Research Article

Data Expression and Protection of Intellectual Property Education Resources Based on Machine Learning

Meng Mei¹ and Hui Tan²

¹Hunan Labor and Human Resources Vocational College, Changsha, Hunan 410100, China

²School of Geosciences and Info-Physics Engineering, Central South University, Changsha 410083, China

Correspondence should be addressed to Hui Tan; 175001013@csu.edu.cn

Received 27 January 2021; Revised 26 May 2021; Accepted 26 June 2021; Published 24 August 2021

Academic Editor: Wei Wang

Copyright © 2021 Meng Mei and Hui Tan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the improvement and growth of instructional informatisation, the contradiction between the open supply of academic resources, information expression, and mental property safety is turning into greater acute. Remedying the relationship between the two is very necessary for the overall performance of records expression of academic assets and the advent of true surroundings for mental property protection. The safety of mental property rights is to shield the rights and pursuits of know-how owners, defend the strength of information producers to produce knowledge, and defend the supply of academic sources sharing. The data expression and protection of intellectual property education resources based on machine learning is a kind of protection tool for the intellectual property of education resources developed using the characteristics of automation, real-time monitoring, and growth of machine learning. It can prevent web crawlers from harming e-commerce websites, prevent them from stealing the intellectual property of e-commerce websites, and analyse web crawlers that visit websites to prevent important website data from being stolen by them. From this point of view, based on the relationship between the fact expression of instructional sources and the safety of mental property rights, this paper advocates to promote the records expression and safety of mental property rights of academic sources from a couple of perspectives.

1. Introduction

With the gradual deepening of the records revolution and the gradual improvement of the open instructional assets movement, the community verbal exchange technological know-how and records digitisation science with technological know-how as the core have significantly enriched and developed the approaches of collection, processing, storage, and transmission of academic resources and opened up an extensive avenue for data exchange, making records sharing end up the groundwork of facts exchange, public conversation, and literature sources development in library and facts circles [1]. However, the sharing of academic assets is based totally on the open and improved circulation of information, which requires the free or affordable use of information, restricts the distinct use of information, and opposes the monopoly of information. Therefore, the

sharing of academic sources inevitably includes the confidentiality, protection, and distinctive use of specific information, particularly the safety of mental property rights of people or units. Intellectual property is the strength based totally on the character or collective mental innovative labour, which is monopolised using the oblige [2]. Nowadays, with the speedy improvement of new data technology, the rights and pursuits of statistics product companies are without difficulty infringed. People are required to reveal entirely in academic sources and enhance the safety of mental property rights of instructional resources. Coordinating the relationship between mental property safety and academic useful resource sharing and attaining each full aid sharing and nice safety of mental property are hassles that need to be significantly solved in the facts age [2].

Intellectual property rights are various forms of legal rights related to specific types of information, ideas, or other

intangible knowledge carriers. The owners of legal rights can enforce the right of exclusive subject documents. Intellectual property rights are protected by law as other properties; intellectual property reflects the idea that knowledge is a product of mind or intelligence. Knowledge should be protected by law as ordinary property [3]; intellectual property law should be used to protect all kinds of intellectual products. Intellectual property law is a highly specialised field requiring professional knowledge, especially when considering the legal differences in terms of authority; it is difficult to grasp accurately. Most open education resources are presented in the form of computer software products, mainly related to the copyright issues in intellectual property rights, and the copyright is composed of five basic rights: the right to copy, the right to adapt, the right to spread, the right to use, and the right to display [4]. With the rapid development of open education resources, the issue of intellectual property rights has become very urgent, which has a very important impact on the sustainable development of open education resources. However, as a new thing born over a decade ago, there is a lack of experience dealing with the intellectual property issues of open educational resources, and it faces many challenges [5]. Educational resources refer to resources with educational information that can be used by teaching practice to help learners master knowledge and skills better; intellectual property is the right given by law. This paper proposes to carry out legal protection for the intellectual science and technology achievements in a certain field and sanction the development, application, and reprint of the intellectual property without the permission of the intellectual property owner [6]. In the process of development and sharing of educational resources, it is inevitable to touch on the issue of intellectual property protection. Therefore, it is very important to explore this issue actively in order to enhance the legitimacy of development and sharing of educational resources.

