Analysis of English Performance Rating Based on Machine Learning and Optimized BP Network Technology

Yuan Liu and Liangfeng Dong

Foreign Studies School of Xuzhou University of Technology, Xuzhou, Jiangsu 221000, China

Correspondence should be addressed to Liangfeng Dong; hannexietuo@163.com

Received 26 May 2021; Revised 21 June 2021; Accepted 29 July 2021; Published 15 August 2021

1. Introduction

English is the universal language of the world. Moreover, in the context of global economic integration, English learning is not only a required course for business elites but also a required course for the general public. At present, in colleges and universities across the country, English is basically presented as a compulsory first foreign language course. Therefore, how to improve the effect of English performance assessment in the context of smart teaching has become an important part of smart English teaching.

Automatic scoring is an important part of evaluating the quality of education in the process of intelligent education [1]. Now, examination is an effective method to grasp the knowledge points of students, and it is also a necessary means to select talents. At the same time, it is inevitable that the daily homework and the grading of test papers in large-scale examinations are inevitable. In the process of manual scoring [2], students’ writing status and answer sheets are affected by the subjective factors of the scorer, which affect the consistency of the score. From the perspective of pedagogy, subjective questions (fill in the blanks, simple answers) will bring more educational benefits. Therefore, the research on automatic scoring methods for subjective problems is particularly urgent [3].

A neurofuzzy logics-based student performance analysis model was given in [4]. The model used the educational data set of students in the time period (1995–2005). It was stated in the work that the analysis has been carried out with the students at higher education levels for improving the results of the students as well as the institutions. In some models, the students’ performances are analyzed with the tutor performances. In [5], a model has been developed for evaluating instructor performance based on the questionnaire model. For that analysis, the model has used the classification models such as decision tree, support vector machine (SVM), discriminant analysis, and artificial neural networks (ANN). In [6], an informative case study has been presented by evaluating EDM-based models for application-oriented services of institutions. The classification models discussed in the paper were multilayer perceptron (MLP), J48 classifier, reduced error pruning decision tree, and sequential minimal optimization (SMO). In [7], Jayasinghe et al. studied the performance analysis of students based on
online and conventional educational patterns. The teachers are directly involved in the evaluation process without considering the learning pattern of the students. For effectively identifying the learning behaviour of students, a k-means clustering approach has been incorporated in [8]. The analysis model included the participation of students in other class activities in addition to their academic achievements. Before the final examinations were carried out, the results are categorized to reduce the failure rate of students. In [9], the authors developed a novel approach called hybrid educational data mining model (HEDM) for analyzing the student performance to effectively enhance the educational quality. The proposed model evaluates the student performance based on distinctive factors that provide appropriate results. Furthermore, the model combines the efficiencies of Naïve Baye’s classification technique and J48 Classifier for deriving the results and categorizing the student performance in precise manner. In [10], it combines artificial intelligence teaching system to build an artificial intelligence sports teaching evaluation model based on neural network. The artificial intelligence model starts from the process evaluation and the final evaluation. Moreover, it uses a recurrent neural network for data training and analysis and introduces a new decoder to perform data processing, and introduces a simplified gated neural network internal structure diagram to build the internal structure of the model. In addition, this study designs a control experiment to evaluate the performance of the model constructed in this study. The research results show that the artificial intelligence model constructed in this paper has a good effect in the performance prediction and evaluation of college sports students. Yong et al. proposed a multiobjective weighted voting ensemble classifier based on differential evolution algorithm for text sentiment classification [11]. This paper assigns greater weights to the classifiers with better performance and proposes a weighted voting approach based on differential evolution. Meanwhile, there are many similar articles, like [12, 13]. Besides, there are many methods based on classification [14–19].

At present, the biggest issue facing automatic scoring is how to correctly describe the similarity of meaning between the student’s answer and the reference answer. The attention mechanism has outstanding effects in the field of natural language. Therefore, this paper attempts to use BP network to solve this problem.

