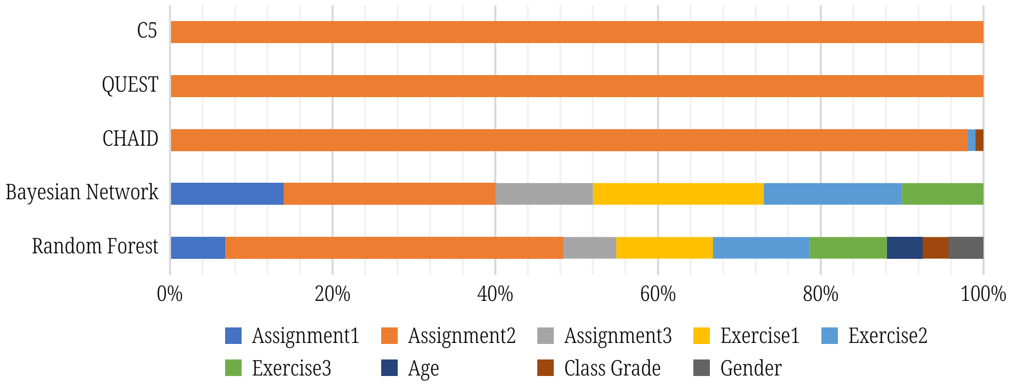

Fig. 3. Predictor importance for each of the five classifiers

Predictor Importance. Figure 3 shows the predictor importance chart, which indicates the significance of each predictor in the algorithm, or in other words the most important factors for predicting the student final grade. In CHAID, the most important predictor is Assignment2, while the least important is Class grade. In QUEST and C5, the most important predictor is Assignment2 and is the only important. In BN, the most important predictor is Assignment2 and the least important is Exercise3. Finally, in RF, the most important predictor is Assignment2 and the least important is Class grade.

From the above, it emerged that all the classifiers showed that the most important predictor factor for predicting the final grade of students is Assignment2, although some of the algorithms did not use all predictors in the process. In particular, the DT algorithms (C5, QUEST, and CHAID) cut off insignificant predictor factors, and for this reason those factors are missing from the chart. Finally, it should be noted that predictor importance is not related to model accuracy, neither is it affected by whether the predictions are correct or not [67].

4 DISCUSSION

In this section we discuss the results of our research. We compare against similar work, and we comment on our research contribution. Finally, we discuss limitations of our work and possible future research directions.

In all discussion, we must keep in mind that our work is focused on the following points: (a) creation of an educational PM for lower secondary schools using student academic data, i.e., quiz and homework grades, as well as demographic data, (b) addressing peculiarities of that education level, i.e., the small data set sizes for non-core courses like Informatics, and (c) teaching using the flipped classroom model, a variation of the jigsaw technique and ER. To the best of our knowledge there is no research work that combines all three points. Therefore, we proceed with discussion and comparison against work that only partially shares the aspects that we touch.

Concerning the first hypothesis of our research, we have established that even with a small data set, it is possible to develop a highly accurate PM to predict the student success rate in a secondary school Informatics course implemented in a flipped classroom. According to the results presented earlier, the CHAID algorithm, a DT machine learning algorithm, provides the maximum accuracy when compared with QUEST, C5, BN and RF, due to its highest AUC value. Nguyen et al. [26], in their research on the prediction of student performance in post-secondary education, also discovered that DT algorithms have a higher accuracy than others such as BN algorithms by 3–12%.

In [68], Ramaswami and Bhaskaran created a PM to predict student performance in higher secondary school education, based on socio-economic and demographic data sets. The machine learning algorithm used was the CHAID algorithm. They found an overall model prediction accuracy of 44.69%, which is much lower than the accuracy that we obtained for the CHAID prediction model. This may be partially due to the different data set types, but also due to the different data encoding used in the PM creation. Although we used a smaller number of attributes, we configured the individual grade attributes to not have all the distinct values from 0 to 20, but values that increased with a step of 4 (quiz exercises) or 5 (homework assignments). This suggests that the overall accuracy of the model is influenced not only by the machine learning algorithm, the number and type of the attributes and the range of the attribute values, but also by the number of individual distinct values an attribute takes within that range of values.

Nouri et al. [63] applied machine learning algorithms in order to predict student performance in a flipped learning course. They used Naïve Bayes (NB), LR, kNN, Neural Network (NN), and Random Forest (RF) as machine learning algorithms. The kNN model exhibited the best prediction accuracy of 81% and AUC level compared to the others. This accuracy rate is consistent with our findings. In fact, the CHAID algorithm in our research exhibits slightly better accuracy. Although they tested the RF algorithm, it ultimately proved to have lower overall accuracy compared to CHAID, which is also consistent with our results.

