JAIEP

AI in Higher Education: Insights from Student Surveys and Predictive Analytics using PSO-Guided WOA and Linear Regression

S.K.Towfek[1,*], Nima Khodadadi[2], Laith Abualigah[3,4,5,6,7,8], Faris H. Rizk[1]

[1]Computer Science and Intelligent Systems Research Center, Blacksburg 24060, Virginia, USA
[2]Department of Civil and Architectural Engineering, University of Miami, Coral Gables, FL, USA
[3]Hourani Center for Applied Scientific Research, Al-Ahliyya Amman University, Amman 19328, Jordan.
[4]Computer Science Department, Al al-Bayt University, Mafraq 25113, Jordan.
[5]Artificial Intelligence and Sensing Technologies (AIST) Research Center, University of Tabuk, Tabuk 71491, Saudi Arabia.
[6]MEU Research Unit, Middle East University, Amman 11831, Jordan.
[7]School of Engineering and Technology, Sunway University Malaysia, Petaling Jaya 27500, Malaysia.
[8]Applied science research center, Applied science private university, Amman 11931, Jordan.

* Corresponding author: sktowfek@jcsis.org

Emails: sktowfek@jcsis.org; nima.khodadadi@miami.edu; aligah.2020@gmail.com; faris.rizk@jcsis.org

Abstract

Artificial intelligence (AI) and machine learning (ML) prediction can change education in a drastic way, where there can be both improvements and regressions concerning the way learning is approached. With individualized learning experiences, being able to spot the students who are falling behind, and customizing the course materials and tests that are fully customized, educators will help students achieve their individualized needs. We at Grand Canyon University conducted our study among 250 students in order to find out how they are interacting with AI in their academic journey. Using the binary Particle Swarm Optimization - Whale Optimization Algorithm for feature selection and the predictive modeling Linear Regression, we came up with vivid findings. For example, the bPSO-Guided WOA algorithm was characterized by a typical Average error of 0.25934, signifying the feat it was in feature selection, and the Linear Regression model particularly stood out in its sustainably low mean squared error (MSE), with a really admirable result of 1.39069E-31. Such evidence indicates the remarkable ability of AI and ML methods to develop true and relevant forecasts by providing teachers with possible and efficient decisions, thus improving the standards and effectiveness of education.

Keywords: Artificial Intelligence; Higher Education; Student Surveys; Predictive Analytics; PSO-Guided WOA; Linear Regression.
MSC: 68T20; 62J05; 93B45

Doi : https://doi.org/10.21608/jaiep.2024.354003
Received: March 19, 2024 Accepted: April 20, 2024

1. Introduction

In the contemporary era, Artificial Intelligence (AI) and machine learning (ML) have come up as transformative forces that change the very fabric of society across diverse domains. Artificial intelligence technology could optimize business processes and customer experiences by transforming healthcare delivery. The potential of these innovations to catalyze a paradigm shift in the field of education on how knowledge is imparted, skills are nurtured, and learning experiences are personalized is poignant [1-3]. With the rise of AI educational technologies, a different reality is now

being opened up in digital learning: adaptive learning systems, intelligent tutoring platforms, and virtual classroom environments. Furthermore, the system is open to integration with human intelligence in educational pedagogy that could address certain persistent issues like student engagement and individualized instruction, as well as many issues that are related to equity of quality education [4-6].
This sets the major context for this study in the background of large-scale technological innovations and educational transformations. Therefore, the main focus of this paper is to discuss the intricate dynamics of the use of AI in a college student population within higher education. In this line, we seek to decipher the underlying factors that determine such multifaceted dimensions of student engagement with AI technologies [7-9]. This work explores students' knowledge, usage, perceptions, and aspirations about AI holistically, which may interest and benefit educational practitioners, policymakers, and industry stakeholders. Finally, we provide empirical evidence and enlightened perspectives in the continuous discussion of whether or how AI will enter education and with what kind of impact, challenges, and opportunities [10-12].

The following questions will guide our inquiry and facilitate a structured investigation:

1. How much awareness have college students gained about the principles and applications of AI, varying in terms of their demographic, disciplinary, and socio-economic dimensions, that might affect their levels of AI literacy?

2. When it comes to college students' efforts, how do AI technologies tend to come into play more often than the college's tasks? What are the major motivators or challenges that encourage or become obstacles, respectively, in the adoption and use of AI among college students in various social and academic activities?

3. What predicts college students' aspirations to work in AI-related fields, and how do these aspirations change or stay stable across demographic and disciplinary borders? How do the educational experiences and extracurricular activities offered to students help shape their perceptions of careers in AI and AI as a potential career path?

4. To what extent do college students perceive certain AI applications and the implications they might have on learning, productivity, and societal advancement? How do students' attitudes and behaviors toward AI adoption, considering concerns about privacy, ethics, and bias?

5. How do personal differences, such as major, prior exposure to AI, and socio-economic background, moderate perceptions, attitudes, and behaviors regarding the adoption and integration of AI technologies by learners at the higher education level? What implications arise for educational equity, access, and inclusion from these different patterns?

This study is important and extends beyond disciplinary limits. It resonates across stakeholders' relevance to the future of education and workforce readiness. Our findings contribute to evidence-based policy development and decision-making aimed at fostering AI literacy, fluency, and competency among current and future generations of learners through a better understanding of the complexities of student engagement with AI technologies.

Furthermore, findings from this study could catalyze renewed innovation in instructional design, curriculum development, and pedagogical practices that reflect a more inclusive, adaptive, and effective educational experience. Moreover, by clarifying socio-cultural, economic, and institutional factors conditioning students' mediated engagement with AI, our work contributes to a more nuanced view of possible societal effects of AI integration into education and other social spheres. Eventually, our study will be useful as a mechanism for dialogue, collaboration, and collective action toward harnessing AI's potential in empowering learners, educators, and communities in the new digital and hyper-connected world.

2. Literature Review

Educational research, having based its work upon a number of factors, has been looking for a long period to determine what influences the academic performance of students. Having examinations has much meaning in education. It teaches you that books present more than the story; they also demonstrate them to you. It is all about how well kids do on tests, whether it is books, videos, or social media. It operates on the principle of putting pieces together into a single comprehensive framework on the scientific basis of such complex factors as difficulty or ease of learning. The key objective here is to examine how and to what extent cognitive and psychological factors, socioeconomic factors, and educational methods affect the results of the test. They also interact with each other.

What is being driven here is that the heavy use of educational data and technology-enhanced learning tools allows us to enhance the teaching process by taking into account how students learn in new ways, improve the classes, and make informed managerial decisions based on this information. It has been a major component of e-learning analysis, which gathers information from online classes [13]. A deep, deep neural network trained on the engineered stream data attributes is used to analyze the risky children and learn how to help them right away. Logistic regression and KNN are better classifiers because they are able to put things into groups more accurately than the decision tree model. Agreeing with AI can be as high as 84-93% of the time. This paper helps to illustrate, as against other research, how old and test data may affect how accurate a model is. With data mining becoming accessible, we have more applications for it, and the desire to study how to perform data mining has increased [14]. These times call upon the most modern directed machine learning techniques to study how much students' scores on tests can be predicted. Gradually, we can study them, and thus, we can make a comparison. To facilitate artificial neural networks to provide successful solutions for classification and regression problems, they must know the level of their interest in the students and the level of their success in past results. The demographic data a person listens to only complements the psychologists' guesses depending on every unique person. The ones who want more from the students' detailed test results are advised about some data collection tools to be used and to involve the students in an interesting learning environment.

