Research Article

Evaluation Method of Innovative Education Model of E-Commerce Video Live Broadcast Based on Big Data Analysis Technology

Huili Liu and **Ou Bai**

Innovation College, North-Chiang Mai University, Chiang Mai 50230, Thailand

Correspondence should be addressed to Ou Bai; g636501030@northcm.ac.th

In the environment of education big data, how colleges and universities make good use of these data not only affects the orderly operation of the whole education and teaching system of colleges and universities. It will also become an inexhaustible driving force to help colleges and universities promote the reform and innovation of the education and teaching system. This paper takes student evaluation data and student online learning data as research objects. Focusing on the teaching operation and students' autonomous learning, this paper uses the improved k-mode algorithm to cluster analyze the classroom teaching operation. This paper uses neural network algorithm based on machine learning to predict and compare students' online course learning. It is hoped that it can provide meaningful reference for the construction of teaching management system and the reform and innovation of teaching management system in colleges and universities. Two research works are mainly carried out through the preliminary analysis and transformation of the data of student evaluation of teaching in a certain university. The improved cosine dissimilarity algorithm is used to eliminate the abnormal teaching evaluation data. The normalization method was used to standardize the teaching evaluation data. The traditional k-mode algorithm is used to cluster the teaching evaluation data. Some problems of k-mode algorithm are pointed out, and the traditional k-mode algorithm is improved. Experimental results show that the improved algorithm is more reasonable and effective.

1. Introduction

Internet technology will be employed in my country, and various social networks and mobile Internet have begun to be popularized in our country, which also gradually broadens the application scope of Internet technology in my country. In recent years, my country's Internet data has also continued to grow, and the era of big data has arrived. The application is in my country, such as market economy, education, and cultural life, and has promoted the development and progress of all aspects of the country. In the era of big data, data processing has three major changes compared with traditional data mining: full volume rather than sampling, efficiency rather than precision, and correlation rather than causality [1]. The real meaning of big data lies not only in the huge data information but also

in the professional modeling and analysis of these data to mine its potential value [2]. Large data volume, low data value density, storage, analysis, and processing [3] are the contained potential values, and it will become the core content of big data research [4].

Two groups on such basis of the minimum error functions. In the scheduled K category, the main idea is to randomly select K cluster center points as the class center of the current cluster (not the final class center point) given the K value of the number of classification categories, and calculate the distance between each point and the class center point, divide it into the clusters with the closest distance until the division of all point clusters is completed, then, recalculate the center point of each cluster (average distance), recalculate ones, and assign each cluster [5], the cluster to which the point belongs, and this cycle iterates until

the center point of the cluster is less than the threshold or reaches the specified number of iterations [6].

The e-commerce video live broadcast course is one of the new contents of the e-commerce major. Because the course content is relatively new, students have more practice in live broadcast, so there are some differences with other majors in the students' live broadcast e-commerce learning process and the teachers' teaching process with disciplinary questions [7].

As an emerging online shopping method, live video streaming has enriched the online sales method of goods and has a strong sense of participation and experience, enabling consumers to have a better consumption experience. Many companies are still optimistic about the development trend of live broadcast e-commerce in the next few years. "Live broadcast with goods" has been a new transformation as well as upgrading of traditional firms [8]. However, live broadcast talents and live broadcast team building are the shortest way for enterprises to expand online sales.

Such positions and division of labor formed around live broadcast are becoming more and more specific, and live broadcast delivery may become the main means of corporate marketing [9]. The new occupations that have emerged so it will indeed put pressure on the training of it and such teaching implementation of new courses represented by live e-commerce courses will put a certain degree of pressure on teachers in lesson preparation, teaching, and tutoring. How to be competent in the teaching work of new courses, perform their duties as teachers, and be responsible as professional teachers of live e-commerce courses. There is a general lack of professional self-confidence and competence [10].