Peraić and Grubišić [69] compared the performance of six machine learning algorithms, namely, LR, support vector machines (SVM), RF, NN, NB, and DT, for classifying student performance in an introductory programming course, using an e-learning platform to distribute course material, lectures, homework, laboratory exercises, and quizzes. They used two different grading systems, binary (Pass/Fail) and three-level grade categorization, to categorize the final grade, as target variable. The LR model achieved the highest accuracy rate of 82.17% on the binary dataset, while the RF model achieved the highest accuracy rate of 65.62% on the three-level dataset. In our work we also use a binary final grading system, and the CHAID algorithm of our PM exhibits about the same accuracy as the LR algorithm of their predictive models.

Martinez-Carrascal et al. [70] addressed the problem of predicting student performance at the early stages of a flipped learning course on the basics of electronics, as a part of a post-secondary engineering program. The course spanned a total of twelve weeks, and the data were collected from the e-learning platform after the fourth week and before the first exam. They used only activity data and not grades, because there were no assessments in that period. The machine learning algorithms modelled were DT, SVC, and KNN, and the accuracy obtained was in the range 72–74%, which is lower compared to what we achieved with the CHAID algorithm. This may be due to the fact that we used grades as input data to train the models instead of activity data.

Regarding our second research hypothesis, we have shown, based on the predictor importance, that the instructor can receive feedback during the school year and proceed to certain class improvement acts. According to the analysis given above we found that the most important predictor factor for predicting the final grade of students is Assignment2. An Informatics instructor using a similar teaching methodology can rely on that information to focus his efforts on upgrading programming material and ER implementations for the lessons covering that homework assignment.

Our findings on the predictor factor importance are in tandem with the results reported by Xu et al. [57], who also found that assignment-related features in an online learning environment, which is part of the learning activities of blended learning like flipped classroom, were important predictor factors in predicting student performance. Such results generally suggest that knowing at an early stage of the course which factors influence student performance, allows the instructor to adjust his teaching strategies immediately, to improve the course as it is being taught [71].

According to our experimental study, demographic attributes do not seem to be important factors that influence student final performance in a flipped classroom, and therefore does not provide any useful information to the instructor for course improvement. This finding is consistent with the results of the research of Egbedokun and Afolabi [72], who also found that the demographic variables such as age and gender have no significance influence on student motivation and academic performance of distance learners in the flipped classroom. On the other hand, according to the research of Swacha and Muszyńska [73], where data were acquired from two programming MOOCs to determine the impact of demographic features on student retention, they concluded that demographic data should be considered in the implementation of a PM in MOOCs. This difference in the findings may be due to the fact that demographic data is an important factor influencing student performance in some educational environments but not in all.

The most significant contribution of this work is that we proved that an accurate educational PM can be created for as few as 140 students. There is no known research work that focuses on the prediction of student performance in lower secondary education on computer programming with a teaching model like the one we used. This alone constitutes another important contribution of the presented work, i.e., the application of predictive modeling on a computer programming teaching environment that combines the flipped classroom, the jigsaw technique and educational robotics. Last contribution, but not least, is the recognition of the importance of prediction factors on the learning process, by proving through predictive modeling, that finding the factors which affect the prediction of student performance more, can help the educator to enhance teaching connected to those factors.

This study provides valuable insights into predicting student performance in Informatics courses in lower secondary education. At this point, it is crucial to discuss certain limitations in these findings. The most significant challenge is the small data sizes, which on one hand are typical in that level of education, forcing us to concentrate on assessing the effectiveness of machine learning algorithms in classroom settings of such sizes. On the other hand, we cannot but acknowledge that such data sizes might restrict the broader applicability of the results.

We must also mention that the dataset was sourced from only one secondary school and pertained solely to a specific Informatics course, which suggests that it might be useful to expand the PM's applicability to a wider educational context through combining data from multiple courses and schools. However, merging data from different Informatics courses and different schools is not straightforward, since separate teachers have their own methodology and techniques, and it is questionable whether such a conglomeration of data sets can produce a meaningful PM.

Nevertheless, if it succeeds, expanding the dataset in this way would not only fortify the model but also ensure its relevance and utility across various educational settings and instructional scenarios.

With these limitations in mind, future research could aim to collect data from a more diverse range of schools and courses, in order to increase data size and create a PM with a broader applicability.

