Proceedings of the 4th International Conference on Computing and Data Science (CONF-CDS 2022)
DOI: 10.54254/2755-2721/2/20220546

A Systematic Study on Intelligent Learning Techniques for Online Education

Jiayi Li[1]

[1]Zhengzhou University, ZhengZhou, China

lijiayi.e@foxmail.com

Abstract. With the increase of the situation and reasons for staying at home, for example, the spread of COVID-19, online education at home has correspondingly been of increasing importance in our daily learning, covering not only academic education but also adult education. Compared with offline face-to-face classroom teaching, the online learning system can capture a lot of students' learning data, such as learning duration, class rate, class completion rate, etc., which can be applied in further education design. Based on these data, one can explore the possibilities to promote the development of online education, using techniques like machine learning and deep learning. Such efforts also provide educational institutions and teachers with more analytical solutions to problems, making intelligent education play a greater role in advanced education, and help students to improve their study effectiveness specifically. Nevertheless, there still lacks of a systematic study on the direction of this work, which hardly depicts the overall development. In this paper, we first try to bridge this gap by sufficient investigation and analysis. We have studied the mainstream efforts on Intelligent learning Online Education (called ILOE) from two different dimensions, i.e., technology and task types. For each dimension, we analyze the characteristics of related works. Furthermore, we also provide useful suggestions for the future improvement of ILOE.

Keywords: intelligent education, machine learning, deep learning, research direction of online education.

1. Introduction

With the increase of the situation and reasons for staying at home, for example, the spread of COVID-19, online education at home has correspondingly been of increasing importance in our daily learning, covering not only academic education but also adult education. The popularity of online education has invoked related application design and effectiveness analysis. Such work naturally requires large-scale data support. Compared with offline face-to-face classroom teaching, the online learning system can capture a lot of students' learning data, such as learning duration, class rate, class completion rate, etc., which can be applied in further education design. Based on these data, one can explore the possibilities to promote the development of ILOE using techniques like machine learning and deep learning. The work targets include prediction and classification for score, study attitude, course evaluation, high-risk students, classroom performance, etc. Such efforts also provide educational institutions and teachers with more analytical solutions to problems, making intelligent education play a greater role in advanced education, and specifically help students to improve their study effectiveness.

Nevertheless, there still lacks of a systematic study on work direction. On the one hand, it hardly depicts the overall development status. It may cause relevant researchers to be unable to quickly grasp the development trend of the research field. It is also difficult to understand the focuses and challenges of related efforts. On the other hand, researchers are also hard to get the future direction of the ILOE research, which has become an obstacle for the development of ILOE.

However, the study on ILOE still faces several challenges. First of all, at the technical level, different educational scenes need different analysis dimensions, different analysis dimensions need different technical means, and different technical means have their own advantages. Online intelligent education needs to be analyzed according to the actual situation, and then take different measures. Therefore, it is difficult to make a difference for online intelligent education. Second, at the economic level, educational institutions will consider whether it is necessary to use existing technology to build a perfect online training system. Although online education is in full swing, most educational institutions still focus on offline training.

In this paper, we first try to bridge this gap by sufficient investigation and analysis. We have studied the mainstream efforts on ILOE from two different dimensions, i.e., technology and task types.

Firstly, in terms of technology, it can be roughly divided into three parts: traditional machine learning technology, deep learning technology, and improved technology. Traditional machine learning technology mainly includes SVM, random forest, logistic regression, linear regression, symbolic regression, etc., while deep learning technology includes deep neural network, sequential neural network, hybrid neural network, recurrent neural network. etc. Improved technology is a research method proposed by researchers to adapt to researchers' educational projects on the basis of the above two technology categories. The advantages and application scenarios of the three technology categories are different, which are described in detail in the following Section 3. Second, in terms of tasks, the mainstream research direction focuses on predicting students' performance, learning enthusiasm, and dropout rate and identifying cheating in exams, learning status, learning behavior, and so on.

For each part, we analyze the characteristics of related works and achieve some interesting conclusions. Traditional machine learning has a better effect on small data. It is cheaper in economy and calculation and easier to explain, but for large data sets, the accuracy is low. Deep learning technology performs well in many studies. It uses data to expand efficiently, does not need feature engineering, has strong adaptability, and is easy to migrate. For the improved learning technology, the adaptation scope is relatively small, which is more suitable for researchers' current research projects, and its universal applicability to another research needs to be further tested.