With the popularity of live streaming, the theoretical learning content of live streaming on the Internet is iterating rapidly, from the development history of internet live streaming, introduction of mainstream live streaming e-commerce platforms, to the prerequisites and preparations for live streaming, introduction to live streaming, live streaming review, data analysis, etc. Relying on textbooks alone is far from enough, and some textbooks are outdated when they are written. Live delivery of goods is a new content and method [9]. At present, such construction of relevant teaching materials and the design of practical links have not completely kept up, and the combination of schools and the live broadcast industry is relatively shallow, which is also a common practice of e-commerce teaching in the mobile Internet era [11].

Course ideology and politics can form a collaborative education mechanism with professional course teaching. First of all, due to the limited ideological and political level of teachers themselves, the excavation of ideological and political elements is not deep enough, so that the ideological and political elements of live broadcast e-commerce have not been naturally and properly integrated into the school's classroom education process [12]. The radiation-driven role of construction needs to be exerted. Secondly, in terms of teaching ability, there is a lack of targeted and demonstrative "course ideological and political" teaching guidance for teachers of live e-commerce professional courses, and an

effective teaching incentive system cannot be formed. Its influence is limited, and it lacks the attractiveness and appeal of effective teaching. Sometimes, too much attention is paid to the professional teaching of live e-commerce courses, and it is neglected to enhance the education of students' "road confidence", "personality confidence", "professional confidence", and "occupational confidence".

2. Literature Review

2.1. Status Quo of Big Data Analysis Technology Education Research. During the "Thirteenth Five-Year Plan" period, Guizhou Province fully implemented the province's big data strategic action, took the development of big data as a new strategic engine for the overall development of the province, established the big data concept that data is a resource, and promoted big data as a government It is a new means of governance capacity, a new way to serve the society and people's livelihood, a new driving force to lead industrial transformation and upgrading, and a new engine to promote mass entrepreneurship and innovation [13]. More than 20 big data scientific research institutions such as Guiyang Big Data Strategic Key Laboratory and Inspur Laboratory have been built successively. The layout of the big data industry has begun to take shape [14]. With the "Cloud Guizhou" system platform as the carrier, intelligent transportation cloud, food safety cloud, and the 7 cloud projects are the first to be applied, and the construction of "N cloud" such as Beidou location cloud and smart education cloud has achieved initial results, and the field of cloud application has been continuously expanded [15].

Foreign countries started earlier in big data research than in China, and foreign research on big data pays more attention to technology and application research, which to a certain extent benefits from the strong support of the government. For example, the U.S. government has established six departments for big data research, and most of its research projects are researching big data analysis algorithms, big data storage technologies, and big data security technologies [16].

Nagoya University in Japan, Columbia University, or Technology Sydney in Australia have all established data science research institutions; a large number of universities such as the University of Dundee and the Chinese University of Hong Kong have newly established data science institutions, Research Orientation Courses [17].

"A Brief Problem Statement", hoping to use educational big data to analyze students' learning behavior, so as to realize students' adaptive learning and improve the target of learning. At the same time, the U.S. Department of Education has carried out a lot of work on the collection of educational data and has collected a large amount of educational data. Its data portal Data.gov has gathered more than 300 large-scale data related to demographics, academic performance, loan status, campus security, etc. In the data set, the data range is very comprehensive. Data.gov also provides data analysts with different formats of data as well as online data analysis and online data visualization capabilities [18]. At the same time, Data.gov also provides an API interface

to facilitate users to obtain data and facilitate external calls and analysis.

2.2. K-Modes Algorithm.

In the ancient traditional taxonomy, the classification problem mainly comes from people's cognition of things. People mainly rely on experience and domain knowledge. The classification of things is mainly in the qualitative sense, and it is difficult to achieve the quantitative sense. However, various industries in various fields have put out problems, and the ancient traditional taxonomy based only on experience and domain knowledge is powerless, so people put mathematics introduced into taxonomy as a tool; a numerical taxonomy with quantitative taxonomic significance was formed [19]. After that, with the further increase of the difficulty of classification problems, people began to gradually introduce the related techniques of multivariate analysis into numerical taxonomy, forming the widely used cluster analysis technology today.