5 CONCLUSIONS

In this work, we explored the possibility to predict the final performance of the students of the Music High School of Lamia in the course of Informatics at an early stage, in order to assist students in achieving a more comprehensive understanding of computer programming in the rest of the school year.

Our study was based on a teaching scenario in a certain framework and with certain data. The framework concerned the application of a novel teaching model with three components, i.e., the use of educational robotics, the application of the flipped classroom model using an asynchronous e-learning platform, and the implementation of a variation of the jigsaw technique. The data we used for modeling purposes were the grades from one third of the total quiz exercises and homework assignments that students had to take or submit during the school year, as well as demographic data.

We found that, despite the typically small size of data sets at the level of lower secondary education, it is possible to develop a quite accurate PM that will show us the highest percentage of success in a course implemented in a flipped classroom. In the development of the PM, we observed that the top five algorithms in performance, as evaluated through the classification accuracy metric, were CHAID, Quest, C5, BN, and RF. The CHAID machine learning algorithm had the best classification accuracy on the test set, compared against Quest, C5, BN and RF, despite the fact that RF exhibited a better performance than the others on the training set. This may be due to the smaller size of the test set compared to the training set, which suggests that the RF algorithm demonstrates better accuracy in classifying rather large than small sets of data. Among the rest of the algorithms, the CHAID algorithm proved to be the overall best to predict student performance, due to its highest AUC value.

We also found that the instructor can receive feedback during the school year to improve the course as it is being taught. More specifically, we found that the most important predictor factor for predicting the final grade of students was the second homework assignment. This means that the respective study material had a significant impact on the final grade and so we could assume that this material was most effective in helping students master the curriculum. On the other hand, if an activity had little to no impact on student grade, it might be removed, or an option could be provided to explore why it did not affect student performance.

Thus, the findings of this research not only demonstrate the viability of using PMs to enhance educational strategies and outcomes in specific learning frameworks, but also underscore the nuanced role of various factors in educational success. By tailoring educational content and methodologies based on predictive analytics, educators can more effectively meet the diverse needs of their students. Overall, this study confirms the efficacy of using predictive models to improve educational outcomes and instructional strategies in informatics, engineering and other fields within flipped classrooms, offering actionable insights for educators to optimize their teaching approaches.

6 REFERENCES

[1] J. M. Sáez-López, M. Román-González, and E. Vázquez-Cano, "Visual programming languages integrated across the curriculum in elementary school: A two-year case study using 'scratch' in five schools," *Comput. Educ.*, vol. 97, pp. 129–141, 2016. https://doi.org/10.1016/j.compedu.2016.03.003

[2] R. J. Segura, F. J. del Pino, C. J. Ogáyar, and A. J. Rueda, "VR-OCKS: A virtual reality game for learning the basic concepts of programming," *Computer Applications in Engineering Education*, vol. 28, no. 1, pp. 31–41, 2020. https://doi.org/10.1002/cae.22172

[3] R. Kadar, N. Abdul Wahab, J. Othman, M. Shamsuddin, and S. B. Mahlan, "A study of difficulties in teaching and learning programming: A systematic literature review," *International Journal of Academic Research in Progressive Education and Development*, vol. 10, no. 3, 2021. https://doi.org/10.6007/IJARPED/v10-i3/11100

[4] E. Fatourou, N. C. Zygouris, T. Loukopoulos, and G. I. Stamoulis, "Teaching concurrent programming concepts using scratch in primary school: Methodology and evaluation," *International Journal of Engineering Pedagogy*, vol. 8, no. 4, pp. 89–105, 2018. https://doi.org/10.3991/ijep.v8i4.8216

[5] A. Ioannou and E. Makridou, "Exploring the potentials of educational robotics in the development of computational thinking: A summary of current research and practical proposal for future work," *Educ. Inf. Technol. (Dordr)*, vol. 23, no. 6, pp. 2531–2544, 2018. https://doi.org/10.1007/s10639-018-9729-z

[6] S. A. Castañeda Rincón, H. Moreno Gudiño, and R. de J. Gil Herrera, "Robotics for inclusive education: Combining active methodologies in a classroom," *Contemp. Educ. Technol.*, vol. 16, no. 3, p. ep522, 2024. https://doi.org/10.30935/cedtech/14939

[7] A. R. El Mohamad, "Educational robotics is a useful tool in education," *ResearchGate*, 2019. [Online]. Available: https://www.researchgate.net/publication/332401229

[8] B. Birgili, F. N. Seggie, and E. Oğuz, "The trends and outcomes of flipped learning research between 2012 and 2018: A descriptive content analysis," *Journal of Computers in Education*, vol. 8, no. 3, pp. 365–394, 2021. https://doi.org/10.1007/s40692-021-00183-y

