Artificial intelligence-enabled prediction model of student academic performance in online engineering education

Pengcheng Jiao¹,² · Fan Ouyang³ · Qianyun Zhang⁴ · Amir H. Alavi⁴,⁵

Published online: 11 August 2022
© The Author(s) 2022

Abstract
Online education has been facing difficulty in predicting the academic performance of students due to the lack of usage of learning process, summative data and a precise prediction of quantitative relations between variables and achievements. To address these two obstacles, this study develops an artificial intelligence-enabled prediction model for student academic performance based on students’ learning process and summative data. The prediction criteria are first predefined to characterize and convert the learning data in an online engineering course. An evolutionary computation technique is then used to explore the best prediction model for the student academic performance. The model is validated using another online course that applies the same pedagogy and technology. Satisfactory agreements are obtained between the course outputs and model prediction results. The main findings indicate that the dominant variables in academic performance are the knowledge acquisition, the participation in class and the summative performance. The prerequisite knowledge tends not to play a key role in academic performance. Based on the results, pedagogical and analytical implications are provided. The proposed evolutionary computation-enabled prediction method is found to be a viable tool to evaluate the learning performance of students in online courses. Furthermore, the reported genetic programming model provides an acceptable prediction performance compared to other powerful artificial intelligence methods.

Keywords
Higher education · Online learning · Collaborative learning · Online engineering education · Prediction model · Artificial intelligence

1 Introduction

With the rapid development of online education in recent years, research direction has been shifted to using data-driven learning prediction models to obtain new insights on how students learn and how to improve their learning performance (Picciano 2014; Siemens and
Baker 2012). With the development of educational data mining (Ahmad et al. 2015; Baker and Yacef 2009) and learning analytics (Siemens and Long 2011), relevant data cleaning, mining, and analytics techniques have been used to understand, report, and optimize online learning and learning environments. For example, the applications of the data-intensive approaches have been used to predict student exam performance (Agudo-Peregrina et al. 2014), create learning prediction models (Paquette et al. 2015), and develop feedback dashboards (Jivet et al. 2018). Those applications have significantly improved the understanding over students’ learning process, performance, and context in online education (Ouyang et al., 2022; Verbert et al. 2012).

Academic performance prediction is one of the most important aspects in online education, which is typically conducted to estimate students’ learning performance using learning information and artificial intelligence (AI) algorithms (Tomasevic et al. 2020). Different types of AI algorithms have been used to develop prediction models, e.g., evolutionary computation (Fung et al. 2004; Takagi 2001), deep learning (Fok et al. 2018), decision trees (Kabra and Bichkar 2011), and Bayesian network (Sharabiani et al. 2014). However, the existing prediction models experience challenges in obtaining quantitative relations between the inputs (i.e., learning information and data) and outputs (i.e., academic performance) due to the following two dilemmas. First, there is a lack of criteria for selecting and transforming learning data (including process and performance data) into explainable parameters due to the complexity of teaching and learning contexts and processes (Hussain et al. 2019; Madhavan and Richey 2016; Uccio et al. 2020). Second, difficulties are found in obtaining high precisions of the relations between the learning inputs and performance outputs (Chassignol et al. 2018; Godwin and Kirn 2020; Tomasevic et al. 2020). As a consequence, it is of research interest to address the challenges and develop accurate performance prediction models in online education contexts.

To address those two gaps, this study uses the advanced AI technique—evolutionary computation (EC)—to develop a prediction model of student academic performance and test the precision of the model. We first identify students’ learning data from the entire learning process and establish the specific criteria to define the variables for characterizing the learning processes. Next, we develop a quantitative prediction model using a robust branch of EC namely genetic programming (GP). The prediction model accurately and efficiently predicts the students’ academic performance. Finally, the analytical and pedagogical implications are provided using the empirical research results to guide the development of performance prediction models and design of online engineering courses. The original contributions of this study can be drawn as:

- Processing learning process and summative data and establishing specific criteria for selecting these data for prediction;
- Developing an AI model to predict students’ academic performance; and
- Design guidance and assistance for online engineering education.

2 Literature review

2.1 Existing studies on academic performance prediction

Academic performance prediction is critical for online education since it helps identify students who are likely to fail, provide student-centered learning pathways, and optimize
Artificial intelligence-enabled prediction model of student…

instructional design and development (Asif et al. 2017; Chen et al. 2020; McArthur et al. 2005; Mozer et al. 2019; Roll and Wylie 2016). Different AI algorithms have been used in the existing studies to predict students’ examination performance using classification and regression (Tomasevic et al. 2020). For example, Kotsiantis et al. (2003) applied multiple machine learning (ML) techniques (e.g., Naïve Bayes, k-nearest neighbors) to categorize students in “pass” or “fail”. Minaei-Bidgoli et al. (2003) used several learning algorithms to classify the student results into different categories, including (a) “pass” or “fail”, (b) high, middle and low levels, and (c) nine classes based on the achieved grades. Marquez-Vera et al. (2012) used genetic programming and other data mining algorithms to predict student failure at school. Marquez-Vera et al. (2015) conducted a prediction model for early dropout in high school using data mining methods. More recently, Cano and Leonard (2019) developed an early warning system to underrepresented student populations. In the domain of linear regression, the research problem is predicting the exact scores that students may earn (Drucker et al. 1997). In summary, most of the existing studies have focused on the performance classification and the regression problems of identifying explicit scores.

Other than applying AI algorithms for the classification and regression, AI-enabled prediction models have been developed to predict academic performance based on the specific input variables that can characterize student learning. A review has been carried out to summarize the performance prediction models into three categories: the similarity-based, model-based and probabilistic approaches (Tomasevic et al. 2020). The review highlighted that the two critical components to develop the prediction models were (a) identifying variables that can characterize the learning processes and performances, and (b) analyzing the learning data using the appropriate AI algorithms. However, there are gaps in the current development of prediction models related to the data identifications and data analytics. First, regarding the data identification, researchers tend to consider all the available student information data (e.g., age, gender, religion, place of living, job, grades, etc.) in the prediction models (Asif et al. 2017), rather than using the data that reflect the specific learning process (Suthers and Verbert 2013). In other words, the identification of student data in most of the existing prediction models is not underpinned by specific standards that characterize the students’ learning process. For example, the most frequently used input data in the prediction models include students’ prior performance, engagement level, and demographic information (Lam et al. 1999; Nicholls et al. 2010; Tomasevic et al. 2020). However, the prediction results usually indicate that no classifier or variable plays a more significant role than the others in academic performance (Asif et al. 2017; Oskouei and Askari 2014; Yehuala 2015). One way to address the issue is to deliberately choose the student data that are underpinned by a learning theory in order to reflect the specific learning process. Because one of the goals of the performance prediction models is to optimize student-centered learning pathways, the choice of students’ input data should be guided by the student-centered learning principle and reflect the student-centered learning processes (Ouyang & Jiao, 2021). There are emerging studies that focus on using online learning behavior data from the process-oriented perspective to accurately predict academic performance, rather than merely using student information data (e.g., demographics) or performance data (e.g., final grades) (Bernacki et al. 2020). Echoing this research trend, this research designs a collaborative learning mode in online courses and deliberately chooses student data from the collaborative process to make academic performance predictions.