In order to solve the problem of English performance rating, this paper proposes a new method which is based on machine learning and optimized BP network technology. The motivation of this paper is to establish a new evaluation model which can help to analyze the English performance rating process and build the corresponding algorithm model. This new strategy can make full use of the advantages of the machine learning and BP neural network methods. This new method combines BP spatial network technology with spatial sampling and spatial statistical algorithms to construct the system kernel algorithm. In addition, this method uses a needs analysis to construct the framework of the entire English performance rating system and evaluates the practical effect of the system constructed in this paper through experimental research.

2. Related Work

At present, many experts and scholars have conducted research on the related technologies of automatic performance rating, and the related research is summarized as follows.

Literature [20] made regular expression rules based on reference answers, and each rule is related to a score. If the student’s answer is consistent with the rules, the corresponding score can be obtained. Literature [21] formulated rules for students’ answers based on dependent grammatical analysis and shallow meaning analysis. Literature [22] uses a model based on computer tagging and computer-aided review to retrieve student responses with specific content, which is specified in the form of multiple tag templates. Literature [23] automatically labels the answer to the student, proving that it is not very accurate through the experiment, which is really unable to replace the manual mode. Literature [24] proposed a method of automatically generating templates. However, due to the limited accuracy of rule acquisition and performance ability, this method has low generalization ability and cannot be used across domains.

Literature [25] proposed a content rating module (CAM), which matches multiple features at a shallow level through machine learning to detect meaning errors. Literature [26] proposed an automatic scoring method that combines character-level and sentence-level features. Moreover, it proposed a partial similarity histogram (HOPS), which correctly confirms the student’s answer through the character-level feature recognition part. The random forest classifier in the literature [27] is constructed using features such as text similarity based on alignment or filling and word item weight. Literature [28] extracted features such as part-of-speech representation, w2v and D2v, keywords, and statistical sentence length to train a random forest classifier for obtaining an explainable automatic scoring system.

Literature [28] used a neural network composed of CNN and LSTM for automatic scoring. The result is better than the nonneural network method. Deep Belief Network (DBN) was used to explore answers, questions, student-based various combination models, and composite models, respectively. The KAGraader model was proposed, which uses the continuous bag of words (CBOW) model and the long-term memory model (LSTM). It has achieved good results in the automatic scoring task of simple answers in Chinese. The literature proposed a partial connection model, which measures the consistency of students’ answers through the weighted average of multiple folded neural networks (CNN). Literature [29] proposed an automatic scoring model based on a deep autoencoder. In order to promote the further development of automatic scoring methods for simple answers, the text is combined with educational technology to improve the degree of educational informationization.
From the above analysis, it can be seen that the existing experts and scholars have mostly implemented theoretical research in the research of intelligent performance rating algorithms, and they all have their own drawbacks. Therefore, this paper analyzes the English performance rating process in combination with the wisdom of teaching and constructs a corresponding algorithm model to improve the effect of English performance rating.

3. Recognition Algorithm of Spatial Trajectory Information of English Text

The algorithm constructed in this paper is mainly applied to the recognition of English text information and can be applied to network examination systems and traditional paper examinations. Therefore, the system needs to have the ability to recognize spatial text information. When
Determining the number of sampling layers and sampling point
Select the core representative point
Calculate feature distance

Whether to identify

Select probability a few identification results
Standard comparison analysis
Digital identification result
The end

Figure 3: Kernel-based sampling block.

Figure 4: Illustration of kernel-based spatial sampling.

Figure 5: Implementation flowchart based on kernel sampling strategy.
Figure 6: Structure of the communication network.

Figure 7: English performance rating system based on machine learning and optimized BP network.

Figure 8: Answer library construction.
constructing the algorithm model, BP spatial network technology is combined with spatial sampling and spatial statistical algorithms.