[9] Ş. Kaya and A. Çebi, "Unleashing the potential of flipped learning in K–12: A review of experimental studies," *Journal of Research on Technology in Education*, pp. 1–15, 2025. https://doi.org/10.1080/15391523.2025.2456052

[10] S. R. Sobral, "Flipped classrooms for introductory computer programming courses," *International Journal of Information and Education Technology*, vol. 11, no. 4, pp. 178–183, 2021. https://doi.org/10.18178/ijiet.2021.11.4.1508

[11] D. C. D. van Alten, C. Phielix, J. Janssen, and L. Kester, "Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis," *Educational Research Review*, vol. 28, p. 100281, 2019. https://doi.org/10.1016/j.edurev.2019.05.003

[12] S. Beltozar-Clemente, O. Iparraguirre-Villanueva, J. Zapata-Paulini, and M. Cabanillas-Carbonell, "Changing mathematical Paradigms at the University level: Feedback from a flipped classroom at a Peruvian University," *International Journal of Engineering Pedagogy*, vol. 13, no. 6, pp. 76–89, 2023. https://doi.org/10.3991/ijep.v13i6.40763

[13] M. H. A. Hassan and N. A. Othman, "Flipped classroom approach in rigid body dynamics: A case study of five-semester observation," *International Journal of Engineering Pedagogy*, vol. 11, no. 1, pp. 87–94, 2021. https://doi.org/10.3991/ijep.v11i1.15005

[14] S. Jacques and T. Lequeu, "The attractiveness of reversing teaching forms feedback on an electrical engineering course," *International Journal of Engineering Pedagogy*, vol. 10, no. 3, pp. 21–34, 2020. https://doi.org/10.3991/ijep.v10i3.12361

[15] D. T. K. Ng, E. H. L. Ng, and S. K. W. Chu, "Engaging students in creative music making with musical instrument application in an online flipped classroom," *Educ. Inf. Technol. (Dordr)*, vol. 27, no. 1, pp. 45–64, 2022. https://doi.org/10.1007/s10639-021-10568-2

[16] J. Cui and S. Yu, "Fostering deeper learning in a flipped classroom: Effects of knowledge graphs versus concept maps," *British Journal of Educational Technology*, vol. 50, no. 5, pp. 2308–2328, 2019. https://doi.org/10.1111/bjet.12841

[17] A. Evseeva and A. Solozhenko, "Use of flipped classroom technology in language learning," *Procedia Soc. Behav. Sci*., vol. 206, pp. 205–209, 2015. https://doi.org/10.1016/j.sbspro.2015.10.006

[18] K. Al-Said, I. Krapotkina, F. Gazizova, and N. Maslennikova, "Distance learning: Studying the efficiency of implementing flipped classroom technology in the educational system," *Educ. Inf. Technol. (Dordr)*, vol. 28, no. 10, pp. 13689–13712, 2023. https://doi.org/10.1007/s10639-023-11711-x

[19] K. Smahi, O. Labouidya, and K. El Khadiri, "Enhancing online assessment quality in higher education: The design of moodle plug-in for personalized exam revision (PER)," *International Journal of Engineering Pedagogy*, vol. 15, no. 5, pp. 127–140, 2025. https://doi.org/10.3991/ijep.v15i5.55055

[20] N. Salma, "Collaborative learning: An effective approach to promote language development," *International Journal of Social Sciences & Educational Studies*, vol. 7, no. 2, 2020. https://doi.org/10.23918/ijsses.v7i2p57

[21] T. V. Dang, P. T. Gia Bao, V. D. Minh, and N. T. Thanh Tu, "The application of problem-based learning in soft skills courses: An experiment in classes with multidisciplinary students in Vietnam," *International Journal of Engineering Pedagogy (iJEP)*, vol. 15, no. 5, pp. 20–41, 2025. https://doi.org/10.3991/ijep.v15i5.53663

[22] D. Chopra, G. Kwatra, B. Bhandari, J. K. Sidhu, J. Rai, and C. D. Tripathi, "Jigsaw classroom: Perceptions of students and teachers," *Med. Sci. Educ*., vol. 33, no. 4, pp. 853–859, 2023. https://doi.org/10.1007/s40670-023-01805-z

[23] A. T. W. Yu, "Using jigsaw method to enhance the learning of research and consultancy techniques for postgraduate students," *Engineering, Construction and Architectural Management*, vol. 24, no. 6, pp. 1081–1091, 2017. https://doi.org/10.1108/ECAM-03-2016-0080