Second, regarding the data analytics, ML algorithms have been widely used in developing the prediction models for the academic performance of students, e.g., artificial neural networks (ANN), support vector machines (SVM), and decision trees (DT) (Chassignol
et al. 2018; Fernandes et al. 2019; Fok et al. 2018). The existing studies have dedicated the efforts on estimating the implicit correlations between the observed learning data and predicted performance (Tomasevic et al. 2020). The ANN prediction models typically consist of series of connected artificial neurons to simulate the neurons in the biological brains, which are critically affected by many factors, e.g., learning rate, objective function, weights initialization, etc. (Chen et al. 2020). SVM has been used to address the challenges in the classification of learning data, i.e., data regression, which is efficient in obtaining the nonlinear classification and implicitly mapping input variables into the higher dimensional space (Drucker et al. 1997). DT has been reported to develop the prediction models using the tree-shaped graph to model the possible decisions and the corresponding consequences (Chaudhury and Tripathy 2017). However, quantitative prediction models have not yet been developed using those AI algorithms to accurately identify the exact relations between the input variables in learning process and output academic performance. This study fills such gap by developing the EC model to explicitly represent the quantitative relations of multiple learning variables in order to predict students’ academic performances in online education.

2.2 Quantitative prediction using evolutionary computation (EC)

This study uses genetic programming (GP), a branch of evolutionary computation (EC), to develop the prediction model. Inspired by the evolution process in the natural world—Survival of the Fittest, EC has been reported as a powerful subdivision in the AI domain, which includes evolutionary strategies (ESs) and evolutionary programming (EP). These techniques are collectively known as evolutionary algorithms (EAs). The EAs are powerful tools to accurately address complex datasets and identify the quantitative relations between the input and output variables (Pena-Ayala 2014). GP is a specialization of EAs that offers high model transparency and knowledge extraction, leading to the conceptualization of the phenomena and derivation of mathematical structures for complex problems. GP evolves prediction models with tree-like structures, which can be recursively evaluated. Therefore, GP typically uses the programming languages that naturally embody tree structures. The nodes in such tree-like computer programs consist of operator functions while all the terminal nodes have operands. This unique model structure facilitates evolving and evaluating millions of mathematical expressions correlating the input variables to the output. GP contains the initial populations of arbitrary variables improved by a series of the genetic operators such as recombination, mutation, or reproduction (Saa 2016). The arbitrary variables are the encoded solutions given by the binary strings of numbers and assessed by certain fitness functions (Xing et al. 2015). The initial populations are randomly selected and the variables are analyzed based on the fitness function. The populations with the most efficient fitness are defined as the higher chance to become the parents for the next generation in GP (Timms 2016). Consequently, improving the initial population of GP algorithms is the key process to obtain the fittest solution with the most efficient convergence.

Compared with their counterparts in AI, GP has the advantages of producing highly nonlinear prediction functions without requirements on predefined existing relations of the variables (Xing et al. 2015). The prediction models evolved by traditional GP are typically in the tree-shaped structures and programmed in the functional programming language (e.g., LISP) (Zaffer et al. 2017). Therefore, GP can be used to develop the quantitative prediction models to establish the exact relations between the input variables and output response in predicting the learning performance of students (Martin and Betser 2020).
However, the application of GP in online education to obtain the quantitative prediction models has yet to be exploited, especially by establishing certain criteria to analyze and convert learning processes into input data. GP and other AI methods are based on the data alone to develop the structure of the models. This implies the important role of collecting datasets with wide ranges for the predictor variables to develop more robust models. For cases where the databases are not large enough, advanced validation methods (e.g., k-fold cross validation) can be deployed to verify the efficacy of these methods.

2.3 Requirements for the quantitative prediction models in online education

The main challenges in this study is to define the main requirements of the GP prediction model (Nguyen et al. 2020). The prediction models should be able to accurately predict the learning performance of students with respect to the five considerations, i.e., interpretability, accuracy, speed, robustness and scalability. In particular, interpretability refers to the criteria established to interpret the learning data from the online courses, which serves as the fundamental stone of the AI predication models. Accuracy refers to the correctness of the AI models in predicting the academic performance of students, which can typically be validated with predication results from other models. Speed refers to the low computational cost used to obtain the prediction results, which is particularly important to the AI models that are designed to address the real-time learning data (e.g., learning-feedback-adjusting process in online education). Robustness refers to the reliability of predicting the learning performance of students from noisy data, which is critical in effectively mining the learning data. Scalability refers to the capability of obtaining prediction results from large volume of multimodal learning data. The prediction model developed in this study aims to achieve these five characteristics. In addition, previous studies usually used the same student dataset to conduct the cross-validation in order to evaluate the prediction results; in other words, the same student cohort was used to develop and validate the prediction models (Asif et al. 2017). But from an educational perspective, it is more appropriate to build the prediction model using the dataset from one group of students, and evaluate the five characteristics of the model by another group of students (Asif et al. 2017). To achieve a more accurate assessment, this study takes different datasets to build and evaluate the prediction model.

3 Research methodology

3.1 Research context, participants and course design

The research context is an online engineering course—Smart Marine Metastructures—offered in Spring 2020 (8 weeks) in the Ocean College at Zhejiang University in China. Smart Marine Metastructures was offered by the first author as a completely online course hosted through the Blackboard online platform (XueZaiZheDa http://course.zju.edu.cn/) and DingTalk (China’s version of Zoom), where the course syllabus, materials, and other resources were uploaded and recorded. Participants were 35 full-time graduate students from the Ocean College at the university; they were all Chinese aged from 22 to 27 years old (female: N = 11; male: N = 24). Prerequisite courses were required for basic background knowledge in structural analysis and artificial intelligence.
This course was designed as a collaborative learning course by the interdisciplinary research team. Grounded upon the social perspectives of learning (Vygotsky 1978), collaborative learning is defined as a small group of people participate in coordinated activities to maintain mutual understandings of problems, to advance joint meaning-making, and to create new knowledge or relevant artifacts (Dillenbourg 1999; Goodyear et al. 2014; Roschelle and Teasley 1995). In engineering education, learning to be an engineer means learning to participate in engineering discourses: the words, discourses, and narratives through which engineers think and communicate (Betser and Martin 2018; Rojas 2001). As a student-centered learning, collaborative learning can foster engineering learners’ cognitive thinking during the intrapersonal process as well as knowledge construction through social interactions (Damşa 2014; Liu and Matthews 2005; Ouyang and Chang 2019). This research followed the ethical, legal requirements from the university’s research ethics committee. All the students were well informed and agreed to participate in the experiments reported in this study.