Based on the machine learning method, this paper analyses and looks up the information of tutorial resources. Firstly, a couple of linear regression methods is used to rank the quantitative factors received from the neighbourhood instructing platform in accordance to the weight of the influence factor on the results. Then, the generalised regression neural neighbourhood is used to model the chosen factors with larger weight and all aspects [7]. After acquiring the features, the crawler awareness module of the device exams whether or not it suits the crawler features. If the behaviour points of the vacationer are judged as "matching" via the module, the warning data will be returned to the front cease, and the corresponding operation will be furnished for the administrator to choose. After the operation, the sense layer will save the records in the database, so that the desktop mastering module can iterate the data, constantly enhance the model, minimise the error, and obtain high-precision crawler recognition. The experimental consequences exhibit that the information illustration and protection model of intellectual property education resources primarily based on machine learning proposed in this paper has proper overall performance and feasibility and performs an effective function in the sensible application.

2. Related Theories and Technologies

2.1. Overview of Machine Learning. The technique primarily based on machine learning makes use of the characteristic vector set extracted from the pattern set to teach the machine learning classification model, which can mechanically predict the classification of unknown malware [8]. The education system of machine learning mostly consists of three processes: (a) modelling process: function vector extraction; (b) preprocessing: optimising characteristic vector set; and (c) classifier: deciding on gorgeous classification algorithm. Yakima and others proposed to extract permissions from many actual malware as features and then instruct the Bayesian classifier to observe unknown applications. The algorithm reduces the evaluation project of virus evaluation professionals and overcomes the barriers of normal signature-based detection methods. The consequences exhibit that the technique has a true detection fee for unknown android malware. Darebin et al. statically analysed Android APK files, usually extracting data such as permissions and aspects of utility requests from occur files and analysing touchy APIs and some community addresses known as through functions from DEX files. Darebin transforms the characteristic facts from static evaluation into a characteristic vector and then makes use of aid vector machine algorithm to realise malware [9]. Rieke et al. proposed a framework to analyse malware behaviour routinely through machine learning method. This technique analyses malware behaviour in an incremental way, warding off the runtime and reminiscence overhead of preceding methods. The process of machine learning is shown in Figure 1.

The specific life cycle is generally such a process: the general cargo will first arrive in a city distribution centre, the distribution centre for preliminary sorting; some cities will be divided into regions, so that the entire distribution centre will be divided only to each district, pulled to each district distribution centre, and then divided into various parts according to the courier's delivery range [10]. The coordination information device of a comprehensive coordination system should integrate the data flow of each different business operation function, together with the management aspects of orders, distribution, inventory, and human resources.

2.2. Classification Algorithm of Machine Learning. At present, the broadly used and mature machine learning classification algorithms at home and abroad are: help vector machine, k-nearest neighbour classification algorithm, Bayesian classification algorithm, decision tree classification algorithm, random forest algorithm, artificial neural network algorithm, and so on. We summarise and analyse the blessings and disadvantages of these classification algorithms and choose the most appropriate algorithm on this basis [11].

The full name of the KNN algorithm is the k-nearest neighbour. It is a nonparametric approach for classification and regression. In both cases, the input is composed of

FIGURE 1: The process of machine learning.

k-nearest training samples in the feature space. The KNN algorithm is case-based learning; the feature is only approximated locally; and all calculations are delayed to classification. The KNN algorithm is the simplest of all machine learning algorithms [12]. For classification and regression, a beneficial technique is to assign weight to the contribution of neighbours, so that the contribution of the nearest neighbours is greater than that of the farther neighbours. For example, a common weighting scheme is to set the weight of each neighbour to $1/D$, and D is the distance from the neighbour. Neighbours are bought from a set of objects whose class (KNN classification) or object attribute values (KNN regression) are known [13]. This can be considered as the training set of the algorithm, but it does no longer need explicit training steps. One of the characteristics of the KNN algorithm is that it is touchy to the local structure of data. One of the advantages of the KNN algorithm is that it can be well utilised for such problems as a large amount of data or many attribute values.