[24] R. Sell, R. Razdan, K. Kase, and T. Rüütmann, "The role of AI chatbots in engineering education: Experimental findings and implementation strategies," *International Journal of Engineering Pedagogy*, vol. 15, no. 5, pp. 4–19, 2025. https://doi.org/10.3991/ijep.v15i5.56681

[25] R. Galici, T. Käser, G. Fenu, and M. Marras, "How close are predictive models to teachers in detecting learners at risk?" in *UMAP 2023 – Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization,* Association for Computing Machinery, Inc, 2023, pp. 135–145. https://doi.org/10.1145/3565472.3595620

[26] N. T. Nghe, P. Janecek, and P. Haddawy, "A comparative analysis of techniques for predicting academic performance," in *2007 37th Annual Frontiers in Education Conference – Global Engineering: Knowledge without Borders, Opportunities Without Passports*, IEEE, 2007, pp. T2G-7–T2G-12. https://doi.org/10.1109/FIE.2007.4417993

[27] IBM, "What is predictive analytics?" n.d. Accessed: Oct. 19, 2023. [Online]. Available: https://www.ibm.com/topics/predictive-analytics

[28] G. Shmueli, "To explain or to predict?" *Statistical Science*, vol. 25, no. 3, pp. 289–310, 2010. https://doi.org/10.1214/10-STS330

[29] W. He, "Examining students' online interaction in a live video streaming environment using data mining and text mining," *Comput. Human Behav*., vol. 29, no. 1, pp. 90–102, 2013. https://doi.org/10.1016/j.chb.2012.07.020

[30] S. Pal, "Mining educational data to reduce dropout rates of engineering students," *International Journal of Information Engineering and Electronic Business*, vol. 4, no. 2, pp. 1–7, 2012. https://doi.org/10.5815/ijieeb.2012.02.01

[31] J. Wong *et al.*, "Educational theories and learning analytics: From data to knowledge: The whole is greater than the sum of its parts," in *Utilizing Learning Analytics to Support Study Success*, Springer International Publishing, 2019, pp. 3–25. https://doi.org/10.1007/978-3-319-64792-0_1

[32] S. Shen, B. Mostafavi, T. Barnes, and M. Chi, "Exploring induced pedagogical strategies through a markov decision process framework: Lessons learned," *Journal of Educational Data Mining*, vol. 10, no. 3, pp. 27–68, 2018. https://doi.org/10.5281/zenodo.3554714

[33] S. Edwards and Z. Li, "Applying recent-performance factors analysis to explore student effort invested in programming assignments," in *International Conference on Frontiers in Education: Computer Science and Computer Engineering (FECS)*, Athens: CSREA, 2018, pp. 3–10. Accessed: Jan. 15, 2023. [Online]. Available: https://www.proquest.com/openview/1344abae126cd4240dfdce3764087786/1?pq-origsite=gscholar&cbl=1976352

[34] Á. Hernández-García, E. Acquila-Natale, J. Chaparro-Peláez, and M. Conde, "Predicting teamwork group assessment using log data-based learning analytics," *Comput. Human Behav*., vol. 89, pp. 373–384, 2018. https://doi.org/10.1016/j.chb.2018.07.016

[35] A. Cano and J. D. Leonard, "Interpretable multiview early warning system adapted to underrepresented student populations," *IEEE Transactions on Learning Technologies*, vol. 12, no. 2, pp. 198–211, 2019. https://doi.org/10.1109/TLT.2019.2911079

[36] J. L. Harvey and S. A. P. Kumar, "A practical model for educators to predict student performance in K-12 education using machine learning," in *2019 IEEE Symposium Series on Computational Intelligence, SSCI 2019*, Institute of Electrical and Electronics Engineers Inc., 2019, pp. 3004–3011. https://doi.org/10.1109/SSCI44817.2019.9003147

[37] D. Scaradozzi, L. Screpanti, and L. Cesaretti, "Machine learning for modeling and identification of educational robotics activities," in *2021 29th Mediterranean Conference on Control and Automation, MED 2021*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 753–758. https://doi.org/10.1109/MED51440.2021.9480309

[38] N. Bousbia and I. Belamri, "Which contribution does EDM provide to computer-based learning environments?" in *Studies in Computational Intelligence*, vol. 524, Springer Verlag, 2014, pp. 3–28. https://doi.org/10.1007/978-3-319-02738-8_1

[39] H. Mi, Z. Gao, Q. Zhang, and Y. Zheng, "Research on constructing online learning performance prediction model combining feature selection and neural network," *International Journal of Emerging Technologies in Learning*, vol. 17, no. 7, pp. 94–111, 2022. https://doi.org/10.3991/ijet.v17i07.25587