Following the collaborative learning mode, the instructor (the first author) designed the three components in this course, including the online lecture, group discussion, and literature review writing. The first component, namely the online lecture, included the basic introduction and advanced learning. The instructor first introduced the main concepts and theories of the field in the first two weeks. Then, during the advanced learning process in the following six weeks, the instructor introduced three gradually-advanced content, i.e., advanced marine metastructures, artificial intelligence in engineering, and structural health monitoring in marine engineering. The former part served as the prerequisite knowledge for the latter part, such that students could progressively build their understandings of the course content. The second component was the group discussion, in which students explored the concepts, exchanged understandings, and constructed meanings in small groups through DingTalk. Students autonomously formed five groups (seven students/group) based on their research interests (see Table 1). The instructor provided prompting questions to guide students’ inquiry of research topics in advance; and the instructor was not engaged in the discussion process. The third component was the group work of literature review writing completed by the groups. The writing process included four steps: first, developing the outline of literature review, second, writing the initial drafts, third, make revisions based on the instructor’s feedback, and finally finalizing the literature review. Students in the same group were graded as the same grade. At the end of the semester, online oral presentations were delivered to the class by the small groups based on their write-up of the literature review. Students made peer-assessment of the oral presentation based on a rubric. The instructor evaluated the groups’ literature review write-up based on a rubric. After the course, the students made a self-reflection to evaluate their acquisition of knowledge in this course comparing to their knowledge prior the course.

3.2 Research purpose and questions

The research purpose is to obtain a quantitative prediction model to predict the learning performance of the students in online learning, which can be used to analyze the contributions of the input variables to the academic performance, and therefore, optimize the design of online courses based on the results. Taken together the existing studies and the requirements for the prediction models, this study addresses the two research challenges,
| Group | Number of students | Research background of participant | Group discussion topic | Write-up topic |
|-------|-------------------|-----------------------------------|------------------------|---------------|
| 1     | 7                 | Ocean Engineering and Fluid Mechanics | Mechanical metamaterials | Mechanical Metamaterials: Principle and Potential Applications in Marine Structures |
|       |                   |                                   | Marine structures       |               |
| 2     | 7                 | Ocean Engineering and Fluid Mechanics | Piezoelectric sensors  | Piezoelectric Sensing Technologies in Ocean Engineering |
|       |                   |                                   | Ocean engineering       |               |
| 3     | 7                 | Ocean Engineering and Fluid Mechanics | Structural design      | Structural Design and Optimization in Advanced Marine Equipment |
|       |                   |                                   | Structural optimization |               |
|       |                   |                                   | Advanced marine equipment |               |
| 4     | 7                 | Information Engineering and Computer Science | Machine learning | Machine Learning (ML)-Based Monitoring Data Analysis in Ocean Engineering |
|       |                   |                                   | Data analysis          |               |
|       |                   |                                   | Ocean engineering      |               |
| 5     | 7                 | Information Engineering and Computer Science | 5G communication techniques | 5G-Enabled Smart Marine Structures |
|       |                   |                                   | Smart marine structures |               |
i.e., the learning data identifications due to the lack of criteria, and the learning data analytics due to the lack of the appropriate AI algorithms. The research questions can be summarized as:

1. How to identify the dominant learning variables that significantly affect the student learning performance?
2. How to develop the robust quantitative prediction model to predict students’ performance with a reasonable accuracy?
3. How to optimize the online course based on the performance prediction results generated by the model?

Resolving the research questions, we propose the characterization criteria to particularly identify the learning processes and obtain the learning data, and then use the learning data to develop the quantitative prediction model using the GP algorithm.

### 3.3 Identification of criteria, variables definition and preliminary data analytics

The criteria are defined to categorize and analyze the students’ learning results obtained from the 35 graduate students in the online ocean engineering course. Student performance variables are identified from both the summative and process perspectives, which include the pre-course prerequisite knowledge, participation performance, procedural performance, summative performance, and post-course knowledge acquisition. Students’ pre-course prerequisite knowledge and post-course knowledge acquisition are self-evaluated by students in terms of a questionnaire. This questionnaire includes ten questions; each question asks students to evaluate their knowledge level of one main topic covered in this course. All responses are measured with a 5-point scale, ranging from 1 point (do not understand the topic), 2 points (understand the topic but need external assistance for clear explanation), 3 points (can directly explain the topic without any assistance), 4 points (can directly explain the topic and its applications in the research and practice without any assistance), to 5 points (understand this topic, can elaborate its applications in research and practice, and are able to apply it in my own study). The student participation includes their participation frequency in the class discussions and in the group discussions. Procedural performance includes students’ performance in group discussions, write-ups and group presentations. Quantitative content analysis is used to evaluate students’ procedural performance. Their oral and written content from discussion, write-ups and presentations were recorded and transcribed into text and two trained raters code those text content in terms of a coding scheme (see Table 2). This coding scheme includes three levels of knowledge contribution, namely superficial-, medium-, and deep-level knowledge. A weighted score \( N_{SK} + 2N_{MK} + 3N_{DK} \) is calculated for each student as the procedural performance. The summative performance is the instructor’s evaluation of students’ final write-up of the literature review. Note that the final learning effectiveness of the students \( L_{neff} \) (i.e., the output variable in the GP prediction) is defined as the total grades of students.

Next, we analyze the learning data to define the variables for developing the GP prediction model. According to the academic performance criteria created in the previous section, a total of 8 input variables are used in the EC model, which are categorized into five types including the prerequisite knowledge (i.e., background of students \( PK_s \)), the participation frequency in the class and group discussions (i.e., \( Par_{class} \) and \( Par_{group} \)), the procedural
Table 2  The content analysis coding scheme (Ouyang and Chang 2019)

| Code                      | Level | Description                                                                 |
|---------------------------|-------|-----------------------------------------------------------------------------|
| Superficial-level Knowledge (SK) | 1     | A participant explores information related to discussion topics, without explicit statements of his/her own ideas, arguments or perspectives |
| Medium-level Knowledge (MK) | 2     | A participant presents his/her own ideas, arguments, or perspectives without detailed elaborations, supports of resources, statistics or personal experiences |
| Deep-level Knowledge (DK)  | 3     | A participant explicitly elaborates his/her own ideas, arguments, or perspectives with detailed explanations, supports of resources, statistics or personal experiences |

A final weighted score was calculated as $N_{SK} + 2N_{MK} + 3N_{DK}$. $N$ represents *number*, namely the frequency of the occurrence of a code in the oral and written content from discussion, write-ups and presentations. Parameters of 1, 2, and 3 are assigned to SK, MK, and DK to represent the gradually increasing weights.
performance (i.e., discussion performance Perf\textsubscript{dis}, write-up performance Perf\textsubscript{write}, and presentation performance Perf\textsubscript{prez}), the summative performance (i.e., summative evaluation of the final write-up Perf\textsubscript{sum}), and the knowledge acquisition (i.e., self-evaluation of knowledge acquisition after the course KA\textsubscript{kn}) (see Table 3).