The data of students is so huge that it has jumpstarted a new field of education study [15]. Education data mining is a way of seeing students' educational data teachers and teases out whether they have excelled or not. While there are some presently applied techniques involving sorting or classification directed toward predictive modeling, a lot of other new ideas continue trickling in. Students from Indian schools and universities are the targeted group who have taken the test with actual data from their diverse classes learning numerous subjects. This method, where grouping and classifications are employed, is, in the majority of cases, more effective than the bulk of the methods of classification. Online connectivity for college students is real-time, which results in not only enhanced learning patterns but also a transformation of their daily lifestyle. As a result, [16] helps to discover the strong bonds between internet use and success so that machine learning could begin to foresee which first-year college students will fall into academic problems. This is enhanced by a focus on the portrayal of some traits watched by real students through observation of how 4000 students behave on the Internet. One can use decision trees, neural networks, or support vector machines as the classification methods. However, in any of the deviations, a certain trend between the rate of online time uses and the rate of performance is applied to some of the data sets. Being a modern dad, one can segregate their kids and forecast how much their performance is by picking on the way that they are using the Internet.

To get the most from freshman college members, the grouping should be based on students' abilities [17]. You could actually clairvoyantly predict how great a particular student would do in school by viewing the data on the final grade at the end of their first school year. The plan is made of two steps, and this is it: The man seeks the same reality, and the world keeps usurping him into the obstacles on his path. The mean grades and the time that it requires to complete the degree are two important tools that we use to monitor students' progress in school. In this study, the segmentation structure designed for an Institute of European Engineers at a public research university based on the data gathered from sample students consisting of 2459 students will be tested. We can likely predict how the child will have academic performance during their whole life in school before they even start kindergarten. It can be observed in the document of [18] that the universities or the colleges need to be accountable for the grades of the students since this shows how well they are running. Machine learning is receiving increasing attention for its ability to sift through big chunks of data very fast and to give good estimates of children under threat. They can find out a lot about how their kids will. Perform in class by collecting information instead of just relying on the evaluations of the teachers. The main problem of the passage lies in the fact that people often do not read books hard enough to understand

Doi : https://doi.org/10.21608/jaiep.2024.354003
Received: March 19, 2024 Accepted: April 20, 2024

what to do and why. It addresses them by the detachable list that is provided, and it can be made with rules of law and grounds for decisions.

When we want to interact with children, we not only discover new things but educate them as well. While the school is responsible for the pupils' overall development, it should also understand the extent of the success of learners in school [19]. The researchers from an Australian university explore different strategies for the utilization of university data in a modeling technique developed to predict how well kids will perform in school. The process of designing student types is quite thorough because the way they are prepared prior to buying can affect how hard they try to learn. Models, wherein we apply tools used by systems that follow rules and tree surfaces for users' ability to comprehend, are very simple. As mentioned in [20], researchers have developed a new teaching strategy that enables both teachers and students to search for information efficiently. Many of the issues that will help people to learn them can be found. The algorithm visualization is a kind of presentation of all the efforts that have been made with the use of active learning as a source. Instructors will have immediate access to scores from built-in testing tools; this will help them see which students are doing well and which are struggling. Moreover, these tools finally allow teachers to save more time during the process of grading and give quick and relevant comments. Instruction, one of which was research that was done in a real class setting and specific in the method, found that it benefited the students. They borrow teachers who are vets in their profession to make comparisons. The data also shows that, because the judges' behavior results from the choices that they make consciously and so do the people's - there are some similarities in the modes these people choose to work with.

According to [21], academic success is, by and large, used as a means to record learning progress in many types of schools in different parts of the world. Therefore, it helps a lot. It points out the fact that AI tools may be employed in order to create quality models that could be helpful in determining how students will examine in the future. By making these processes easy, it seems that these two steps can predict the future performance of students in school. Two models are presented in a mixed approach, including a feature-weighted support vector machine (SVM) as well as an artificial neural network (ANN). Both models can validate the effectiveness of the applied approach. Tests, an ablation study, as well as a data set from two Portuguese secondary schools provide extra evidence for this.

Besides the other tests, this piece is also known for the tests research that assesses how well students do on tests. Following along with these, it is possible to realize that school is hard and needs much to be put into consideration. The study examined a variety of areas, from early education and secondary training to upper education. The type of person, social setting, educational system, and desire to learn are all critical components in the formation of a person. These researches are very revolutionary; they introduce many new academic outlooks, hence changing the policies and programs run by the institution. Experts, teachers, and politicians are the people who are mostly used to searching for the best solutions for kids to be successful at school and to learn well.

3. Proposed Methodology

3.1 Dataset

1. **Survey Design:** The survey used in this study was meticulously crafted to provide a comprehensive understanding of college students' engagement with AI technologies in academic settings. It was designed based on existing literature and program inputs, incorporating purposeful investigative questions to gauge respondents' understanding, knowledge, attention, and expectations of AI. The questionnaire was developed using a controlled approach to measure students' understanding of AI basics, frequency of AI use for personal and academic tasks, interest in pursuing AI-related careers, specific AI applications, and their demographic and academic backgrounds.

2. **Participant Recruitment:** The ethical principles concerning participant recruitment, along with IRB approval, were ensured throughout the whole research process. The participants for the study were chosen from Arizona-based Grand Canyon University, which utilizes both the convenience sampling method and the targeted outreach. It strived to embrace students coming from different disciplines of academia, demographic situations, and the whole variability of knowledge of AI. Potential participants were comprehensively informed about the study purpose, engagement on a voluntary basis, confidentiality and methods of data use. Informing the subject(s), all participants had signed the consent form with the privacy and confidentiality materials in place.

3. **Data Collection Period:** We have conducted data collection starting from March 31, 2023, to January 4, 2024, which included a multi-wave survey pattern for collecting a large sample size and assuring the proportionate representation of the study population. An extended period was another opportunity for participants from different semesters, cohorts, and course tracks to be present and included in the dataset. It will again make the diversity of it richer. Communication techniques such as follow-ups, prompting tools, and incentives were used to ensure that participants were fully engaged and responded to the study. Constant quality control and data verification exercises with this in mind remained central for the accuracy, completeness and integrity of the collected records.

4. **Sample Description:** The pool for this analysis was made up of students who were enrolled at the Grand Canyon University in Phoenix, Arizona. Students from the engineering and arts departments were present, which is a testimony to the presence of students from different prominent departments of the student body. Attention was also paid to making sure we were balanced and included more than just one or two specific groups, such as age, gender, ethnicity, and academic year students. The final sample size of over 250 was, however, faithfully diverse and more than enough for thorough analysis [22]. The reporting phase of this data collection approach was extremely strict because the goal was to provide as much reliable data as possible for future analysis, where this data was all used at a conceptual level.