We assume that the number of uncertain parameters of the analyzed structural model is \( k \) and random sampling is performed \( n \) times in the \( k \)-dimensional uncertain parameter space, and the sampling is repeated twice. Moreover, we extract the rows of the two matrices \( A \) and \( B \) in the following way to represent the sampling points, and the columns from left to right represent the dimensional coordinates of the \( k \) parameters

\[
A = \begin{bmatrix}
\begin{array}{cccc}
\mathbf{x}_{11} & \cdots & \mathbf{x}_{1i} & \cdots & \mathbf{x}_{1k} \\
\mathbf{x}_{21} & \cdots & \mathbf{x}_{2i} & \cdots & \mathbf{x}_{2k} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
\mathbf{x}_{n1} & \cdots & \mathbf{x}_{ni} & \cdots & \mathbf{x}_{nk}
\end{array}
\end{bmatrix}
\]

\[
B = \begin{bmatrix}
\begin{array}{cccc}
\mathbf{x}_{11} & \cdots & \mathbf{x}_{1i} & \cdots & \mathbf{x}_{1k} \\
\mathbf{x}_{21} & \cdots & \mathbf{x}_{2i} & \cdots & \mathbf{x}_{2k} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
\mathbf{x}_{n1} & \cdots & \mathbf{x}_{ni} & \cdots & \mathbf{x}_{nk}
\end{array}
\end{bmatrix}
\]

Next, we process the matrices \( A \) and \( B \) of the above formula, exchange the \( i \)-th column of the two matrices, and keep the remaining columns as they are. The obtained matrices are denoted as \( C_i \) and \( C_{-i} \), as shown in the following formula:
If all the sample points in the processed matrix can be used for simulation calculation to obtain the model output response value of the corresponding sample point, then the GSA parameter sensitivity index corresponding to the dispersion estimate of the model response is calculated:

$$\tilde{f}_0^2 = \frac{1}{n} \sum_{i=1}^{n} f(x_{r_1}, x_{r_2}, \ldots, x_{r_k}) f'(x_{r_1}', x_{r_2}', \ldots, x_{r_k}'),$$

$$\tilde{V}(y) = \frac{1}{n} \sum_{i=1}^{n} f^2(x_{r_1}, x_{r_2}, \ldots, x_{r_k}) - \tilde{f}_0^2,$$

$$\tilde{U}_i = \frac{1}{n} \sum_{r=1}^{n} f(x_{r_1}, x_{r_2}, \ldots, x_{r_k}) f'(x_{r_1}', x_{r_2}', \ldots, x_{r_k}').$$  (3)

The first-order GSA index $S_{x_i}$ of the model parameter $x_i$ is

$$\tilde{S}_{x_i} = \frac{\tilde{U}_i - \tilde{f}_0^2}{\tilde{V}(y)}.$$  (4)

The full-order GSA index $\tilde{S}_{x_i}^V$ of model parameter $x_i$ is

$$\tilde{S}_{x_i}^V = 1 - \frac{\tilde{U}_i - \tilde{f}_0^2}{\tilde{V}(y)}.$$  (5)

The second-order interactive GSA index $S_{x_i x_j}$ of model parameters $x_i$ and $x_j$ is

$$\tilde{S}_{x_i x_j} = \frac{\tilde{U}_{ij} - \tilde{f}_0^2}{\tilde{V}(y)} - \tilde{S}_{x_i} - \tilde{S}_{x_j}.$$  (6)

The above sampling simulation is more suitable for the application of actual text recognition problems. The actual problem is mainly the finite element model of the structure, and there is no specific analytic function. Through the continuous simulation of the sampling points of the sampling generator matrix $A, B, C$, and $C_{-i}$ to calculate the output, the sensitivity indicator of all the corresponding parameters can be calculated. Illustration of the trajectory method design is shown in Figure 1.

Figure 1 shows an example of a single sample trajectory generated for the situation. In each trajectory, the element effect of the input parameter variable can be calculated for two adjacent points in the trajectory. The specific process of trajectory design through matrix operation is shown in the following.