We use a traditional linear regression approach to make preliminary data analytics. It can be seen that similar performance (i.e., patterns that either evenly distributed or critically fluctuated) are observed on every variable for the five groups, which indicates that the students in the entire class are learning with similar effectiveness (see Fig. 1). However, the results show that the linear regression analysis approach is ineffective in analyzing the complex learning data in the online engineering course (Multiple $R = 0.479$, $R^2 = 0.230$, Std. error = 6.240). According to the linear regression of Lrn\textsubscript{eff} in the five groups, the deviation of the first group G1 is the smallest while the fifth group G5 is the largest. Consequently, a more advanced approach is needed to develop the quantitative prediction model. In next section, an advanced AI technique, i.e., GP, is applied to develop the predictive model for the learning data.

| Table 3 Input and output variables |
|-----------------------------------|
| Role | Type | Variable | Symbol |
| Input | Prerequisite knowledge | Background | PK\textsubscript{bg} |
| Participation frequency | In class discussions | Par\textsubscript{class} |
| | In group discussions | Par\textsubscript{group} |
| Procedural performance | Discussion | Perf\textsubscript{dis} |
| | Write-up | Perf\textsubscript{write} |
| | Presentation | Perf\textsubscript{prez} |
| Summative performance | Instructor’s summative evaluation of the final write-up | Perf\textsubscript{sum} |
| Knowledge acquisition | Self-evaluation of knowledge acquisition after the course | KA\textsubscript{kn} |
| Output | Learning effectiveness | Final learning performance | Lrn\textsubscript{eff} |

4 Development of the performance prediction model using the GP algorithm

4.1 Definition of the GP model

According to the previous discussion, the output of the prediction models is defined as the learning effectiveness Lrn\textsubscript{eff} of the graduate students in the online engineering course. The input variables are classified into five categories, i.e., the prerequisite, participation, procedural performance, summative performance, and knowledge acquisition, as summarized in Table 3. As a consequence, the quantitative prediction model (i.e., objective function) can be written as
Artificial intelligence-enabled prediction model of student learning effectiveness

where $f$ represents the unknown highly nonlinear relationships between the output and inputs, which can only be determined using the AI techniques. The GP prediction model presented in Eq. (1) is defined in terms of 8 input variables, that includes both process and summative learning data.

$$L_{\text{rn\_eff}} = f(PK_{bg}, Par_{\text{class}}, Par_{\text{group}}, Perf_{\text{dis}}, Perf_{\text{write}}, Perf_{\text{prez}}, Perf_{\text{sum}}, KA_{kn})$$ (1)

Fig. 1 The input and output variables of the student groups in the online engineering course. (All the input scores are normalized to the full score as 100 to ensure that the influence of certain variables will not be eliminated)
4.2 Training and testing of the GP model

The GP algorithm directly learns from the learning data and extracts the subtle functional relations between the independent input variables (see Table 3). The GP model efficiently considers the interactions between the output variable Lrn\textsubscript{eff} and the input variables. Due to the lack of the learning data (i.e., only 35 datapoints) in the online engineering course, the k-folder cross validation has been applied for the model development, where k = 5 in this study. A series of preliminary runs revealed the influences of PK\textsubscript{bg}, Par\textsubscript{class}, Par\textsubscript{group}, Perf\textsubscript{dis}, Perf\textsubscript{write}, Perf\textsubscript{prez}, Perf\textsubscript{sum}, and KA\textsubscript{kn} on improving the prediction performance of the GP model. Extensive preliminary analyses were performed to tune the GP parameters including the crossover rate, initial population size, mutation rate, program head size, etc. In particular, more than fifty different combinations of the factors were considered for deriving the best prediction GP model for Lrn\textsubscript{eff}. Details of training parameters and ranges used in this study are listed in Table 4. Parameter ranges were selected based on a trial study, and according to previous studies (Roy et al. 2010; Gandomi and Alavi 2012; Zhang et al. 2021). Three replications were conducted for every factor combination, and the GP algorithm was until no significant improvements was observed in Lrn\textsubscript{eff}. The simulations were conducted on a desktop computer with CPU: Intel® Xeon® CPU E5-1650 v4 @ 3.60 GHz, GPU: NVIDIA Quadro K420, RAM: 31.9 GB. The total time for training the current dataset was 15 min and 32 s for the optimal model. The best prediction model was selected from all models with highest accuracy and lowest loss. The gene trees of the best model are illustrated in Fig. 2. Individual gene expression and final simplified model are shown in Eqs. (2)-(9).

\begin{align*}
\text{Gene}_1 &= -6.5 \cos (\text{Perf}_{\text{prez}}), \quad (2) \\
\text{Gene}_2 &= -104200 \frac{\cos (\text{Perf}_{\text{sum}})}{\text{Perf}_{\text{dis}}^3}, \quad (3) \\
\text{Gene}_3 &= 7.1 \cos \left( \cos \left( \sqrt{\text{PK}_{\text{bg}}} \right) \right) \cos \left( \text{KA}_{\text{kn}} \cos \left( \text{Perf}_{\text{sum}} \right) \right), \quad (4)
\end{align*}

Table 4  GP training parameter settings

| MGGP training parameter combinations | Settings |
|-------------------------------------|----------|
| Function set                        | +, −, ×, ÷, log, sqrt, cubic, square, cos, sin, tanh, power |
| Population size                     | [15, 25, 50, 100] |
| Maximum number of genes allowed in an individual | [4, 6] |
| Maximum tree depth                  | [4, 6] |
| Tournament size                     | [15, 20] |
| Tournament type                     | Pareto (Probability = 1) |
| Elite fraction                      | [0.25, 0.5] |
Artificial intelligence-enabled prediction model of student…

The simplified GP model is given as:

\[
\text{Gene}_4 = -676.4 \cos (\cos (\log (\text{Par}_{\text{class}}))) \cos (\text{Perf}_{\text{prez}}),
\]

(5)

\[
\text{Gene}_5 = -0.9KA_{kn} \cos (\cos (\log (KA_{kn}))) \cos (KA_{kn}),
\]

(6)

\[
\text{Gene}_6 = 0.7\text{Par}_{\text{class}} \cos (KA_{kn}),
\]

(7)

and

\[
\text{Bias} = 754.5.
\]

(8)

The simplified GP model is given as:

\[
\text{Lrn}_{\text{eff}} = 754.5 - 6.5 \cos (\text{Perf}_{\text{prez}}) - \frac{104200 \cos (\text{Perf}_{\text{sum}})}{\text{Perf}_{\text{dis}}^3} \\
+ 7.1 \cos (\sqrt{PK_{bg}}) \cos (KA_{kn} \cos (\text{Perf}_{\text{sum}})) \\
- 676.4 \cos (\cos (\log (\text{Par}_{\text{class}}))) \cos (\text{Perf}_{\text{prez}}) \\
- 0.9KA_{kn} \cos (\cos (\log (KA_{kn}))) \cos (KA_{kn}) + 0.7\text{Par}_{\text{class}} \cos (KA_{kn}),
\]

(9)

Fig. 2 Individual gene trees for developed GP model
where $\text{Par}_{\text{group}}$ and $\text{Perf}_{\text{write}}$ are omitted due to the similarity in the learning data. Equation (2) defines the quantitative relations between the input variables and the learning effectiveness of the students in the online engineering course. It can be seen that the GP model is a complex combination of the variables and operators to predict $\text{Lrn}_{\text{eff}}$. The fittest solutions obtained by the GP model after controlling the millions of preliminary linear and nonlinear running via the evolutionary process.