[40] J. Bravo-Agapito, C. F. Bonilla, and I. Seoane, "Data mining in foreign language learning," *WIREs Data Mining and Knowledge Discovery*, vol. 10, no. 1, pp. 1–16, 2020. https://doi.org/10.1002/widm.1287

[41] K. I. Chan, P. I. S. Lei, and P. C. I. Pang, "A literature review on educational data mining with secondary school data," in *ACM International Conference Proceeding Series, Association for Computing Machinery*, 2023. https://doi.org/10.1145/3599640.3599659

[42] A. Çetinkaya, Ö. K. Baykan, and H. Kırgız, "Analysis of machine learning classification approaches for predicting students' programming aptitude," *Sustainability (Switzerland)*, vol. 15, no. 17, p. 12917, 2023. https://doi.org/10.3390/su151712917

[43] OECD, "Education at a Glance 2024: OECD Indicators," 2024. https://doi.org/10.1787/c00cad36-en

[44] S. Hadjerrouit, "Exploring the effect of teaching methods on students' learning of school informatics," in *Proceedings of Informing Science & IT Education Conference (InSITE) 2015*, Tampa, Florida, 2015, pp. 201–219. https://doi.org/10.28945/2183

[45] A. Kullberg, Å. Ingerman, and F. Marton, "The variation theory of learning," in *Planning and Analyzing Teaching*, London: Routledge, 2024, pp. 20–29. https://doi.org/10.4324/9781003194903-3

[46] I. Koprinska, J. Stretton, and K. Yacef, "Predicting student performance from multiple data sources," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer, Verlag, 2015, pp. 678–681. https://doi.org/10.1007/978-3-319-19773-9_90

[47] A. Almalawi, B. Soh, A. Li, and H. Samra, "Predictive models for educational purposes: A systematic review," *Big Data Cogn. Comput.*, vol. 8, no. 12, p. 187, 2024. https://doi.org/10.3390/bdcc8120187

[48] S. Van Goidsenhoven, D. Bogdanova, G. Deeva, S. Vanden Broucke, J. De Weerdt, and M. Snoeck, "Predicting student success in a blended learning environment," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, 2020, pp. 17–25. https://doi.org/10.1145/3375462.3375494

[49] RbtsInMath, "Content development for robotics applications in flipped learning literature review," 2025. Accessed: Apr. 10, 2025. [Online]. Available: https://www.rbtsinmath.eu/wp-content/uploads/2025/03/WP2.A1_Literature-Review_ENG.pdf

[50] E. M. Bakheet and A. M. Gravell, "Would flipped classroom be my approach in teaching computing courses: Literature review," in *2021 9th International Conference on Information and Education Technology, ICIET 2021*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 166–170. https://doi.org/10.1109/ICIET51873.2021.9419631

[51] J. López-Zambrano, J. A. L. Torralbo, and C. Romero, "Early prediction of student learning performance through data mining: A systematic review," *Psicothema*, vol. 33, no. 3, pp. 456–465, 2021. https://doi.org/10.7334/psicothema2021.62

[52] L. Sandra, F. Lumbangaol, and T. Matsuo, "Machine learning algorithm to predict student's performance: A systematic literature review," *TEM Journal*, vol. 10, no. 4, pp. 1919–1927, 2021. https://doi.org/10.18421/TEM104-56

[53] T. Agasisti and A. J. Bowers, "Data analytics and decision making in education: Towards the educational data scientist as a key actor in schools and higher education institutions," in *Handbook of Contemporary Education Economics*, J. Geraint, J. Jill, and A. Tommaso, Eds., Edward Elgar Publishing Limited, 2017, pp. 184–210. https://doi.org/10.4337/9781785369070.00014

[54] K. Kim, H.-S. Kim, J. Shim, and J. S. Park, "A study in the early prediction of ICT literacy ratings using sustainability in data mining techniques," *Sustainability*, vol. 13, no. 4, p. 2141, 2021. https://doi.org/10.3390/su13042141

[55] D. A. Varsos, G. Dimitriou, and N. C. Zygouris, "The implementation of the flipped classroom model in the teaching of educational robotics: A study in secondary school students," in *2022 IEEE Global Engineering Education Conference (EDUCON)*, IEEE, 2022, pp. 1429–1436. https://doi.org/10.1109/EDUCON52537.2022.9766812