Support vector computing device (SVM) is 90 based on statistical learning theory, a machine gaining knowledge of algorithm for pattern awareness and pattern classification developed in the mid-1980s; improves the generalisation capacity of the learning machine by seeking the minimum structural risk; and realises the minimisation of empirical hazard and confidence range to gain the purpose of acquiring good statistical guidelines with less statistical samples [14]. According to the constrained sample information, the satisfactory trade-off between the complexity of the model and the learning capacity is sought in order to obtain the high-quality generalisation ability. It is a two-class classification model; its basic mannequin is defined as the linear classifier with the biggest interval in the feature space, that is, the getting to know the strategy of guide vector machine is to maximise the interval, which can subsequently be transformed into a convex quadratic programming problem [15]. It indicates many unique blessings in solving small sample, nonlinear and high-dimensional sample recognition, and can be applied to different machines getting to know problems such as feature fitting.

The Bayesian classification algorithm is a statistical classification method; it is a sort of classification algorithm using knowledge of chance and facts. These algorithms are primarily based on the Bayesian theorem, so they are referred to as the Bayesian classification. Given that the samples have distinct posterior possibilities in every category, the Bayesian classification technique calculates the posterior possibilities of every class of taking a look at data and classifies the samples into the class with the greatest posterior probability. The prior chance of every class can be recognised in strengthening using the statistical method. When estimating the conditional chance of attributes, it can be got with the aid of a statistical approach or assumed distribution model [16]. Naive Bayes algorithm wishes to anticipate that the attribute attributes of information need to be conditionally unbiased or essentially impartial. When this circumstance is established, the accuracy of naive Bayes classification is the highest; however, unluckily every characteristic attribute is regularly no longer conditional independent but has a sturdy correlation, which limits the capacity of naive Bayes classification, and the classification impact will be considerably affected [17].

The decision tree is a type of tree primarily based on strategic choice. It is a prediction model, which represents a mapping relationship between object attributes and object values. Each node in the tree represents an object, and every department direction represents a viable attribute value, whilst every leaf node corresponds to the cost of the object represented with the aid of the course from the root node to the leaf node [18]. The selection tree has solely a single output; if you desire to have complicated output, you can set up an impartial selection tree to deal with distinct outputs. Building a selection tree normally consists of two steps: choice tree technology and choice tree pruning. After the selection tree is generated, it frequently faces the trouble of overfitting. Decision tree algorithm has the benefits of small complexity, quick classification pace, and robust resistance to noise data. It is the sole classification algorithm that can deal with each statistics kind and traditional kind attributes at an identical time and has a huge vary of applications, but

when there are many records attributes, it is convenient to produce over becoming a phenomenon. When the quantity of pattern statistics is inconsistent, the statistics reap tends to have greater numerical characteristics.

2.3. Analysis of Machine Learning Technology. There are many classification methods of machine learning; the most common one is to divide machine learning into supervised learning, unsupervised learning, and reinforcement learning. Supervised learning is the most widely used one, which can be used for most classification and prediction problems, such as anomaly detection, house price prediction, and so on; unsupervised learning is mostly used to discover potential patterns in data automatically, such as consumer reinforcement learning is mostly used in the field of robot control. The following will choose these three types of machine learning algorithms that are representative of the introduction and analysis.

Supervised learning means that the data used for the training model is marked, which will be used to adjust the parameters of the machine learning estimation function in the process of machine learning training [19]. Supervised learning can be divided into regression and classification models according to the different types of variables. When the prediction output is continuous variables, it is called regression model, when the output is discrete variables, it is called classification model. The essence of the two learning models is the same, which can be transformed by discretising the output of the regression model or continuous the output of the classification model. Supervised learning can be divided into generative and discriminant methods [20]. Logistic regression is one of the most simple and practical supervised learning models. It is widely used because it has the advantages of fast training, easy understanding, and easy realisation. Logistic regression is a kind of linear two classifiers, which is essentially based on the linear regression, through a logistic function transformation of the calculation results. The original form of logistic regression can only be applied to the linear binary classification problem, and it can be applied to the non-linear classification problem after transforming the non-linear features. The derived softmax regression can solve the multiclassification problem. The advantage of logistic regression is that it is easy to implement and widely used in the industrial environment. The disadvantage is that it is easy to underfit and the accuracy is not high. Support vector machine (SVM) is a linear binary classifier similar to logistic regression. When it works out the hyperplane that can segment two class samples, it also requires the maximum interval between the hyperplane and two classes. SVM is a linear classifier with maximum interval. The maximum interval hyperplane is completely determined by the sample points close to it, which are called support vectors. The schematic diagram of the support vector machine is shown in Figure 2.

