



Recent trends in computational intelligence for educational big data analysis

Anayo Chukwu Ikegwu¹ · Henry Friday Nweke² · Chioma Virginia Anikwe¹

Received: 22 May 2023 / Accepted: 20 August 2023 / Published online: 7 September 2023
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2023

Abstract

Educational big data analytics and computational intelligence have transformed our understanding of learning ability and computing power, catalyzing the emergence of Education 4.0. However, educators and researchers still struggle to identify appropriate methods to analyze the diverse data generated within educational environments. The complexity and uncertainty inherent in heterogeneous and homogeneous data often compound these challenges. This study aims to explore the potential applications of computational intelligence methods to support educational big data analysis. We begin by discussing the processes involved in educational big data analytics (EDA), including data collection, data preprocessing, feature extraction, modeling, and evaluation. We then provided an extensive review of computational intelligence and its methods, including artificial intelligence approaches, machine learning methods, deep learning methods, meta-heuristic optimization approaches, ensemble techniques, and the Markov model, as applied to educational big data analysis. Furthermore, we discussed novel application areas for computational intelligence in education, including predicting academic performance, social network analysis, detecting undesirable student behaviors, adaptive curriculum sequencing and personalization, courseware development, and decision support systems. We also mapped various computational intelligence methods to these novel application areas. Despite the progress made in educational big data analytics implementation, challenging research areas still require further investigation. These research areas include enhanced academic performance prediction, data-driven intelligent tutoring systems, adversarial machine learning, student engagement, personalized learning, and more. In this paper, we briefly discussed these ten important research directions.

Keywords Computational intelligence · Machine learning · Educational data mining · Big data analytics · Artificial intelligence · Deep learning

1 Introduction

Implementation of educational big data analytics and computational intelligence methods in education have changed

our learning ability, and computing power, and revolutionized education 4.0, industry 5.0, and the entire education sector. Educational big data are important to the formulation of institutional curriculum development and improving the learning mechanism and processes. Educational big data analytics largely depends on data to enhance learning experiences. Data are viewed as a new gold [1] as education is the bedrock of national sustainability and development. The homogeneous and heterogeneous education-based data are sourced from convergence and divergence ways that include web portals, sensors, mobile apps, student records, administrative records, academic historical records, demographic data, financial transaction data, etc. [2]. The generated data appeared in inform of structured, semi-structured, and unstructured formats, and were analyzed using educational data mining (EDM) methods for decision-making. EDM is an emergent area charged with the responsibility to extract

✉ Anayo Chukwu Ikegwu
anayoikegwuwchukwu@gmail.com

✉ Henry Friday Nweke
henry.nweke@ebsu.edu.ng

Chioma Virginia Anikwe
chiomavanikwe@gmail.com

¹ Computer Science Department, Alex Ekwueme Federal University, Ndufu-Alike Ikwo, Abakaliki, Ebonyi, Nigeria

² Centre for Research in Machine Learning, Artificial Intelligence and Network Systems, Computer Science Department, Ebonyi State University, Abakaliki, Nigeria

and analyze the hidden knowledge from the education data through the use of data mining techniques, such as classification, association rule mining, clustering, regression, etc. [3]. Therefore, big data (BD) is a concept that deploys computational methods to analyze diverse, numerous data sets, and reveal patterns and associations inherent in the data [4]. Big data analytics is a phenomenon that reveals the hidden pattern and correlations within the data set [4–6]. Consequently, educational big data analytics (EBDA) is the process of inferring meaning through definite stages, such as data collection, data preprocessing, feature extraction, modeling, and evaluation methods [7–9]. Educational big data analytics have helped to process a large significant amount of data in educational environments for effective decision-making and learning processes. Educational big data analytics processes (EBDAP) such as data collection, data preprocessing, feature extraction, modeling, and evaluation measure involve sequences undergo to achieve a comprehensive and efficient data analysis in the educational data mining domain [4]. EBDAP is facilitated by computational intelligence (CI).

Furthermore, computational intelligence is the use of computational concepts such as application, design, and biological and linguistic development that are highly motivated by computational paradigms [10] to learn intrinsic characteristics from data. Hence, computational intelligence is integrated into educational big data analytics and forms a binding force to intelligently discover the pattern, and implement significant data analysis, prediction, and decision-making. The implementation of computational intelligence methods is set to analyze large data sets to reduce cost and time and improve smart decision-making. The CI methods for educational big data analytics entail various fundamental elements that help to process and analyze significant data peculiar to education. The fundamental elements include artificial intelligence (AI), machine learning, deep learning, meta-heuristic optimizations, ensemble, and the Markov model for the decision-making approach [11–14]. Moreover, computational intelligence methods for educational big data analysis have been applied in numerous novel areas, such as prediction of academic performance, social network analysis, detection of undesirable student behaviors, adaptive curriculum sequencing and personalization, building courseware, and decision support system [7, 15–20].

However, educators and researchers are still faced with the challenges of choosing appropriate methods to analyze data collected from educational sectors. Some of the challenges center on the complexity and uncertainty of heterogeneous and homogeneous data, such as high dimensionality degree (difficulty in processing and analyzing data set), imprecision, or missing data [21]. Another major issue inherent in educational big data analysis is the algorithm interpretability to proffer solutions and aid the decision-making process in a contemporary real-life scenario. A quiet number of research

studies have been undertaken in computational intelligence methods for educational big data analysis [16, 22, 23]. Some of these studies reported varied performance results, while others reviewed recent studies in computational intelligence methods. Therefore, it is difficult to understand these computational intelligence methods and their intricate natures for educational data analysis.

For instance, Shu [23] surveyed six techniques for big data analytics, and outlined exhaustive theories and practices for big data analysis. However, they failed to comprehensively discuss different issues related to data sets analysis, and the strength and weaknesses of the big data methods were not provided. Recently, Iqbal et al. [22] reviewed big data analytics, computational intelligence, and its application areas. This study only focuses on smart cities, the strength and weaknesses of the identified techniques were not stated. Moreover, the methods did not focus on educational sectors. In addition, Bakhshinategh [16] surveyed educational data mining applications and their tasks for the past 10 years. This systematic review only covers some application areas in EDM; however, appropriate computational intelligence methods for educational data analysis and their weaknesses were not discussed. Hernández-Blanco et al. [24] discussed the applications, current state-of-the-art, and future directions of deep learning in educational data mining. Nonetheless, the review focused on only deep learning-based computational intelligence methods. Other methods such as meta-heuristics, machine learning, etc., and their strengths and weaknesses were not comprehensively discussed. A review by Romero & Ventura [2] presented a review of educational data mining and learning analytics for educational data analysis. The review discussed various tools, data sets, methods, and future research areas. Nonetheless, computational intelligence techniques, strengths, and weaknesses were not covered. Recently, Limna et al. [25] presents a review covering the benefits and challenges of artificial intelligence in education during the digital era. Nevertheless, the studies did not discuss computational intelligence techniques, strengths, and weaknesses. More recently, a systematic review by Kadoura et al. [26] reviewed machine learning models for online learning and examination management. This covers the overview and roles of machine learning phases and analysis including issues and limitations of machine learning in the exam management. However, it does not discuss CI techniques and its application domain.

Conversely, from our study and to the best of our knowledge, none of these studies have extensively discussed computational intelligence methods, strengths, and weaknesses of the techniques in the educational sector.

Conversely, the current review differs from other reviews and surveys in the literature in several ways. First, the review presents a big data analysis process and framework for educational data analysis. Second, the present review

outlines different computational intelligence methods for educational big data analysis, their strengths, and their weaknesses. Finally, the paper reviews various areas of application of computational intelligence for educational big data analytics. These academic performance predictions, contents personalization, detection of undesirable behavior in students, curriculum sequencing, etc., and recent studies in these areas. From available studies in the literature, there are no comprehensive reviews or surveys that provide significant discussion on these all-inclusive computational intelligence methods for educational big data analysis.

Accordingly, the contributions of this studies to the current body of knowledge in computational intelligence, big data analytics, and educational data mining are as follows:

- To provide an in-depth discussion on educational big data analytics and the processes for analyzing significant educational data sets;
- To comprehensively explore the prospective computational intelligence methods for educational big data analysis, including strengths, and weaknesses;
- To identify and discuss novel recent areas of applications of computational intelligence for educational big data analysis and map computational intelligence methods in the identified novel areas;

The other part of the paper is the breakdown as follows: Sect. 2 provides educational big data analytics and its processes. Moreover, computational intelligence is also explained in Sect. 2.2. An in-depth discussion of computational intelligence methods for educational big data analysis is presented in Sect. 3. Section 4 presents the novel application areas of computational intelligence for educational big data analysis. The discussion is presented in Sect. 5. Finally, Sect. 6 provides the concluding remarks arising from this study.

2 Educational big data analytics

Big data analytics has played an exhaustive role in educational data mining (EDM) as it infers meaning through a definite process. Educational big data analytics (EBDA) measures, collects, analyzes, and processes data within the educational learning environments to enhance and optimize teaching and learning outcomes. It further helps with effective decision-making and enhances educational curriculum developments, program customization, and students' active participation and engagement, and offers a high level of academic personalization. In addition, educational sector administration and industrial workforce are improved through the technological knowledge of EBDA [27]. There is a definite process involved in applying educational big data



Fig. 1 Educational big data analytics process

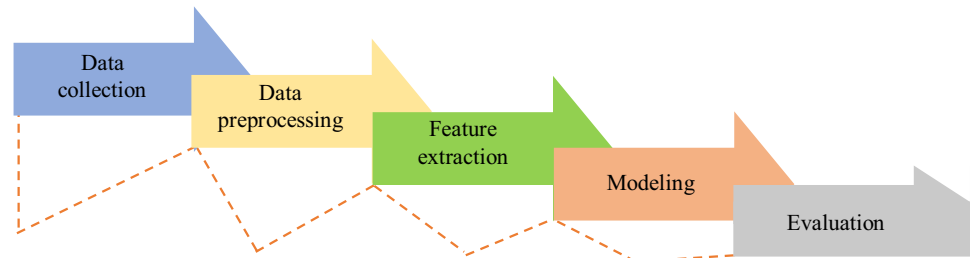
analytics to analyze huge data within the educational environment. These processes are explained in the next section of the review.

2.1 Educational Big data analytics process

The educational big data analytics process (EBDAP) involves a sequence undergo to achieve a comprehensive and efficient data analysis in the educational data mining domain. Therefore, EBDAP processes include data collection, data preprocessing, feature extraction, modeling, and evaluation measure. The five stages are depicted in Fig. 1 and discussed below.

2.1.1 Data collection

The systematic process by which data or data sets are extracted or crawled from online, historical education data, and institutional education agencies are referred to as data collection [4]. Significant data sets are collected from educational environments, such as university administration, students' data or affiliates bodies, social media, web portals, server logs, etc. The format of data collected from students includes personal profiles and institutional data sources, such as attendance, performance scores, sports activities, and assessment reports for extracurriculars, alumni, etc. [28]. Other sources through which data can be collected in the education system include an iris scanner, Radio Frequency Identification data (e.g., livestream tracking during online lectures, address unique identifiers (e.g., IP address, domain names, UUIDs, GUIDs, etc.), physics model (e.g., games, CAD, etc.), internet of things (IoT), and wearable sensor like using smart gadgets, such as smart watches, smart glasses, smartphones to record lecture materials and delivery presentation [7, 29, 30]. Recently, various authors have utilized this approach to collect or extract a significant amount of data or crawl data from their institutions for big data analysis. Pierrakeas et al. [31] collected available data from 3,882 students' details from the course setup module to predict student

Fig. 2 Data collection

dropout with machine learning models (e.g., decision trees, rule learners, Bayesian algorithms, neural networks, etc.). A data set of more than 130,000 log events was parsed from a custom plugin for Moodle and was further collected from several massive online open courses (MOOCs) self-regulated learning and a custom plugin for Moodle [15, 32]. Massive data collected from educational sectors enable big data analytics for various applications, such as students' dropout prediction, curriculum sequencing, content personalization, etc. The media for data collection in education data mining are diagrammatically shown in Fig. 2

2.1.2 Data preprocessing

Data preprocessing is the second stage in educational big data analytics. Data preprocessing involves the art of cleaning the collected data, data integration, and data compression. In furtherance to the data preprocessing, it encompasses fixing missing values, removing duplicate data, locating irrelevant data parts, and altering the spotted noisy data. The essence of data preprocessing is the big volume, duplication, and uncertainty features of the collected data which require an intelligent technique before storage [33]. Moreover, data preprocessing undergoes the processes of refinement, orchestration, virtualization, and blending operations [29]. Refinement converts structured and semi-structured data using statistical techniques into an organized format. Orchestration involves the composition of the data, and blending approaches to fuse big data from multiple data sets. A multidimensional view of underlying big data is envisaged by blending the operations. However, after the tedious tasks of data collection and preprocessing, it is stored in the database as text files, video or audio files, image files, or various other file formats. It is an essential component as it filters, integrates, transforms, reduces, and cleans the actual data ready for further analysis in an educational environment. For instance, during students' enrolment and examination result computation, incorrect value or mismatch data are found inconsistent, grades or results might be missing, duplication and incomplete candidate's records, etc. may be found and clean to a well-formed data. There are inherent challenges faced in data preprocessing. Various studies have attempted to solve

these problems which include data types (text, sensor, audio, and video), erroneous and noisy data, and data duplication. In their recent study, Gracie-Gil et al. [34] utilized homogeneous and heterogeneous ensemble filter methods to remove noise issues in big data classification. The methods produce an easy smart data set from any significant data classification issue. However, multiclass data and cost-sensitive filters still need to be addressed for the effective removal of instances from the majority classes. Furthermore, data preprocessing was applied in the data set that contained 300 instances with 24 attributes from student academic's performance for well-formed data using WEKA to predict student dropout [35].

Figure 3 shows the preprocessing stage.

2.1.3 Feature extraction

Feature extraction [35] determines the most significant attributes to train the data mining algorithms. Some important feature values include correlation-based, gain-ratio, information-gain, relief, and proportioned uncertainty attribute evaluation. In other words, greedy search remains the only correlation-based attribute as such as are search method ranked. Feature extraction tackles the problem of finding the most informative, distinctive, and reduced set of features tackled by the feature extraction phenomena. It improves the optimum practice of data storage and processing. Therefore, it is essential in the sense that a quiet number of attributes can be effectively reduced. It enhanced the speed of data mining using classification type of supervised learning techniques. In addition, feature extraction (FE) is important, because it represents and analyzes the valuable attributes or variables that are most relevant after data cleaning. For precise processing, the cleaned data in the database are reduced to more manageable attributes. The significant data sets from the preprocessed database are used for prediction modeling. It encompasses normalization mathematical metrics that usually capture the appearance of attributes and data sets. Furthermore, Feature extract has played a significant role in educational data mining (EDM). For example, during student academic performance (SAP) and student drop prediction (SDP), learning analytics (LA), massive

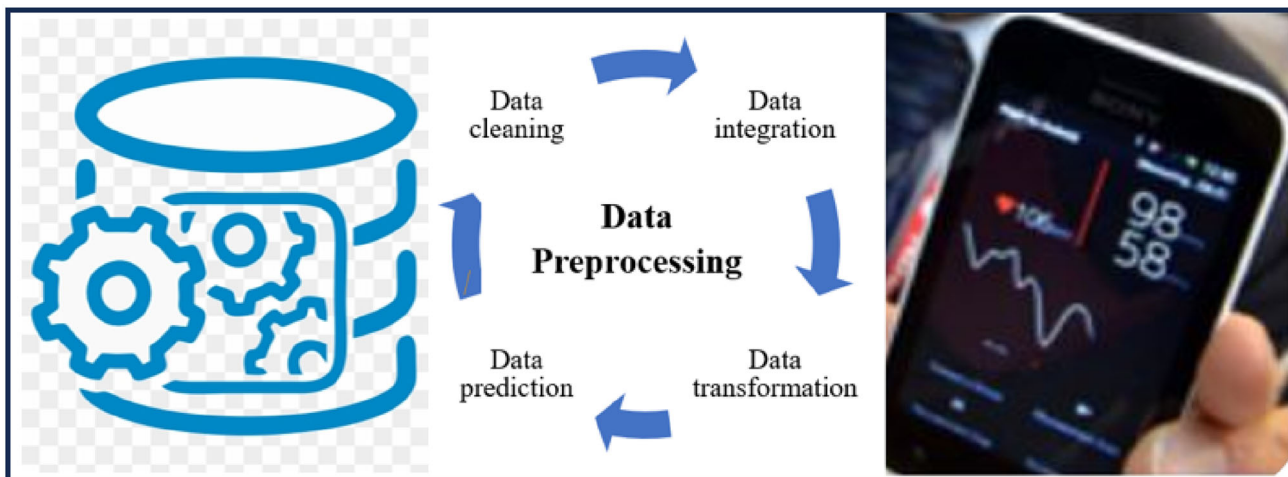


Fig. 3 Data pre-processing

online open courses (MOOCs) training, education curriculum review (ECR), etc. so many variables or data sets are collected and preprocessed and not all preprocessed variables will be used for modeling and prediction. The feature extraction approach reveals or selects the appropriate variables that are mostly needed. Various studies have utilized this approach for analyzing large educational data. For instance, Chen et al. [36] extracted learning behavior records such as effective learning time features from weekly course launches to predict MOOC dropout. It improves the accuracy of the decision tree and extreme learning machine algorithms (DT-ELM). The proposed features achieved 12.78%, 22.19%, and 6.87% which is higher when comparing the accuracy to the baseline algorithms, such as AUC and F1-score. In addition, important features such as grades, dates, results, and number of attempts were extracted from a data set containing 487 students’ details to predict student drop using machine learning algorithms. The proposed algorithms achieved high prediction accuracies of 95% after 3 semesters than the classification of 83% accuracy [37]. The feature extraction process is illustrated in Fig. 4.