[56] E. Vives, C. Poletti, A. Robert, F. Butera, and P. Huguet, "Learning with jigsaw: A systematic review gathering all the pieces of the puzzle more than 40 years later," *Rev. Educ. Res.*, 2024. https://doi.org/10.3102/00346543241230064

[57] Z. Xu, H. Yuan, and Q. Liu, "Student performance prediction based on blended learning," *IEEE Transactions on Education*, vol. 64, no. 1, pp. 66–73, 2021. https://doi.org/10.1109/TE.2020.3008751

[58] P. R. Perez, C. Romero, and S. Ventura, "A java desktop tool for mining moodle data," in *4th International Conference on Educational Data Mining*, Eindhoven: Technische Universiteit Eindhoven, 2011, pp. 319–320. Accessed: Jan. 22, 2023. [Online]. Available: https://educationaldatamining.org/EDM2011/wp-content/uploads/proc/edm11_proceedings.pdf

[59] P. M. Moreno-Marcos, P. J. Muñoz-Merino, C. Alario-Hoyos, I. Estévez-Ayres, and C. Delgado Kloos, "Analyzing the predictive power for anticipating assignment grades in a massive open online course," *Behavior and Information Technology*, vol. 37, nos. 10–11, pp. 1021–1036, 2018. https://doi.org/10.1080/0144929X.2018.1458904

[60] P. M. Moreno-Marcos, T. C. Pong, P. J. Munoz-Merino, and C. D. Kloos, "Analysis of the factors influencing learners' performance prediction with learning analytics," *IEEE Access*, vol. 8, pp. 5264–5282, 2020. https://doi.org/10.1109/ACCESS.2019.2963503

[61] E. Alyahyan and D. Düştegör, "Predicting academic success in higher education: Literature review and best practices," *Int. J Educ. Technol. High. Educ*. vol. 17, no. 3, 2020. https://doi.org/10.1186/s41239-020-0177-7

[62] D. P. Kroese, Z. I. Botev, T. Taimre, and R. Vaisman, *Data Science and Machine Learning*, 1st ed. New York, NY: Chapman and Hall/CRC, 2019. https://doi.org/10.1201/9780367816971

[63] J. Nouri, M. Saqr, and U. Fors, "Predicting performance of students in a flipped classroom using machine learning: Towards automated data-driven formative feedback," *Systemics, Cybernetics and Infor-Matics*, vol. 17, no. 2, pp. 17–21, 2019. Accessed: Jan. 15, 2023. [Online]. Available: https://www.iiisci.org/journal/sci/FullText.asp?var=&id=EB614LI19

[64] E. D. Evangelista, "A hybrid machine learning framework for predicting students' performance in virtual learning environment," *International Journal of Emerging Technologies in Learning*, vol. 16, no. 24, pp. 255–272, 2021. https://doi.org/10.3991/ijet.v16i24.26151

[65] M. O. Edeh *et al.*, "Bootstrapping random forest and CHAID for prediction of white spot disease among shrimp farmers," *Sci. Rep.*, vol. 12, no. 1, p. 20876, 2022. https://doi.org/10.1038/s41598-022-25109-1

[66] S. Huang and J. Wei, "Student performance prediction in mathematics course based on the random forest and simulated annealing," *Sci. Program*, vol. 2022, pp. 1–9, 2022. https://doi.org/10.1155/2022/9340434

[67] IBM Corporation, "IBM SPSS Modeler 18.3 Algorithms Guide," 2021. [Online]. Available: http://www.ibm.com/spss

[68] M. Ramaswami and R. Bhaskaran, "A CHAID based performance prediction model in educational data mining," *IJCSI International Journal of Computer Science Issues*, vol. 7, no. 1, 2010. [Online]. Available: www.IJCSI.org

[69] I. Peraić and A. Grubišić, "Predicting academic performance of students in a computer programming course using data mining," *International Journal of Engineering Education*, vol. 39, no. 4, pp. 836–844, 2023. [Online]. Available: https://www.researchgate.net/publication/372237042

[70] J. A. Martínez-Carrascal, D. Márquez Cebrián, T. Sancho-Vinuesa, and E. Valderrama, "Impact of early activity on flipped classroom performance prediction: A case study for a first-year Engineering course," *Computer Applications in Engineering Education*, vol. 28, no. 3, pp. 590–605, 2020. https://doi.org/10.1002/cae.22229

[71] Z. Balogh and M. Kuchárik, "Predicting student grades based on their usage of LMS moodle using Petri nets," *Applied Sciences (Switzerland)*, vol. 9, no. 20, p. 4211, 2019. https://doi.org/10.3390/app9204211