Since there is a certain overlap in their respective characteristics between various clustering analysis algorithms, it is difficult to find a clear classification plan to give a concise classification of clustering analysis methods. Currently, the commonly used classification methods are mainly divided based on the idea of clustering, as follows:

(i) Partitioning Methods: for a data set containing N samples, the event element specifies the number of classification families K $(K < N)$, first randomly gives the initial K cluster centers and records according to the same class The closer the distance between different records, the farther the distance between heterogeneous records is. Clustering is performed based on the K centers, and the K cluster centers are relocated and repeated iteratively to continuously improve the clustering effect until the optimal clustering is finally obtained

(ii) Method (Hierarchical Methods): for a given sample data set, the decomposition conditions are continuously decomposed according to the hierarchical nature of the sample data until the termination conditions are met. This method can be further subdivided into "bottom-up" hierarchical method and "top-down" hierarchical method according to different situations of the problem. The basic idea of the hierarchical method: when the method is hierarchically divided, it can be divided according to distance, density, and connectivity, or it can be extended to subspace for hierarchical division. That is, a bottom-up or top-down strategy is selected for a given data set, and iteratively divides by distance or density or connectivity until the decomposition satisfies a given condition

(iii) Ways: this method divides according to the density of data points according to a preset threshold. It will be divided into the same in close clusters

(iv) Grid-based Methods: it adopts such idea of space-driven, also known as fuzzy minimum and the several (finite) cell grids; all processing is carried out on the unit grid as the object. Such processing time has nothing to do with the number of data objects to be processed, and its processing time is only related to the number of unit grids in the quantization space

(v) Methods: the basic idea of this method is to assume a mathematical model for a certain cluster, and find the sample data that matches the mathematical model, so that the sample data and the mathematical model form an optimal fit combine. The usual practice is to determine sometimes based on statistical results. The idea of change is to achieve optimal clustering by optimizing the adaptability between a given data object and a mathematical model. The representative algorithms based on this idea are: PARTICLE FILTERS algorithm, MRKD-TREES algorithm, SOON algorithm, and hybrid algorithm

2.3. Neural Network Algorithm.

Since the neural network and mathematical has gone through decades. Although its development has experienced several ups and downs, some researchers with unique insights are still working on neural networks The research of network will undergo the research, laid such solid foundation for the wide application and rapid development of today's neural network in various fields.

Since the simulating of the biological nervous system, after decades of research, the work units have such characteristics of linearity, nonlimitation, very qualitative, nonconvexity, etc. It has the advantages of them.

At present, such simulation of the human brain, the method will lead to the processing problem into several processing units. Through the distributed parallel processing mechanism, the processing of unstructured information and some perceptual information is realized. Compared with the past, there has been a qualitative change, opening up a new space for the application of neural networks in many fields.

The course practice of e-commerce video live broadcast, in addition to the preparatory work and later maintenance, is mainly the live broadcast business. To develop a professional live broadcast business, you need to prepare a lot of professional props, such as sound cards, independent microphones, multitasking and multithreaded computers, cameras, background walls, arranged and designed scenes, etc. Most of these props can be simplified by various methods, such as computers and cameras can be directly replaced by mobile phones, and the background wall can be a solid color and simple wall. However, to achieve smooth live broadcast practice, a fast network environment is inseparable, and this is a hardware support that cannot be replaced by any method. This has also become a major problem for many teachers when implementing tasks in traditional classrooms.

To sum up, they incorporated the live video major into their talent training programs, and have carried out various forms of e-commerce live video education model innovation to varying degrees. However, the innovative education model of most colleges and universities e-commerce video live broadcast is not perfect, which restricts the realization of teaching objectives and the development of students'

comprehensive quality. Colleges and universities need such a complete evaluation system to evaluate the professional education mode of e-commerce video live broadcast to be perfected. Therefore, the analysis evaluates education innovation mode of e-commerce video live broadcast to enhance the innovation for video live broadcast education model in colleges and universities, cultivate more high-quality video broadcast talents, and promote economic development.