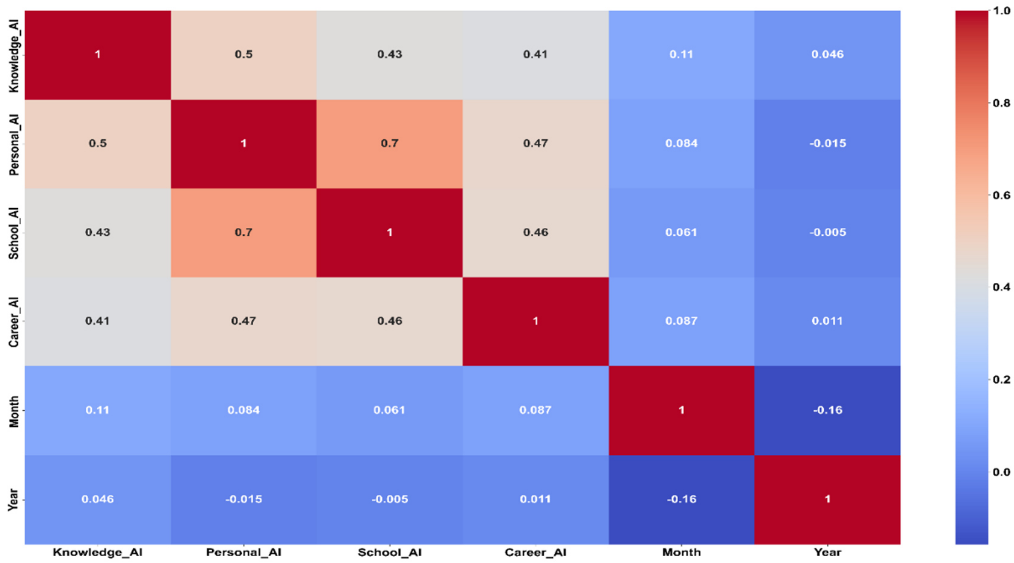

Figure 1: Heatmap of the Dataset

Figure 1 is a graphic representation of a heat map that gives a quick pictorial view of interconnections between distinct components. Every cell represented by a heat map shows a correlation coefficient between pairs of variables, with the warmest colors indicating a strong correlation. This heatmap, therefore, correlates the underlying relationship and interdependency among the stated variables and can guide the exploration and analysis for meaningful interpretation of the data.

Doi : https://doi.org/10.21608/jaiep.2024.354003
Received: March 19, 2024 Accepted: April 20, 2024

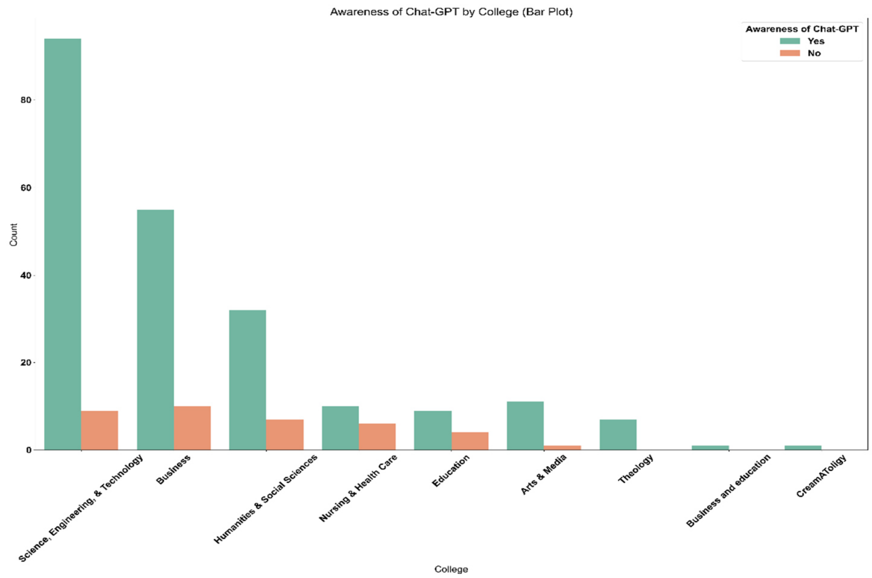

Figure 2: Histogram of Awareness of Chat-GPT by College

The popularity of the Chat-GPT software among different colleges is captured in Figure 2. The histogram above reveals the frequency distribution of cognizance of Chat-GPT among different academic levels of college students, thus disclosing disparities of being aware among different disciplines of studies. The section on the on-campus experience introduces Chat-GPT beginners to the diversity among the college student demographics and familiarizes them.

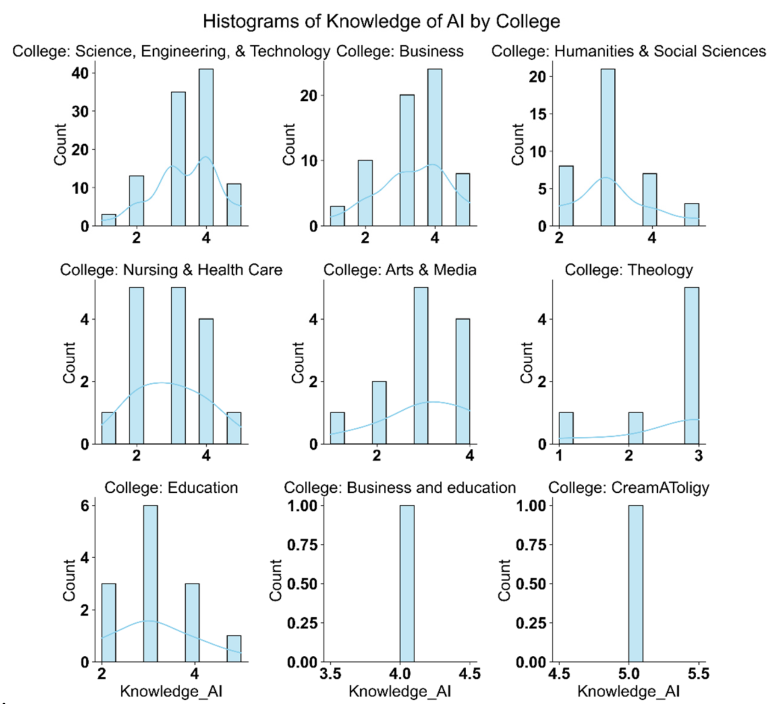

Figure 3: Histograms of Knowledge of AI by College

Figure 3 is a graphic representation of the level of AI knowledge students from different colleges have. This figure shows the distribution of AI knowledge among the academic departments; thus, such a diagram can be used to understand the diversity of AI literacy across various subjects of study. From these histograms, it is clear that patterns and disparities in AI knowledge among college students can be largely distinguished.

Figure 4 pairs up relationships between college students' intentions to pursue AI careers, their AI knowledge, and AI uses on a personal level. On the scatter plots diagonal and on the correlation off-diagonal, the system will highlight possible connections among these variables. This is because they are mutual influencers that reveal new student behaviors, attitudes, and interactions with AI.