2.1.4 Modeling educational big data analytics process

Modeling is a very important stage in analyzing educational data. After feature extraction and selection, the important attributes or variables are utilized for effective analysis of educational data. The essence of the modeling stage is to train and test the extracted feature for result orientation and to support effective decision-making. There are various data mining models or techniques deployed in modeling extracted features from educational data. Some of the data mining models include decision trees (DTs), logistics regression (LR), Naïve Bayes (NBs), support vector machine (SVM), k-nearest neighbors (k-NN), and random forest (RF), etc.

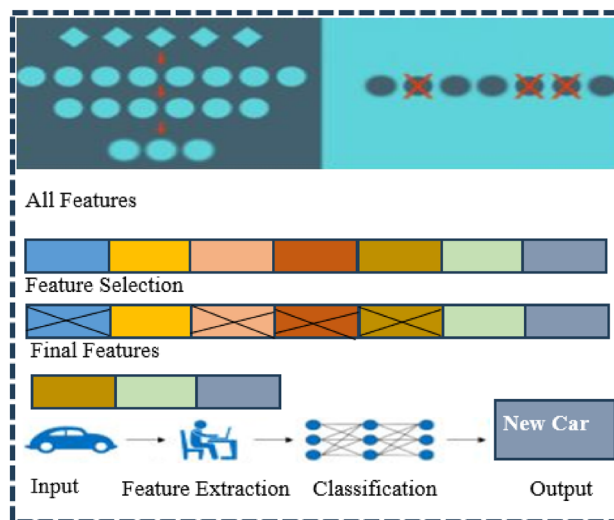


Fig. 4 Feature extraction

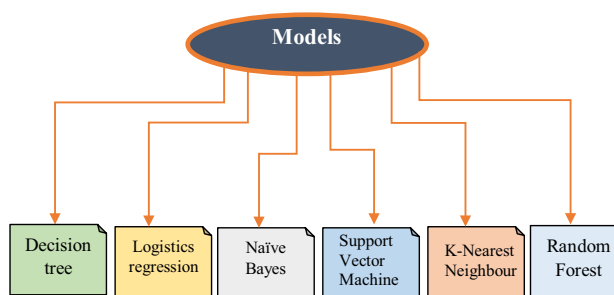


Fig. 5 Modeling algorithms

[38–41]. Some of these models are discussed below, while the diagram represented is shown in Fig. 5.

A decision tree [40] is one of the predictive modeling created by recursively splitting extracted features based on

observed inputs. Decision trees (DTs) select features with considerable classification ability. This is called a non-cyclic flowchart [42]. DTs are employed to visually exemplify and inform decision-making. The importance of decision tree knowledge is to create a real working model that will predict the value of a target label based on input attributes in educational data mining. Hence, the attributes of the data that form observation are depicted by the branches, and the conclusions (target value of the data) are represented in the sheets. A decision tree is made up of three nodes. These include (i) a root node (initiates the tree), (ii) an internal node (associated by a decision rule and split into two child nodes), and (iii) a terminal node (no child nodes). In the same view, the root and internal nodes; the binary decision rule are defined by a cut point on an identified input. It is now established that recursive partitioning D into smaller subsets $D_1, D_2 \dots D_k$ until attaining the specified stopping criterion which is aided by DTs. For clarity purposes, all the identified subsets belong to a single class. Similarly, using a single feature split is recursively defined for all nodes of the tree using nearly criterion. Hence, DTs knowledge contribution has led to effective prediction in the field of educational data mining. For instance, Chen et al. [36] combine a decision tree with an extreme learning machine to determine student dropout in massive open online courses (MOOCs) with an accuracy of 91.48%. In addition, potential dropout in higher education was determined with DTs using student demographic data and the characteristics of their academic progression [41].

Logistic regression [41] is a predictive classification model based on the probability of data attributes and is widely used in statistics for prediction. The coefficients of the model are utilized for predictive and descriptive reasons. Logistic regression (LR) birthed the linear regression model which is only used to predict numeric target attributes [43]. Though, the set target attribute can be nominal. Going forward, determine whether a customer is interested in the on-sale product, know if a student is engaged or lost interest, and whether a student will get a question correct. The logistic regression model assumes a linear relationship among variables and is incorporated in most machine learning libraries [31] and mahout [44]. A recent study utilized logistic regression to predict the probability of student dropout [41] and achieve a precision of 83.47%. However, it requires more models to detect potential dropouts.

Naïve Bayes [45, 46] is a Bayesian classifier that determines the probabilistic relationship between the classes and their attributes. Naïve Bayes (NBs) utilize the probability phenomenon bestowed on the fact that a new instance of the data set is correctly classified by offering a probabilistic measure in the classification results. The NB classifier is a simple classification model based on Bayes' theorem (class variable value is assumed to be independent) [47, 48]. Bayes' theorem is mathematically depicted [49]

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \text{ where } A \text{ and } B \text{ are two different events,}$$

$P(A)$ and $P(B)$ are the probability of A and B occurring, respectively.

$P(A|B)$ is the probability of A occurring given that B has already occurred

On this basis, the idea if the starting hypothesis is powered only by the NBs, is “that all the attributes are independent of each other”. Naïve Bayes has produced excellent performance results in the prediction of collected data based on the probabilistic dependence and independence relation between the data. For instance, Vilorio et al. [50] utilized Naïve Bayes in the dropout analysis using socioeconomic data obtained from students' records, Engineering sciences faculty at Mumbai University from 2017 and 2018. Hence, the study obtained an 87% accuracy result. In addition, [31] applied Naïve Bayes to achieve 73.45% accuracy after the analysis of why students drop out in distance learning and predicted dropout-prone students at the end of the session. Nonetheless, the applicability of more models is needed to advance toward the tale end of student studies.

Random forest [51, 52] is a dominant non-parametric statistical method or machine learning-based model that has two-class and multi-class classification and regression problems in a single and multipurpose framework. The framework is included some libraries, such as Mahout Library, Spark Library, R Package, and Python Library. Random forests (RFs) are used widely for high-performance data analysis and require few parameter tuning to achieve optimal results. RFs is also an ensemble model that uses the results from many different models to compute responses based on the available data. Random forest builds a series (forest) of CART (classification and regression tree) that recursively partitions a group of reference variables to predict definite responses, such as dropout rate in higher institutions, or disease states in the healthcare system [51].

Various authors have utilized random forests and their components such as a latent class for intelligent computing. For instance, the latent class forest was utilized to recursively partition observations into groups to help identify at-risk students of failing the course offered in the higher institution [40]. The RFs formed a series of decision-making body's stand in as weak categorizers, which are poor predictors but strong general predictions. Hence, RFs help smart learning analytics (SLA) to unveil all ages student with a knowledge-established framework to take full advantage of all categories

of resources as well as intelligent tools and achieved a result of 79% accuracy [13]. In addition, the RFs algorithm was applied during the analysis of undergraduate academic performance, which achieved 71.15% accuracy in the overall result [42, 49].

K-nearest neighbors is one of the classifications of algorithm-based computational intelligence that has been utilized for educational big data mining. K-nearest neighbors (k-NN) model selects first the k closest samples for a test sample from all the training samples, and the rest of the samples are predicted with a simple classifier called the majority classification rule [53]. In a simple definition, when an unknown data point is categorized upon which its class is already known is termed as K-nearest neighbors [54]. Moreover, it is built on learning through close examples in the space of the elements. The essence of KNN is to monitor sample data classified by the majority class. KNN models are established to close the gap in training data between new features and products-based. Therefore, it depends on the principle of finding k products by identifying training data closest to the testing algorithm available. This is to say, KNN determines test documents for training data that is the nearest range to that of the test available information prediction for the majority of KNN. K-nearest neighbors is implemented in machine learning based platform, such as Spark, Python, R package etc. [34]. Several studies have utilized the KNN models in educational big data analysis. For example, Kausar et al. [13] applied KNN models in smart learning analytical data to stimulate students' in-depth understanding of the perspective of course content sequencing and organization. The implemented algorithms achieved 74% accuracy. In addition, KNN was used in predicting the potential dropouts of students in higher institutions using demographic and academic data and achieved a cumulative detection percentage and performance precision of 37.88% [41].

Support Vector Machine (SVM) is a computational intelligence technique deployed in the analysis of educational big data. In a simpler form, the SVM model segregates two classes (hyper-plane or line), which is applied to solve regression or classification problems [7]. Support Vector Machine [55, 56] is based on VC-dimensional (that is $\{f(a)\}$) and uses a maximum number of training points) principle of statistical learning theory. The algorithm is based on two-group classification problems which categorize new text after labelled training data. Most importantly, the reason for SVM in educational data mining is to solve the linear constraint quadratic programming issue and finds the global optimal solution that invariably resolves the challenges of overfitting inherent in the conventional neural network [56]. SVM has been utilized in several areas, such as pattern classification recognition, prediction time sequence, function regression estimation, and stock trend. Furthermore, various authors have applied the SVM model for effective prediction. For

instance, pre-warning student academic performance was predicted using SVM and obtained 48.6% accuracy [57]. More so, the study by Vidhya & Shanmugalakshmi [58] utilized SVM in the multiple analysis of students' performance in mathematics-related subjects with regard to a modified adaptive neuro-fuzzy inference system, hence, achieving 92% accuracy.

2.1.5 Evaluation measure

The evaluation measure is the final stage in education data mining analysis. There is a need for an assessment or appraisal of the model deployed with the baseline model. Evaluation methods check the quality of the system before deployment to achieve maximum results. For example, in the analysis of student academic performance (SAP), student dropout prediction (SDP), and analysis of MOOCs learner's state, etc. The evaluation measures are carried out to obtain accurate results of implemented models. Performance metrics for the evaluation of data analytics models are categorized into several methods depending on the types of models used. These include mean absolute error (MAE), receiver operating characteristic (ROC), root mean square error (RMSE), relative absolute error (RAE), and root relative squared error (RRSE). In addition, performance metrics such as KAPPA (KAP) statistics, accuracy (ACC), sensitivity (SEN), specificity (SPC), and precision [7, 8, 15, 59] are used to evaluate data analytics models. The MAE, RMSE, RAE, and RRSE run a comparison of the used models, hence, the essence is to assess discrepancies between the predictions or forecasts provided by the models and the real-world results that are actual outcomes [60].

Different studies have utilized some of the models for the effective evaluation of education data mining algorithms. For instance, Tsiakmaki et al. [15] utilized the accuracy of the confusion matrix, mean absolute error (MAE), and receiver operating characteristic (ROC) to evaluate the performance of the classification models. In addition, [59] evaluated distributed machine learning using f-measure as a weighted sum of the precision and recall measures to solve the difficulty of clustering at scale. In addition, five evaluation measures such as accuracy, MAE, RMSE, RAE, and RRSE were utilized as a performance measure to analyze and predict the academic performance of a student. In terms of student dropout and academic performance measure of students, different evaluation measures such as accuracy, sensitivity, specificity, precision, and KAPPA has been deployed to evaluate the rate of dropout of students in the higher institution [7, 8, 15]. The ratio of correct predictions to the total number of prediction cases is termed accuracy. Sensitivity is responsible for the correct use of probability means to classify student dropout. The probability that a successful student is correctly described is determined by the specificity approach. Then,

precision ensures the conditional probability of dropout classified students in the correct order is achieved. In addition, KAPPA predicts the majority category in all observations. This is mathematically explained as Accuracy (ACC) = $\frac{(TP+TN)}{(TP+TN+FP+FN)}$, Sensitivity (SEN) = $\frac{TP}{(TP+FN)}$, Specificity (SPC) = $\frac{TN}{(TN+FP)}$, Precision = $\frac{TP}{(TP+FP)}$, KAPPA (KAP) = $\frac{ACC}{(1-ACC_r)}$ where ACC_r is the expected accuracy.

It is worth noting that True Positives (TP) count non-successful events. True Negatives (TN) determine successful events which are correctly represented as successful. False Positives (FP) focus on successful events that are also misrepresented. False Negatives (FN) represent non-successful events that are misrepresented as successful. In addition, ACC_r is the expected accuracy.

2.2 Computational intelligence

Computational intelligence (CI) is deemed as the theory, design, application, and development of biologically and linguistically motivated computational paradigms [10]. Intelligence simply means the ability and capability to familiarize ourselves with the environment and to solve problems associated with secular environments [11]. Computational intelligence has played a significant role in developing an intelligent system that aided decision-making and reshaped the phenomenon of educational development in Nigeria and beyond. CI consists of some components which are considered traditional pillars for developing intelligent systems. These components include neural networks and fuzzy systems as well as evolutionary computation and ambient intelligence [61, 62]. Others include artificial life, cultural learning, artificial endocrine networks, social reasoning, and artificial hormone networks. In addition, more recent paradigms such as deep learning methods, artificial intelligence, machine learning, meta-heuristic optimization, etc. have emerged [22, 63, 64]. Computational intelligence [65] is developed to discover new patterns, modeling, effective prediction, and data analysis. In the subsequent section, the main fundamentals of computational intelligence techniques will be discussed. These fundamentals include artificial intelligence (AI) such as deep learning, fuzzy and machine learning methods, meta-heuristic optimization, ensemble techniques, and the Markov model for the decision-making approach and how these methods are utilized to analyze big educational data.

3 Computational intelligence methods for educational big data analysis

The computational intelligence method for education big data analysis (CIAEBDA) entails the various elements that

help to process and analyze significant data peculiar to education. We categorize computational intelligence into six (6) main fundamental elements. These include artificial intelligence (AI), machine learning, deep learning, meta-heuristic optimization, ensemble techniques, and the Markov model for the decision-making approach as earlier outlined.

3.1 Artificial intelligence method

Artificial intelligence is a concept that plays a significant role in implementing an intelligent system in computing paradigms and data analytics. Artificial intelligence is the new normal and key power in today's business development, classroom training, and education 4.0 in general. Hence, education 4.0 plays a unique role through the use of contemporary learning strategies that include games, research-based, and other related problem-based learning that is linked to collaborative teaching and learning. In addition, brainstorming-based techniques centered on cybernetics theory, and information–communication technologies applications have created open research areas [11] in education. AI in educational data analysis is utilized majorly for student evaluation, individualization of learning processes, improvement of seminars, workshops, and presentations, and searching for useful information that supports teaching and learning [11]. AI further provide support to education builder and learners through academic sustainability. This concept is essential in future education for curriculum development, policy formulation, training enhancement, customized learning that runs through natural language processing, speech recognition, speech synthesis interaction, student evaluation, prediction of learning levels, and decision-making [38]. Therefore, AI is further breakdown into three categories. These include the neuro-fuzzy method, genetic fuzzy, and evolutionary fuzzy cognitive map. The categories are discussed in the subsection, and their strengths and weaknesses are provided in Table 1.