Furthermore, we also provide useful suggestions for the future improvement of ILOE. In the field of intelligent education research, many studies focus on how to identify cheating, predict students' performance, etc. In the future, more attention should be paid to the concrete ability improvement and learning facilitation for students. The specific ideas are as follows: in the early stage of learning, students should be warned in advance and reminded of learning situations. In the middle stage of learning, there should be interventions to assist learning. In the final stage, through the analysis of multi-dimensional learning data, we can get students' personalized problems and give personalized improvement suggestions.

Overall, in this paper, we aim to answer the following three research questions:

RQ1: What are the mainstream works focusing on the Intelligent Learning-based Online Education (ILOE)?

RQ2: What are the typical methodologies and characteristics within the existing ILOE works?

RQ3: How to improve existing ILOE strategies to achieve better performance and more influence?

2. RQ1: A study of intelligent learning for online education

To answer RQ1, we have investigated numbers of recently-published research works for the ILOE. Although the investigated works target at diverse topics, their methods can be typically divided into the following four types.

2.1. Traditional machine learning

Traditional machine learning-based methods, such as SVM, Logistic regression, and decision tree are widely used in the area of ILOE, due to their simplicity and accuracy. For instance, Parthiban et al. [1] applied a typical K-means clustering algorithm and seven machine learning classifiers to conduct the classification for a quite practical prediction problem, i.e., to evaluate the effects of full shutdown throughout the COVID-19pandemic on learners' education and stress. To this end, they collected 647 understudies' reactions to online course suggestions as data to train and test, where five types of responses were treated as labels for further training. The experimental comparison results show that SVM achieves the best performance in the classification. Furthermore, the works verify that students believe an online class can be used to supplement knowledge but cannot replace classroom learning due to one-on-one interaction. Jos´e A. et al. [2] constructed a random forest classifier to compare the effect of online education and analyze the problem characteristics of students using Copying Answers using Multiple Existences Online (CAMEO), a common cheating method on the Internet. Through the learning-based analysis of independent IP, the use of CAMEO when submitting homework answers can be effectively detected. Such random forest classifier is able to mark the cheating answers and make the detection results significantly improved, especially at sensitivity and specificity levels. Hsiang-Yu Chien et al. [3] employed two methods, stepwise logistic regression and random forest, to predict the grade outcomes of online learners. To this end, 225 students' learning data were collected and trained by several machine learning models. The experimental result demonstrates that random forest is more accurate than other methods in such grade prediction tasks. ALEXANDER J. STIMPSON et al. [4] found that online classroom provides new learning data compared with traditional educational data. Such data includes learning duration, learning process, etc. Based on this, they aimed to explore whether the information at temporal and process levels was useful for predicting student performance without exam results. To achieve the goal, they compared simple learning algorithms (linear regression and logistic regression) with relatively complex machine learning models (support vector machine and artificial neural network). The results show that the process data is very helpful for the early prediction of online courses, and for small course data sets, simpler algorithms (linear regression and logistic regression) are superior to most complex models. Fedora Duzhin et al. [5] proposed that in order to control the initial preparation level of educational quasi-experiments, the effectiveness of lecturers' teaching methods and students' prior knowledge should be comprehensively evaluated. To this end, they developed a machine learning-based symbolic regression method that uses non-experimental data collected by universities as input to predict changes in students' test scores. The proposed method was then used to assess the impact of different teaching methods. The experimental results provide positive suggestions for teaching strategies. For instance, online homework with immediate feedback is more effective than clickers, while clickers are more effective than traditional handwritten homework. The algorithm has now been integrated into an app and shared with the educational community.