First, the algorithm generates random starting points for each trajectory in the $k$-dimensional design space and converts each starting point into $m$ ($m = k + 1$) points by selecting $m \times k$ sampling matrix $B_j$, where $B_j$ is defined as follows [21]:

$$B_j = \left( I_{m,1} X_j + \frac{\Delta}{2(2B - J_{m,k})D + J_{m,k}} \right).$$  (7)
Among them, the subscript $j$ represents the $j$-th trajectory, $J_{mk}$ is a $m \times k$ matrix with all matrix elements being 1 and $X_j = [x^1_j, x^2_j, \ldots, x^k_j]$ is the standardized starting point of the $j$-th trajectory. $\Delta$ is the change step of $x^i_j$, usually equal to $1 / (p - 1)$ or $p / [2 (p - 1)]$, where $P$ is the selected number of layers of $x^i_j$. $B$ is the original sampling matrix of $m \times k$, and its $i$-th row and $(i+1)$-th row are only different in the $i$-th element:

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$  \hspace{1cm} (8)

$B_j$ inherits the characteristics of $B$; that is, any one of the adjacent rows is different only in its $v$-th element, and $v$ is determined by the random permutation matrix $P$. Moreover, each row of $B_j$ is a design point generated from the starting point $X_j$.

Therefore, $r$ trajectories require the computational cost of a total of $r \times m$ model simulations. We define the meta-effect of the $v$-th input parameter in the $j$-th trajectory as the following formula:

$$d^v_j = \frac{y(x^1_j, \ldots, x^{v-1}_j, x^v_j + \Delta, x^{v+1}_j, \ldots, x^k_j) - y(x^1_j, \ldots, x^{v-1}_j, x^v_j, x^{v+1}_j, \ldots, x^k_j)}{\Delta}.$$  \hspace{1cm} (9)

The trajectory-based sampling strategy is an effective method to estimate the global sensitivity metric. However, the random sampling process of trajectory design may cause

| Number | Rating effect |
|--------|---------------|
| 1      | 98.8          |
| 2      | 98.2          |
| 3      | 98.2          |
| 4      | 98.6          |
| 5      | 97.3          |
| 6      | 98.4          |
| 7      | 97.1          |
| 8      | 97.2          |
| 9      | 97.5          |
| 10     | 98.2          |
| 11     | 97.5          |
| 12     | 97.0          |
| 13     | 98.9          |
| 14     | 98.1          |
| 15     | 98.0          |
| 16     | 98.4          |
| 17     | 97.7          |
| 18     | 97.2          |
| 19     | 98.6          |
| 20     | 98.7          |
| 21     | 98.9          |
| 22     | 99.0          |
| 23     | 98.1          |
| 24     | 97.4          |
| 25     | 97.4          |
| 26     | 98.6          |
| 27     | 97.6          |
| 28     | 98.7          |
| 29     | 98.2          |
| 30     | 98.7          |
| 31     | 97.1          |
| 32     | 98.3          |
| 33     | 97.9          |
| 34     | 98.1          |
| 35     | 97.4          |
| 36     | 98.5          |
| 37     | 98.1          |
| 38     | 98.6          |
| 39     | 97.1          |
| 40     | 98.4          |
| 41     | 97.2          |
| 42     | 98.3          |
| 43     | 97.9          |
| 44     | 98.5          |
| 45     | 98.9          |
| 46     | 97.8          |
| 47     | 98.9          |
| 48     | 97.7          |
| 49     | 97.8          |
| 50     | 98.2          |
| 51     | 98.7          |
| 52     | 97.0          |
| 53     | 97.3          |
| 54     | 98.4          |
| 55     | 99.7          |
| 56     | 98.4          |
| 57     | 97.3          |
| 58     | 98.5          |
| 59     | 97.8          |
| 60     | 97.3          |
| 61     | 97.5          |
nonoptimal coverage of the input space, especially for complex models with a large number of input variables. In order to solve the problem of unsatisfactory coverage, we propose an improved Morris sampling strategy, which ensures that the parameter input space of the model can be scanned better without increasing the number of simulations of the model. The specific steps are as follows. First, the algorithm randomly generates a large number of different trajectories. The distance between the $m$-th trajectory and the $l$-th trajectory is

$$d_{ml} = \sqrt{\sum_{i=1}^{k} \sum_{j=1}^{k} \left[ X^{(j)}_z(m) - X^{(j)}_z(l) \right]^2}, \quad m \neq l,$$  \hspace{1cm} (10)$$