3.1.1 Neuro-fuzzy method

Neuro-fuzzy method [58] is a computational intelligence method that combines the artificial neural network (ANN) with that fuzzy logic. It utilizes a learning algorithm obtained by neural network theory to determine its parameters (that is the fuzzy rules and fuzzy sets) by processing sample data. This method is important in diverse fields, such as the health-care system, agriculture, and especially in education [66, 67]. Neuro-fuzzy method has gained momentum recently and the method is used to enhance electronic learning, student categorization, and prediction. For example, Khodke et al. [68] utilized an interactive educational development tool powered by AI planning for robotics otherwise known

Table 1 Artificial intelligence methods, strengths, and weakness

Methods	Description	Strength	Weakness	References
Neuro-fuzzy	Neural network theory determines its parameters by processing sample data through a learning algorithm	Simplicity, robustness, and easy extension of the knowledge base by new rules formation. This is represented based on linguistic variables that inherent uncertainties of human knowledge	This relied on the knowledge of an expert to make inferences and logical rules that cannot be generalized. In addition, suitable membership values are not easy to find	[58, 71, 83, 84]
Genetic fuzzy	This identifies structure and parameter through the search process (e.g., laws of natural evolution selection)	Flexible with good interpretability in terms of if-then rules. In addition, avoidance of over-general rules	The method has high computational complexity, difficulty in choosing a membership, and is time-consuming	[72, 85–88]
Evolutionary fuzzy cognitive map	Represents causal reasoning for decision support. In addition, integrating prior experience and knowledge regarding a subject matter	The method offers optimal solutions and involves easy systematic development of a clear human-readable structure for decision-making	Time-consuming for large data sets and operating under multi-stimulus dynamic environments. In addition, it is difficult to analyze and interpret	[62, 80–82, 89]

as REACT. This tool detects human expression using psychological signals based on neural networks. However, it requires more features of fuzzification and hybridization of the technique for efficient results in the e-learning tutorial systems. To reduce the issue of fuzzification and optimization of millions of parameters, a recent study integrated deep learning with neuro-fuzzy (deep neuro-fuzzy classifier) for educational data analysis [21].

Furthermore, the factors associated with student academic performance and psychological prediction of course selection through in-depth data analysis is vital especially when selecting courses for postgraduate students having considered undergraduate performance. Hence, going by these challenges, Petković et al. [69] proposed a modify hybrid neuro-fuzzy inference system (M-HNFIS) to predict Postgraduate course selection. The proposed method analyzes a questionnaire that focused on demographic and educational factors. The proposed hybrid neuro-fuzzy inference system achieved 97.5% accuracy in the course selection process. Nonetheless, more investigation on the parameters and techniques based on curriculum or graduation attributes of the course is required. In addition, the computational intelligence of the adaptive neuro-fuzzy inference system (ANFIS) method for data analysis has been deployed for the effective evaluation of pupil performance based on e-learning platforms which produced high performance [70]. A recent study by Naaj et al. [71] utilized a neuro-fuzzy approach to predicting students' academic performance with much emphasis on the factors affecting the rate of academic performance of students, such as course attendance rate, course category, school type, and grade point average (GPA). Nonetheless, the study failed to consider some important influences, such as scores

on English fluency tests, age, finance, health challenges, and other economic crises, etc.

3.1.2 Genetic fuzzy method

The genetic fuzzy method [72] is one of the soft computing intelligence that utilizes genetic programming designed to identify structure and parameters through the search process grounded on the laws of natural evolution selection. This concept is derived from the fuzzy logic principle. Fuzzy logic could be best described as a unique intelligent computing paradigm that unilaterally and concurrently processes numerical data and linguistic knowledge. A genetic Algorithm contains categorically three operations, such as selection, genetic operations, and replacement. Hence, the uncertainty representation features of fuzzy logic with the learning ability of the genetic algorithms are combined by the genetic fuzzy. Going by the deviations in measurement deltas, a fuzzy system is produced and GA optimized the rule base and membership functions. Moreover, the genetic fuzzy system is used for solving utilizing multi-modal data, where fuzzy logic presents a robust and flexible inference mechanism in subject challenges to maintain imprecision and uncertainty [72, 73]. A fuzzy system is easy to interrelate based on the knowledge of linguistic representation. Nonetheless, the hybridization of the GA and fuzzy logic accelerate an advanced soft computing approach known as the genetic fuzzy system (GFS), in which a GA is deployed to develop a fuzzy system by turning the fuzzy mechanism of membership functions and understanding the rules of the fuzzy.

Furthermore, the application of genetic fuzzy systems has been widely utilized in different fields to solve problems in

such areas that include bio-inspiration in engineering, software testing, biology, real-time system identification, mining classification rules in significant data sets, quantum computing, vector processing, and mostly in education. For instance, the optimization problem of examination timetabling [74] was handled with genetic algorithms. The candidate generates solutions through the chromosomes. Hence, the number of exams to be taken by the candidate measures the length of the chromosome length. Each genome is made of a timeslot, a list of classrooms, and an observer's indices list for each classroom index. The GFS was utilized in the education development process and analysis of qualitative data to balance the curriculum thereby improving the rate of student success and increasing the standard of the higher institution [75]. In addition, the genetic fuzzy approach was applied to educational data and data sets of students of Shahid Rajaee University of Tehran to predict the student academic performance (SAP), where 1.09–24.39% and 0.29–6.57% accuracy were obtained [76]. More so, a study by Cuzzocrea et al. [77] developed a framework to determine the degrees of membership of the places of articulation in phonetics. The genetic–fuzzy algorithms were utilized to analyze the transitions between phonemes gotten from the IPA (International Phonetic Alphabet) to allow simpler adaption to various languages. The proposed hybridization achieved an efficient subjective evaluation result of up to 89%. In the same manner, the genetic fuzzy algorithm analyzes and solves the problem of academic staff planning, allocation, and optimization, thereby providing an excellent reduction of computational complexities that emanated as a result of reorientation [78]. Recently, Shokouhifar and Pilevari [79] integrated a genetic algorithm and adaptive neuro-fuzzy (ANFIS-GA) approach to enhance the assessment of e-learning resilience for the period of coronavirus disease (COVID-19) outbreak and post-COVID-19. This approach helped to evaluate the improvement of resilience toward e-learning during the COVID-19 pandemic. However, due to the multiplicity of the desired result, it requires more approaches such as fuzzy machine learning to ensure accurate result and generality.

3.1.3 Evolutionary fuzzy cognitive map method

Evolutionary fuzzy cognitive map [62] is one of the artificial intelligence computational tools that represents casual reasoning for decision support which was proposed by Bart Kosko in 1986. The cognitive map (CM) is a visual representation of a person's or group of students' mental models for a given concept. CM has no visual or special rules that must follow for analysis to take place. For example, cognitive map representation is experienced when a new student is making inquiring for directions to the faculty, you can

create a mental image in your mind of the roads, remarkable places to turn, etc., along the way to the faculty from the point him/her is starting from. Then, the fuzzy cognitive map (FCM) is a highly influential fuzzy-graph tool for the simulation of dynamic complex tasks, predicting learning systems that have many dependent variables in uncertain academic environments [80, 81]. The importance of FCM to the education learning environments is vital as it mimics a system with the use of basic concepts and causal relationships among them, modeling and carryout decision-making concerning data uncertainty. In addition, FCM is represented by integrating prior knowledge and experience about the subject matter or course study.

Furthermore, the evolution of fuzzy cognitive maps has created unprecedented achievement in various aspects of learning. For instance, FCM reconstructs the premises associated with the behavior of a given agent, assists decision-makers to consider effective representation in any assigned task, and generates a more accurate description of a complex strategic function [80, 82]. An earlier study by Poczeta et al. [62] utilized a multi-layer fuzzy cognitive maps approach to analyze complex problems related to different concepts that are comprised of other parameters based on the learning criterion output or decision concepts. In addition, FCM was utilized in decision-making to simulate students learning for human activity simulation in an ambient intelligence model [81].

3.2 Machine learning method

The machine learning method [12] is one of the computational intelligence methods that have the ability and capability to analyze, inform meaning, and predict the outcome from significant education data and other generated data sources. The benefit of machine learning algorithms in computational intelligence includes: providing fast processing and real-time predictions. In addition, determine the probability of award-winning academic research, enables feature learning, helps in parameter optimization, dynamically handles multi-dimensional and multi-variety data, and provides complex process educational environments [90]. Machine Learning (ML) beckoned on two major aspects that is classification and prediction. It is crucial, because it supports computers with the ability to learn without being necessarily programmed and targets predictions dependent on established properties learned from training data sets. The ML methods focus on initially established properties from the training data to make predictions for future or unknown properties. The machine is classified into three and includes supervised, unsupervised, and re-enforcement learning. The supervised machine learning algorithms include regression models, decision tree, random forest, while unsupervised

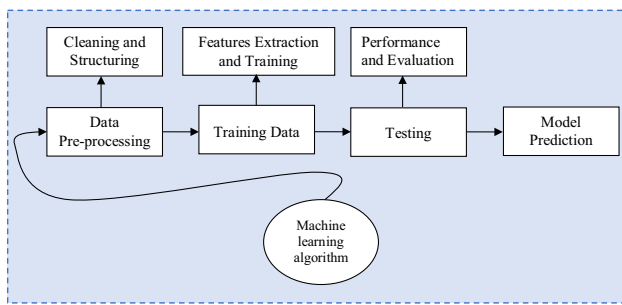


Fig. 6 Machine learning analysis procedure

machine learning is made up of clustering, association analysis, Hidden Markov Model, etc. [91, 92]. As supervised learning predict value based on labeled data, then, unsupervised learning base prediction value on unlabeled data. In addition, predictions based on the feedback process is learn is achieved by the re-enforcement learning approach. ML algorithms typically consist of two phases that are training and testing. Progressively, training data develop a model, whereas the testing phase validates models that are developed. ML follows some stringent steps in solving problems, especially in educational big data. These steps include (a) identifying class features/attributes and classes from training data, (b) locating a necessary subset of the attributes for classification (i.e., feature/dimensionality reduction), (c) learn the model using training data, and (d) utilize trained model to classify the unknown data.

Furthermore, each feature in the training phase, class is learned using the appropriate algorithms from the training set. In like manner, new data in the testing phase are run via the model as well as the algorithm classifies. This is known whether it belongs to one of the classes identified in the training data sets. For instance, using a case of educational data sets, such as student academic performance predictions (e.g., MOOC, dropout, etc.), automatic reading of handwriting, and automatic text classification (like spam filtering) [93]. The machine learning analysis procedure is depicted in Fig. 6 and the strengths and weaknesses are shown in Table 2.

In the figure, we identify four major procedures [91] in machine learning analysis; hence, the data are the first step which is pre-processing that includes data collection, cleaning, and structuring, through problem definition and assembling a data set. Step two focuses on training data by features extraction and training, third steps concentrate on testing the desired data sets through performance and evaluation. Finally, the last step is the model prediction based on the resulting model that may be predicted. Furthermore, machine learning methods have been notably utilized in numerous areas of educational data analysis and have proven to be efficient. For instance, Maestrales et al. [94] utilized machine learning methods to analyze and score multi-dimensional assessments of courses, such as physics

and chemistry. In their implementation, 26,000 samples and 6,700 students undertook the responses, whose result (Cohen's $k = 0.40\text{--}0.75$) to excellent (Cohen's $k > 0.75$) was significant and improved the scoring placed on teachers. Nevertheless, there are still inherent challenges in scoring three-dimensional assessments compare to two-dimensional assessments, hence, requiring further research. Monllaó et al. [95] deployed a supervised machine learning framework to assess and identify students at risk of dropping out of massive open online courses (MOOC). The study utilized attributes, such as student enrolment details, courses, and site users' information for the prediction. The researchers used a Python machine learning backend alongside with supervised learning framework for the assessment. The student enrolments of 46,895 in MOOCs eight times from 2013 to 2018 were analysed. The results showed that the average accuracy achieved is 88.81% having a 0.9337 F1 score and an area ROC curve of 73.12%. However, the assessment model used a few attributes instead of multi-variant attributes in the prediction and might be difficult to generalize to most of cases dropping out risk in MOOCs and other learning analytics. A recent work by Tsiakmaki et al. [15] presented automated machine learning (autoML) techniques such as classification and regression to predict the outcome of student participation in online learning using decision trees and rule-based classifiers. The data were collected through a custom plugin for Moodle. Nevertheless, such a method is error-prone with high computational time. AutoML purely lacks basic features such as interpretability and decision-making of the produced results and may be prone to failure. An early warning system based on educational big data has also been implemented to solve student performance challenges [96, 97]. In addition, Ciolacu et al. [98] implemented Education 4.0 with learning analytics and machine learning algorithms to predict the final score from the final examination taken by the students. Machine learning helps in gifted education [92]. This was demonstrated using neural networks and supervised learning to achieve human elevation. Atkinson [99] utilizes machine learning models to analyze data streams having an intuitive parameterization by deploying the use of linear regression and K-means clustering to balance streaming data. Furthermore, machine learning methods have significantly changed and improved education development and school administration through its excellent libraries [12, 93, 100]. In addition, students' behavioral and academic data were used to ensure effective prediction of the performance of student success using machine learning models [101] and effective access to mobile learning platforms, especially during the period of the COVID-19 pandemic using J48 machine learning classifier, and achieved 89.37% accuracy [102].

A recent study by Sreenivasulu et al. [103] utilized machine learning methods to predict the low performance of pupils and dropout rates using student grades from the

Table 2 Machine learning technique, strength, and weakness

Methods	Description	Strength	Weakness	References
Decision tree	Recursively splitting extracted features based on observed inputs. It predicts the value of a target label based on input attributes. In addition, visually exemplify and inform decision-making	Good at handling general features and requiring few parameter tuning. It supports linear solutions	Interpretability of an ensemble of trees is difficult	[40, 106]
Logistics regression	This is based on the probability of data attributes and is widely used in data analysis for prediction	Has good performance in small data sets and probability-based interpretability	Only provide linear solutions and data are compiled based on assumptions	[31, 41]
Naïve Bayes	This determines the probabilistic relationship between the classes and their attributes	Provide independent assumptions for data	Require data to be transformed into multiple binaries to obtain enhanced performance	[45, 46, 50]
Support Vector machine	It segregates two classes (hyper-plane or line), which is applied to solve regression or classification problems	Ability to handle large dimensional data and does require strong assumption	A long time is taking to train the support vector machine model	[13, 51, 52]
K-nearest neighbors	It selects first the k closest samples for a test sample from all the training samples, and the rest of the samples are predicted with a simple classifier known as the majority classification rule	It is simple and can handle linear complex data intuitively	The user must define the number of variables and record poor performance for data with irregular shape/object data	[13, 41, 53, 54]
Random forest	Widely for high-performance data analysis and require few parameter tuning to achieve optimal results	Reduce errors over-fitting, and missing values, outperform on imbalanced data sets	It requires complex decision tree integration and has high computation time	[7, 55, 56]

department. Nevertheless, in-depth and more machine learning models analysis are professed to train the data set and attributes for efficient prediction of student academic performance. In addition, Kanetaki et al. [104] implemented a hybrid machine learning model to solve grade prediction of student performance in online courses using grade as a predictor, which achieved almost fitting success. However, more dependable variables such as attendance, GPA, subjective, etc. for prediction of student academic success. More recently, Chen and Ding [105] used machine learning models such as decision trees, random forest, logistic regression, support vector machine, and neural networks to obtain testing accuracies of 48%, 54%, 50%, 51%, and 60%, respectively, to predict student academic performance in Pennsylvania's school. These, however, required more data analysis, training, and testing of the models for consistent output.