In order to benchmark the prediction power of GP against other conventional AI methods, ANN and SVM prediction models are developed using the same database. MATLAB version 2019a is used to develop the ANN and SVM models. The model performance evaluation of GP, ANN and SVM models are carried out. Figure 3 presents the comparisons of the measured and predicted learning effectiveness using the GP, ANN and SVM prediction models. This figure also presents the correlation coefficient (R) and mean squared error (MSE) performance indexes for the model for the training and testing data. It is seen that GP outperforms SVM on the training and testing data. GP outperforms ANN on the training data and provides a comparable performance with ANN on the testing data. However, it should be noted that ANN suffer from some major shortcomings. The first issue is that the knowledge extracted by ANNs is stored in a set of weights that cannot be properly interpreted. In ANN-based simulations, the weights and biases are randomly assigned for each run. These assignments considerably change the performance of a newly trained network even all the previous parameter settings and the architecture are kept constant. This leads to extra difficulties in selection of optimal architecture and parameter settings. Moreover, the structure and network parameters of ANNs should be identified in advance. This is usually done through a time-consuming trial and error procedures. The GP method presented in this study overcomes these

![Fig. 3 Comparisons of the measured and predicted learning effectiveness using the GP and ANN models: (a) GP model on the training data, (b) ANN model on the training data, (c) SVM model on the training data, (d) GP model on the testing data, (e) ANN model on the testing data, and (f) SVM model on the testing data](image-url)
shortcomings by introducing the completely new characteristics and traits. One of the major distinctions of EC lies in its powerful ability to model the learning behavior without requesting prior form of the existing relationships. The numbers and combination of terms are automatically evolved during the model calibration in GP, which is different from that in ANNs. The other superiority of the proposed GP method over ANN and almost all other AI methods pertains to its ability to extract the complex functional relationships for the investigated system.

A parametric analysis is performed to ensure the robustness of the developed models. This analysis is based on varying one parameter within a practical range, while other parameters are kept constant value. Figure 4 presents the parametric analysis results for the learning effectiveness $L_{\text{rnen}}$ with respect to the input variables $PK_{bg}$, $Par_{class}$, $Perf_{dis}$, $Perf_{prez}$, $Perf_{sum}$, and $KA_{kn}$. It can be seen that $L_{\text{rnen}}$ is heavily sensitive to participation in class $Par_{class}$, summative performance $Perf_{sum}$ and knowledge acquisition $KA_{kn}$. As we can see, knowledge acquisition affects the learning effectiveness in a sinusoidal trend. The

![Fig. 4 Sensitivity analysis obtained using the GP prediction model for the learning effectiveness $L_{\text{rnen}}$ with respect to the variables of a $PK_{bg}$, b $Par_{class}$, c $Perf_{dis}$, d $Perf_{prez}$, e $Perf_{sum}$, and f $KA_{kn}$ $Par_{group}$ and $Perf_{write}$ are omitted)](image-url)
learning effectiveness is increasing with higher participations in class and summative performance. Also, increasing performance of presentation ($\text{Perf}_{\text{prez}}$) and discussion ($\text{Perf}_{\text{dis}}$) result in higher learning effectiveness. On the other hand, the learning effectiveness is decreasing with higher prerequisite knowledge ($\text{PK}_{bg}$).

To simplify the analysis while investigating the most dominant variables (i.e., $\text{Par}_{\text{class}}$, $\text{Perf}_{\text{sum}}$, and $\text{KA}_{kn}$) in the optimal design for the online education course, $\text{PK}_{bg}$, $\text{Perf}_{dis}$ and $\text{Perf}_{\text{prez}}$ are fixed as the mean values of the 35 students (i.e., $\text{PK}_{bg} = 43$, $\text{Perf}_{\text{dis}} = 55$, and $\text{Perf}_{\text{prez}} = 89$). Substituting the constants to Eq. (9), the GP prediction model is reduced to:

$$
\text{Lrn}_{\text{eff}} = 754.4 - 0.6 \cos(\text{Perf}_{\text{sum}}) + 7.1 \cdot 0.6 \cos(\text{KA}_{kn} \cos(\text{Perf}_{\text{sum}})) \\
- 676.4 \cos\left(\cos(\log \text{Par}_{\text{class}})\right)^{0.02} \\
- 0.9 \text{KA}_{kn} \cos\left(\cos(\log \text{KA}_{kn})\right) \cos(\text{KA}_{kn}) + 0.7 \text{Par}_{\text{class}} \cos(\text{KA}_{kn}),
$$

(10)

4.3 Validation of the GP model

To validate the proposed GP prediction model, we apply the model to another online course—*Information Technologies and Education*—designed by the same research team (taught by the second author), using the same pedagogy (i.e., collaborative learning mode) and technologies (i.e., *XueZaiZheDa* and *DingTalk*) as the online engineering course. The validation course is a graduate-level, 8-week course offered in 2020 summer semester by the Educational Technology (ET) program at the same university. This course focuses on learning theories, instructional design, educational technologies, emerging tools, and trending topics related to the application of information technologies in education. 19 graduate students (female: 10; male: 9) from the College of Education enrolled in this course. Learning data collection is the same as the online engineering course.

Since the validation course used the same pedagogy and technology, we use the learning effectiveness $\text{Lrn}_{\text{eff}}$ of the 19 students in this course to validate the proposed GP prediction model. Note that the learning effectiveness $\text{Lrn}_{\text{eff}}$ in the validation course was measured as the total grades of the students in the class. Figure 5 compares the learning effectiveness $\text{Lrn}_{\text{eff}}$ between the actual performance of the validation course and result generated by the GP model. It can be seen that the GP model accurately obtains the

![Fig. 5 Comparison of the learning effectiveness $\text{Lrn}_{\text{eff}}$ between the GP prediction model and the existing validation course](image-url)
distribution pattern of $Lr_{n\text{eff}}$ over the entire 19 students in the validation course. The maximum difference between the GP model and validation course is 5%, which demonstrates the accuracy and efficiency of the developed prediction model. In addition, the GP model accurately predicts the learning effectiveness of the students below the average of the validation course, i.e., $Lr_{n\text{eff}} \leq Lr_{n\text{eff}}^{\text{mean}}$. Therefore, the GP model is able to find the students with inadequate $Lr_{n\text{eff}}$; the information can be provided to the instructor for providing further interventions of those low-performing students.

4.4 Optimization of online learning using the GP model

Here, we study $Lr_{n\text{eff}}$ function in Eq. (10) to obtain the optimal online course design (i.e., the maximum learning effectiveness) for the online engineering education. Investigating the contributions of the dominant variables (i.e., $Par_{\text{class}}$, $Perf_{\text{sum}}$, and $KA_{\text{kn}}$) to $Lr_{n\text{eff}}$, we find the most important variable to affect the learning effectiveness. Therefore, $Lr_{n\text{eff}}$ of students can be effectively improved by increasing the most important variable, which provide helpful guidance to instructors in online engineering education.