Among them, $k$ is the number of model input variables, and $X^{(j)}_z(m)$ represents the $z$-th coordinate of the $i$-th point in the $m$-th trajectory. The optimal $m$ trajectory in $m$ is selected by maximizing the total Euclidean distance of $r$ trajectories. The total Euclidean distance of $r$ trajectories is defined as follows:

$$D_r = \sqrt{\sum_{m=1}^{r} \sum_{l=m+1}^{r} d_{ml}}.$$  \hspace{1cm} (11)$$

In order to select the best $r$ trajectory, all the combinations of $r$ trajectories are arranged starting from the initial $M$ trajectory, and the total Euclidean distance of each combination is compared. Finally, the combination of the best $r$ Euclidean trajectories with the largest total distance is selected. The optimization process is based on a computationally intensive optimization operation of trajectory design. However, the fatal problem of the Euclidean distance optimization based on the trajectory method is that the entire process must consider all points on each trajectory, which leads to a sharp increase in the optimization calculation time of the high-order original problem. An illustration of the radial method design is shown in Figure 2.

The new sampling design strategy is mainly proposed based on the two shortcomings of the above-mentioned traditional sampling strategy:

1. The optimization of the Euclidean distance degradation method needs to consider the huge optimization calculation burden of all sampling points on the sampling trajectory or radial sampling block.

2. Since the number of random samples is very small, it is not enough to eliminate the parameter $k$-dimensional hypercube input space.

In order to reduce the computational burden of Euclidean distance optimization as one of the shortcomings, a femto-based design is proposed, which reduces the number of consideration points required for Euclidean distance optimization. As shown in Figure 3, the core point is the initial control point of each sample cube and is located at the center of all sample points of each cube, so it serves as a representative point for optimization of the Euclidean distance of all sample points in each cube. When the center point of the $k$-dimensional hypercube is obtained, that is, $X = (x_1, x_2, \ldots, x_k)$, the other two points in the direction of the $v$th ($v = 1, 2, \ldots, k$) coordinate axis can be obtained by adding and subtracting the step size $s$, namely,

$$X^+ = (x_1, x_2, x_{v+1}, \ldots, x_k),$$

$$X^- = (x_1, x_2, x_{v-1}, \ldots, x_k).$$  \hspace{1cm} (12)$$

Figure 10: Statistical diagram of the effect of English performance rating.
We define the meta-effect EE of the \( v \)-th input parameter in the \( j \)-th sampling box as

\[
EE^v_j = \frac{y(x^1_j, \cdots, x^{r-1}_j, x^r_j + \text{step}, x^{r+1}_j, \cdots, x^k_j) - y(x^1_j, \cdots, x^{r-1}_j, x^r_j - \text{step}, x^{r+1}_j, \cdots, x^k_j)}{2 \times \text{step}}.
\]  

(13)

Kernel-based sampling block is shown in Figure 3. Illustration of kernel-based spatial sampling is shown in Figure 4.

For the newly proposed Euclidean distance optimization of the kernel-based sampling strategy, all the core points in the multidimensional sampling cube are control points, so as to greatly reduce the number of spatial points considered in the Euclidean distance optimization based on the nuclear sampling design. In order to avoid the problem of random sampling at the defect, the kernel-based sampling design uses a multilayer LHS method to generate excellent coverage samples in the entire multidimensional input space to generate enough sample cube center points. The new sampling design greatly reduces the idea of representative points, the number of points that should be considered when optimizing the Euclidean distance, and the computational burden of this optimization process. On the other hand, the total Euclidean distance of the \( r \) cube is defined by the following formula. Compared with the initial design of the Euclidean distance optimization method, the kernel-based design saves more than ten times of calculation time in the Euclidean distance optimization process, especially when considering the large number of model input parameters:

\[
D = \sum_{m=1}^{r} \sum_{i=m+1}^{r} \left( \frac{\sum_{j=1}^{k} (X^m_j - X^i_j)^2}{2} \right). 
\]  

(14)

Here, \( X^m_i \) represents the coordinate value of the \( i \)-th parameter of the \( m \)-th sampling core point.