2.2. Deep learning

Deep learning is widely used in online education and is often compared to traditional machine learning and statistical methods. For example, Waheed H et al. [6] deployed a deep artificial neural network to predict at-risk students by extracting manual features from clickstream data in a virtual learning environment. The results show that the classification accuracy of this model is higher than the baseline logistic regression and support vector machine models. Shuo-Chang Tsai et al. [7] took 3,552 students from a university in Taiwan as samples and conducted an early diagnosis of students with high dropout risk through the statistical learning method and the deep neural network learning method, so as to provide early intervention measures. Sadiq Hussain et al. [8] used the data of 10,140 students from three universities in Assam, India, and performed deep learning on their records using sequential neural models and Adam optimization methods. It is found that the deep learning technology achieves the highest classification accuracy, which reaches 95.34%. Chao Li et al. [9] found that in numerous platforms for classes, most of the online course completion rates were below 13% and a number of platforms for course professors were in the same way for all learners. The lack of considering different requirements and characteristics of learners, such as the prior knowledge and learning mode, negatively impacts the individualized teaching and learning. Under this insight, the authors proposed a hybrid neural network model, which integrates the convolutional neural network (CNN) and the gated recurrent unit (GRU). The model dynamically detects the learning style of learners by observing and identifying their behavioral states and conducts training according to the behavioral data of learners. After predicting the learning style of learners, the MOOC platform can provide learners with personalized learning paths and push relevant learning content according to their characteristics, so as to improve the learning efficiency of learners in a large-scale online learning environment.

In many educational programs, most universities, for example, provide a curriculum plan for each major, and students can choose different required courses and electives to form a personal learning path. Sosinskaya et al. [10] used recurrent neural network (RNN) to train the data of individual learning paths, classify the learning path, and evaluate the quality of such training. They took the individual learning path accumulated in the university information system as the observation value of the input of the deep learning network. The method has been applicable to most colleges and universities.

2.3. Improved learning

Many scholars use deep learning technology to continuously improve the specific application of online education. They put forward new models based on existing technologies to better adapt to the improvement field. For example, Dias S B et al. [11] demonstrated that deep learning technology can be used to analyze LMS interactive data, and proposed a novel prediction model DeepLMS, which can predict the quality of interaction with LMS. The model employed the Long Short-term Memory (LSTM) network to test the same database before and during an epidemic. DeepLMS provides a new way for educators to evaluate the holistic view of learner motivation and engagement in the learning process. Similarly, Kim et al. [12] exploited the bidirectional long short-term memory (BLSTM) to predict the future performance of students according to their online classroom interaction performance. The authors proposed a new GritNet algorithm and applied the algorithm to real Udacity student graduation projections. The results show that the GritNet algorithm not only consistently outperforms standard logistic regression methods, but also significantly improves the performance prediction during the most challenging first few weeks.

MOOC and other large online education platforms provide massive data for the research and development of intelligent education. Lee et al. [13] believed that intelligent learning analysis based on such educational data is the key direction of the development of intelligent education. In this data-driven learning analysis, the authors compared the prediction results of collaborative filtering algorithm with logistic regression algorithms and proposed a novel test response model associating collaborative filtering with enhanced human interpretability for the learning status prediction. The experimental results show that the predictive test results of the new fully data-driven collaborative filtering algorithm significantly outperform the traditional learning method such as logistic regression.

Tang et al. [14] designed a real-time classroom evaluation system using computer vision target recognition technology upon the FER-2013 facial expression dataset. The system improves the conventional CNN by concretely optimizing the trained parameters. The experimental results show that the proposed system can effectively achieve a real-time evaluation of students' classroom performance and fast feedback to teachers. Furthermore, the system is verified to have a higher average precision (AP) with notable lower training time. It could provide teachers with rich states, including concentration, positive, negative, surprise, and so on.

In order to realize the intelligent evaluation of the practical teaching process, He et al. [15] analyzed the accuracy of students' operation progress and task completion in practical courses. They then proposed an improved deep learning model based on the existing CNN, called LCNN, where the redundancy of parameters was effectively reduced. The experimental results demonstrate that the new method can reduce the model parameters by more than 50%, making the intelligent evaluation be more applicable in the teaching process.

3. RQ2: Methodologies and characteristics

3.1. Methodologies

Table 1. Details of mainstream methodologies

Technical Category	Model/Methods	Application Domain/Object
Traditional Machine Learning	SVM Random Forest Logistic regression Linear regression Symbolic regression	Classification of learners' stress responses Preventing from cheating Prediction of learning outcomes Prediction of student performance Effectiveness of teaching methods
Deep Learning	Deep artificial neural network Deep neural network Sequential neural network Hybrid neural network Recurrent neural network and its variants	Predicting students at risk of dropping out Predicting students at risk of dropping out Predicting student achievement Pushing personalized learning paths Categorizing and evaluating individual learning paths
Improved Learning	DeepLMS Grit Net algorithm New fully data-driven collaborative filtering algorithm The modified model of Convolutional Neural Network Lightweight Convolutional Neural Network	Predict users' online learning interactions Predicting Student Performance Analyze students' learning status Evaluation of students' performance Intelligent evaluation

To sum up, various technologies of traditional machine learning and deep learning have been extensively studied in the field of intelligent education, and actively promoted, applied, and improved by recent related efforts. Especially during the COVID-19 pandemic, online education has become a critical substitution of traditional education and attracted various related studies. Among them.