Support vector machine (SVM) directly predicts the result by judging that the sample is on one side of the hyperplane. The expression is as follows:

FIGURE 2: Schematic diagram of support vector machine.

$$\zeta_{\varphi}(x) = \begin{cases} 1, & \varphi^T x \geq 0, \\ 0, & \varphi^T x < 0. \end{cases} \quad (1)$$

When training the support vector machine, the loss function with the loss value of hinge loss function is as follows:

$$\psi(\varphi) = \sum_{i=1}^m \lambda_i \zeta_{\varphi}(x^{(i)}) + \sum_{j=1}^n \varphi_j^2 + c. \quad (2)$$

The optimisation objectives are as follows:

$$\min_{\varphi} \psi(\varphi) = \frac{1}{2} \sum_{i=1}^n (\xi_i + \varphi_i). \quad (3)$$

When using the support vector machine to deal with nonlinear problems, the data can be mapped into a linearly separable space by selecting an appropriate kernel function. The commonly used kernel functions are linear kernel, polynomial kernel, radial basis function (RBF) kernel, sigmoid kernel, and so on. The mapping diagram of the kernel function is shown in Figure 3.

Unsupervised learning means that the data used to train the model are not labeled, but the potential patterns in the data are automatically discovered through training. The most common pattern is data classification. The process of automatic data classification through unsupervised learning is called clustering. The result of clustering is called cluster analysis, which can be divided into partition-based method, hierarchy-based method, density-based method, model-based method, and grid-based method. The following will be the most used clustering algorithm k-means principal analysis. The goal of the k-means algorithm is to divide n samples into k clusters. In each cluster, the mean value w of the samples is used as the cluster centre, and each sample

FIGURE 3: Schematic diagram of kernel function mapping of support vector machine.

belongs to the family corresponding to the nearest cluster centre. The mathematical description of k-means algorithm is the known sample set. The cluster has to be found to satisfy the following equation:

$$\arg \min \sum_{i=1}^m \frac{\sum_{x \in H_i} \|x - \mu_i\|}{\lambda_i}. \quad (4)$$

It is difficult to solve the problem directly, so iterative optimisation is usually used to solve the problem. Firstly, random values in the same dimensional space as the samples are randomly selected as the initial clustering centre. Each sample is assigned to the cluster corresponding to the nearest cluster centre as follows:

$$H_i^{(t)} = \left\{ x_p : \frac{\|x_p - \mu_i^{(t)}\|^2}{\lambda_i} \leq \frac{\|x_p - \mu_j^{(t)}\|^2}{2}, 1 \leq j \leq m \right\}. \quad (5)$$

The new mean value of each cluster after allocation has to be updated and calculated as follows:

$$\mu_i^{(t+1)} = \frac{\sum_{x_j \in H_i} x_j}{|H_i^{(t)}|} + \zeta_i. \quad (6)$$

The two steps of allocation and update are repeated. When the allocation is no longer changed, the algorithm converges. The iteration diagram of k-means is shown in Figure 4.