The extremum of the $Lr_{n\text{eff}}$ function in Eq. (10) can be determined with respect to $Par_{\text{class}}$, $Perf_{\text{sum}}$, and $KA_{\text{kn}}$ as

\[
\begin{bmatrix}
\frac{\partial Lr_{n\text{eff}}}{\partial Par_{\text{class}}} = 0 \\
\frac{\partial Lr_{n\text{eff}}}{\partial Perf_{\text{sum}}} = 0 \\
\frac{\partial Lr_{n\text{eff}}}{\partial KA_{\text{kn}}} = 0
\end{bmatrix},
\]

and the Hessian matrix is used to determine the maximum $Lr_{n\text{eff}}$ as

\[
HM_{Lr_{n\text{eff}}} = \begin{bmatrix}
\frac{\partial^2 Lr_{n\text{eff}}}{\partial Par_{\text{class}}^2} & \frac{\partial^2 Lr_{n\text{eff}}}{\partial Par_{\text{class}} \partial Perf_{\text{sum}}} & \frac{\partial^2 Lr_{n\text{eff}}}{\partial Par_{\text{class}} \partial KA_{\text{kn}}} \\
\frac{\partial^2 Lr_{n\text{eff}}}{\partial Perf_{\text{sum}} \partial Par_{\text{class}}} & \frac{\partial^2 Lr_{n\text{eff}}}{\partial Perf_{\text{sum}}^2} & \frac{\partial^2 Lr_{n\text{eff}}}{\partial Perf_{\text{sum}} \partial KA_{\text{kn}}} \\
\frac{\partial^2 Lr_{n\text{eff}}}{\partial KA_{\text{kn}} \partial Par_{\text{class}}} & \frac{\partial^2 Lr_{n\text{eff}}}{\partial KA_{\text{kn}} \partial Perf_{\text{sum}}} & \frac{\partial^2 Lr_{n\text{eff}}}{\partial KA_{\text{kn}}^2}
\end{bmatrix}.
\]

Substituting Eq. (10) into Eqs. (11) and (12), however, we encounter the difficulty in analytically solving for the maximized $Lr_{n\text{eff}}$ due to the complexity nature of the objective function. As a consequence, numerical method is used to maximize $Lr_{n\text{eff}}$ by discretizing the objective function and variables. Figure 6 demonstrates the flowchart of maximizing the learning effectiveness function $Lr_{n\text{eff}}$ using the analytical and numerical methods. Eventually, we obtain the maximum learning effectiveness as $Lr_{n\text{eff}}^{\text{max}} = 132.3$, and the corresponding optimal variables are: $Par_{\text{class}} = 80$, $Perf_{\text{sum}} = 78.6$, and $KA_{\text{kn}} = 91.2$.

Figure 7 presents the distributions of the learning effectiveness $Lr_{n\text{eff}}$ with respect to $Par_{\text{class}}$, $Perf_{\text{sum}}$ and $KA_{\text{kn}}$. Figure 7a indicates the influences of $KA_{\text{kn}}$ and $Perf_{\text{sum}}$ on $Lr_{n\text{eff}}$ at $Par_{\text{class}} = 80$. It can be seen that $Lr_{n\text{eff}}$ is critically fluctuated with $KA_{\text{kn}}$, while it is not significantly affected by $Perf_{\text{sum}}$ to the same level. Therefore, we find that $KA_{\text{kn}}$ plays a much more significant role on $Lr_{n\text{eff}}$. Figure 7b indicates the influence of $KA_{\text{kn}}$ and $Par_{\text{class}}$ on $Lr_{n\text{eff}}$ at $Perf_{\text{sum}} = 78.6$. Although similar findings are obtained that $KA_{\text{kn}}$ greatly affects the learning effectiveness, $Par_{\text{class}}$ is likely to play a role as well. Figure 7c shows the influences of $Perf_{\text{sum}}$ and $Par_{\text{class}}$ on $Lr_{n\text{eff}}$ at $KA_{\text{kn}} = 91.2$. Comparing between $Par_{\text{class}}$ and $Perf_{\text{sum}}$, it is found that the class participation is more important. As a consequence,
the significance of the three variables to the learning effectiveness can be characterized as $KA_{kn} > Par_{class} > Perf_{sum}$. 

Springer
5 Discussions

5.1 Addressing the research questions

To answer the three research questions of this study, we first identify the dominant variables that significantly affect the student learning performance. The learning effectiveness (i.e., academic performance) function $L_{\text{rneff}}$ obtained by the GP prediction model demonstrates that the dominant variables that affect student learning in the online engineering course are the knowledge acquisition $K_{\text{Akn}}$, following by the participation in class $P_{\text{arclass}}$, and the summative performance $P_{\text{fsum}}$. Furthermore, the prerequisite knowledge $P_{\text{Kbg}}$ tends not to play a key role, which indicates that the students with different levels of background knowledge could reach similar learning effectiveness after course learning. Regarding the second research question, we apply the developed prediction model to another online course and the results indicate that a reasonable accuracy of the model for predicting students' learning performance (with a maximum difference between of 5%). Therefore, the results demonstrate the accuracy and efficiency of the developed prediction model. According to the results, students’ self-evaluation of knowledge acquisition, class-level participation frequency and the instructor’s summative evaluation serve as the critical indicators of particularly good or poor performance. Finally, the results indicate that we can optimize the online course based on the performance prediction results generated by the reported model.

5.2 Pedagogical implications

Academic performance prediction is a difficult problem to solve due to the large number of factors or characteristics that can influence students’ performance (Romero et al. 2013). Based on the empirical research results, we conclude that instructional design of online course should take into consideration students’ self-evaluation, discussion participation, and instructor’s summative evaluation. First of all, because students’ self-evaluation serves as a key to predict student performance, the online instructors can use self-evaluation as formative rather than a summative tool to foster student motivation for high achievement in course design (Arthur 1995). Secondly, consistent with previous research (Ouyang et al. 2020; Romero et al. 2013), our research results show that students’ participation in class discussions is a critical indicator of their learning performance and effectiveness. It is reasonable to conclude that students who obtain higher scores in the course are those who participate in a more active fashion in the class discussions, while those students who obtain lower scores in the course are the ones who participate less active. However, our results indicate that group discussion participation is not a critical indicator for student performance, which could be explained by Chinese students’ cultural tendency to put more emphasis on their performance under the instructor presence rather than peer collaborative learning (Ouyang et al. 2021; Zhang 2013). Thirdly, like previous research indicates, the instructor’s summative evaluation plays an important role to predict student performance; however, it is less important than students’ self-evaluation of their own learning. Taken together, this research provides pedagogical implications of online course design related to critical factors of student evaluation, instructor presence, and cultural background.
5.3 Analytical implications

Previous research indicates that the difficulties in identifying the dominant variables from learning process are likely to significantly affect the student performance, which are the obstacles in the development of quantitative prediction models. Most previous studies rely on the data that are not directly generated from the learning process (e.g., demographic data or other personal information); therefore, the data collection and analytics significantly decrease the effectiveness of the data-driven supports and increase the time and personnel required to manage such initiatives (Bernacki et al. 2020). This study proposes the AI-enabled prediction model using the advanced EC technique to accurately predict the student performance based on the learning variables generated from the students’ collaborative learning process. We argue that the identification of the data variables should be grounded upon the theoretical underpinnings rather than the non-malleable student factors or fixed information. Therefore, we obtain the robust prediction model and bridge the gap between the data-driven approaches and learning theories (Suthers and Verbert 2013). A challenge we face during the prediction model development is the data cleaning and analytics: we manually code students’ oral and written content as one element of the input variable; future work should integrate automatic content analytics to provide real-time prediction models. In addition, the generalizability of the prediction model should be strengthened in two ways: first, taking one dataset to build the model and evaluate the model with another dataset; second, enlarging the validation of the model through multiple iterations of different online courses.