Support Vector Machine (SVM) is a representative linear classifier that classifies data by the supervised learning method, which was proposed in 1964. After the 1990s, it developed rapidly and derived a series of improved and extended algorithms, which have been applied in pattern recognition such as image recognition and text classification [16]. In the field of intelligent education, SVM plays a critical role in educational prediction and classification.

Apart from SVM, random forests are another effective classifier that uses multiple trees to train and predict samples. Random forests are built upon many classification trees. To classify a new object from an input vector, the input vector is put forward to each tree in the forest to conduct the classification using a "voting" fashion. The forest chooses the classification which has the most votes as the final result. Random forests are a very flexible and practical method with many advantages, such as excellent accuracy, effectiveness on large-scale datasets, and great compatibility for high-dimensional features [17].

Other traditional machine learning techniques are also applied in different prediction scenarios. For instance, logistic regression is well suited for describing and testing hypotheses about the relationship between categorical and continuous outcome variables. Logistic regression can be adapted to multivariable classification results. In the field of intelligent education, it works well to solve dichotomies (0 or 1) to evaluate the likelihood of some specific tasks, such as whether a student is at risk of dropping out and whether a student should be classified as disabled. [18]

Regression analysis, especially multiple linear regression is a common statistical and machine learning algorithm. The focus of general regression analysis is to predict continuous results by analyzing the relationship between dependent variables and independent variables.[19]

As a variant of traditional regression analysis, symbolic regression uses characteristic variables to predict target variables by discovering some hidden mathematical formula. The advantage of symbolic regression is that it can establish a symbolic model for the nonlinear system without relying on prior knowledge or models. Symbolic regression is based on evolutionary algorithm and its main goal is to synthesize the best possible solution to the custom problem by using an evolutionary method. Different from regression analysis, symbolic regression does not assume the form of function at all, the goal of which is to automatically discover patterns and laws within the inner educational data.

Deep neural network (DNN) is based on the extension of perceptron, and DNN can be understood as a neural network with many hidden layers. The neural network layer inside DNN can be divided into three categories, input layer, hidden layer, and output layer. The parameter training of DNN relies on the backpropagation algorithm and batch-based gradient descent strategy. Compared to traditional machine learning models, the DNN and further deep learning-based models have significantly better performance in the feature extraction, especially for large-scale training data such as educational data.

As an extension of the classical deep neural network, the sequential neural network model is proposed according to the characteristics of neural networks that sequentially construct high-level features through their continuous layers. In contrast to classical multi-layer networks, sequential neural networks allow learning of a set of local transformations without considering global transformations. The prediction process can be regarded as a sequential process with different forms by processing data with different characteristics through specific local transformation sequences, which greatly improves the ability to solve complex problems. Typical sequential neural networks include the recurrent neural network and its variants (e.g., LSTM and GRU). In the intelligent education prediction tasks, the sequential neural networks can well capture the temporal features within the collected educational data, and thus improve the prediction performance significantly.

Hybrid neural network combines diverse types of existing neural network models, such as CNN, RNN, and LSTM. Such hybrid neural networks integrate the advantages of each component model to be applied in some specific prediction scenarios. Furthermore, the hybrid neural network commonly improves the educational prediction results compared with a single learning model.

DOI: 10.54254/2755-2721/2/20220546

3.2. *Characteristics*

Table 2. Characteristics comparison of techniques in ILOE

Characteristics	Traditional Machine Learning	Deep Learning	Improved Learning
Advantages	Works better on small data Cheaper financially and computationally Easier to explain	Superb performance Efficient scaling with data No feature engineering required Adaptable and easy to migrate	Fit existing projects
Shortages	For large data sets, the accuracy is low Hard to extract fine-grained features	A large scale of training data is required Low training efficiency and high resource cost	The universality needs to be tested

In recent years, deep learning has become an important technique choice for intelligent education prediction problems and has shown remarkable performance in many tasks, including speech, natural language, visual, and gaming. However, despite the excellent performance of deep learning, classical machine learning methods still present performance in terms of both effectiveness and efficiency in some specific scenarios such as linear regression or decision trees, compared with using the deep network with large-scale datasets. In this section, the typical characteristics of both aforementioned mainstream methodologies are comprehensively concluded and systematically analyzed.