3.3 Deep learning method

Deep learning [24, 107] is one of the new concepts associated with computational intelligence derived from the machine learning phenomenon, where high-level features in data are

modeled and have been recently applied in educational data mining (EDM). Deep learning (DL) as an approach for analyzing educational data encompasses various categories of neural networks, such as deep neural networks (DNN) [108], recurrent neural networks (RNN) [109], convolutional neural networks (CNN) [110, 111], and artificial neural network (ANN) [112]. These represent features from low to high levels structures. Deep learning has become very popular among researchers to tackle issues associated with flawed data, partially missing values, and a large volume of data collected from online education environments and click streams, etc. [113]. Deep learning methods are utilized to resolve various issues in the area, such as machine translation, environment monitoring, natural language processing, and mostly in educational data mining. In addition, deep learning methods have also been deployed for diverse tasks in education data mining through automatic feature representation and modeling. The increase in deep learning approach in data mining is driven by its ability to extract salient features from varied raw data without necessary high dependence on strenuously hand-crafted features. Contrary to conventional feature modeling algorithms such as support vector machine (SVM), k-nearest

neighbor (k-NN), etc. which are incapable of handling automatic feature representation from the raw data sets, deep learning-based models can automatically extract important feature sets from raw data with little preprocessing [114].

In addition, various researchers have utilized deep learning methods to build a knowledge block and created exceptional results in education data mining due to its exhibited characteristics, such as flexibility, robustness, and system performance enhancement. For instance, detecting cheating in an online examination electronically [108] using deep learning. The results showed an accuracy of 68% for DNN (deep neural network), 92% for LSTM (long–short-term memory), and 95% for RNN (recurrent neural network). In addition, Hand and Xu [115] utilized exclusive learning paths based on deep learning and image detection to develop ecological evolution paths for a smart education platform. The DL provides algorithm-based smart education level analysis with in-depth instruction on how to review and learn from a thread on image detection. In addition, Waheed et al. [116] analyzed academic data from virtual learning environments (VLE) using deep learning models (DLM) to predict students at-risk of dropping out to provide early intervention. The proposed model achieved an accuracy of 84–93% with a deep artificial neural network (DANN). Furthermore, the deep learning method facilitated the analysis of difficult topics derived from education text mining and identification of what makes the learning of a subject matter problematic and achieved 90% accuracy [117]. More so, Artificial Neural Networks (ANNs) models were utilized in the analysis with data sets (academic data, socio-economic status data, etc.) of 162,030 Colombian students to predict their students' academic performance. The proposed model achieved a high-performance accuracy of 82% and a low performance of 75% accuracy [112, 118].

In a recent study by Kishore et al. [119] implemented learning analytics for educational institutions using a deep learning approach with a university data set from the University of California Irvine (UCI) Machinery. Nonetheless, the result fall short for higher institution administration and required more in-depth analysis. More recently, Pei and Lu [120] also implemented Intelligent Educational Evaluation System using the Deep Neural Networks (DNN) method to achieve a training time of 54.92%. However, more critical training is required to improve the performance of unbalanced data classification problems. The strengths and weaknesses of this method are provided in Table 3.

3.4 Meta-heuristic optimization method

This method includes a genetic algorithm, ant colony, particle swarm, and artificial immune system. Therefore, the methods are discussed in the subsection, and strengths and weaknesses are presented in Table 4.

3.4.1 Genetic algorithms method

Genetic algorithms [75] are search techniques dedicated to computing complex spaces. Mostly deployed in different stages of educational research processes. It further helps to determine the degree of success impact recorded in courses offered earlier based on the curriculum of short courses. Categorically, genetic algorithms (GA) are subjected to the targeted method and random search of space solutions during a search for a global optimum. For example, in the case where needed information or optimum path is not gotten from the stored data then GA can be applied. This is because the whole set of conceivable solutions is not explored by GA. In addition, GA may be viewed in terms of the stochastic method which means that it is not dependent on any possible initial value; however, it is based on its application. That is to say, it is possible to find a global extreme of a specific target function using a certain probability.

Furthermore, various tasks have been accomplished with the aid of the genetic algorithm method. These include very large small integration (VLSI) circuit layout, solving systems of nonlinear equations, quantum computing, vector processing, mining classification rules in significant data sets, software testing, and educational data. For instance, Dedic et al. [75] utilized the GA in balancing educational curriculum thereby drastically improving the quality of teaching in the higher institution. Nevertheless, students who did not develop skills and competencies based on abstract thinking due to the goal set for the course are likely to underperform. An attempt was made to proffer solutions to this recurrence decimal by splitting the course modules to achieve desired goals. Nonetheless, more research is needed to investigate the use of GA to solve complex structures in other fields of study and to develop more flexible knowledge in terms of the delivery process. In addition, GA was utilized to solve the problem of student result grouping [124]. More important note, the global interest in educational research centered on how to improve and enhance teaching, learning, and decision-making process in educational environments. Several research efforts have been made from a different perspective including a methodological approach during the academic planning in the institution. The genetic algorithm [78] in its nature has been utilized to solve the inherent uncertainties such as academic planning problems as a result of the resignation and retention of resourceful persons engage in teaching, planning academic programs, staff allocation, and optimizing academic course activities in the educational environments. More genetic algorithms approaches were utilized during the personalized remedial learning system for understudying the object-oriented aspects of Java programming language, which resulted in significant overall learner's pre- and post-test scores changes [125]. In addition, a recent study by Sendari et al. [126] utilized GA to optimize

Table 3 Deep learning methods, strengths, and weaknesses

Methods	Description	Strengths	Weaknesses	References
ANN	These are interconnected nodes called neurons that receive input from the initial layers, execute the mathematical operation, and pass the result/output to the next layer	Capable of learning any non-linear functions	Involve error risk, no specific rules for designing the algorithm. The algorithms depend on black boxes	[112, 121, 122]
RNN	This is dedicated to processing sequential data (e.g., time series, speech, natural language, etc.)	Captures the sequential information present in the input data, also, suitable for temporal dynamic behavior exhibition	Has a problem of vanishing and exploding gradient and the likelihood of long dependency likelihood	[24, 109]
CNN	This is normally deployed for image and video processing jobs. It utilized convolutional filters to the input image, then, extract features and proceed to learn spatial patterns	A convolutional neural network is robust and automatically learns features from raw data. Furthermore, it is invariant to changes in the learning data and education methods	Require large data sets to be collected from massive online and open courses to ensure efficient extraction of relevant data. However, it is difficult to collect such data to reduce overfitting	[110, 111]
DNN	This is utilized for more complex jobs, such as speech recognition, computer vision, natural language processing, etc. In addition, during training, it updates the weights and biases of the neurons with backpropagation	Uses canonical (simple representation of data entities and their relationship) data of student learning. In addition, has fewer neurons and higher accuracy	It has higher tendencies of overfitting and computational complexity	[120, 123]

educational game applications through K-means and fuzzy C-means approaches for clustering questions. This achieved some level of optimization in data grouping to solve the problem of clustering questions. Nonetheless, more approaches or methods are needed to setup and validate the desired results.

3.4.2 Ant colony optimization method

Ant colony optimization [18] is a meta-heuristic optimization approach deployed for solving combinatorial optimization problems. Ant colony optimization (ACO) was coined from the behavior and pattern of an ant while sourcing food cooperatively. Ants crave way like a trailing manner in auspicious paths for other ants to apply. Going by the definite paths created by ants to a full solution to the ant optimization problem. Then, the nature of the solution is envisaged, and based on the nature of the solution found, the trial path is simplified. The ACO trail cooperation is guided by iterations and traversals to follow the paths that have the likelihood and maximum pheromone trail [124]. The essence of ACO is to focus on finding a definite path of solutions toward the predefined problem, since it is a meta-heuristic. In the educational data analysis approach, ACO normally turns in a knowledge domain that is displayed as a graph based on their attributes in which learning materials represent nodes and students represent ants walking on the edges finding the

best path. Hence, the knowledge of the ACO method for solving combinatorial optimization problems for improving e-learning systems and finding optimal learning paths based on self-organization has become part of the educational system. In their recent study, Rastegarmoghadam and Ziarati [127] utilized ant colony optimization to improve the modeling of adaptive tutoring systems. The proposed method offers personalized learning objects based on learning and style of problem-solving. Hence, ACO is deployed to generate the adaptive optimal learning paths, thereby achieving high performance and optimal learning paths based on learning styles. However, a real-life scenario experiment is further required in a learning environment to fully analyze and assess the performance of ant colony optimization and learners' behavior.

3.4.3 Particle swarm optimization method

Particle swarm optimization [128] is a meta-heuristic algorithm normally applied to solve problems of combinatorial optimization, discrete, and continuous. It is a nature-inspired algorithm that is utilized in the search for optimal neural network architectures. Moreover, the optimization methods are designed for simulating the behavior and social conduct of fish schools and animals, such as a bird flock or fish, where each animal moves in different directions while trying to interact with each other as both positions and velocities

Table 4 Meta-heuristic optimization methods, strengths, and weakness

Methods	Description	Strength	Weakness	References
Genetic algorithms	It is a search technique dedicated to computing complex spaces. It helps to determine the degree of success impact recorded in courses offered earlier based on the curriculum of short courses	Support multi-objective and stochastic optimization. It does not require derivatives, and the search is based on population and payoff information	It is sometimes trapped in local optima with minimal information for training. In addition, the method is computationally intensive	[75, 125, 126, 137, 138]
Ant colony optimization	This solves combinatorial optimization problems. Its focus is on finding a definite path of solutions toward the predefined problem, since it is a meta-heuristic	Easy adapt to a new change, guaranteed convergence, involve parallel search in a population, and rapid discovery of the good condition	A probability distribution can change for each iteration, theoretical analysis is difficult	[18, 138, 139]
Particle swarm	This solves problems of discrete, continuous, and combinatorial optimization. It is a nature-inspired algorithm that is utilized in the search for optimal neural network architectures	Simple to implement with fewer parameters to adjust. It is robust and capable of running parallel computation with shorter computation time	Difficult to define initial design parameters, scattering optimization problems, and non-coordinate system	[128, 131, 133]
Artificial immune system	It involves biologically inspired algorithms based on existing knowledge functions of the immune system in animals	The iterative process adjusts the initial set of the detector to best cover the search space, and eliminate the need for the preliminary learning phase in the protected learning environment	Hard to analyze, and poor practical security tool or real-world application issue	[134, 140, 141]

are being updated for the better position that could help in achieving an optimal solution. The new velocity is applied to modify the position of the particle. Therefore, the particle means the bird or fish which represents a candidate solution, hence, the candidate solution is evaluated concerning an objective function [18]. Note that PSO performance is sensitive to the fact that values assigned to its control parameters and normal tuning of the referred parameters improve the routine performance. The characteristics of PSO such as robustness, simplicity, and global search capabilities qualify its uniqueness to be tagged as one of the best intelligence-based algorithms. Though, it is still faced with susceptibility problems [18].

Different PSO methods such as hybrid K-Means PSO (KMPSO), sequential approximation optimization assisted particle swarm optimization (SAOPSO), modified particle swarm optimization (MPSO), and particle swarm optimization convolutional neural network (psoCNN) have been deployed for optimal data computation and analysis. For instance, Sherar and Zulkernine [129] utilized the KMPSO method to achieve better clustering output compared to

Spark's built-in clustering algorithms and handle large and complex workloads through efficient scaling of resources. In the aspect of complex problems, the SAOPSO method was utilized for improving the optimization efficiency which has a good balance between the searchability of the SAO and the PSO [130]. In tackling the problem of high dimensional variables, the MPSO method with Gaussian mutation was deployed for minimal modification and updating of an already prevailing laid-down structure with 13 variables. Much significant time and computational resources were still spotted with the current system. However, the psoCNN method was introduced to overhaul the envisaged challenges by quickly finding good CNN architectures that achieve quality performance [128, 131]. Another study by Yang [132] implemented particle swarm optimization to solve the problem of imbalance of educational resources, thereby allocating computer teaching management resources effectively. The process enhances and prevents resource waste and the utilization rate of teaching resources.

Recently, a study by Sheng et al. [133] attempted to find optimal curriculum sequencing in the educational management system. The authors proposed a novel metaheuristic algorithm known as the group–theoretic PSO method that partially solved the adaptive curriculum sequencing (ACS) in an online teaching and learning system development. However, mapping functions to transfer discrete and continuous representations, so that the continuous optimization problems are tackled is not covered.

3.4.4 Artificial immune system method

The artificial immune system [134] is one of the metaheuristic approaches that involved biologically inspired algorithms based on existing knowledge functions of the immune system in animals. The artificial immune system (AIS) shares similar patterns to that of machine learning methods. Moreover, the theory of AIS is based on animal body reactions against the antigen. There are infinite numbers of white blood cells ‘*B*’ and ‘*T*’ cells and each ‘*B*’ or ‘*T*’ cell has selective properties and is subtle to a specific type of antigen [135]. AIS operates with four major theories. These theories [136] include clonal selection, negative selection, danger theory, and immune network. The theoretical application of AIS has been deployed in educational data analysis for solving complex computational and even engineering problems. For instance, the artificial immune system [134] played a significant role in routing protocol concerning mobile ad-hoc networks. The optimized link state routing (OLSR) protocol mechanism selects the shortest flow between source and destination by Dijkstra’s algorithm. Therefore, AIS–OLSR is derived to enhance the energy efficiency using some concepts, such as hop count, remaining energy in the intermediate nodes, and distance among nodes, which is realized by negative selection and clonal algorithms of AIS. More so, in solving the issue of size and location optimization of distributed generation within distribution system, AIS was utilized by combining clonal selection (CS) principle and particle swarm optimization (PSO). Table 4 outlines the strengths and weaknesses of meta-heuristic optimization methods.

3.5 Ensemble techniques method

Ensemble techniques combine multiple models into one and tend to produce superior performance results compared to standard methods [142]. Ensemble techniques create better accurate classification tree models through a weak combination of the tree models to ensure stronger versions are created. The traditional statistical models are characteristically not good for the analysis of enormous and complex data sets, hence, the integration of multiple classification algorithms

would increase the robustness, accuracy, and overall generalization of the data mining models [143].

Therefore, ensemble techniques [144] provide a diversity of opinions on multiple classification algorithms that are aggregated to provide a final decision. Ensemble learning combines several models, such as Bayesian [46, 145], Bayesian networks [146, 147], naïve Bayes classifier [48, 54], etc., to improve machine learning results in educational big data. These machine learning techniques are fused into a single classification model to decrease variance, and algorithm bias and improve prediction. The methods are categorized into two; (i) sequential ensemble methods are how learners are generated sequentially (e.g., Adaboost) and (ii) parallel ensemble methods are the types, where the base learners are generated in parallel (e.g., is Random Forest) [13]. Similarly, the entire data in supervised learning and image classification are usually classified into three segments; (i) training, (ii) testing, and (iii) validation data. The testing data are applied after training data using an image classifier. This is considered to achieve the global optimal performance on the entire set of the three segments. Since algorithms are required to extract information from a data subset, the whole data set, the ensemble algorithm utilizes methods, such as bootstrapping, boosting, bagging, and stacking to design the ensemble algorithm. And also, preferred performance estimation can be achieved through the integration of multiple learning algorithms, especially when compared to any other baseline constituent learning algorithms in most scenarios, and are best suited for analyzing a complex educational data set. In recent times, assemble techniques have been utilized in solving educational large data set issues. For instance, a study by Zhang et al. [3] applied an ensemble model based on weighted voting to predict student achievement from two secondary education schools in Portugal. The proposed model uses five classifiers, such as random forest, adaptive boosting, gradient boosting decision tree, extreme gradient boosting, and decision tree, whose prediction performance was impressive. This, however, needs more baseline outlier detection and prediction methods. Furthermore, the value of extracted knowledge is proportional to the quality of the data used in the context of the knowledge discovery process; however, noise affects the data quality. For example, the recent work by García-Gil et al. [34] proposed two methods to address these issues. These methods include homogeneous ensemble and heterogeneous ensemble filters, with much stress on the scalability and performance traits which proved efficient in the smart data set obtained from educational big data. More recently, ensemble techniques have been utilized in educational big data analysis over smart learning improvement [13], predicting student academic performance [144], dropout in education [50], effective data integration [146], and detecting trustworthy users in the social media networks communication

[147]. In addition, a rapid hybrid clustering algorithm was implemented in a recent study to tackle the problem of large volumes of high-dimensional data [148].