5.4 Limitations

Limitations of the reported GP prediction model can be summarized with respect to the data, algorithm, ethics and generalizability. First, the number of participants used to develop the GP prediction model is relatively small. The proposed model can be enhanced by taking into account more data (e.g., more participants or multiple-round of experiments). Second, the GP algorithm, similar to other supervised AI methods (e.g., ANNs), cannot be deployed to learn in real-time, incrementally or interactively. However, a GP model well-trained with a larger database can be a viable option for offline analysis. Third, the ethical issue, such as the potential influence of student learning outcome performed by AI-enabled models, should be considered by researchers. Future work needs to focus on providing real-time predictions, timely warnings, and advice during online engineering education to ensure students obtain positive influences by AI prediction models (e.g., Asif et al. 2017). Finally, future work should deepen the generalizability of the prediction model through multiple iterations of empirical research in different educational contexts and consider the influence of other external factors such as holidays, family events, social relations, etc.

6 Conclusions

In higher education, the design of prediction models and early warning systems has become a critical enterprise (Bernacki et al. 2020). However, the prediction models suffer the issues related to the learning data identification and analytics. This study addressed the issues by
developing the AI model for the quantitative prediction of the academic performance in online engineering education. Like the emerging performance prediction approaches (Bernacki et al. 2020), the learning data identified in the current prediction model overcame the weaknesses of the prior approaches that rely on the non-malleable factors (e.g., student demographics) or the confounded factors (e.g., early performance). In addition, the current prediction model analyzed the quantifiable contributions of the dominant variables in the online engineering course to make an accurate prediction. The main findings indicated that the dominant variables in the online engineering course were the knowledge acquisition, following by the participation in class and the summative performance, while the prerequisite knowledge tended not to play a key role. Based on the prediction results, we provided the pedagogical and analytical implications for the online course design and prediction model development. The AI-based quantitative prediction model can be used to evaluate and predict the learning performance in online engineering education.

Acknowledgements This research is supported by National Natural Science Foundation of China (62177041) and Graduate Education Research Project of Zhejiang University (20210308). P. J. and F. O. acknowledge the financial supports from the startup foundations from the Zhejiang University, China. A.H.A. acknowledges the startup fund from the Pittsburgh University. P.J. thanks the graduate students’ participation in the Course of Smart Marine Metastructures (Spring 2020) offered in the Ocean College at Zhejiang University.

Declarations

Conflict of interest The authors have no conflicts of interest to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Agudo-Peregrina ÁF, Iglesias-Pradas S, Conde-González MÁ, Hernández-García Á (2014) Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. Comput Hum Behav 31:542–550
Ahmad F, Ismail NH, Aziz AA (2015) The prediction of students’ academic performance using classification data mining techniques. Applied Mathematical Science 9(129):6415–6426
Arthur H (1995) Student self-evaluations: How useful? How valid? Int J Nurs Stud 32(3):271–276
Asif R, Merceron A, Ali SA, Haider NG (2017) Analyzing undergraduate students’ performance using educational data mining. Comput Educ 113:177–194
Baker RS, Yacef K (2009) The state of educational data mining in 2009: a review and future visions. JEDM J Educ Data Min 1(1):3–17
Bernacki ML, Chavez MM, Uesbeck PM (2020) Predicting achievement and providing support before STEM majors begin to fail. Comput Educ103999
Betser S, Martin LM (2018) Engineering discourse development in an informal youth-driven maker club. Int Soc Learn Sci
Cano A, Leonard JD (2019) Interpretable multiview early warning system adapted to underrepresented student populations. IEEE Trans Learn Technol 12(2):198–211
Chassignol M, Khoroshavin A, Klimova A, Bilyatdinova A (2018) Artificial intelligence trends in education: a narrative overview. 7th International young scientist conference on computational science. Proc Comput Sci 136:16–24

Chaudhury P, Tripathy HK (2017) An empirical study on attribute selection of student performance prediction model. Int J Learn Technol 12(3):241–252

Chen L, Chen P, Lin Z (2020) Artificial intelligence in education: a review. IEEE Access 8:75264–75278

Damsga CI (2014) The multi-layered nature of small-group learning: Productive interactions in object-oriented collaboration. Int J Comput-Support Collab Learn 9:247–281

Dillenbourg P (1999) What do you mean by “collaborative learning”? In: Dillenbourg P (ed) Collaborative learning: cognitive and computational approaches, vol 1. Elsevier, Oxford, pp 1–15

Drucker H, Burges CJC, Kaufman L, Smola A, Vapnik V (1997) Support vector regression machines. In: Mozer MC, Jordan MI, Petsche T (eds) Advances in neural information processing systems. MIT Press, Cambridge, MA, pp 155–161

Fernandes E, Holanda M, Victorino M, Borges V, Carvalho R, Erven GV (2019) Educational data mining: Predictive analysis of academic performance of public school students in the capital of Brazil. J Bus Res 94:335–343

Fok WWT, He YS, Yeung A, Law KY, Cheung KH, Ai YY, Ho P (2018) Prediction model for students’ future development by deep learning and tensorflow artificial intelligence engine. 4th IEEE international conference on information management, pp 103–106

Fung CC, Li JB, Wong KW, Wong KP (2004) A java-based parallel platform for the implementation of evolutionary computation for engineering applications. Int J Syst Sci 35:741–750

Gandomi AH, Alavi AH (2012) A new multi-gene genetic programming approach to nonlinear system modelling. Part I: materials and structural engineering problems. Neural Comput Appl 21:171–187

Godwin A, Kirn A (2020) Identity-based motivation: connections between first-year students’ engineering role identities and future-time perspectives. J Eng Educ 1–22

Goodyear P, Jones C, Thomson K (2014) Computer-supported collaborative learning: instructional approaches, group processes and educational designs. In: Spector JM, Merrill MD, Elen J, Bishop MJ (eds) Handbook of research on educational communications and technology, 4th edn. Springer, New York, pp 439–451

Hussain M, Zhu W, Zhang W, Abidi SMR, Ali S (2019) Using machine learning to predict student difficulties from learning session data. Artif Intell Rev 52(1):381–407

Jivet I, Scheffel M, Specht M, Drachsler H (2018) License to evaluate: preparing learning analytics dashboards for educational practice. Proceedings of the 8th international conference on learning analytics & knowledge (pp 31–40). ACM.

Kabra RR, Bichkar RS (2011) Performance prediction of engineering students using decision trees. Int J Comput Appl 36(11):8–12

Kotsiantis S, Pierrakeas C, Pintelas P (2003) Preventing student dropout in distance learning systems using machine learning techniques. 7th International conference on knowledge-based intelligent information & engineering systems, lecture notes in artificial intelligence, vol 2774 (pp 267–274). Springer-Verlag

Lam PC, Doverspike D, Mawasha RP (1999) Predicting success in a minority engineering program. J Eng Educ 88(3):265–267

Liu CH, Matthews R (2005) Vygotsky’s philosophy: constructivism and its criticisms examined. Int Electron J 6:386–399

Madhavan K, Richey MC (2016) Problems in big data analytics in learning. J Eng Educ 105(1):6–14

Marquez-Vera C, Cano A, Romero C, Ventura S (2012) Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data. Appl Intell 38:315–330

Marquez-Vera C, Cano A, Romero C, Noaman AYM, Fardoun HM, Ventura S (2015) Early dropout prediction using data mining: a case study with high school students. Expert Syst 33(1):107–124