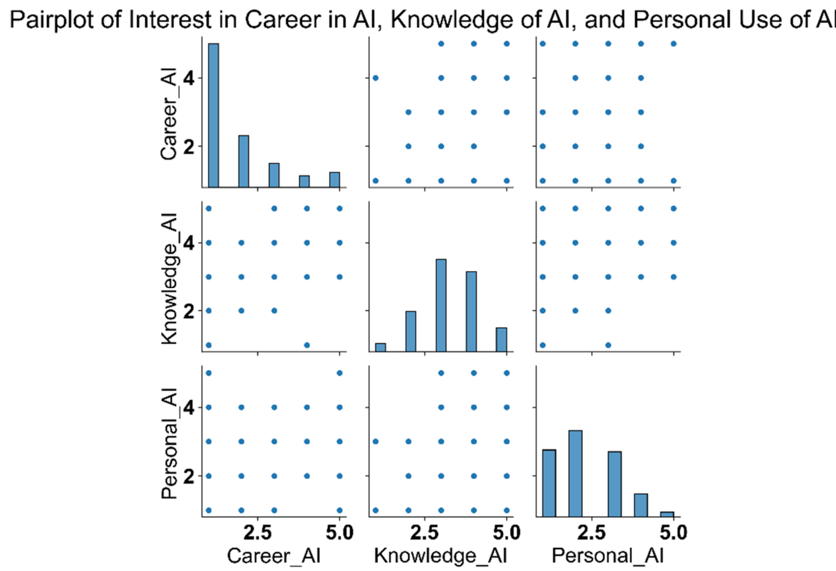

Figure 4: Pairplot of Interest in Career in AI, Knowledge of AI, and Personal Use of AI

3.2 Data Preprocessing

1. **Cleaning and Formatting:** Next, we went through the process of performing a substantial cleaning and rearrangement of the initial raw data to make it more consistent, accurate and reliable. The initial step was a process of identifying and correcting any major differences and inconsistencies in the retrieved information. Outliers and untrusted entries were adjusted or excluded, and data integrity tests were done to verify and ensure data quality. The processed data are homogenous, and the standardized structure and format enable the data to seamlessly fit into the overall need, whether it is the analysis or the interpretation.

2. **Handling Missing Values:** Missing data in the dataset was treated systematically by substitutions if necessary and by excluding the incomplete data. Patterns and the reasons for misread observations were revealed, and the points were matched into clusters with a particular pattern. A case in point is to address the type and amount of missing data, and most of the time, the mean, median, and regression imputation methods were used to estimate and fill in the missing values. When it is not possible, or it is not an "issue" to impute, records with missing values are not utilized to avoid causing bias and maintaining result integrity.

3. **Encoding Categorical Variables:** Different categorical variables in the dataset were encoded through the techniques of their transformation into the form of numbers or binaries that were suitable for study rather than branches of statistics or models. The binary assignment setup entailed devising a unified numerical code or introducing dummy variables to stand for different categories in each given categorical variable. The design of the data encoding scheme aimed to provide a framework for preserving the integrity of data categories while at the same time facilitating the smooth conduct of analytical tasks.

We properly arranged the required tasks of the data preprocessing, such as managing cleaning, handling of missing values, and conversion of categorical variables, ensuring that the dataset is complete and ready for analysis and production of conclusions that can be related to. This step is crucial to the creation of a good platform for students' active involvement in technology use in classes to study.

3.3 Feature Selection Techniques

Doi : https://doi.org/10.21608/jaiep.2024.354003
Received: March 19, 2024 Accepted: April 20, 2024

Feature selection is an inalienable step of machine learning that is one of the most important steps at the same time. The goal of the latter is to define features that are informative enough to resolve the problem in the best way possible. Here is a breakdown of the techniques employed:

- bPSO-Guided WOA (Binary Particle Swarm Optimization - Guided Whale Optimization Algorithm): This hybrid algorithm is a mix of bPSO principles and the WOA. The former is a binary Particle Swarm Optimization algorithm, and the latter is a Guided Whale Optimization algorithm. It greatly relies on particles' intelligence and whales' behavior to facilitate an effective feature space exploration and ultimately contribute to a successful subset of optimal predictive features.

- bGWO (Binary Grey Wolf Optimizer): Concentrating on grey wolves' leadership and hunting conduct, the bGWO leads to the update of the binary feature vectors, which are encouraged to enhance the accuracy of classification. Through a Carol Hunting simulation, it strikes the balance of exploration and exploitation so as to catapult to the discovery of applicable feature subsets.

- bGuided-WOA (Binary Guided Whale Optimization Algorithm): The signal-based Guided Whale Optimization Algorithm integrates the binary coding and guiding processes into the WOA. This technique easily deals with the exploitation-exploration paradox, arriving at the optimal subsets through the repeated application of the binary feature vector adjustments.

- bPSO (Binary Particle Swarm Optimization): The binary variant of Particle Swarm Optimization (PSO) is a population-based technique in which the functions are optimized by generating new findings once the positions are updated according to the local and global best solutions of a search area. It investigates the uniform feature space with the goal of discovering important facts.

- bBA (Binary Bat Algorithm): performance by replicating their cation procedure used to deflect/filter various selected feature vectors through frequency and volume param. parameters. Byterhis, by the way, covers the search field and points out the best features that contribute to each possible classification.

- bWAO (Binary Whale Optimization Algorithm): The binary Whale Optimization Algorithm gets ideas from humpback whales based on their social behaviors and hunting tactics. By going into an iterative process under this algorithm, feature vectors are being actively modified to improve their classification accuracy. It uses migration and contouring to find out its shape-based key features.

- bBBO (Binary Biogeography-based Optimizer): BBO, a Meta-heuristic based on natural habitat distribution, inspired it. A binary form of this algorithm where migration and mutation occur has a notable impact on picking an appropriate vector in a feature space, thus providing the desired subset of features.

- bPSO_GA (Binary Particle Swarm Optimization—Genetic Algorithms): This suggests using both PSO and GA techniques in a hybrid manner but with feature selection for the best results. It attains both the exploration and exploitation of the bPSO framework with the evolutionary advantage of GA to contribute to the effective search of informative feature subsets.
- bFA (Binary Firefly Algorithm): Firefly is a Firefly Algorithm (FA) designed by imitating the flashing motion of fireflies. In its binary form, it mimics attractiveness and motion as a modifying feature set and plots the fitness landscape of our model for further classification by iteration through subsets.

- bGA (Binary Genetic Algorithm): The Binary Genetic Algorithm is a metaheuristic inspired by natural selection and evolutionary principles. It uses binary feature vectors as a framework for selection, crossover, and mutations, which increase the opportunity to identify the best features and improve prediction.