A recent study by Ajibade et al. [149] implemented ensemble algorithms such as Bagging, Boosting, and Voting Algorithms in an e-learning system to predict student academic performance using behavioral features. This obtained a certain level of performance accuracy of 96%. This, however, did not cover demographic and academic features for prediction. In addition, a more recent study by Safarov et al. [150] proposed a deep learning approach using a recommendation system in an e-learning business platform to find the most appropriate material for educationalists. Thus, achieved equal precision rates of 0.626 and 0.492 in the cases of Top-1 and Top-5 courses, respectively. However, it requires more in-depth data analysis to overcome resource overload during learner's generation and materials retrieval. Table 5 presents the strength and weaknesses of these methods.

3.6 Markov model for decision-making method

The Hidden Markov Model [151] is a computational intelligence model that involves random sequence generation observation of the chains of the Hidden Markov Model (HMM). The random process presents a set of states in such a way that the probability of the identified states at the next observation depends solely on the process of the state history [152]. It was innovatively first applied in the ecological field and spread to other fields of study. This is essential in the sense that the recent computational intelligence utilizes the unique chains for forecasting time series data of any event. In temporal cyber security big data, Teoh et al. [153] outlined the successes recorded in various areas such as finance, bio-informative, artificial intelligence, healthcare, agriculture, and signal recognition by the applications of the Hidden Markov model. For instance, Hidden Markov is implemented for the identification of major factors, such as learning behavior data, basic information, curriculum scores, and data participation in massive online courses (MOOC) learner classification and analysis [151]. Hence, it helps MOOC developers to develop an improved curriculum and provide enough references for teaching development. In addition, the Hidden Markov model accelerates the motivation in MOOC learning [14]. The learning behavior properly analyzes from the learning activities and continuous learning which tremendously use for the development of MOOC learning. Moreover, the hidden Markov Model was successfully applied in mechanisms for family formation [152]. The two-course models analysis such as sequence analysis and event history analysis were targeted for the decision-making process. The course events were explained as the experiences of an individual's narrative of all events. HMM are extensively utilized in the analysis of educational data to

inform decisions according to a current study by [154], the author utilized Markov Model to evaluate the physical education teaching quality in colleges. The experimental study achieved an accuracy of 90.3% compared to other baseline analyses. Furthermore, the two-layer Markov model (TL-MHH) produces high output during the modeling of MOOC student behavior patterns through unsupervised learning on large data sets of an observed sequence of the student [155]. Nevertheless, further studies are required to generalize the analysis of the log of any course to ensure an in-depth understanding of student behaviors and the correlations concerning behaviors and other variables like grades.

Recently, [156] adopted the Markov neural network–BP model and countermeasures to develop a scale of vocational education. The model obtained MSE (mean squared error) and MRE (mean relative error) of 10.184 and 5.017, respectively, which is fair in solving the enrolment and scale issues in Higher Vocational Education (HVE) in China. However, more models are required to carry out effective future predictions of HVE for better strategy and educational planning. A more recent study by Xu & Xia [157] implemented a speech recognition system to accelerate remote vocal music teaching using the Markov model. The study attempted to construct and analyze the development of distance learning systems. However, the proposed system lacks details analysis and training modeling approaches to accurately test and validate the claim. Table 5 presents the strength and weaknesses of these methods.

4 Educational big data analytics computational intelligence applications areas

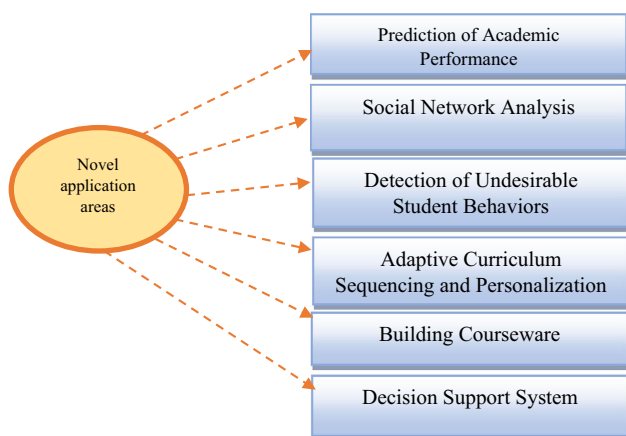
Educational big data analytics and computational intelligence methods have changed the perception of our learning ability, and computing power, and revolutionized Industry 5.0, and the entire education sector. This section provides a novel area of applications of computational intelligence methods to tackle modern challenges because of the complexities of data acquisition and its analysis. The vivid areas include prediction of academic performance, social network analysis, detection of undesirable student behaviors, adaptive curriculum sequencing and personalization, building courseware, and decision support system. Figure 7 shows the novel application areas.

4.1 Prediction of academic performance

Recently, computational intelligence for the analysis of educational big data analytics has been widely utilized in the areas of student academic performance (SAP), student dropout prediction (SDP), and curriculum development

Table 5 Ensemble and Markov model, strength, and weakness

Methods	Description	Strengths	Weaknesses	References
Ensemble technique	Multiple models are combined to produce superior performance results compared to standard methods through the learning method	Easy handling of complex and high-dimensional data is	Suffer from a lack of interpretability and it is computationally expensive	[3, 142, 144, 158]
Markov model	It is a random sequence generation observation of the chains of the Hidden Markov Model (HMM). Uniquely use chains for forecasting time series data of any event	It has a strong statistical foundation and flexible generation of sequence profiles. The computation methods are easier to implement and analyze and eliminate the label bias problem	The method is computationally expensive both in memory and time. In addition, it does not match the predicted target function, and a huge number of parameters are required for evaluation. Other weaknesses include issues patterning lack of defining joint probability over observation, label sequence, and use of only positive data to train	[151, 152, 159, 160]

**Fig. 7** Computational intelligence methods application areas

(CD). Computational intelligence, such as machine learning techniques, deep learning algorithms, artificial intelligent approaches, etc., have been applied to enhance academic performance and improve learning interactivity, and decision processes in higher institutions. Consequently, computational intelligence methods are deployed to resolve issues related to the lack of curriculum upgrades, and data analysis for informed decision-making by educational administrators. In recent times, quite a lot of researchers have developed various computational intelligence methods to predict academic student performance. For instance, Yu et al., [144] applied ensemble classifiers using learner data to predict student performance which helps educational administrators and policymakers to improve the higher institution and other

mediums of learning. In addition, Tsiakmaki et al. [15] implemented machine learning techniques to predict the student's academic performance by exploring students' learning outcomes concerning engagement in online learning platforms. However, an indebt analysis is requested to further comprehend the reasons a student is prone to fail to enhance learning activities. More so, Márquez-Vera et al. [161] used interpretable classification rule mining (ICRM) algorithm to predict the early dropout of students in the first 4–6 weeks of undertaking a course. Nonetheless, the development of an early warning system (EWA) intervention is required to evaluate the effect of the various intervention mechanism adopted to ensure students at risk of dropout are supported. Furthermore, the intelligence mechanism for analyzing educational big data has been widely adopted by various studies. This can be viewed in the application of a genetic algorithm which improves accuracy up to 1.09–24.39%, especially when compared to other baseline methods between 0.29% and 6.57% for student academic performance prediction [76], low- and high-achieving students' performance [35, 42]. In addition, the methods have been applied in temporal analysis for dropout prediction exhibited in self-paced MOOCs through self-regulated learning strategies [32].

4.2 Social network analysis

Social network analysis (SNA) is a technique that views social relationships that depends on social network theory [162]. SNA model application [16] looks forward to achieving student models to appear in graph form to X-ray various forms of relationships between them. Therefore, SNA has

attracted importance in educational data mining and intelligent computing, hence to this effect; it helps to analyze the relationships amongst students or individuals, tutors, and the learning environments. SNA is implemented in educational data mining (EDM) to ensure the structure and relations in collaborative tasks and interactions with communication tools as applied in [19]. The major essence of SNA is to focus on measuring the relationships amongst entities in an information-networked environment for effective learning and to aid decision-making. The main goal [163] of social network analysis are: (a) to ensure how a student from different populations form ties with outsiders, (b) to find a particular student within a group influence, (c) to locate the minimum required that is directly tied to link/connect two or more students or individuals, and (d) to accelerate the social structure comprehension of a student resides in an institutional environment. Moreover, social network analysis has recently performed creditably in different aspects of education, for instance, it detects hidden learning groups and communities [164] and measures emotional responses by Facebook [165]. In addition, SNA was utilized in the evaluation of student learning performance [166].

4.3 Detection of undesirable student behavior

Detection of undesirable student behavior (DUSH) is similar to student academic performance but differs in terms of contexts and concepts. DUSH [16, 19, 24] is important in student behavior, such as low-level motivation, preventing and advising against gang formation, academic poor performance, drug abuse, cheating in an examination hall, dropping out of courses, etc., to attain their educational career and aspiration. The computational intelligence techniques leveraged in this aspect are artificial intelligence, deep learning, and machine learning. Several studies have utilized these techniques for the detection of unusual student behavior. For instance, Wang et al. [167] proposed computational intelligence such as DNN, CNN, and RNN from KDD-2015 data sets to predict the chances of dropout and completion in MOOCs competition which achieved good results compared to baseline feature engineering methods. Student behavior was modeled in virtual learning in an educational environment using co-embedded DNN and RNN to predict the dropout from MOOCs and identify the knowledge tracing in an intelligent tutor system [168]. In addition, Xing and Du [169] proposed deep learning for personalized intervention for students that are likely to drop out of MOOC. Recently, improved quantum particle swarm optimization (IQPSO) based computational intelligence algorithm was applied to predict student dropout based on their learning behavior in MOOCs [170]. More recently, Park et al., [17] presented a framework to detect descriptive talk in student chat-based

discussion and collaboration in a game-based learning environment using machine learning and deep learning models. The framework was developed to investigate how it affects the learning process. In addition, learning style was identified using artificial neural network (80.7%) and particle swarm optimization (79.1%) that enhance automatic learning style identification and have better accurate personalization toward learning styles in school premises [171].

4.4 Adaptive curriculum sequencing and personalization

The computational intelligence approaches improve adaptive curriculum sequencing and personalization (ACSP) in educational environments. Personalization ensures learning content is tailored to accommodate specific students or individual users based on the traits of the users, learning suggestions, and outcomes [172]. Adaptive Curriculum Sequencing [18, 173] involves an arrangement of an important thought for personalization, in which it points out the sequence of recommended learning materials that has to match student models with a particular process as related. The need for ACSP is to solve the problems derived from the diversity of sequences that are usually experienced through big data repositories of learning materials. Moreover, the particle swarm method and genetic algorithm were applied in ACSP to identify different parameters such as intrinsic and extrinsic contents that could be of interest to students and knowledge domain modeling [18]. Various studies have applied computational intelligence models for ACSP in different areas. For instance, machine learning techniques were applied in ACSP in the improvement of gamification design for the student to achieve maximum learning objectives [174]. Similarly, a personalized learning system was developed using K-NN classification that provides feedback to students as they play and design games for educators [175, 176]. In addition, Moon et al. [177] proposed a conceptual framework to ensure the best way to promote learners' and tutors' personalized open provision of education resources (POER) is located. Furthermore, research by Lee and Ferwerda [178] proposed personalization in an online learning environment and engagement tracing using machine learning methods, which in furtherance detect undesired learning behaviors.

4.5 Building courseware

Building courseware [16, 19] involves the phenomenon by which educational software is designed to provide articles, audio and videos, and some other substantive learning materials in an educational environment. Here, computational intelligence is applied to assist educators in automatically generating courseware (course and software) materials and learning contents for student usage information. In their recent study, Singh and Sunil [179] developed courseware for MOOC and other web-based learning management systems using machine learning methods. The machine learning methods were used to improve and validate the courseware and support the personalization of students' learning in the evolving e-learning landscape. Educational technology-based MOOC platform on courseware resource storage system was also developed with the knowledge of computational intelligence [180].

4.6 Decision support system

A decision support system [16] is another crucial application area of computational intelligence methods. It involves the process of learning to assist stakeholders or educators to make an effective decision. Specific support areas include resource planning, enhanced courseware development, feedback mechanism from learners, recommendation generation, etc., which is targeted more at tutors, school administrators, and students. Educators or education administrators utilized computational intelligence techniques to create an alert system in real-time in case of low motivation, misuse, dropout, cheating, etc., by the students. This was observed in a recent study by [15] that developed automated machine learning (autoML) techniques such as classification and regression to predict the performance outcome of the students while participating in online learning using decision trees and rule-based classifiers. The data was collected through a custom plugin for Moodle. Nevertheless, such a method is error-prone with high computational time. AutoML purely lacks basic features such as interpretability and decision-making of the produced results. More so, an early warning system has also been implemented to solve student performance challenges [96, 97]. Nonetheless, there are still inherent issues with the existing system, such as result orientation and decision-making, etc. Furthermore, in planning and recommendation generation, a work by Sugiyarti et al. [181] predicted the final grades of the candidate using the decision tree C4.5 (94.73%) based computational intelligence method. The grades help the facilitator in selecting candidates eligible for a scholarship. However, other CI techniques such as GA, ACO, and PSO require leverage in further study.

5 Discussion

This study extensively surveys academic articles on computational intelligence for educational big data analysis (CIEBDA). Precisely, the study is split into six (6) sections. These include an introduction, educational big data analytics, computational intelligence (CI) for educational data analysis and its methods, novel application areas, discussion, and conclusion. Therefore, from the study, we observed that big data analytics played a unified role in educational data mining (EDM) as it infers meaning through a definite process. Our finding observed educational big data analytics are classified into five (5) processes, namely, data collection, data preprocessing, feature extraction, modeling, and evaluation measure, which are sequentially followed for efficient data analysis and optimized learning [4, 8, 13, 34, 41]. The findings of this extensive survey clearly show that computational intelligence methods for educational big data analysis and its applications have been significantly utilized in education. We discovered that computational intelligence methods played a vital role in intelligent data analysis that aids effective decision. The findings of this paper show that Computational Intelligence methods can be categorized into five (5) methods, such as artificial intelligence [11, 58], deep learning [24, 112], machine learning [12, 95] methods, meta-heuristic optimization [18, 75], ensemble techniques [3, 13], and Markov model for decision-marking approach [14, 154]. This study further outlines novel application areas, where computational intelligence methods are utilized for effective implementation, educational development, and optimal decision-making. These areas include prediction of academic performance, social network analysis, detection of undesirable student behavior, adaptive curriculum sequencing and personalization, building courseware, and decision support systems [15–20].

6 Conclusion

This study focused on a survey of computational intelligence methods and its potentials applications to support educational big data analysis. We discussed educational big data analytics (EDA) and its processes, such as data collection, data preprocessing, feature extraction, modeling, and evaluation. We further review extensively computational intelligence and its methods such as artificial intelligence (AI) approaches, machine learning methods, deep learning methods, meta-heuristic optimization approaches, ensemble techniques, and the Markov model for educational big data analysis. Moreover, the novel application areas, which include prediction of academic performance, social network analysis, detection of undesirable student behaviors, adaptive curriculum sequencing and personalization, courseware development,

and decision support system were also explained. In addition, various Computational Intelligence methods were mapped to the novel application areas.

Nevertheless, research in educational big data analytics implementation has led to a more challenging research direction that can be further pursued. Some of these challenging research areas include enhanced academic performance prediction, data-driven intelligent tutoring systems, adversarial machine learning, student engagement, personalized learning, etc. Here, these important research directions are briefly discussed in ten (10) important themes.