Using matrix \( P \) and matrix \( R \), the sample space \( S \) can be obtained, as shown in the following formula:

\[
S = \frac{1}{N} \times (P - R). 
\]  

(15)

Each element in \( S \) is the corresponding cumulative distribution probability, and then the parameter sample value can be obtained by using the inverse function of the cumulative probability distribution to invert:

\[
x_{ij} = F^{-1}_{x_i}(s_{ij}). 
\]  

(16)

Here, \( i = 1, 2, \ldots, N, j = 1, 2, \ldots, K \). \( F^{-1}_{x_i} \) represents the inverse function of the cumulative distribution function of the \( j \)-th random variable, and \( x_{ij} \) represents the element in the \( i \)-th row and \( j \)-th column in the sample space \( S \). Then, \( X = [x_{ij}] \) represents the original sample space matrix extracted by LHS. Figure 5 shows the operation process.

In order to realize the parallel evaluation of multiple students' performance, we can build a data collection and performance evaluation system.

The structure of the communication network can be found in Figure 6.

### 4. System Model Construction and Testing

After the optimized BP network algorithm is obtained through the above analysis, the English performance rating system is constructed. This article mainly extracts students' English answers through text recognition when the system is constructed and compares them with the standard library. Among them, objective questions are mainly scored directly through standard answer comparison, and subjective questions are to compare student scores with standard answers. After that, a certain prediction is made through the system and finally a subjective question score is given. The system model constructed in this paper is shown in Figure 7.

The system database constructed in this paper is mainly an answer bank of English test questions. Since the system in this paper is suitable for every English exam, it is necessary to build a special answer bank, which is convenient for teachers to operate the system, and the answer bank can be reset every time. The system can also collect the answer information of different students and compare it with the standard answer library. In this process, the English answer information of all students is first obtained through systematic spatial statistics and intelligent recognition and compared and analyzed through the standard library. Clustering and screening are also involved in this process, as shown in Figure 8.

On the basis of the above analysis, the performance of the system is analyzed. The system in this paper needs to perform text recognition on students' English answer sheets and then performs system score determination on the basis of text recognition. Therefore, firstly collect test paper information through text recognition, multiple sets of information through experimental research, and statistics on the accuracy of information. Since the answer recognition of English objective questions is relatively simple and the recognition results can be directly compared with standard answers, the system constructed in this paper mainly scores English subjective questions. In order to verify the performance of the method proposed in this paper, we compare the performance with the other two methods proposed in references [9, 10]. The results are shown in Table 1 and Figure 9.

From the above research and analysis, it can be seen that the accuracy rate of text information recognition of test papers of the method proposed in this paper is almost around 99%, which means that this method has good performance. Meanwhile, the English performance rating system based on machine learning and optimized BP
network proposed in this paper has a good effect on the text information extraction of English test papers. On this basis, the effect of systematic English performance rating is evaluated, and multiple experiments are carried out. The results obtained are shown in Table 2 and Figure 10.

Through the above experimental research, it can be seen that the effect of English performance rating of the method proposed in this paper is almost around 98%, which means that this method has good performance. Meanwhile, the English performance rating system based on machine learning and optimized BP network proposed in this paper has a good performance rating effect of English subjective questions.