Traditional Machine Learning

Work better on a small scale of data. For small-scale datasets, classical ML algorithms are generally superior to deep learning. Both the selection of datasets and the scenarios of task analysis have a great influence on the performance of algorithm technology. The work of STIMPSON et al. [4] is such a typical example. The analysis results of both simple and complex machine learning algorithms are overall consistent. To achieve attractive performance, very large-scale datasets have to be required by deep learning, which is not always applicable.

Low cost of computation. With tremendous data and a large number of parameters to train, deep learning requires high-level support from both software and hardware, such as GPUs and corresponding learning frameworks. Even though, the overall training time for the deep learning model is still significantly longer than traditional machine learning. In specific ILOE scenarios, machine learning, as a light and fast prediction strategy, can be more effective than deep learning models.

Easy to explain. Since classical machine learning involves direct feature engineering, these algorithms are easy to explain and understand. In addition, it is easier to tune parameters and change model designs as we gain a deeper understanding of the data and underlying algorithms. Deep learning, on the other hand, is a "black box," and even now, researchers don't fully understand the "inside" of the deep network. Due to the lack of theoretical foundation, hyper-parameter and network design are also a big challenge. Since the aims of ILOE tend to provide useful suggestions to educators, the interpretability is a critical requirement, which demonstrates more effectiveness by machine learning in interpretability-related practice.

Deep learning-based Models

Superb performance. Deep networks have achieved better prediction precision than classical ML methods in many ILOE tasks. The reason is quite evident: deep models are commonly constructed with more complex structures and more parameters, which makes them have a stronger capability for feature representation. The variants of typical deep learning models further improve the prediction performance. Under this insight, deep learning-based ILOE models are naturally prone to attract more research focuses.

No feature engineering required. Classical ML algorithms typically require complex feature engineering which increases manual costs. To this end, semantic data analysis needs to be firstly performed on the dataset; then dimensionality prune and numerical representation should be conducted to facilitate further model learning; finally, the represented features should be carefully selected to feed into the target ML algorithm. When using deep learning, such feature engineering is not required because good performance can be achieved immediately by simply passing the data directly to the network. Deep models can extract features in an automated and deep way with their natural computation mechanism. Such an advantage eliminates the onerous challenges brought by feature engineering.

Adaptable and easy to migrate. Compared with classical ML algorithms, deep learning techniques can be more easily adapted to different domains and applications. Some deep learning technique such as transfer learning enable pre-trained deep networks to work for different applications in the same domain. Using these pre-trained networks as a front end simplifies training for the entire model and often helps achieve higher performance in less time.

4. RQ3: Improvement strategies

The development of intelligent education technology aims to facilitate the education design for educators as well as improve the learning strategy for learners. From the aforementioned investigation, existing works employ methods of machine learning or deep learning to conduct the prediction for course score, study attitude, course evaluation, high-risk students, classroom performance, etc. Although existing efforts have obtained noticeable achievements on ILOE and have resolved typical problems, they still face challenges in technology development, and the application of these efforts is heavily limited by existing task types.

4.1. Technology

From the analysis in Section 3, the main learning methodologies applied in intelligent online education prediction rely on either traditional machine learning or deep learning. Both of them have advantages in specific application scenarios, but also expose shortages in computation or resource cost. In future work, a single type of intelligent learning techniques can hardly cover the upcoming prediction requirements of ILOE. Instead, the hybrid tactic could be a better way to balance different prediction requirements. The model can be improved by novel algorithm structures, such as the hierarchical structure and graph structure. And the model can be enhanced by combining diverse types of networks to complement the shortages of every single network. Apart from that, some other intelligent technology such as the recommendation system, generation adversarial network, and reinforcement learning, can be improved and used in ILOE. Such intelligent technologies in a new direction can cover more prediction and classification tasks in the future requirements.

4.2. Target task

The task types studied in existing ILOE-related efforts mainly include the prediction for course score, study attitude, course evaluation, high-risk students, classroom performance, and so on. With the development of online education and the emergence of new education methods like the flipped classroom and blended teaching, ILOE should be extended to adapt to such new requirements. For instance, the interaction effectiveness is quite important for the evaluation of the flipped classroom, which however lacks sufficient studies in existing works. Thus, the types of target tasks should be treated as a research focus for the future development of ILOE.