- (i) *Enhancement of student academic performance prediction*: prediction of student academic performance (SAP) and student dropout prediction [15] have become generic problem that challenges the current algorithms. A combination of computational intelligence methods is sorted. The effective performance evaluation is purely based and dependent on the nature of educational data sets extracted and features selected. In addition, enhanced models such as the ensemble classifier and machine learning algorithms [15, 144] are necessitated to ensure optimal results.
- (ii) *Hybrid courseware development: the divergence view of educational course materials is on the rise*: Further research is required to ensure course materials are hybridized for better interactivity, user-friendly, and informative [179]. In addition, the development of scalable storage systems, retrieval, and integrated platforms to aid data analysis [180] is sorted.
- (iii) *Intelligent tutoring system*: integrating intelligence into our educational course curriculum development, instructing and learning process would make understanding of the data set generation, computation, and analysis. In addition, using AI and ML algorithms [11, 91, 92], in intelligent tutors will enhance the learner's progress and tailor instruction provision. More so, the learner's knowledge and skill level assessment and possibly identify areas, where the learners are finding difficult to understand, thereby providing additional resources, practice, and feedback in the identifiable areas.
- (iv) *Natural language processing*: processing natural language [38] has become a generative problem. The mechanism for developing algorithms that can generate human language and easily be understood should be increased. Researching for enhanced tools and techniques that best match to aid effective communication between AI systems and humans is required. This will further accelerate the analysis of student writing skills and ensure feedback on grammar, syntax, etc. thereby improving the academic performance of students.
- (v) *Adversarial machine learning*: this is an emerging area that required attention. It is essential to develop algorithms that are capable of detecting and defending against attacks on AI systems especially applying in related areas, such as finance, healthcare, etc.
- (vi) *Social network analysis*: analysing social network data [164, 166] has become a big challenge. As big data sets are generated and stored via social platforms. It is imperative to analyze the relationships between individuals in social media interaction and engagements. Hence, it could identify groups of students who intend to collaborate and develop a means for improving social engagement and collaboration activities.
- (vii) *Predictive analytics*: this encompasses the idea of using historical data to make a prediction concerning future events [41]. This, when developed, will assist to identify students on verge of dropping out from academic learning and participation.
- (viii) *Student engagement*: this area mainly focuses on academic success. This focus on the development of student behaviors and learning activities. Developing effective and enhanced strategies such as adaptive learning and gamification to improve engagement should be intensified.
- (ix) *Personalized learning*: development of appropriate educational resource materials through personalized learning techniques [175, 176] using efficient algorithms, such as data mining and machine learning. This would assist in tailoring educational experiences and unique learning styles to student
- (x) *Learning analytics*: this is an important area of computational intelligence toward achieving educational optimal data sets to aid effective big data analysis. Learning analytics [13] is the most effective approach to improving learning skills, design, and delivery of educational programs using data. This will help student performance data and gain insights into teaching methods and resource materials.

Author contributions All the authors contributed equally in conducting the research. The authors drafted, proof-read, and reviewed the manuscript.

Funding There is no external funding received by the authors.

Availability of data and materials (data transparency) Not applicable.

Code availability (software application or custom code) Not applicable.

Declarations

Conflict of interest The authors declare no conflict of interest for this article.

Ethical statement The submitted work should be original and has not been published elsewhere in any form or language to our knowledge.

References

- Bart, C., Karolina, L., Magdalena, J., Daniel, B., Michael, F., Stefania, A.: Lists of ethical, legal, societal and economic issues of big data technologies. Report. Leiden: Leiden University, pp 1–109 (2017)
- Romero, C., Ventura, S.: Educational data mining and learning analytics : an updated survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **10**(3), e1355 (2020). <https://doi.org/10.1002/widm.1355>
- Zhang, S., Chen, J., Zhang, W., Xu, Q., Shi, J.: Education data mining application for predicting students' achievements of portuguese using ensemble model. *Sci. J. Educ.* **9**(2), 58–62 (2021). <https://doi.org/10.11648/j.sjedu.20210902.16>
- Ikegwu, A.C., Nweke, H.F., Anikwe, C.V., Alo, U.R., Okonkwo, O.R.: Big data analytics for data-driven industry: a review of data sources, tools, challenges, solutions and research directions. *Clust. Comput.* **25**, 3343–3387 (2022). <https://doi.org/10.1007/s10586-022-03568-5>
- Bao, R., Chen, Z., Obaidat, M.S.: Challenges and techniques in big data security and privacy : a review. *Secur. Priv.* **1**(4), e13 (2018). <https://doi.org/10.1002/spy2.13>
- Verma, S., Sekhar, S., Kumar, S.: An extension of the technology acceptance model in the big data analytics system implementation environment. *Inf. Process. Manag.* (2018). <https://doi.org/10.1016/j.ipm.2018.01.004>
- Ikegwu, A. C., Nweke, H. F., Alo, U. R., Okonkwo, O. R.: HMC-PAED: a new framework for students' dropout prediction. In: *ICT4NDS2021: ICT and Sustainability in the 5th Industrial Revolution*, pp. 131–140. Available: lorinuijipc.com.ng (2021)
- Anikwe, C.V., et al.: Mobile and wearable devices for health monitoring: review of sensors, components modules, applications and future prospects. *Expert Syst. Appl.* **202**, 117362 (2022)
- Jain, P., Gyanchandani, M., Khare, N.: Big data privacy : a technological perspective and review. *J. Big Data* **1**(3), 1–25 (2016). <https://doi.org/10.1186/s40537-016-0059-y>
- Iqbal, R., Doctor, F., More, B., Mahmud, S., Yousuf, U.: Big data analytics and computational intelligence for cyber-physical systems: recent trends and state of the art applications. *Futur. Gener. Comput. Syst.* **105**, 766–778 (2017). <https://doi.org/10.1016/j.future.2017.10.021>
- Andrej, F., Boris, A.: *Artificial Intelligence in Education*. IntechOpen (2021)
- Belmonte, J.L., Segura-Robles, A., Moreno-Guerrero, A.J., Parra-González, M.E.: Machine learning and big data in the impact literature. A bibliometric review with scientific mapping in web of science. *Symmetry (Basel)* (2020). <https://doi.org/10.3390/SY112040495>
- Kausar, S., et al.: Mining smart learning analytics data using ensemble classifiers. *Int. J. Emerg. Technol. Learn.* **15**(12), 81–102 (2020). <https://doi.org/10.3991/ijet.v15i12.13455>
- Chen, Y., Han, D., Xia, L.: A hidden Markov model to characterise motivation level in MOOCs learning. *Int. J. Comput. Sci. Eng.* **23**(1), 42–49 (2020). <https://doi.org/10.1504/ijcse.2020.110189>
- Tsiakmaki, M., Kostopoulos, G., Kotsiantis, S., Ragos, O.: Implementing autoML in educational data mining for prediction tasks. *Appl. Sci.* **10**(1), 1–27 (2020). <https://doi.org/10.3390/app10010090>
- Bakhshinategh, B., Zaiane, O.R., ElAtia, S., Ipperciel, D.: Educational data mining applications and tasks: A survey of the last 10 years. *Educ. Inf. Technol.* **23**(1), 537–553 (2018). <https://doi.org/10.1007/s10639-017-9616-z>
- Park, K., et al.: Detecting disruptive talk in student chat-based discussion within collaborative game-based learning environments. *ACM Int. Conf. Proc. Ser.* (2021). <https://doi.org/10.1145/3448139.3448178>
- de Machado, M.O.C., Bravo, N.F.S., Martins, A.F., Bernardino, H.S., Barrere, E., de Souza, J.F.: Metaheuristic-based adaptive curriculum sequencing approaches: a systematic review and mapping of the literature. *Artif. Intell. Rev.* (2021). <https://doi.org/10.1007/s10462-020-09864-z>
- Romero, C., Ventura, S.: Data mining in education. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **3**(1), 12–27 (2013). <https://doi.org/10.1002/widm.1075>
- Ellison, S. F., et al.: Costs of managerial attention and activity as a source of sticky prices. Structural estimates from an online market (No. w24680). National Bureau of Economic Research. **24680**, 1–53 (2018)
- Talpur, N., Abdulkadir, S.J., Hasan, M.H.: A deep learning based neuro-fuzzy approach for solving classification problems. *Int. Conf. Comput. Intell. (ICCI)* (2020). <https://doi.org/10.1109/icci51257.2020.9247639>
- Iqbal, R., Doctor, F., More, B., Mahmud, S., Yousuf, U.: Big data analytics : Computational intelligence techniques and application areas. *Technol. Forecast. Soc. Change* **153**, 119253 (2018). <https://doi.org/10.1016/j.techfore.2018.03.024>
- Shu, H.: Big data analytics: six techniques. *Geo-Spat. Inf. Sci.* **19**(2), 119–128 (2016). <https://doi.org/10.1080/10095020.2016.1182307>
- Hernández-Blanco, A., Herrera-Flores, B., Tomás, D., Navarro-Colorado, B.: A systematic review of deep learning approaches to educational data mining. *Complexity* (2019). <https://doi.org/10.1155/2019/1306039>
- Limna, P., Jakwatanatham, S., Siripipattanakul, S., Kaewpuang, P., Sriboonruang, P.: A review of artificial intelligence (AI) in education during the digital era. *Adv. Knowl. Exec.* **1**(1), 1–9 (2022)
- Kaddoura, S., Popescu, D.E., Hemanth, J.D.: A systematic review on machine learning models for online learning and examination systems. *Peer Comput. Sci.* (2022). <https://doi.org/10.7717/peerj-cs.986>
- Pejic-Bach, M., Bertonce, T., Meško, M., Krstić, Ž.: Management text mining of industry 4.0 job advertisements. *Int. J. Inf. Manag.* **50**, 416–431 (2020)
- Khan, S., Shakil, K., Alam, M.: PABED a tool for big education data analysis. *IEEE Int. Conf. Ind. Technol. (ICIT)* (2019). <https://doi.org/10.1109/icit.2019.8755178>
- Rehman, M.H.U., Ahmed, E., Yaqoob, I., Hashem, I.A.T., Imran, M., Ahmad, S.: Big data analytics in industrial IoT using a concentric computing model. *IEEE Commun. Mag.* **56**(2), 37–43 (2018). <https://doi.org/10.1109/MCOM.2018.1700632>
- Tortonesi, M., Govoni, M., Morelli, A., Riberto, G., Stefanelli, C., Suri, N.: Taming the IoT data deluge : an innovative information - centric service model for fog computing applications. *Futur. Gener. Comput. Syst.* **93**, 888–902 (2018). <https://doi.org/10.1016/j.future.2018.06.009>
- Pierrakeas, C., Koutsonikos, G., Lipitakis, A.D., Kotsiantis, S., Xenos, M., Gravvanis, G.A.: The Variability of the Reasons for Student Dropout in Distance Learning and the Prediction of Dropout-Prone Students, Vol 158. Springer International Publishing (2020). https://doi.org/10.1007/978-3-030-13743-4_6
- Moreno-Marcos, P.M., Muñoz-Merino, P.J., Maldonado-Mahauad, J., Pérez-Sanagustín, M., Alario-Hoyos, C., Delgado Kloos, C.: Temporal analysis for dropout prediction using self-regulated learning strategies in self-paced MOOCs. *Comput. Educ.* (2020). <https://doi.org/10.1016/j.compedu.2019.103728>

33. Dai, H.N., Wong, R.C.W., Wang, H., Zheng, Z., Vasilakos, A.V.: Big data analytics for large-scale wireless networks: Challenges and opportunities. *ACM Comput. Surv.* **52**(5), 1–29 (2019). <https://doi.org/10.1145/3337065>
34. García-Gil, D., Luengo, J., García, S., Herrera, F.: Enabling smart data: noise filtering in big data classification. *Inf. Sci. (NY)* **479**, 135–152 (2019). <https://doi.org/10.1016/j.ins.2018.12.002>
35. Hussain, S., Dahan, N.A., Ba-Alwib, F.M., Ribata, N.: Educational data mining and analysis of students' academic performance using WEKA. *Indones. J. Electr. Eng. Comput. Sci.* **9**(2), 447–459 (2018). <https://doi.org/10.11591/ijeecs.v9.i2.pp447-459>
36. Chen, J., Feng, J., Sun, X., Wu, N., Yang, Z., Chen, S.: MOOC dropout prediction using a hybrid algorithm based on decision tree and extreme learning machine. *Math. Probl. Eng.* (2019). <https://doi.org/10.1155/2019/8404653>
37. Kemper, L., Vorhoff, G., Wigger, B.U.: Predicting student dropout: a machine learning approach. *Eur. J. High. Educ.* **10**(1), 28–47 (2020). <https://doi.org/10.1080/21568235.2020.1718520>
38. Lee, H.S., Lee, J.: Applying artificial intelligence in physical education and future perspectives. *Sustainability* **133**, 351 (2021). <https://doi.org/10.3390/su13010351>
39. Olson, D.L., Lauhoff, G.: *Descriptive Data Mining*. University of Nebraska, Lincoln (2019). https://doi.org/10.1007/978-981-13-7181-3_8
40. Pelaez, K., Levine, R.A., Guarcello, M.: Using a latent class forest to identify at-risk students in higher education. *J. Educ. Data Min.* **11**(1), 18–46 (2019)
41. Amelec, V., Alexa, N.S., Hugo, P.H., William, N.N., Leonardo, N.N.: Using big data to determine potential dropouts in higher education. *J. Phys.* (2020). <https://doi.org/10.1088/1742-6596/1432/1/012077>
42. Asif, R., Merceron, A., Ali, S.A., Haider, N.G.: Analyzing undergraduate students' performance using educational data mining. *Comput. Educ.* **113**, 177–194 (2017). <https://doi.org/10.1016/j.compedu.2017.05.007>
43. Da Xu, L., Duan, L.: Big data for cyber physical systems in industry 4.0: a survey. *Enterp. Inf. Syst.* **7575**, 22 (2018). <https://doi.org/10.1080/17517575.2018.1442934>
44. Landset, S., Khoshgoftaar, T.M., Richter, A.N., Hasanin, T.: A survey of open source tools for machine learning with big data in the Hadoop ecosystem. *J. Big Data* **2**(1), 1–36 (2015). <https://doi.org/10.1186/s40537-015-0032-1>
45. Mueen, A., Zafar, B., Manzoor, U.: Modeling and predicting students' academic performance using data mining techniques. *Int. J. Mod. Educ. Comput. Sci.* **11**, 36–42 (2016). <https://doi.org/10.5815/ijmecs.2016.11.05>
46. Hu, J.: A Bayesian statistics course for undergraduates: Bayesian thinking, computing, and research. *J. Stat. Educ.* (2020). <https://doi.org/10.1080/10691898.2020.1817815>
47. Huang, A.Y.Q., Lu, O.H.T., Huang, J.C.H., Yin, C.J., Stephen, J.H.: Predicting students' academic performance by using educational big data and learning analytics: evaluation of classification methods and learning logs. *Interact. Learn. Environ.* (2019). <https://doi.org/10.1080/10494820.2019.1636086>
48. Khalaf, A., Majeed, A., Akeel, W., Salah, A.: Students' success prediction based on bayes algorithms. *Int. J. Comput. Appl.* **178**(7), 6–12 (2017). <https://doi.org/10.5120/ijca2017915506>
49. Pojon, M.: Using machine learning to predict student performance. *Univ. Tampere*, pp. 1–28. <https://tampub.uta.fi/bitstream/handle/10024/101646/GRADU-1498472565.pdf?sequence=1> (2017)
50. Vilorio, A., Lezama, O.B.P., Varela, N.: Bayesian classifier applied to higher education dropout. *Proced. Comput. Sci.* **160**, 573–577 (2019). <https://doi.org/10.1016/j.procs.2019.11.045>
51. Briec, M.S.O., Waters, C.D., Drinan, D.P., Naish, K.A.: A practical introduction to random forest for genetic association studies in ecology and evolution. *Mol. Ecol. Resour.* **18**(4), 755–766 (2018). <https://doi.org/10.1111/1755-0998.12773>
52. Genuer, R., Poggi, J.M., Tuleau-Malot, C., Villa-Vialaneix, N.: Random forests for big data. *Big Data Res.* **9**, 28–46 (2017). <https://doi.org/10.1016/j.bdr.2017.07.003>
53. Zhang, S., Cheng, D., Deng, Z., Zong, M., Deng, X.: A novel kNN algorithm with data-driven k parameter computation. *Pattern Recognit. Lett.* **109**, 44–54 (2018). <https://doi.org/10.1016/j.patrec.2017.09.036>
54. Amra, I.A.A., Maghari, A.Y.A.: Students performance prediction using KNN and Naïve Bayesian. *Int. Conf. Inf. Technol. Proc.* (2017). <https://doi.org/10.1109/ICITECH.2017.8079967>
55. Yu, K., Luo, S., Zhou, X., Wang, R., Sun, L.: A novel method of applying big data for analysis model of library user behavior. *Adv. Econ. Bus. Manag. Res.* **100**, 742–745 (2019)
56. Yang, R., Yu, L., Zhao, Y., Yu, H., Xu, G., Wu, Y.: Big data analytics for financial market volatility forecast based on support vector machine. *Int. J. Inf. Manage.* **50**, 452–462 (2020). <https://doi.org/10.1016/j.ijinfomgt.2006.01.003>
57. Guo, B., Zhang, R., Xu, G., Shi, C., Yang, L.: Predicting students performance in educational data mining. *Int. Symp. Educ. Technol.* (2015). <https://doi.org/10.1109/ISET.2015.33>
58. Vidhya, K., Shanmugalakshmi, R.: Modified adaptive neuro-fuzzy inference system (M-ANFIS) based multi-disease analysis of healthcare Big Data. *J. Supercomput.* **76**(11), 8657–8678 (2020). <https://doi.org/10.1007/s11227-019-03132-w>
59. Caton, S., Venugopal, S., Tn, S.B., Velamuri, V.S., Katrinis, K.: Dynamic model evaluation to accelerate distributed machine learning. *IEEE Int. Congr. Big Data* (2018). <https://doi.org/10.1109/BigDataCongress.2018.00027>
60. Atsalakis, G.S., Atsalaki, I.G., Pasiouras, F., Zopounidis, C.: Bitcoin price forecasting with neuro-fuzzy techniques. *Eur. J. Oper. Res.* **276**(2), 770–780 (2019). <https://doi.org/10.1016/j.ejor.2019.01.040>
61. Mohamed, A., Najafabadi, M.K., Wah, Y.B., Zaman, E.A.K., Maskat, R.: The state of the art and taxonomy of big data analytics: view from new big data framework. *Artif. Intell. Rev.* (2019). <https://doi.org/10.1007/s10462-019-09685-9>
62. Poczeta, K., Kubuś, Ł., Yastrebov, A.: Analysis of an evolutionary algorithm for complex fuzzy cognitive map learning based on graph theory metrics and output concepts. *BioSystems* **186**(2018), 39–47 (2019). <https://doi.org/10.1016/j.biosystems.2019.104068>
63. Gomedede, E., Gaffo, F.H., Briganó, G.U., de Barros, R.M., de Mendes, L.M.: Application of computational intelligence to improve education in smart cities. *Sensors (Switzerland)* **18**(1), 1–26 (2018). <https://doi.org/10.3390/s18010267>
64. Molina-perez, E., Esquivel-flores, O.A., Zamora-maldonado, H.: Computational intelligence for studying sustainability challenges: tools and methods for dealing with deep uncertainty and complexity. *Front. Robot. AI* **7**, 1–18 (2020). <https://doi.org/10.3389/frobt.2020.00111>
65. Rahat, I., Doctor, F., More, B.: Big data analytics: computational intelligence techniques and application areas. *Technol. Forecast. Soc. Change* (2018). <https://doi.org/10.1016/j.techfore.2018.03.024>
66. Zhou, G., Moayedi, H., Bahiraei, M., Lyu, Z.: Employing artificial bee colony and particle swarm techniques for optimizing a neural network in prediction of heating and cooling loads of residential buildings. *J. Clean. Prod.* (2020). <https://doi.org/10.1016/j.jclepro.2020.120082>
67. Mat, U.B., Buniyamin, N.: Using neuro-fuzzy technique to classify and predict electrical engineering students' achievement upon graduation based on mathematics competency. *Indones. J. Electr. Eng. Comput. Sci.* **5**(3), 684–690 (2017). <https://doi.org/10.11591/ijeecs.v5.i3.pp684-690>