3.4 Machine Learning Models

1. Description of Models Used:

In order to analyze college students' engagement with AI in the educational context, we employed a whole array of machine learning models that are among the most powerful in terms of the identification of enigmatic patterns as well as relationships in the dataset [27-30]. These models employ computational techniques that analyze input data in a defined manner so that they may be trained to make predictions or classifications. Here is an overview:

1. **Linear Regression:** This model is based on this simple fact: It is linear, interpretable, and does well with the use of the target variable and input features. It is better when the linearity of the relationship is reasonable and has simple conclusions about the influences of variables on variables.
2. **Support Vector Regression (SVR):** SVR, otherwise known as Support Vector Regression, can be utilized to sort data by identifying an optimal hyperplane that distinguishes between different data points. It is efficient in tackling non-linear relationships and handling immense-size data sets.
3. **Pipeline:** The model is a combination of both preprocessing and machine learning algorithms; this system is a basis for quick algorithm development. This will make workflows more consistent and scalable by singing they combine multiple phases into one process.
4. **MLPRegressor (Multi-layer Perceptron Regressor):** This is an artificial neural network that mimics biological neurons. This model not only can catch complex non-linear relationships but also uses interconnected layers of neurons to learn more quickly and achieve better performance.
5. **XGBoost (Extreme Gradient Boosting):** Efficient and accurate sequence is one of the main features of this method that uses the decision trees as the weak learners. In the following process, these weak learners are refined to minimize the error and heighten the predictive precision.
6. **ExtraTreesRegressor:** This approach generates multiple decision trees by using random subsets of features and data, leading to less overfitting and increasing robustness. This feature makes the model efficient in handling noisy or high-dimensional data.
7. **GradientBoostingRegressor:** Similar to XGBoost, it builds accuracy by reinforcing the regression ensemble of decision trees using gradient descent optimization.
8. **RandomForestRegressor:** Through this process, the learning components create multiple decision trees, and voting or averaging aggregates their predictions. It provides crowd intelligence so that the search does not overfit and improves performance on out-of-sample issues.
9. **CatBoost:** CatBoost is geared up to effectively transform categorical features with novel optimization techniques, which in turn leads to performance and minimal preprocessing.
10. **KNeighborsRegressor:** This algorithm is not parametric. It calculates predictions considering the k-nearest neighbors in the feature space. This sort of non-parametric method is suitable for capturing complex data patterns.
11. **DecisionTreeRegressor:** This tree-based model recursively sifts features into limited components, which makes the procedure applicable and understandable while producing outputs from the average of the leaf nodes.

2. Evaluation Metrics

The models that we are evaluating need to have a set of metrics that are appropriate for them to pull out the performance of these models thoroughly. As stated below, Table 1 presents the metrics that cover the model's accuracy, precision, and generalization aspects.

Table 1: Criteria for Evaluating Regression Results

Doi : https://doi.org/10.21608/jaiep.2024.354003
Received: March 19, 2024 Accepted: April 20, 2024

Metric	Formula
RMSE	$\sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\hat{V}_n - V_n\right)^2}$
RRMSE	$\frac{RMSE}{\sum_{n=1}^{N}\hat{V}_n} \times 100$
MAE	$\frac{1}{N}\sum_{n=1}^{N}\left\lvert\hat{V}_n - V_n\right\rvert$
MBE	$\frac{1}{N}\sum_{n=1}^{N}\left(\hat{V}_n - V_n\right)$
NSE	$1 - \frac{\sum_{n=1}^{N}\left(V_n - \hat{V}_n\right)^2}{\sum_{n=1}^{N}\left(V_n - \hat{V}_n\right)^2}$
WI	$1 - \frac{\sum_{n=1}^{N}\left\lvert\hat{V}_n - V_n\right\rvert}{\sum_{n=1}^{N}\left\lvert V_n - \bar{V}_n\right\rvert + \left\lvert\hat{V}_n - \bar{V}_n\right\rvert}$
R^2	$1 - \frac{\sum_{n=1}^{N}\left(V_n - \hat{V}_n\right)^2}{\sum_{n=1}^{N}(\sum_{n=1}^{N}V_n) - V_n)^2}$
r	$\frac{\sum_{n=1}^{N}\left(\hat{V}_n - \hat{V}_n\right)\left(V_n - \bar{V}_n\right)}{\sqrt{\left(\sum_{n=1}^{N}\left(\hat{V}_n - \bar{\hat{V}}_n\right)^2\right)\left(\sum_{n=1}^{N}(V_n - \bar{V}_n)^2\right)}}$

1. Mean Squared Error (MSE): It determines the arithmetic mean of the square of the difference between the outputs and predictions, which informs about the model's accuracy. The higher the accuracy, the lower the MSE.
2. Root Mean Squared Error (RMSE): The square root of RSME is the RMSE, which is used to show the distance between the average prediction and the real value. The smaller the range, the broader the range of solutions that are being presented.
3. Mean Absolute Error (MAE): Computes the mean absolute deviation from the observed and fitted values, thus showing the extent to which the prediction is off the mark.
4. Mean Bias Error (MBE): This measure calculates the average difference between the best prediction's forecast and the actual outcome. Positive values are generally considered overestimated, whereas negative values mean underestimations.
5. R (Correlation Coefficient): Shows the degree and specific direction of linear correlations. A value close to 1 May indicates a high positive correlation.
6. R^2 (Coefficient of Determination): This is the percentage of the explained variation in a dependent variable through the set of independent variables. Bigger values measure better model fit.
7. Relative Root Mean Squared Error (RRMSE): This measures RMSE as a fraction of the data range, thereby providing a readability of the magnitude of observations.
8. Nash-Sutcliffe Efficiency (NSE): The results of the metrics used to observe model accuracy are examined. Numbers approaching 1 mean higher good performance.
9. Willmott Index (WI): This index requires evidence to substantiate whether the given values and predicted values match. An indicator value of 1 implies perfect agreement, while 0 indicates completed work or improvement is needed.
10. Fitted Time: The time taken to train the model using the data in a shorter period is presented with fewer numbers, which is equivalent to faster training.

Such measures are basic factors that are formal components of our evaluation standards. Recognizing their importance is monumental, and it helps us determine model performance and the best possible models for predictive tasks.

4. Results

4.1 Feature Selection Results:

Table 2 gives the results of the feature selection process employing the different binary optimization algorithms. Metrics differ across algorithms, but the most common representation is an algorithm's ability to identify relevant features for predictive modeling of these metrics:

Table 2: Feature Selection Results

	bPSO-Guided WOA	bGWO	bGuided WOA	bPSO	bBA	bWAO	bBBO	bPSO_GA	bFA	bGA
Average error	0.25934	0.27654	0.31584	0.31034	0.31994	0.31014	0.27854	0.29664	0.30874	0.29014
Average Select size	0.21214	0.41214	0.54544	0.41214	0.55154	0.57554	0.57594	0.33494	0.44664	0.35454
Average Fitness	0.32254	0.33874	0.34704	0.33714	0.36004	0.34494	0.34284	0.34484	0.38904	0.35014
Best Fitness	0.22434	0.25904	0.30054	0.31744	0.24974	0.30904	0.33254	0.32264	0.30774	0.25344
Worst Fitness	0.32284	0.32594	0.41054	0.38514	0.35134	0.38514	0.41904	0.39884	0.40534	0.36854
Standard deviation Fitness	0.14484	0.14954	0.16774	0.14894	0.15884	0.15114	0.19384	0.15014	0.18574	0.15114

- bPSO-Guided WOA: This algorithm showcases a significant potential to solve a task, with a notably lower error value of 0. 25934. Its performance, when compared to other metrics, leads to compelling findings that suggest it is a promising feature in selection, not to mention the considerable reduction in error rate.

- bGWO: Despite its slightly higher root mean square (RMS) error of 0. 27654, the bGWO algorithm's average value is small enough to warrant a comprehensive comparison with bPSO-Guided WOA. This comparison brings to light the complexity hypothesis, which posits a potential trade-off between the weighted average of the feature fitness of the main algorithm and the selection size in terms of feature selection quality.