Martin L, Betser S (2020) Learning through making: the development of engineering discourse in an out-of-school maker club. J Eng Educ 109(2):194–212

McArthur D, Lewis M, Bishay M (2005) The roles of artificial intelligence in education: current progress and future prospects. I-Manager’s J Educ Technol 1(4):42–80

Minaei-Bidgoli B, Kashy DA, Kortmeyer G, Punch WF (2003) Predicting student performance: an application of data mining methods with an educational web-based system. 33rd Annu Front Educ 1:T2A–T18

Mozer MC, Wiseheart M, Novikoff TP (2019) Artificial intelligence to support human instruction. Proc Natl Assoc Sci 116(10):3953–3955
Nguyen H, Wu L, Fischer C, Washington G, Warschauer M (2020) Increasing success in college: examining the impact of a project-based introductory engineering course. J Eng Educ. https://doi.org/10.1002/jee.20319

Nicholls GM, Wolfe H, Besterfield-Sacre M, Shuman LJ (2010) Predicting STEM degree outcomes based on eighth grade data and standard test scores. J Eng Educ 99(3):209–223

Oskouei RJ, Askari M (2014) Predicting academic performance with applying data mining techniques (generalizing the results of two different case studies). Comput Eng Appl J 3(2):79–88

Ouyang F, Chang YH (2019) The relationship between social participatory role and cognitive engagement level in online discussions. Br J Educ Technol 50(3):1396–1414

Ouyang F, Jiao P (2021) Artificial intelligence in education: The three paradigms. Comput Educat: Artif Intell 100020:1–6. https://doi.org/10.1016/j.caei.2021.100020

Ouyang F, Li X, Sun D, Jiao P, Yao J (2020) Learners’ discussion patterns, perceptions and preferences in a China’s massive open online course (MOOC). Int Rev Res Open Distrib Learn 21(3):264–284

Ouyang F, Hu Y, Zhang Y, Guo Y, Yang Y (2021) In-service teachers’ knowledge building during face-to-face collaborative learning. Teach Teach Educ 107:103479. https://doi.org/10.1016/j.tate.2021.103479

Ouyang F, Zheng L, Jiao P (2022) Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. Educat Inform Technol 1–33. https://doi.org/10.1007/s10639-022-10925-9

Paquette L, Baker RS, de Carvalho A, Ocumpaugh J (2015) Cross-system transfer of machine learned and knowledge engineered models of gaming the system. In: Proceedings of the 23rd conference on user modelling, adaptation and personalization (pp 183–194). Dublin, Ireland

Pena-Ayala A (2014) Educational data mining: a survey and a data mining-based analysis of recent works. Expert Syst Appl 41:1432–1462

Picciano AG (2014) Big data and learning analytics in blended learning environments: benefits and concerns. Int J Artif Intell Interactive Multimedia 2(7):35–43

Rojas EM (2001) Fostering collaborative education through Internet technologies. J Eng Educ 90(4):623–626

Roll I, Wylie R (2016) Evolution and revolution in artificial intelligence in education. Int J Artif Intell Educ 26:582–599

Romero C, Lopez MI, Luna JM, Ventura S (2013) Predicting students’ final performance from participation in on-line discussion forums. Comput Educ 68:458–472

Roschelle J, Teasley S (1995) The construction of shared knowledge in collaborative problem solving. In: O’Malley C (ed) Computer-supported collaborative learning. Springer Verlag, Berlin, pp 69–197

Roy R, Bhatt P, Ghoshal SP (2010) Evolutionary computation based three-area automatic generation control. Expert Syst Appl 37(8):5913–5924

Saa AA (2016) Educational data mining & students’ performance prediction. Int J Adv Comput Sci Appl 7(5):212–220

Sharabiani A, Karim F, Sharabiani A, Atanasov M, Darabi H (2014) An enhanced Bayesian network model for prediction of students’ academic performance in engineering programs. IEEE global engineering education conference (EDUCON), 2014/04/3-5. Harbiye, Istanbul, Turkey

Siemens G, Baker R (2012) Learning analytics and educational data mining: towards communication and collaboration. Proceedings of the second international conference on learning analytics & knowledge (pp 252–254). ACM

Siemens G, Long P (2011) Penetrating the fog: analytics in learning and education. EDUCAUSE Rev 46(5):30–37

Suthers D, Verbert K (2013) Learning analytics as a middle space. In: Suthers D, Berbert K, Duval E, Ochoa X (eds) Proceedings of the 3rd international conference on learning analytics & knowledge (pp 1–4). ACM, New York, NY

Takagi H (2001) Interactive evolutionary computation: fusion of the capabilities of EC optimization and human evaluation. Proc IEEE 89(9):1275–1296

Timms MJ (2016) Letting artificial intelligence in education out of the box: educational cobots and smart classrooms. Int J Artif Intell Educ 26:701–712

Tomasovic N, Gvozdenovic N, Vranes S (2020) An overview and comparison of data mining techniques for student tests examination performance prediction. Comput Educ 143:103676

Uccio US, Colantonio A, Galano S, Marzoli I (2020) Development of a construct map to describe students’ reasoning about introductory quantum mechanics. Phys Rev Phys Educ Res 16:010144

Verbert K, Manouselis N, Drachsler H, Duval E (2012) Dataset-driven research to support learning and knowledge analytics. Educ Technol Soc 15(3):133–148

Vygotsky LS (1978) Mind in society: The development of higher psychological processes. Harvard University Press, Cambridge
Xing W, Guo R, Petakovic E, Goggins S (2015) Participation-based student final performance prediction model through interpretable genetic programming: Integrating learning analytics, educational data mining and theory. Comput Hum Behav 47:168–181
Yehuala MA (2015) Application of data mining techniques for student success and failure prediction (The case of debre_Markos university). Int J Sci Technol Res 4(4):91–94
Zaffer M, Hashmani MA, Savita KS (2017) Performance analysis of feature selection algorithm for educational data mining. 2017 IEEE conference on big data and analytics (ICBDA), pp 7–12
Zhang J (2013) Chapter 28: collaboration, technology, and culture. In: Hmelo-Silver CE, Chinn CA, Chan CKK, O’Donnell AM (eds) The international handbook of collaborative learning. Routledge, New York, pp 495–508
Zhang Q, Vandenbossche JM, Alavi AH (2021) An evolutionary computational method to formulate the response of unbonded concrete overlays to temperature loading. Engineering Computations, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/EC-11-2020-0641

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Pengcheng Jiao1,2 · Fan Ouyang3 · Qianyun Zhang4 · Amir H. Alavi4,5

Pengcheng Jiao
pjiao@zju.edu.cn
Qianyun Zhang
qiz89@pitt.edu
1 Institute of Port, Coastal and Offshore Engineering, Ocean College, Zhejiang University, Zhoushan 316021, Zhejiang, China
2 Engineering Research Center of Oceanic Sensing Technology and Equipment, Ministry of Education, Zhejiang University, Zhoushan, China
3 College of Education, Zhejiang University, Hangzhou 310058, Zhejiang, China
4 Department of Civil and Environmental Engineering, University of Pittsburgh, Pittsburgh, PA 15261, USA
5 Department of Computer Science and Information Engineering, Asia University, Taichung, Taiwan