5. Conclusion
The algorithm in this paper is mainly applied to the recognition of English text information and can be applied to the network examination system and the traditional paper examination. Therefore, the system needs to have the ability to recognize spatial text information. When constructing the algorithm model, the BP spatial network technology is used to combine spatial sampling and spatial statistical algorithms. Moreover, when the system is constructed, the students’ English answers are extracted through text recognition and compared with the standard library. Among them, objective questions are mainly scored directly through standard answer comparison, and subjective questions are for comparing student scores with standard answers. After that, a certain prediction is made through the system and finally a subjective question score is given. In addition, the system needs to perform text recognition on students’ English answer sheets and then perform system score determination on the basis of text recognition. Finally, this paper collects test paper information through text recognition, collects multiple sets of information through experimental research, and collects statistical information accuracy and performance rating accuracy. From the experimental research results, it can be seen that the English performance rating system constructed in this paper has a certain effect. The method proposed in this paper is based on the machine learning and BP neural network. Therefore, the application of the algorithm requires high computing power. So, in the future, we should further study an algorithm which can ensure efficiency and low computational complexity.

Data Availability
The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest
The authors declare that they have no conflicts of interest.

Acknowledgments
This study was supported by the Jiangsu Higher Education Philosophy and Social Science Project—Research on Internationalization of Undergraduates Cultivation in Jiangsu Universities under Background of “the Belt and Road” Initiatives (2019SJA0979).

References
[1] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: a new perspective," Neurocomputing, vol. 300, no. 2, pp. 70–79, 2018.
[2] J. N. Goetz, A. Brenning, H. Petschko, and P. Leopold, "Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling," Computers & Geosciences, vol. 81, no. 1, pp. 1–11, 2015.
[3] H. Darabi, B. Choubin, O. Rahmati, A. Torabi Haghighi, B. Pradhan, and B. Klove, "Urban flood risk mapping using the GARP and QUEST models: a comparative study of machine learning techniques," Journal of Hydrology, vol. 569, no. 5, pp. 142–154, 2019.
[4] A. Altaher and O. BaRukab, "Prediction of student’s academic performance based on adaptive neuro-fuzzy inference," International Journal of Computer Science and Network Security, vol. 17, pp. 165–169, 2017.
[5] M. Agaoglu, "Predicting instructor performance using data mining techniques in higher education," IEEE Access, vol. 4, pp. 2379–2387, 2016.
[6] P. Kaur, M. Singh, and G. S. Josan, "Classification and prediction based data mining algorithms to predict slow learners in education sector," Procedia Computer Science, vol. 57, pp. 500–508, 2015.
[7] U. Jayasinghe, A. Dharmaratne, and A. Atukorale, "Students' performance evaluation in online education system vs traditional education system," in Proceedings of the IEEE 2015 12th international conference on remote engineering and virtual instrumentation (REV), Bangkok, Thailand, February 2015.
[8] S. Ayesha, T. Mustafa, A. Sattar, and I. Khan, "Data mining model for higher education system," European Journal of Scientific Research, vol. 43, no. 1, pp. 24–29, 2010.
[9] V. G. Karthikeyan, P. Thangaraj, and S. Karthik, "Towards developing hybrid educational data mining model (HEDM) for efficient and accurate student performance evaluation," Soft Computing, vol. 24, no. 5, 2020.
[10] Y. . Zhang, "An AI based design of student performance prediction and evaluation system in college physical education," Journal of Intelligent and Fuzzy Systems, vol. 40, no. 6, pp. 1–9, 2020.
[11] Z. Yong, H. Zhang, J. Cai et al., "A weighted voting classifier based on differential evolution," Abstract and Applied Analysis, vol. 2014, Article ID 376950, 6 pages, 2014.
[12] Y. Zhang, S. Wang, P. Phillips, and G. Ji, "Binary PSO with mutation operator for feature selection using decision tree applied to spam detection," Knowledge-Based Systems, vol. 64, pp. 22–31, 2014.
[13] A. Rajkomar, J. Dean, and I. Kohane, "Machine learning in medicine," New England Journal of Medicine, vol. 380, no. 14, pp. 1347–1358, 2019.
[14] Y. Xin, L. Kong, Z. Liu et al., "Machine learning and deep learning methods for cybersecurity," IEEE Access, vol. 6, no. 1, Article ID 53565, 2018.
[15] L. Ward, A. Agrawal, A. Choudhary et al., "A general-purpose machine learning framework for predicting properties of inorganic materials," npj Computational Materials, vol. 2, no. 1, pp. 1–7, 2016.
[16] P. Feng, B. Wang, D. L. Liu, C. Waters, and Q. Yu, "Incorporating machine learning with biophysical model can
improve the evaluation of climate extremes impacts on wheat yield in south-eastern Australia,” *Agricultural and Forest Meteorology*, vol. 275, no. 3, pp. 100–113, 2019.