- bGuided WOA: With an average error of 0%, that is a remarkable achievement. 31584. When we compare the results of WOA-Guided, bGWO-Guided, uniWOA, and bPSO-Guided WOA, the latter ones demonstrate better performance. However, EQUAL (Equilibrated ENsemble of Classifiers) stands out with its ability to achieve competitive average fitness and selection size, underscoring its validity as a distinct feature selection technique.

- bPSO: Much like the bGuided WOA, the bPSO demonstrates the mean of its error to be 0. 31034. Nevertheless, it maintains in other metrics, still indicating that it is a real professional in the filtering process.

- bBA, bWAO, bBBO: These algorithms exhibit practically identical probabilistic errors, with all the mean error intervals covering the value 0. 31014 to 0. 31994. Although they demand a little larger computation power than other algorithms, the feature selection of GAs is always preferred to be used based on many other factors such as average fitness and selection size.

- bPSO_GA, bFA, bGA: As algorithms differ, they have different levels of performance. The errors have been discovered to be in an average range between nothing and 0. 29014 to 0. 30874. In particular, bFA and bGA have been shown to achieve competitive fitness values with tunable precision to identify progenitor cells.

The highlighted outcome in Table 2 shows that the choice of suitable feature-selection methods is crucial and may depend on the dataset characteristics and modeling goals. With error minimization being core to some algorithms and other directions on features subset size and fitness features, a thorough assessment concerning different feature selections must be done in order to come up with the best method for the task.

Doi : https://doi.org/10.21608/jaiep.2024.354003
Received: March 19, 2024 Accepted: April 20, 2024

4.2 Performance of Machine Learning Models

Table 3 describes the results from the different types of machine learning models used to work with the dataset. Several metrics, such as predictive accuracy, robustness and complexity, evaluate all model models:

Table 3: Performance of Machine Learning Models

Models	mse	rmse	mae	r	R2	NSE	WI	Fitted Time
LinearRegression	1.39E-31	3.73E-16	2.84E-16	1	1	1	1	0.01711
SVR	4.15E-07	0.000644	0.0005	0.999997	0.999994	0.999988	0.998355	0.315974
pipeline	1.46E-06	0.001209	0.000144	0.99998	0.99996	0.999959	0.999525	0.151072
MLPRegressor	4.38E-05	0.006617	0.005408	0.999662	0.999325	0.998785	0.982212	4.850946
XGBoost	0.000398	0.019953	0.012878	0.995218	0.990458	0.988946	0.957644	22.69078
ExtraTreesRegressor	0.000556	0.023586	0.011073	0.992398	0.984854	0.984555	0.963581	0
GradientBoostingRegressor	0.000636	0.025221	0.01117	0.991208	0.982493	0.98234	0.963259	0
RandomForestRegressor	0.000779	0.027919	0.015275	0.989856	0.979815	0.978359	0.949758	0.464773
CatBoost	0.001071	0.032724	0.011586	0.986496	0.973174	0.97027	0.961893	12.69663
KNeighborsRegressor	0.002988	0.054666	0.039652	0.962476	0.926359	0.917034	0.869581	0.020573
DecisionTreeRegressor	0.00312	0.055856	0.038087	0.957404	0.916622	0.913381	0.874727	0.210368

- Linear Regression: The mentioned model has achieved near-zero MSE, RMSE, and MAE values with very fast time representation. Here, the straight line shows that the values are in complete correlation (r = 1) and determination ($R^2 = 1$) with the data. This implies that the data fit exactly with the regression line. However, it calculates workload using minimum computation time, which may be an important solution to tasks of predictive modeling.
- SVR, Pipeline: These models work flawlessly, as they show low MSE, RSM, and MAE. They ranked high on the correlation and determination coefficients, furthering the argument that their trends tend to be predictable. The computation time of the SVR is half a million nanoseconds more than that of the Pipeline.
- MLPRegressor, XGBoost, ExtraTreesRegressor, GradientBoostingRegressor, RandomForestRegressor: These results are validated by the ensemble and deep learning models that showcase the competitive performance, which exhibits low to moderate MSE, RMSE, and MAE values. They manifest strong correlations and are characterized by high determination coefficients, which proves the models to have good predictive power. On the one hand, they take greater computation time that can match that of SVR and Linear Regression as examples of less complex models.
- CatBoost, KNeighborsRegressor, DecisionTreeRegressor: Unlike these models, which might, therefore, produce moderately accurate solutions, the others generally produce fewer errors. Hence, this evidence indicates an existing correlation and determination coefficient, which means a descriptive capacity. In addition to this, they utilize different amounts of computer time, with Cold Boost taking more time than the rest.

Altogether, the showings in Table 3 point to multiple successful and different outputs from machine learning algorithms, which help to define the level of engagement of college students with AI in an educational context. For models such as linear regression, the predictions are made in a very swift manner, are accurate, and do not consume so much of the computational power of the computer since they are simpler. However, more complicated models such as MLPRegressor and XGBoost may take longer computational time, but they will have more predictive power. In this regard, it is necessary to accord significant consideration to the predictive accuracy, model complexity, and computational efficiency in the process of model selection.

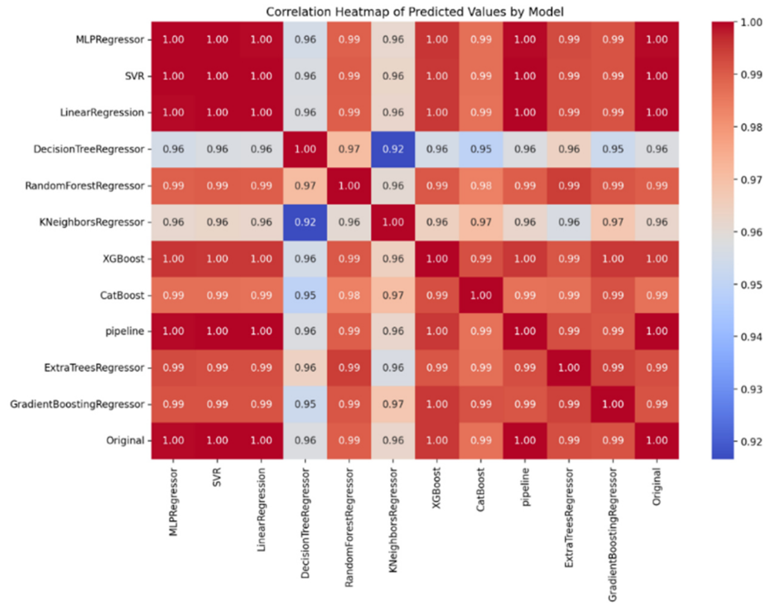

Figure 5: Correlation Heatmap of Predicted Values by Model

Figure 5 illustrates a correlation heatmap where the number of predicted values of the two models are compared. Different models can be tuned to provide more accurate insights into weather conditions. Visualization shows the consistency and reliability of predictions across these models.