68. Khodke, P.A., Tingane, M.G., Bhagat, A.P., Chaudhari, S.P., Ali, M.S.: Neuro fuzzy intelligent e-learning systems. *IEEE*. (2017). <https://doi.org/10.1109/GET.2016.7916766>
69. Patel, P.S., Undavia, J., Bhatti, D.: Master course selection prediction model using modify hybrid neuro-fuzzy inference system. *ICTACT J. Soft Comput.* **11**(01), 2205–2212 (2020). <https://doi.org/10.21917/ijsc.2020.0314>
70. Petković, D., Denić, N.: Neuro-fuzzy assessment of pupil performance based on e-learning platform implementation. *J. Inst. Electron. Comput.* **2**(1), 12–27 (2020). <https://doi.org/10.33969/ijec.2020.21002>
71. Naaj, M.A., Mehdi, R., Mohamed, E.A.M., Nachouki, M.: Analysis of the factors affecting student performance using a neuro-fuzzy approach. *Educ. Sci.* (2023). <https://doi.org/10.3390/educsci13030313>
72. Je, S.M., Huh, J.H.: Estimation of future power consumption level in smart grid: application of fuzzy logic and genetic algorithm on big data platform. *Int. J. Commun. Syst.* (2019). <https://doi.org/10.1002/dac.4056>
73. Singh, S.P., Nayyar, A., Kumar, R., Sharma, A.: Fog computing: from architecture to edge computing and big data processing. *J. Supercomput.* **75**(4), 2070–2105 (2019). <https://doi.org/10.1007/s11227-018-2701-2>
74. Aki, O.: University exam timetabling using genetic algorithms. *Int. Sci. Conf.* **1**, 395 (2020)
75. Dedic, F., Bijedi, N., Gaspar, D.: Genetic algorithms as tool for development of balanced curriculum. *Interdiscip. Descr. Complex Syst.* **18**(2-B), 175–193 (2020). <https://doi.org/10.7906/indecs.18.2.8>
76. Rohani, Y., Torabi, Z., Kianian, S.: A novel hybrid genetic algorithm to predict students' academic performance. *J. Electr. Comput. Eng. Innov.* **8**(2), 219–232 (2020). <https://doi.org/10.22061/JECEI.2020.7230.373>
77. Cuzzocrea, A., Mumolo, E., Grasso, G.M.: An effective and efficient genetic-fuzzy algorithm for supporting advanced human-machine interfaces in big data settings. *Algorithms* (2020). <https://doi.org/10.3390/a13010013>
78. Gupta, S., Sinha, S.: Academic staff planning, allocation and optimization using genetic algorithm under the framework of fuzzy goal programming. *Proced. Comput. Sci.* **172**(2019), 900–905 (2020). <https://doi.org/10.1016/j.procs.2020.05.130>
79. Shokouhifar, M., Pilevari, N.: Combined adaptive neuro-fuzzy inference system and genetic algorithm for e-learning resilience assessment during COVID-19 pandemic. *Concurr. Comput. Pract. Exper.* **34**, e791 (2022). <https://doi.org/10.1002/cpe.6791>
80. Poczeta, K., Papageorgiou, E.I.: Fuzzy cognitive maps optimization for decision making and prediction. *Mathematics* (2020). <https://doi.org/10.3390/math8112059>
81. Nachazel, T.: Fuzzy cognitive maps for decision making in dynamic environments. *Genet. Progr. Evolvable Mach.* (2020). <https://doi.org/10.1007/s10710-020-09393-2>
82. Papageorgiou, E.I.: Review study on fuzzy cognitive maps and their applications during the last decade. *IEEE*. (2011). <https://doi.org/10.1109/FUZZY.2011.6007670>
83. Nitin, K., Sunita, S.: Comparison of ANNs, fuzzy logic and neuro-fuzzy integrated approach for diagnosis of coronary heart disease: a survey. *Int. J. Comput. Sci. Mob. Comput.* **2**(6), 216–224 (2013)
84. Shihabudheen, K., Pillai, G.N.: Recent advances in neuro-fuzzy system: a survey. *Knowl.-Based Syst.* (2018). <https://doi.org/10.1016/j.knosys.2018.04.014>
85. Adams, S., Beling, P.A., Lambert, J.H., Scherer, W.T., Cody, H.F.: *Systems Engineering in Context*. Springer (2019)
86. Mousa, H.M.: Chaotic genetic-fuzzy encryption technique. *Int. J. Comput. Netw. Inf. Secur.* **4**, 10–19 (2018). <https://doi.org/10.5815/ijcnis.2018.04.02>
87. Herrera, F.: Genetic fuzzy systems : taxonomy, current research trends and prospects. *Evol. Intell.* (2008). <https://doi.org/10.1007/s12065-007-0001-5>
88. Majid, M., Saeed, H., Fatemeh, D., Azra, M.: The challenges and advantages of fuzzy systems applications. *A Preprint* (2020). <https://doi.org/10.13140/RG.2.2.22310.96328>
89. Chi, Y., Liu, J.: Learning of fuzzy cognitive maps with varying densities using a multi-objective evolutionary algorithm. *IEEE Trans. Fuzzy Syst.* (2015). <https://doi.org/10.1109/TFUZZ.2015.2426314>
90. Tyagi, A.K., Rekha, G.: Machine learning with big data. *SSRN Electron. J.* (2019). <https://doi.org/10.2139/ssrn.3356269>
91. Mathur, S., Badone, A.: A methodological study and analysis of machine learning algorithms. *Int. J. Adv. Technol. Eng. Explor.* **6**(51), 45–49 (2019). <https://doi.org/10.19101/ijatee.2019.650020>
92. Hodges, J., Mohan, S.: Machine learning in gifted education: a demonstration using neural networks. *Gift. Child Q.* **63**(4), 243–252 (2019). <https://doi.org/10.1177/0016986219867483>
93. Huang, A.Y.Q., Lu, O.H.T., Huang, J.C.H., Yin, C.J., Yang, S.J.H.: Predicting students' academic performance by using educational big data and learning analytics: evaluation of classification methods and learning logs. *Interact. Learn. Environ.* **28**(2), 206–230 (2020). <https://doi.org/10.1080/10494820.2019.1636086>
94. Maestrales, S., Zhai, X., Touitou, I., Baker, Q., Schneider, B., Krajcik, J.: Using machine learning to score multi-dimensional assessments of chemistry and physics. *J. Sci. Educ. Technol.* **30**(2), 239–254 (2021). <https://doi.org/10.1007/s10956-020-09895-9>
95. Monllaó, O.D., Huynh, D.Q., Reynolds, M., Dougiamas, M., Wiese, D.: A supervised learning framework: using assessment to identify students at risk of dropping out of a MOOC. *J. Comput. High. Educ.* (2019). <https://doi.org/10.1007/s12528-019-09230-1>
96. Wang, Z., et al.: Design and implementation of early warning system based on educational big data. In: 2018 5th International Conference on Systems and Informatics (ICSAI), IEEE, pp 549–553 (2018)
97. Fischer, C., et al.: Mining big data in education: affordances and challenges. *Rev. Res. Educ.* **44**(1), 130–160 (2020). <https://doi.org/10.3102/0091732X20903304>
98. Ciolacu, M., Tehrani, A.F., Binder, L., Svasta, P.M.: Education artificial intelligence assisted higher education: early recognition system with machine learning to support students success. *IEEE Int. Symp. Des. Technol. Electron. Packag. SIIEME* (2019). <https://doi.org/10.1109/SIIEME.2018.8599203>
99. Atkinson, K.: Big data real time ingestion and machine learning. In: 2018 IEEE Second International Conference on Data Stream Mining & Processing, pp 25–31 (2018)
100. Villegas-Ch, W., Román-Cañizares, M., Palacios-Pacheco, X.: Improvement of an online education model with the integration of machine learning and data analysis in an LMS. *Appl. Sci.* (2020). <https://doi.org/10.3390/AP10155371>
101. Zeineddine, H., Braendle, U., Farah, A.: Enhancing prediction of student success: automated machine learning approach. *Comput. Electr. Eng.* **89**, 1–22 (2021). <https://doi.org/10.1016/j.compeleceng.2020.106903>
102. Akour, I., Alshurideh, M., Al Kurdi, B., Al-Ali, A., Salloum, S.: Using machine learning algorithms to predict people's intention to use mobile learning platforms during the COVID-19 pandemic: Machine learning approach. *JMIR Med. Educ.* **7**(1), 1–17 (2021). <https://doi.org/10.2196/24032>
103. Sreenivasulu, D., Devi, S., Arulprakash, P., Venkataramana, S., Kazi, K.: Implementation of latest machine learning approaches for students grade prediction department of CSE. *Int. J. Early*