In addition to the above description and analysis of machine learning methods, this section will also elaborate on the generalisation ability and evaluation index of machine learning, which will lay a theoretical foundation for the subsequent research, calculation, and analysis of machine learning algorithms. After the estimation function is obtained by machine learning, the most important thing is whether it can get the correct prediction results when it is applied to new data [21]. This ability to predict unknown data is called the generalisation ability of machine learning. When discussing generalisation ability, the most common problems are underfitting and overfitting, which are the main reasons for poor performance of machine learning algorithms. Underfitting is that the model does not fit the data enough, which leads to the failure to capture the

FIGURE 4: Iterative diagram of the algorithm.

structure of data features [22, 23]. When new data comes, it cannot effectively predict according to the features. Underfitting problem is also called high deviation problem, and overfitting means that the model fits the data too thoroughly, so that the noise in the data is also considered to be the structure of the features, which leads to the decline of the prediction effect of new data; overfitting problem is often called high variance problem. The diagram of underfitting and overfitting is shown in Figure 5.

In Figure 5, the graph on the left is underfitted due to too few model parameters, while the graph on the right is overfitted due to too many model parameters, and the graph in the middle is neither underfitted nor overfitted, which has good generalisation ability. In the case of underfitting, the fitting ability of the model can be improved by increasing the feature dimension to reduce the regularisation parameter. In the case of overfitting, the generalisation ability of the model can be improved by increasing the number of samples, reducing the feature dimension, or increasing the regularisation parameters [24]. In order to improve the generalisation ability of machine learning, cross-validation is usually used to train the model, that is, to group the original sample set: one part as the training set and the other part as the test set. After the model training is completed by the training set, the model is used to predict the test set, and the generalisation ability of the model is measured by the prediction results of the test set. The advantage of cross-

FIGURE 5: Schematic diagram of overfitting and underfitting: (a) underfitting, (b) balanced, and (c) overfitting.

validation is that we can get as much effective information as possible from the limited data, so that we can learn samples from multiple perspectives and avoid falling into the local extremum. In this process, both training samples and test samples are learned as much as possible.

2.4. Design of Intellectual Property Protection Technology Framework of Educational Resources Based on Machine Learning. In this paper, three core technologies of machine learning, decision tree, cart algorithm, and supervised learning are used to connect machine learning with anti-crawler mechanism. The algorithm processes the training data, generates a visual decision tree, uses the existing decision tree to analyse the new data, and keeps the analysis results in the training data set of machine learning in the analysis, so as to realise supervised learning and continuously improve the accuracy of recognition. Framework design of tool system is as follows. (1) Crawler features have to be collected, and anticrawler mechanism has to be used. (2) After obtaining the training data, the cart algorithm is used to generate a readable decision tree. (3) In the identification module, the access data is identified; the probability analysis is carried out by using the decision tree algorithm; the website is monitored; and the access IP data is collected, analysed, and scored to judge whether the access is a crawler operation and return the analysis results and analysis basis. (4) Real-time monitoring access data, based on the polling principle, can achieve real-time monitoring function after it is applied to practice. (5) In order to improve the accuracy of the machine learning module in identifying web crawlers, we use the supervised learning module to update the training data set. The design of the protection technology framework of educational resources based on machine learning is shown in Figure 6.

The anti-crawler mechanism is a mechanism to reduce the error rate of judgment. Antispider is a set of anticrawler measures. Crawler and anticrawler are a process of the continuous game. The so-called anticrawler mechanism is the response mechanism based on the means of the crawler robot. It is a strategic application; anticrawler mechanism achieves the purpose of blocking the crawler by pre-processing request header, blocking IP, asynchronous loading, using encrypted JS algorithm, setting verification code, and so on. The anticrawler mechanism of this tool is

embodied in the following operations: blocking the crawler IP and adding a “learning library” after using machine learning to identify the crawler [25]. The decision tree is composed of decision points, state nodes, scheme branches, and probability branches. The decision tree is a method to approximate the value of the discrete function; it is a typical classification method. Firstly, it processes the data, generates readable rules and decision tree by induction algorithm, and then analyses the new data by the decision tree. In essence, the decision tree is a process of data analysis through a series of rules. Supervised learning ability realises the complete access process to a client. After testing and judging whether it is a web crawler, the corresponding data is retained in the machine learning training data set, which can realise the updating and growth of training data. It can reduce the error rate of judgment and improve the accuracy of learning.