5. Conclusion

With the development of machine learning and deep learning, intelligent learning techniques for online education (ILOE) have received increasing focus in recent years. In this paper, we conduct the

DOI: 10.54254/2755-2721/2/20220546

first systematical study for such efforts. To this end, we investigated typical ILOE works from two dimensions, i.e., technology and task types. We also analyzed the main characteristics of these efforts. To further enhance this direction of works, we also provide improvement suggestions based on the main characteristics of existing works for future development of this field.

References

[1] Parthiban K, Pandey D, Pandey B K. Impact of SARS-CoV-2 in Online Education, Predicting and Contrasting Mental Stress of Young Students: A Machine Learning Approach[J]. Augmented Human Research, 2021, 6(1): 1-7.
[2] Ruiperez-Valiente J A, Muñoz-Merino P J, Alexandron G, et al. Using machine learning to detect 'multiple-account'cheating and analyze the influence of student and problem features[J]. IEEE transactions on learning technologies, 2017, 12(1): 112-122.
[3] Chien H Y, Kwok O M, Yeh Y C, et al. Identifying At-Risk Online Learners by Psychological Variables Using Machine Learning Techniques[J]. Online Learning, 2020, 24(4): 131-146.
[4] Stimpson A J, Cummings M L. Assessing intervention timing in computer-based education using machine learning algorithms[J]. IEEE Access, 2014, 2: 78-87.
[5] Duzhin F, Gustafsson A. Machine learning-based app for self-evaluation of teacher-specific instructional style and tools[J]. Education Sciences, 2018, 8(1): 7.
[6] Waheed H, Hassan S U, Aljohani N R, et al. Predicting academic performance of students from VLE big data using deep learning models[J]. Computers in Human behavior, 2020, 104: 106189.
[7] Tsai S C, Chen C H, Shiao Y T, et al. Precision education with statistical learning and deep learning: a case study in Taiwan[J]. International Journal of Educational Technology in Higher Education, 2020, 17(1): 1-13.
[8] Hussain S, Muhsion Z F, Salal Y K, et al. Prediction Model on Student Performance based on Internal Assessment using Deep Learning[J]. iJET, 2019, 14(8): 4-22.
[9] Li C, Zhou H. Enhancing the efficiency of massive online learning by integrating intelligent analysis into MOOCs with an application to education of sustainability[J]. Sustainability, 2018, 10(2): 468.
[10] Sosinskaya S, Dorofeev R, Rogacheva S, et al. Generating Data on Individual Learning Paths for Classification Using Deep Learning Networks[C]//8th Scientific Conference on Information Technologies for Intelligent Decision Making Support (ITIDS 2020). Atlantis Press, 2020: 369-374.
[11] Dias S B, Hadjileontiadou S J, Diniz J, et al. DeepLMS: a deep learning predictive model for supporting online learning in the Covid-19 era[J]. Scientific reports, 2020, 10(1): 1-17.
[12] Kim B H, Vizitei E, Ganapathi V. GritNet: Student performance prediction with deep learning[J]. arXiv preprint arXiv:1804.07405, 2018.
[13] Lee K, Chung J, Cha Y, et al. Machine learning approaches for learning analytics: Collaborative filtering or regression with experts[C]//NIPS Workshop, Dec. 2016: 1-11.
[14] Tang J, Zhou X, Zheng J. Design of Intelligent classroom facial recognition based on Deep Learning[C]//Journal of Physics: Conference Series. IOP Publishing, 2019, 1168(2): 022043.
[15] He Y, Li T. A lightweight CNN model and its application in intelligent practical teaching evaluation[C]//MATEC Web of Conferences. EDP Sciences, 2020, 309: 05016.
[16] Qin, J. and He, Z.S., 2005, August. A SVM face recognition method based on Gabor-featured key points. In Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on (Vol. 8, pp. 5144-5149). IEEE.
[17] Random Forest's homepage (by Leo Breiman and Adele Cutler)
[18] Peng C Y J, Lee K L, Ingersoll G M. An introduction to logistic regression analysis and reporting[J]. The journal of educational research, 2002, 96(1): 3-14.
[19] Maulud D, Abdulazeez A M. A Review on Linear Regression Comprehensive in Machine Learning[J]. Journal of Applied Science and Technology Trends, 2020, 1(4): 140-147.