[17] K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, and D. I. Fotiadis, “Machine learning applications in cancer prognosis and prediction,” *Computational and Structural Biotechnology Journal*, vol. 13, no. 5, pp. 8–17, 2015.

[18] S. Amershi, M. Cakmak, W. B. Knox, and T. Kulesza, “Power to the people: the role of humans in interactive machine learning,” *AI Magazine*, vol. 35, no. 4, pp. 105–120, 2014.

[19] V. Rodriguez-Galiano, M. Sanchez-Castillo, M. Chica-Olmo, and M. Chica-Rivas, “Machine learning predictive models for mineral prospectivity: an evaluation of neural networks, random forest, regression trees and support vector machines,” *Ore Geology Reviews*, vol. 71, no. 3, pp. 804–818, 2015.

[20] C. W. Coley, R. Barzilay, T. S. Jaakkola, W. H. Green, and K. F. Jensen, “Prediction of organic reaction outcomes using machine learning,” *ACS Central Science*, vol. 3, no. 5, pp. 434–443, 2017.

[21] A. Chowdhury, E. Kautz, B. Yener, and D. Lewis, “Image driven machine learning methods for microstructure recognition,” *Computational Materials Science*, vol. 123, no. 8, pp. 176–187, 2016.

[22] S. Basith, B. Manavalan, T. H. Shin, and G. Lee, “SDM6A: a web-based integrative machine-learning framework for predicting 6mA sites in the rice genome,” *Molecular Therapy - Nucleic Acids*, vol. 18, no. 6, pp. 131–141, 2019.

[23] C. Voyant, G. Notton, S. Kalogirou et al., “Machine learning methods for solar radiation forecasting: a review,” *Renewable Energy*, vol. 105, no. 2, pp. 569–582, 2017.

[24] C. Folberth, A. Baklanov, J. Balković, R. Skalský, N. Khabarov, and M. Obersteiner, “Spatio-temporal downscaling of gridded crop model yield estimates based on machine learning,” *Agricultural and Forest Meteorology*, vol. 264, no. 4, pp. 1–15, 2019.

[25] J. Sieg, F. Flachsenberg, and M. Rarey, “In need of bias control: evaluating chemical data for machine learning in structure-based virtual screening,” *Journal of Chemical Information and Modeling*, vol. 59, no. 3, pp. 947–961, 2019.

[26] F. Thabtah and D. Peebles, “A new machine learning model based on induction of rules for autism detection,” *Health Informatics Journal*, vol. 26, no. 1, pp. 264–286, 2020.

[27] F. A. Narudin, A. Feizollah, N. B. Anuar, and A. Gani, “Evaluation of machine learning classifiers for mobile malware detection,” *Soft Computing*, vol. 20, no. 1, pp. 343–357, 2016.

[28] Q. Yao, H. Yang, R. Zhu et al., “Core, mode, and spectrum assignment based on machine learning in space division multiplexing elastic optical networks,” *IEEE Access*, vol. 6, no. 6, pp. 15898–15907, 2018.

[29] D. Bzdok and A. Meyer-Lindenberg, “Machine learning for precision psychiatry: opportunities and challenges,” *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, vol. 3, no. 3, pp. 223–230, 2018.