3. Experiment and Analysis

3.1. Experimental Data Preparation. This paper takes the basic computer course of the university as an example. The database contains 14 quantitative characteristics of educational resources extracted from the network teaching platform. The database contains multiple data sets. After effective data set division, the generalised regression neural network is used to model all features, and some features are obtained after feature selection through multiple linear regression. Through the analysis of the results, we can get the prediction error of the model and a series of evaluation indexes. According to the results of feature selection, the reasonable division of the data set will directly affect the stability of the regression model. The data of students’ learning behaviour obtained by the network teaching platform is divided into a training set and a test set according to the ratio of 4:1. According to the above-mentioned multiple linear regression method for feature selection, the weight of each feature on the response variable is shown in Table 1.

3.2. Analysis of Prediction Ability and Fitting Degree. In order to measure the prediction ability and fitting degree of the model more intuitively, we use a scatter diagram to show the experimental results of the sample test set and the training set respectively. The prediction results of the sample test set and the training set are shown in Figure 7.

FIGURE 6: Design of protection technology framework of educational resources based on machine learning.

From Figure 7, it can be concluded that for most samples, both the test set and the training set tend to be linear $y = x$. Among them, the samples with the error range of 5 in the test set are 67.6%, and the samples with the error range of 5 in the training set are 65.8%, which proves that the generalisation ability of the model is strong, and there is no overfitting phenomenon. In addition, by observing the scatter distribution, it can be seen that for the samples with scores between 70 and 90, the prediction results of the model are more accurate, while for the samples with scores above 90 and below 60, the prediction with high scores is lower, and the prediction with low scores is higher. This is because there are few samples with scores above 90 and below 60 in the data set, especially only one sample with scores below 40. We notice that the predicted value of the model is 45.2 for the sample with the observed value of 26.6. Although the error is large, the predicted result is relatively low, which has a high reference value for failing the examination.

3.3. Recognition Accuracy Analysis of Different Algorithms. k-Nearest neighbour classification algorithm and support vector machine algorithm are robust feature selection methods based on optimal mean, so it is necessary to test the optimal mean and robustness of the algorithm because the two methods are the same in robustness and optimal mean. The Bayesian classification algorithm is represented by norm-based feature selection method; the decision tree classification algorithm is represented by norm-based robust feature selection method, and the k-nearest neighbour classification algorithm is represented by norm-based and optimal mean robust feature selection method [26]. By comparing k-nearest neighbour classification algorithm, support vector machine algorithm, Bayesian classification algorithm, and decision tree classification algorithm, the

TABLE 1: Weight of each feature.

Feature number	Weight	Feature number	Weight
1	0.157	8	0.682
2	0.024	9	0.785
3	2.585	10	1.078
4	1.353	11	0.742
5	1.676	12	0.634
6	0.089	13	0.587
7	0.162	14	0.985

norm robustness and optimal mean are studied. In the simulation data, different SNR is added to test the robustness of the norm. The experiments of k-nearest neighbour classification algorithm, support vector machine algorithm, Bayesian classification algorithm, and decision tree classification algorithm are repeated 30 times, and the average recognition accuracy is taken for comparison. The specific results are shown in Figure 8. In all cases, the Bayesian classification algorithm has higher and more stable recognition accuracy than decision tree classification algorithm, which is attributed to the robustness of norm; k-nearest neighbour classification algorithm has higher recognition accuracy than support vector machine algorithm, Bayesian classification algorithm, and decision tree classification algorithm, which shows that the optimal mean theory can improve the feature recognition performance of the algorithm.

3.4. Analysis of Simulation Data Set. In this experiment, the block diagonal simulation data set is used for the simulation experiment. The data set is a 100×100 matrix and contains four 25×25 block matrices arranged diagonally. Each block represents a subspace. The data in each block represents the

FIGURE 7: Prediction results of: (a) sample test set and (b) sample training set.

FIGURE 8: Relationship between recognition degree of different algorithms and NSR: (a) NSR=0.3 and (b) NSR=0.6.

association degree of two corresponding points in a cluster and is generated randomly in the range of 0 and 1. All data outside the block represent noise and are generated randomly in the range of 0 to Q , where Q is between 0 and 1. In addition, in order to make the clustering task more challenging, 25 noise data points are randomly selected and set to 1.