[72] A. O. Egbedokun and O. M. Afolabi, "Influence of distance learners' demographic variables on motivation and performance in flipped classroom," *Asian Journal of Education and Social Studies*, vol. 40, no. 4, pp. 1–9, 2023. https://doi.org/10.9734/ajess/2023/v40i4878

[73] J. Swacha and K. Muszyńska, "Predicting dropout in programming MOOCs through demographic insights," *Electronics (Switzerland)*, vol. 12, no. 22, p. 4674, 2023. https://doi.org/10.3390/electronics12224674

7 AUTHORS

Dimitrios A. Varsos is a PhD candidate in the Department of Computer Science and Telecommunications at the University of Thessaly, focusing on Teaching of Computer Science and ICT in Secondary Education. He holds an Integrated Master's

Degree in Electrical and Computer Engineering and a MSc in Informatics and Computational Biomedicine. He is currently a Computer Science teacher in the Music School of Lamia, Greece. His interests include Teaching methods, Educational Technology, Educational Robotics, Programming, Artificial Intelligence, Machine Learning (E-mail: dvarsos@uth.gr).

Nikolaos C. Zygouris completed his undergraduate studies in Psychology at the University of Crete and obtained a Ph.D. in Clinical Neuropsychology from the Department of Special Education at the University of Thessaly. He pursued advanced training in repetitive Transcranial Magnetic Stimulation (rTMS) at the Temerty Center for Therapeutic Brain Intervention, Centre for Addiction and Mental Health, Toronto, Canada, and in the use of Event-Related Potentials (ERPs) at the EBneuro Academy in Florence, Italy. He has previously served as a Research Associate at the Neuropsychology Laboratory of the University of Thessaly. He currently serves as an Associate Professor in Neuropsychological Assessment of Children and Adolescents, with a specialization in Informatics and New Technologies, at the Department of Computer Science and Telecommunications, University of Thessaly. He is also the Director of the Laboratory of Digital Neuropsychological Assessment (edu.cs.uth.gr). His research interests include the application of electrophysiological methods for brain function assessment and rehabilitation, as well as the design of digital tools and web-based platforms for the evaluation and intervention of cognitive functions, learning difficulties, ADHD, anxiety, and mood disorders. He has authored or co-authored over one hundred scientific publications in peer-reviewed journals, conference proceedings, and academic volumes.

Kostas M. Kolomvatsos received his B.Sc. in Informatics from the Department of Informatics at the Athens University of Economics and Business, his M.Sc. in Computer Science – New Technologies in Informatics and Telecommunications and his Ph.D. from the Department of Informatics and Telecommunications at the National and Kapodistrian University of Athens (UoA). Currently, he serves as an Associate Professor in the Department of Informatics and Telecommunications, University of Thessaly. He is the founder of the Intelligent Pervasive Systems (iPRISM) research group (http://www.iprism.eu/) and a research collaborator of the Pervasive Computing Research Group, Department of Informatics and Telecommunications, University of Athens and the Essense research group, School of Computing Science, University of Glasgow. He was a Marie Sklodowska-Curie Research Fellow (Individual Grant) in the MSCA-2016 call. He has participated in and coordinated several European and National research projects. His research interests are in the definition of Intelligent Systems and techniques adopting Artificial Intelligence, (Deep) Machine Learning, Computational Intelligence and Soft Computing for Pervasive (Data/Edge) Computing, Distributed Systems, Internet of Things, and the management of Large Scale Data. In these areas, he has published over 170 articles.

Antonios N. Dadaliaris is an Assistant Professor in the Department of Informatics and Telecommunications at the University of Thessaly. He obtained his degree in Computer, Telecommunications and Network Engineering in 2005, his master's degree in Computer Science and Technology in 2008, and his PhD from the same department in 2012. His research interests primarily focus on floorplanning and placement of integrated circuits, as well as the broader field of VLSI physical design. Additionally, he is interested in the development and implementation of EDA/CAD tools within standard integrated circuit design flows.

Georgios Dimitriou received a PhD degree from the University of Illinois, USA. Since September of 2001 he has been teaching undergraduate and graduate courses at the Departments of Electrical and Computer Engineering, and Informatics and Telecommunications of the University of Thessaly. Since October 2017 he holds an Assistant Professor position at the Department of Informatics and Telecommunications of the University of Thessaly. He is currently the Director of the Computer Architecture, Compilers and Security Lab of the department. His research interests include computer architecture, focusing on processors and parallel systems, compilers, focusing on high-level synthesis and optimizations, and also teaching methods of informatics and engineering (E-mail: dimitriu@uth.gr).