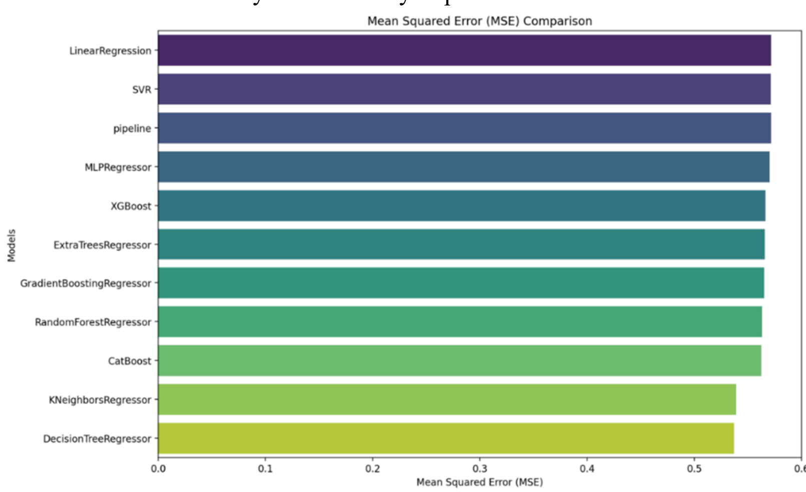

Figure 6: Mean Squared Error (MSE) ML Comparison

As can be seen in Figure 6, comparing different models' performance by means of Mean Squared Error shows a variation among the models. When MSE is drawn for every model, it constitutes an objective method for determining the most accurate prediction. Those metrics of iterative model evaluation allow us to determine which models have the highest quality and accuracy for this data.

Doi : https://doi.org/10.21608/jaiep.2024.354003
Received: March 19, 2024 Accepted: April 20, 2024

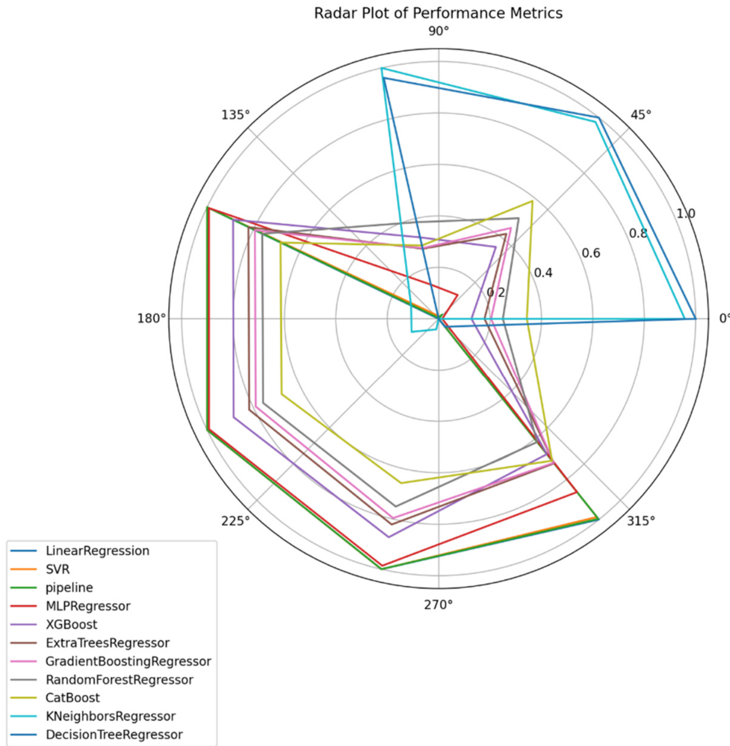

Figure 7: Radar Plot of ML Performance Metrics

As Figure 7 depicts, a radar plot is drawn to measure the machine learning model performance achieved on many criteria. Every model built is represented by a polygon, the vertices of which stand for model accuracy, precision, recall and F1 measures. Included in this presentation is a detailed analysis of every model, with particular attention paid to the good and bad aspects of each model. This gives more choice than selecting just one model among many of them.

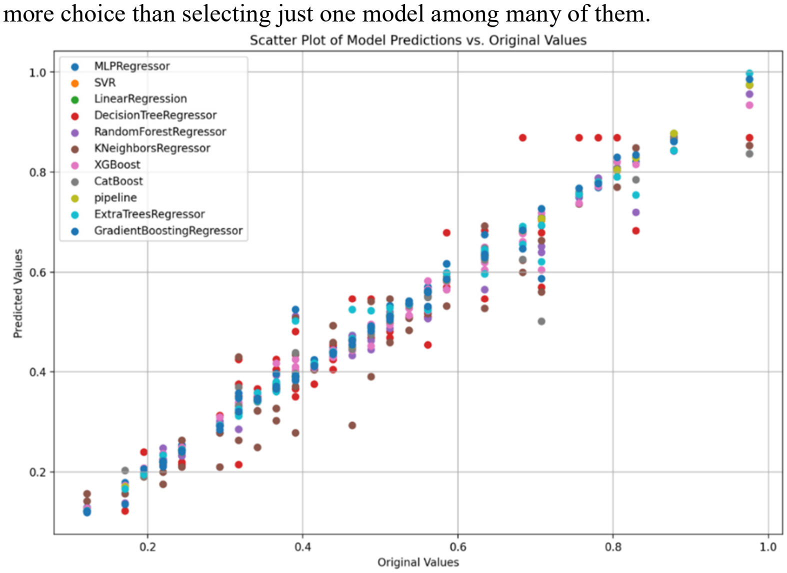

Figure 8: Scatter Plot of Model Predictions vs. Original Values

Figure 8 is a scatter plot depicting the link between the model's outputs and actual readings of a target number. Through the use of charting observed as well as predicted numbers, the visualization tool helps assess the prediction accuracy and ratio of errors. This plot makes this metric more easily interpretable; it can determine the authenticity of the patterns that models discovered in the dataset.

5. Conclusion

This study investigated student enrollment in AI technology with educational systems in which their survey answers were drawn from Grand Canyon College. Through our research, we succeeded in revealing critical elements of the student's choice of AI use in school tasks, career motivations in this field, and the strong and weak points of various machine learning models as predictors of such choice and behavior. Our Area of study is such an important aspect of educational technology and predictive analytics industries. We give you useful evidence concerning the extent and pattern of AI use among college students, which in turn testifies to the fact that AI is playing a great part in improving educational participation and output. Moreover, our study of the feature selection methods and machine learning models can help develop a practical guide that educationalists, policymakers, and researchers can use to achieve success using data-driven research to enhance student learning and innovation. The results of the research should be considered by educational institutions, politicians, and industry as they make decisions on how to make education more flexible. Educators can tailor instructional methods and course layouts according to the phenomenon that makes AI useful intelligence and other factors that drive student engagement. Education officials not only can leverage this information to design AI literacy training and innovation policies but also to provide an impetus for improvement in education. Educators, students and other stakeholders in academia can use this data to build AI-based tools and platforms that will reshape the way people receive and consume education. Thus, it is argued that AI in education is a potent factor of transformation, and clarity through data-driven approaches comes in handy to improve the learning process and help students achieve their goals. Putting the predictive analytics and machine learning capabilities into action, we would be able to deliver even more tailored and successful educational experiences and empower students as they adapt to the dynamic world of growing digital connectivity. As the relationship between AI and youthful training keeps on changing, dedicating to adequacy, inclusivity, and moral matters will be fundamental to empowering most students to secure the benefits of AI innovation, which will also aid learning.