The structure learned from the simulation data is shown in Figure 9. The values of Q are set to 0.7 and 0.8. When noise=0.7 and 0.8, the graphical representation of the original simulation data is, respectively, shown in Figures 9(a) and 9(b), and the structures after learning machine learning are shown in Figure 9(a). When

noise=0.7, the clustering accuracy of machine learning is 100%; when noise=0.8, the clustering accuracy of machine learning is 85%. Therefore, DHLs can achieve good subspace clustering results. Because the LRS method is the basic version of machine learning, machine learning is compared with other methods. It is obtained by random initialisation, which leads to the instability of clustering results. Therefore, 10 experiments are repeated on the same simulation data for comparative analysis. Under different noise settings, the clustering accuracy is shown in Figure 10. Different clustering evaluation criteria are considered, including mean, median, minimum, and standard deviation.

FIGURE 9: The result of learning from simulation data: (a) raw data matrix and (b) classification accuracy curve.

FIGURE 10: Continued.

FIGURE 10: Clustering accuracy analysis of different methods in simulation data set: (a) noise = 0.70, (b) noise = 0.75, and (c) noise = 0.8.

4. Conclusion

In the process of development and sharing of educational resources, the protection of intellectual property rights is a problem we need to face positively. We should face up to the infringement of intellectual property rights encountered in the current development and sharing of educational resources and actively take corresponding measures to improve and adjust: correctly understand the relationship between the two, establish a scientific concept of educational resources development, and pay attention to the guidance of public policies. We should give full play to the role of government functional departments, emphasise the cultivation of professional legal spirit talents, ensure the compliance of work, establish the development mechanism of documents, and create a good development environment. This is very important for building a more efficient education resource platform. With the help of machine learning technology, the influence weights of quantitative features obtained from the network teaching platform are sorted, and then the generalised regression neural network is used to model the selected features with higher weight and all features. In the future, it is necessary to increase the coverage of sample data set, especially for the expansion of low- and high-score data, so as to realise the data expression and protection model of intellectual property education resources based on machine learning, which has good performance and feasibility and plays a positive role in practical application.

Data Availability

No data sets were generated or analysed during the current study.

Conflicts of Interest

The authors declare that there is no conflict of interest.

References

- [1] G. Schneider, "European intellectual property and data protection in the digital-algorithmic economy: a role reversal(?)," *Journal of Intellectual Property Law & Practice*, vol. 13, no. 3, pp. 229–237, 2018.
- [2] J. Cabrera Medaglia and F. Perron-Welch, "The benefit-sharing principle in international law," *Journal of Intellectual Property Law & Practice*, vol. 14, no. 1, pp. 62–76, 2019.
- [3] M. Veale, R. Binns, and J. Ausloos, "When data protection by design and data subject rights clash," *International Data Privacy Law*, vol. 8, no. 2, pp. 105–123, 2018.
- [4] F. Li, "Effective analysis on circular economy development based on entropy method and data envelopment analysis (DEA): case study of Liaoning province," *Advanced Materials Research*, vol. 86, no. 9, pp. 883–886, 2014.
- [5] S. Yufeng and G. Quanying, "Ecological efficiency analysis of the circular economy system in mining area based on energy analytic approach," *Acta Ecologica Sonica*, vol. 34, no. 3, pp. 710–717, 2014.
- [6] A. Janich and M. Safranine, "Circular economy performance of EMAS organizations in Poland based on an analysis of environmental statements," *Multiciliary Aspects of Production Engineering*, vol. 2, no. 1, pp. 536–547, 2019.
- [7] B. Kuldeep, "Statistical and machine-learning data mining: techniques for better predictive modelling and analysis of big data," *Computing Reviews*, vol. 60, no. 7, p. 281, 2019.
- [8] H. Toivanen and O. Gross, "Data mining and machine learning in computational creativity," *Wiley Interdisciplinary Reviews: Data Mining & Knowledge Discovery*, vol. 5, no. 6, pp. 265–275, 2015.
- [9] A. Yosef, O. E. Nahum, A. Y. Anderson et al., "Data mining and machine learning tools for combinatorial material science of all-oxide photovoltaic cells," *Molecular Informatics*, vol. 34, no. 6, pp. 367–379, 2015.
- [10] A. Agnihotri, V. Yadav, and V. D. Kaushik, "Role of data mining and machine learning techniques in medical imaging," *International Journal of Advanced Intelligence Paradigms*, vol. 15, no. 3, pp. 340–345, 2020.