- Child. Spec. Educ. **14**(03), 1308–5581 (2022). <https://doi.org/10.9756/INT-JECSE/V14I3.1141>
104. Kanetaki, Z., Stergiou, C., Bekas, G., Troussas, C., Sgouropoulou, C.: A hybrid machine learning model for grade prediction in online engineering education. *Int. J. Eng. Pedagog* **12**(3), 4–24 (2022). <https://doi.org/10.3991/ijep.v12i3.23873>
 105. Chen, S., Ding, Y.: A machine learning approach to predicting academic performance in Pennsylvania's schools. *Soc. Sci.* (2023). <https://doi.org/10.3390/socsci12030118>
 106. Juarez-orozco, L.E., Martinez-manzanera, O., Nesterov, S.V., Kajander, S., Knuuti, J.: The machine learning horizon in cardiac hybrid imaging. *Eur. J. Hybrid Imaging* (2018). <https://doi.org/10.1186/s41824-018-0033-3>
 107. Nweke, F.H., Wah, Y., Al-garadi, M.A., Alo, R.U.: Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: state of the art and research challenges. *Expert Syst. Appl.* **105**, 233–261 (2018)
 108. Tiong, L.C.O., Lee, H.J.: E-cheating prevention measures: detection of cheating at online examinations using deep learning approach: a case study. *J. Latex Cl. Files* **XX**(1–9), 2021 (2021)
 109. Lorenzo-trueba, J., Henter, G.E., Takaki, S.: Investigating different representations for modeling and controlling multiple emotions in DNN-based speech synthesis. *Speech Commun.* (2018). <https://doi.org/10.1016/j.specom.2018.03.002>
 110. Fadlillah, U., Mahamad, A.K., Handaga, B.: The development of android for indonesian sign language using tensorflow lite and CNN: an initial study. *J. Phys.* (2021). <https://doi.org/10.1088/1742-6596/1858/1/012085>
 111. Wu, Y., Qin, X., Pan, Y., Yuan, C.: Convolution neural network based transfer learning for classification of flowers. In: *IEEE 3rd International Conference on Signal and Image Processing*, pp. 562–566 (2018)
 112. Felipe, C., Hernández, R.: An artificial neural network analysis of academic performance in higher education. Thesis: Faculty Of Psychology And Educational Sciences/ Ku Leuven. Center for professional learning & development, corporate training and lifelong learning. **2020**, 1–264 (2020)
 113. Perrotta, C., Selwyn, N.: Deep learning goes to school: toward a relational understanding of AI in education. *Learn. Media Technol.* **45**(3), 251–269 (2020). <https://doi.org/10.1080/17439884.2020.1686017>
 114. Nweke, F.H., Al-Garadi, M.A., Mujtaba, G., Alo, U.R., Waqas, A.: Deep learning fusion conceptual frameworks for complex human activity recognition using mobile and wearable sensors. *Int. Conf. Comput. Math. Eng. Technol. Innov. Integr. Socioecon. Dev.* (2018). <https://doi.org/10.1109/ICOMET.2018.8346364>
 115. Han, Z., Xu, A.: Ecological evolution path of smart education platform based on deep learning and image detection. *Microprocess. Microsyst.* (2020). <https://doi.org/10.1016/j.micpro.2020.103343>
 116. Waheed, H., et al.: Predicting academic performance of students from VLE big data using deep learning models. *Learn. Model. Comput. Hum. Behav.* (2019). <https://doi.org/10.1016/j.chb.2019.106189>
 117. Araujo, L., Lopez-Ostenero, F., Martinez-Romo, J., Plaza, L.: Deep-learning approach to educational text mining and application to the analysis of topics' difficulty. *IEEE Access* **8**, 218002–218014 (2020). <https://doi.org/10.1109/ACCESS.2020.3042099>
 118. Adejare, S.A.: Can online discussions facilitate deep learning for students in General Education? *J. Heliyon* **7**, e06414 (2021)
 119. Kishore, R., Patra, I., Naved, M., Veera, V., Arcinas, M.M.: Learning analytics using deep learning techniques for efficiently managing educational institutes. *Mater. Today* (2022). <https://doi.org/10.1016/j.matpr.2021.11.416>
 120. Pei, Y., Lu, G.: Design of an intelligent educational evaluation system using deep learning. *IEEE Access* **11**, 29790–29799 (2023). <https://doi.org/10.1109/ACCESS.2023.3260979>
 121. Heriz, H.H., Salah, H.M., Bashir, S., Abdu, A., El Sbihi, M.M.: English alphabet prediction using artificial neural networks. *Int. J. Acad. Pedagog. Res.* **2**(11), 8–14 (2018)
 122. Muhammad, S.H., Lukito, E.N., Paulus, I.S.: Model detecting learning styles with artificial neural network. *J. Technol. Sci. Educ.* **9**(1), 85–95 (2019)
 123. Sun, Y., Haghghat, F., Fung, B.C.M.: Energy and buildings a review of the-state-of-the-art in data-driven approaches for building energy prediction. *Energy Build.* **221**, 110022 (2020). <https://doi.org/10.1016/j.enbuild.2020.110022>
 124. Patrick, K., Fawei, B.: Meta-heuristic solutions to a student grouping optimization problem faced in higher education institutions. *J. Adv. Math. Comput. Sci.* **35**(7), 61–74 (2020). <https://doi.org/10.9734/jamcs/2020/v35i730304>
 125. Lin, C., Liu, Z., Chang, C., Lin, Y.: A genetic algorithm-based personalized remedial learning system for learning object-oriented concepts of java. *IEEE Trans. Educ.* **64**, 237 (2018)
 126. Sendari, S., Bella, A., Utama, P., Susetyo, N., Putri, F., Jumadil, R.: K-means and fuzzy C-means optimization using genetic algorithm for clustering questions. *Int. J. Adv. Sci. Comput. Appl.* **1**(1), 1–10 (2022). <https://doi.org/10.47679/ijasca.v1i1.2>
 127. Rastegarmoghadam, M., Ziarati, K.: Improved modeling of intelligent tutoring systems using ant colony optimization. *Educ. Inf. Technol.* **22**(3), 1067–1087 (2017). <https://doi.org/10.1007/s10639-016-9472-2>
 128. Erivaldo, F.F., Yen, G.G.: Particle swarm optimization of deep neural networks architectures for image classification. *Swarm Evol. Comput.* **49**, 62–74 (2019). <https://doi.org/10.1016/j.swevo.2019.05.010>
 129. Sherar, M., Zulkernine, F.: Particle swarm optimization for large-scale clustering on apache spark. *IEEE* (2017)
 130. Cai, X., Gao, L., Li, F.: Sequential approximation optimization assisted particle swarm optimization for expensive problems. *Appl. Soft Comput. J.* (2019). <https://doi.org/10.1016/j.asoc.2019.105659>
 131. Juan, D., Wei, Y.H.: Particle swarm optimization neural network for research on artificial intelligence college English classroom teaching framework. *J. Intell. Fuzzy Syst.* (2020). <https://doi.org/10.3233/jifs-189400>
 132. Yang, X.: An effective allocation model of computer teaching management resources based on particle swarm optimization. *Int. J. Emerg. Technol. Learn.* **14**(18), 4–15 (2019). <https://doi.org/10.3991/ijet.v14i18.11189>
 133. Sheng, X., Lan, K., Jiang, X., Yang, J.: Adaptive curriculum sequencing and education management. *Systems* (2023). <https://doi.org/10.3390/systems11010034>
 134. Sarkohaki, F., Fotohi, R., Ashrafian, V.: An efficient routing protocol in mobile Ad-hoc networks by using artificial immune system. *ArXiv* (2020). <https://doi.org/10.14569/ijacsa.2017.080473>
 135. Bhadoria, V.S., Pal, N.S., Shrivastava, V.: Artificial immune system based approach for size and location optimization of distributed generation in distribution system. *Int. J. Syst. Assur. Eng. Manag.* **10**(3), 339–349 (2019). <https://doi.org/10.1007/s13198-019-00779-9>
 136. Farzadnia, E., Shirazi, H., Nowroozi, A.: A novel sophisticated hybrid method for intrusion detection using the artificial immune system. *J. Inf. Secur. Appl.* **58**, 102721 (2020)
 137. Beg, A., Zahidul, M.I.: Advantages and limitations of genetic algorithms for clustering records. *IEEE Conf. Ind. Electron. Appl. (ICIEA)* (2016). <https://doi.org/10.1109/ICIEA.2016.7604009>

138. Abdmouleh, Z., Gastli, A., Ben-brahim, L., Haouari, M.: Review of optimization techniques applied for the integration of distributed generation from renewable energy sources. *Renew. Energy* (2017). <https://doi.org/10.1016/j.renene.2017.05.087>
139. Selvi, V., Tamilnadu, S.: Comparative analysis of ant colony and particle swarm optimization techniques. *Int. J. Comput. Appl.* **5**(4), 1–6 (2010)
140. Rasli, R.M., Aziz, N.A.A., Razali, F.M., Basir, N., Norwawi, N.M.: A preliminary survey on artificial immune systems (AIS): a review on their techniques, strengths and drawbacks. *Int. J. Acad. Res. Bus. Soc. Sci.* **9**(14), 121–144 (2019). <https://doi.org/10.6007/IJARBS/v9-i14/6835>
141. Fernandes, D.A.B., Freire, M.M., Fazendeiro, P.A.P., Inácio, P.R.M.: Applications of artificial immune systems to computer security: a survey. *J. Inf. Secur. Appl.* **35**, 138–159 (2017)
142. Kaviyarasi, R., Balasubramanian, T.: Exploring the high potential factors that affects students' academic performance. *Int. J. Educ. Manag. Eng.* **8**(6), 15–23 (2018). <https://doi.org/10.5815/ijeme.2018.06.02>
143. Yu, C.H., Lee, H.S., Lara, E., Gan, S.: The ensemble and model comparison approaches for big data analytics in social sciences. *Pract. Assess. Res. Eval.* **23**, 17 (2018)
144. Adejo, O.W., Connolly, T.: Predicting student academic performance using multi-model heterogeneous ensemble approach. *J. Appl. Res. High. Educ.* **10**(1), 61–75 (2018). <https://doi.org/10.1108/JARHE-09-2017-0113>
145. Gebretekle, T.K.: Bayesian analysis of retention and graduation of female students of higher education institution: the Case of Hawassa University (HU), Ethiopia. *Am. J. Theor. Appl. Stat.* **8**(2), 47 (2019). <https://doi.org/10.11648/j.ajtas.20190802.12>
146. Zhang, B., Sanner, S., Bouadjene, M.R., Gupta, S.: Bayesian networks for data integration in the absence of foreign keys. *IEEE Trans. Knowl. Data Eng.* **32**(4), 803–808 (2020). <https://doi.org/10.1109/TKDE.2019.2940019>
147. Chen, X., Yuan, Y., Orgun, M.A.: Using Bayesian networks with hidden variables for identifying trustworthy users in social networks. *J. Inf. Sci.* **46**(5), 600–615 (2020). <https://doi.org/10.1177/0165551519857590>
148. Rathore, P., Kumar, D., Bezdek, J.C., Fellow, L., Rajasegarar, S., Palaniswami, M.: A rapid hybrid clustering algorithm for large volumes of high dimensional data. *IEEE Trans. Knowl. Data Eng.* **31**(4), 641–654 (2018). <https://doi.org/10.1109/TKDE.2018.2842191>
149. Ajibade, S., Dayupay, J., Oyebo, O.: Utilization of ensemble techniques for prediction of the academic performance of students. *J. Optoelectron. Laser* **41**(6), 48 (2022)
150. Safarov, F., Kutlimuratov, A., Abdusalomov, A.B., Nasimov, R., Cho, Y.I.: Deep learning recommendations of e-education based on clustering and sequence. *Electron* (2023). <https://doi.org/10.3390/electronics12040809>
151. Chen, H., Dai, Y., Gao, H., Han, D., Li, S.: Classification and analysis of MOOCs learner's state: the study of hidden Markov model. *Comput. Sci. Inf. Syst.* **16**(3), 849–865 (2019). <https://doi.org/10.2298/CSIS181002030C>
152. Han, S.Y., Liefbroer, A.C., Elzinga, C.H.: Mechanisms of family formation: an application of Hidden Markov Models to a life course process. *Adv. Life Course Res.* **43**, 100265 (2020). <https://doi.org/10.1016/j.alcr.2019.03.001>
153. Teoh, T.T., Nguwi, Y.Y., Elovici, Y., Cheung, N.M., Ng, W.L.: Analyst intuition based Hidden Markov Model on high speed, temporal cyber security big data. *Int. Conf. Nat. Comput. Fuzzy Syst. Knowl. Discov.* (2018). <https://doi.org/10.1109/FSKD.2017.8393092>
154. Zeng, Y.: Evaluation of physical education teaching quality in colleges based on the hybrid technology of data mining and hidden markov model. *Int. J. Emerg. Technol. Learn.* **15**(1), 4–15 (2020). <https://doi.org/10.3991/ijet.v15i01.12533>
155. Geigle, C., Zhai, C. X.: Modeling MOOC student behavior with two-layer hidden Markov models. In: Proceedings of the 10th International Conference on Educational Data Mining, EDM, p. 6 (2017)
156. Zhang, Y., Yang, X.: Prediction of the development scale of vocational education using markov algorithm and countermeasures. *Mob. Inf. Syst.* **2022**, 1–10 (2022)
157. Xu, F., Xia, Y.: Development of speech recognition system for remote vocal music teaching based on Markov model. *Soft Comput.* **27**(14), 10237–10248 (2023). <https://doi.org/10.1007/s00500-023-08277-8>
158. Zhang, J., Wang, Y., Sun, Y., Li, G.: Strength of ensemble learning in multiclass classification of rockburst intensity. *Int. J. Numer. Anal. Methods Geomech.* **44**(13), 1833–1853 (2020). <https://doi.org/10.1002/nag.3111>
159. Tadayon, M., Pottie, G.: Predicting student performance in an educational game using a hidden markov model. *IEEE Trans. Educ.* (2020). <https://doi.org/10.1109/TE.2020.2984900>
160. Kayte, S., Mundada, M., Gujrathi, J.: Hidden Markov model based speech synthesis: a review. *Int. J. Comput. Appl.* **130**(3), 35–39 (2015). <https://doi.org/10.5120/ijca2015906965>
161. Márquez-Vera, C., Cano, A., Romero, C., Noaman, A.Y.M., Fardoun, H.M., Ventura, S.: Early dropout prediction using data mining: a case study with high school students. *Expert Syst.* **33**(1), 107–124 (2016). <https://doi.org/10.1111/exsy.12135>
162. Yaqoob, I., et al.: Big data: From beginning to future. *Int. J. Inf. Manage* **36**(6), 1231–1247 (2016). <https://doi.org/10.1016/j.ijin.fomgt.2016.07.009>
163. Chang, V.: A proposed social network analysis platform for big data analytics. *Technol. Forecast. Soc. Change* **130**, 57–68 (2018). <https://doi.org/10.1016/j.techfore.2017.11.002>
164. He, K., Li, Y., Soundarajan, S., Hopcroft, J.E.: Hidden community detection in social networks. *Inf. Sci. J.* **425**, 92–106 (2017). <https://doi.org/10.1016/j.ins.2017.10.019>
165. Bayer, J., Ellison, N., Schoenebeck, S., Falk, E.B.: Facebook in context (s): Measuring emotional responses across time and space. *New Media Soc.* (2018). <https://doi.org/10.1177/14614448166681522>
166. Putnik, G., Costa, E., Alves, C., Castro, H., Varela, L., Shah, V.: Analysing the correlation between social network analysis measures and performance of students in social network-based engineering education. *Int. J. Technol. Des. Educ.* **26**(3), 413–437 (2016). <https://doi.org/10.1007/s10798-015-9318-z>
167. Wang, W., Yu, H., Miao, C.: Deep model for dropout prediction in MOOCs. *ACM Int. Conf. Proc. Ser.* **1306**, 26–32 (2017). <https://doi.org/10.1145/3126973.3126990>
168. Teruel, M., Alemany, L.A.: Co-embeddings for student modeling in virtual learning environments. *Proc. Conf. User Model. Adapt. Pers.* (2018). <https://doi.org/10.1145/3209219.3209227>
169. Xing, W., Du, D.: Dropout prediction in MOOCs: using deep learning for personalized intervention. *J. Educ. Comput. Res.* **57**(3), 547–570 (2019). <https://doi.org/10.1177/0735633118757015>
170. Jin, C.: MOOC student dropout prediction model based on learning behavior features and parameter optimization. *Interact. Learn. Environ.* (2020). <https://doi.org/10.1080/10494820.2020.1802300>
171. Bernard, J., Chang, T.W., Popescu, E., Graf, S.: Learning style Identifier: improving the precision of learning style identification through computational intelligence algorithms. *Expert Syst. Appl.* **75**, 94–108 (2017). <https://doi.org/10.1016/j.eswa.2017.01.021>
172. Lang, C.: Handbook of learning analytics. *Handb. Learn. Anal.* (2017). <https://doi.org/10.18608/hla17>

173. Sentance, S., Csizmadia, A.: Computing in the curriculum: challenges and strategies from a teacher's perspective. *Educ. Inf. Technol.* **22**(2), 469–495 (2017). <https://doi.org/10.1007/s10639-016-9482-0>
174. Meder, M., Till, P., Sahin, A.: A primer on data-driven gamification design. <http://ceur-ws.org> (2017)
175. Jing, S., Tang, Y., Liu, X., Gong, X., Cui, W., Liang, J.: A parallel education based intelligent tutoring systems framework. *IEEE Int. Conf. Networking, Sens. Control. ICNSC* (2020). <https://doi.org/10.1109/ICNSC48988.2020.9238052>
176. Tang, Y., Liang, J., Hare, R., Wang, F.Y.: A personalized learning system for parallel intelligent education. *IEEE Trans. Comput. Soc. Syst.* **7**(2), 352–361 (2020). <https://doi.org/10.1109/TCSS.2020.2965198>
177. Moon, J., Do, J., Lee, D., Choi, G.W.: A conceptual framework for teaching computational thinking in personalized OERs. *Smart Learn. Environ* (2020). <https://doi.org/10.1186/s40561-019-0108-z>
178. Lee, M., Ferwerda, B.: Personalizing online educational tools. *Proc. ACM Work. Theory-Informed User Model. Tailoring Pers. Interfaces* **5**, 5 (2017). <https://doi.org/10.1145/3039677.3039680>
179. Singh, S., Sunil, P.L.: Educational courseware evaluation using machine learning techniques. *IEEE Conf. e-Learn. e-Manag. e-Serv.* (2013). <https://doi.org/10.1109/IC3e.2013.6735969>
180. Jing, L.: Construction of modern educational technology MOOC platform based on courseware resource storage system. *Int. J. Emerg. Technol. Learn.* **12**(9), 105–116 (2017). <https://doi.org/10.3991/ijet.v12.i09.7491>
181. Sugiyarti, E., Jasmi, K.A., Basiron, B., Huda, M., Shankar, K., Maseleno, A.: Decision support system of scholarship grantee selection using data mining. *Int. J. Pure Appl. Math.* **119**(15), 2239–2249 (2018)

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.