Funding: “This research received no external funding”
Conflicts of Interest: “The authors declare no conflict of interest.”

References

[1] Ahmad, S. F., Rahmat, M. K., Mubarik, M. S., Alam, M. M., & Hyder, S. I. (2021). Artificial Intelligence and Its Role in Education. Sustainability, 13(22), Article 22. https://doi.org/10.3390/su132212902
[2] Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. AI and Ethics, 2(3), 431–440. https://doi.org/10.1007/s43681-021-00096-7
[3] Bates, T., Cobo, C., Mariño, O., & Wheeler, S. (2020). Can artificial intelligence transform higher education? International Journal of Educational Technology in Higher Education, 17(1), 42. https://doi.org/10.1186/s41239-020-00218-x
[4] Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: A narrative overview. Procedia Computer Science, 136, 16–24. https://doi.org/10.1016/j.procs.2018.08.233
[5] Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. IEEE Access, 8, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510
[6] Chen, X., Xie, H., & Hwang, G.-J. (2020). A multi-perspective study on Artificial Intelligence in Education: Grants, conferences, journals, software tools, institutions, and researchers. Computers and Education: Artificial Intelligence, 1, 100005. https://doi.org/10.1016/j.caeai.2020.100005
[7] Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two Decades of Artificial Intelligence in Education: Contributors, Collaborations, Research Topics, Challenges, and Future Directions. Educational Technology & Society, 25(1), 28–47.
[8] Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. Computers and Education: Artificial Intelligence, 4, 100118. https://doi.org/10.1016/j.caeai.2022.100118
[9] Cope, B., Kalantzis, M., & Searsmith, D. (2021). Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies. Educational Philosophy and Theory, 53(12), 1229–1245. https://doi.org/10.1080/00131857.2020.1728732

Doi : https://doi.org/10.21608/jaiep.2024.354003
Received: March 19, 2024 Accepted: April 20, 2024

[10] Guan, C., Mou, J., & Jiang, Z. (2020). Artificial intelligence innovation in education: A twenty-year data-driven historical analysis. International Journal of Innovation Studies, 4(4), 134–147. https://doi.org/10.1016/j.ijis.2020.09.001
[11] Hwang, G.-J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. Computers and Education: Artificial Intelligence, 1, 100001. https://doi.org/10.1016/j.caeai.2020.100001
[12] Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable Artificial Intelligence in education. Computers and Education: Artificial Intelligence, 3, 100074. https://doi.org/10.1016/j.caeai.2022.100074
[13] Alam, A., & Mohanty, A. (2023). Predicting Students' Performance Employing Educational Data Mining Techniques, Machine Learning, and Learning Analytics. In R. S. Tomar, S. Verma, B. K. Chaurasia, V. Singh, J. H. Abawajy, S. Akashe, P.-A. Hsiung, & R. Prasad (Eds.), Communication, Networks and Computing (pp. 166–177). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-43140-1_15
[14] Francis, B. K., & Babu, S. S. (2019). Predicting Academic Performance of Students Using a Hybrid Data Mining Approach. Journal of Medical Systems, 43(6), 162. https://doi.org/10.1007/s10916-019-1295-4
[15] Grivokostopoulou, F., Perikos, I., & Hatzilygeroudis, I. (2017). An Educational System for Learning Search Algorithms and Automatically Assessing Student Performance. International Journal of Artificial Intelligence in Education, 27(1), 207–240. https://doi.org/10.1007/s40593-016-0116-x
[16] Helal, S., Li, J., Liu, L., Ebrahimie, E., Dawson, S., Murray, D. J., & Long, Q. (2018). Predicting academic performance by considering student heterogeneity. Knowledge-Based Systems, 161, 134–146. https://doi.org/10.1016/j.knosys.2018.07.042
[17] Huang, C., Zhou, J., Chen, J., Yang, J., Clawson, K., & Peng, Y. (2023). A feature weighted support vector machine and artificial neural network algorithm for academic course performance prediction. Neural Computing and Applications, 35(16), 11517–11529. https://doi.org/10.1007/s00521-021-05962-3
[18] Miguéis, V. L., Freitas, A., Garcia, P. J. V., & Silva, A. (2018). Early segmentation of students according to their academic performance: A predictive modelling approach. Decision Support Systems, 115, 36–51. https://doi.org/10.1016/j.dss.2018.09.001
[19] Tomasevic, N., Gvozdenovic, N., & Vranes, S. (2020). An overview and comparison of supervised data mining techniques for student exam performance prediction. Computers & Education, 143, 103676. https://doi.org/10.1016/j.compedu.2019.103676
[20] Waheed, H., Hassan, S.-U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R. (2020). Predicting academic performance of students from VLE big data using deep learning models. Computers in Human Behavior, 104, 106189. https://doi.org/10.1016/j.chb.2019.106189
[21] Xu, X., Wang, J., Peng, H., & Wu, R. (2019). Prediction of academic performance associated with internet usage behaviors using machine learning algorithms. Computers in Human Behavior, 98, 166–173. https://doi.org/10.1016/j.chb.2019.04.015
[22] College Student AI Use in School. (n.d.). Retrieved May 5, 2024, from https://www.kaggle.com/datasets/trippinglettuce/college-student-ai-use-in-school
[23] Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J. H., Ogata, H., Baltes, J., Guerra, R., Li, P., & Tsai, C.-C. (2020). Challenges and Future Directions of Big Data and Artificial Intelligence in Education. Frontiers in Psychology, 11. https://doi.org/10.3389/fpsyg.2020.580820
[24] Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. Computers and Education: Artificial Intelligence, 2, 100020. https://doi.org/10.1016/j.caeai.2021.100020
[25] Roll, I., & Wylie, R. (2016). Evolution and Revolution in Artificial Intelligence in Education. International Journal of Artificial Intelligence in Education, 26(2), 582–599. https://doi.org/10.1007/s40593-016-0110-3
[26] Schiff, D. (2021). Out of the laboratory and into the classroom: The future of artificial intelligence in education. AI & SOCIETY, 36(1), 331–348. https://doi.org/10.1007/s00146-020-01033-8

[27] Timms, M. J. (2016). Letting Artificial Intelligence in Education Out of the Box: Educational Cobots and Smart Classrooms. International Journal of Artificial Intelligence in Education, 26(2), 701–712. https://doi.org/10.1007/s40593-016-0095-y

[28] Yang, S. J. H., Ogata, H., Matsui, T., & Chen, N.-S. (2021). Human-centered artificial intelligence in education: Seeing the invisible through the visible. Computers and Education: Artificial Intelligence, 2, 100008. https://doi.org/10.1016/j.caeai.2021.100008

[29] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education, 16(1), 39. https://doi.org/10.1186/s41239-019-0171-0

[30] Sanusi, I. T., Olaleye, S. A., Agbo, F. J., & Chiu, T. K. F. (2022). The role of learners' competencies in artificial intelligence education. Computers and Education: Artificial Intelligence, 3, 100098. https://doi.org/10.1016/j.caeai.2022.100098

Doi : https://doi.org/10.21608/jaiep.2024.354003
Received: March 19, 2024 Accepted: April 20, 2024