- [11] S. A. Ahmadi, G. Vicar, J. Frei et al., "Towards computerized diagnosis of neurological stance disorders: data mining and machine learning of post urography and sway," *Journal of Neurology*, vol. 26, no. 1, pp. 108–117, 2019.
- [12] H. I. Shusha, A. H. Awed, D. A. Moran et al., "Data mining and machine learning algorithms using il28b genotype and biochemical markers best predicted advanced liver fibrosis in chronic hepatitis," *Japanese Journal of Infectious Diseases*, vol. 71, no. 1, pp. 51–57, 2018.
- [13] M. L. Williams, N. Mac Parthaláin, P. Brewer, W. P. J. James, and M. T. Rose, "A novel behavioral model of the pasture-based dairy cow from GPS data using data mining and machine learning techniques," *Journal of Dairy Science*, vol. 99, no. 3, pp. 2063–2075, 2016.
- [14] A. Buck and E. Given, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1153–1176, 2017.
- [15] C. Hilma, T. Cramer, S. Kramer et al., "Data mining and machine learning techniques for the identification of mutagenicity inducing substructures and structure activity relationships of no congeneric compounds," *Journal of Chemical Information and Computer Sciences*, vol. 35, no. 4, pp. 1402–1411, 2018.
- [16] X. Liu, X. Wang, W. Zhang et al., "Parallel data: from big data to data intelligence," *Pattern Recognition and Artificial Intelligence*, vol. 30, no. 8, pp. 673–681, 2017.
- [17] F. Malmo and V. Senna, "Data intelligence for local government? assessing the benefits and barriers to use of big data in the public sector," *Policy & Internet*, vol. 9, no. 1, pp. 7–27, 2017.
- [18] N. Dey, A. E. Sassanian, C. Bhatt et al., "Internet of things and big data analytics toward next-generation intelligence," *Evaluation of Big Data and Business Intelligence*, vol. 10, no. 6, pp. 147–176, 2018.
- [19] K. Lu, "Experimental research of English teaching management model based on personalized multimedia technology," *Open Access Library Journal*, vol. 7, no. 4, pp. 1–13, 2020.
- [20] B. Xu, S. Song, and D. Wang, "Application of smart safety training and education in network teaching management," *Safety Emcee*, vol. 12, no. 4, pp. 104–108, 2020.
- [21] T. Ajam and D. J. Fourie, "Public financial management reform in south African provincial basic education departments," *Public Administration and Development*, vol. 36, no. 4, pp. 263–282, 2016.
- [22] Z. Xinmin, "Construction of financial auditing teaching mode based on artificial intelligence expert system," *Bolotin Technicon*, vol. 55, no. 17, pp. 743–747, 2017.
- [23] A. Karimzadehfini, R. Ahmadi, V. Kolkata et al., "2. Colonizing Kolkata," *Colonizing, Decolonizing, and Globalizing Kolkata*, vol. 14, no. 6, pp. 37–76, 2017.
- [24] X. Ge, "Application research of computer artificial intelligence in college student sports autonomous learning," *Educational Sciences Theory & Practice*, vol. 18, no. 5, pp. 135–137, 2018.
- [25] C. E. Bravo, L. Capitelli, F. Rivas et al., "State of the art of artificial intelligence and predictive analytics in the industry: a technology survey," *Spa Journal*, vol. 19, no. 4, pp. 547–563, 2014.
- [26] Z. Lina and L. Zhidong, "A improved method for teaching system innovation design of physical education major in colleges based on flipped classroom," *Revest de la Faulted de Ingenerate*, vol. 32, no. 9, pp. 630–635, 2017.