3. Methodology

3.1. Application of the Improved K-Modes Algorithm in the Evaluation of Teaching Conditions of College Teachers. As a regular part of higher education, the evaluation of teachers' teaching quality plays a guiding role in promoting the reform of higher education mode and realizing the high-quality development of education. At present, there are some problems in the evaluation of teaching quality of college teachers, such as the separation of evaluation from reality, excessive procedural justice, and so on. It does not reflect the humanistic concept and growth value of teachers' teaching quality evaluation, which leads to problems such as low effectiveness of evaluation results and low validity of evaluation indicators. The current classroom teaching quality evaluation system in colleges and universities in China mainly evaluates students' learning effect and teachers' teaching effect. Based on qualitative analysis and k means clustering analysis, this paper evaluates students' learning effects from a quantitative perspective. And according to the different types of students and the characteristics of the curriculum, improve teaching strategies, select scientific teaching methods, enhance the teaching effect of teachers, and improve the teaching quality of teachers, the classic one of the partition method. Its algorithm implementation is basically the same. From the introduction above, it can be seen that it is a simple and practical clustering method, but it cannot handle data sets containing categorical variables. Therefore, Huang et al. improved the one which can solve the problem of categorical data. The algorithm uses the common SMD (Simple Matching Distance) method to process categorical variables. The mode replaces the mean, the Hamming between the sample data points, and their corresponding cluster centers.

Through the analysis, processing and mining of teaching evaluation data, exploring a scientific, reasonable and effective evaluation strategy has far-reaching significance for the improvement of teachers' teaching ability, the improvement of student training quality, the improvement of teaching management, and the sustainable and healthy development of schools.

The multidimensional classification problem, due to the difference between categorical and numerical data, the degree of difference between the data is difficult to measure by grading. At the same time, for the convenience of expression, without changing its interpretation, this paper makes corresponding transformations for each evaluation value of students' evaluation of teaching, namely, assign each distinct type of data.

$f(x)$ is the transformation function of the evaluation value;

x is the original evaluation value of PJ; $x \in$ (excellent, good, moderate, failed), and $i = 1, 2, 3, 4$.

$$f(x) = \begin{cases} 5, x = \text{excellent}, \\ 4, x = \text{good}, \\ 3, x = \text{medium}, \\ 2, x = \text{pass}, \\ 1, x = \text{failed}. \end{cases} \quad (1)$$

Some abnormal data will inevitably exist. It will remove them from the sample set, which singular data to ensure the validity and authenticity for such evaluation.

To make an objective and fair evaluation when evaluating a course taught by a teacher, but made an evaluation with a strong personal color, which caused their evaluation to be different from other students. There is a large deviation in the evaluation of the students. Therefore, these abnormal data need to be removed from the evaluation data.

The improvement of the calculation formula of cosine distance similarity, directly using the cosine distance method to remove abnormal data for each classification file obtained above will cause the problem that the elimination result does not match the actual situation. Example 1: if there are two sample data x and y, their evaluation values are (1, 1, 1, and 1) and (5, 5, 5, and 5), respectively, and the cosine distance similarity is calculated as follows:

$$\cos(x, y) = \frac{1 \times 5 + 1 \times 5 + 1 \times 5 + 1 \times 5}{\sqrt{1^2 + 1^2 + 1^2 + 1^2} \times \sqrt{5^2 + 5^2 + 5^2 + 5^2}} = 1. \quad (2)$$

From the calculation results of the above formula, it can be seen that x and y are very similar, which is obviously contrary to the actual situation. There is a big problem in directly adopting this method. Therefore, this method needs to be improved. The improvement strategy specifically includes the following three aspects (still take Example 1 as an example):

The improved dissimilarity calculation formula is shown in Formula (3).

$$\begin{aligned} \text{Sim}(X, Y) &= \cos(X, Y) \\ &= \frac{\sum_{i=1}^{m}\left((x_i - p_x) \times \left(y_i - p_y\right)\right)}{\sqrt{\sum_{i=1}^{m}(x_i - p_x)^2} \times \sqrt{\sum_{i=1}^{m}\left(y_i - p_y\right)^2}} \\ &= \frac{X - Y}{\|X\| \times \|Y\|}. \end{aligned} \quad (3)$$

Recalculate the cosine distance similarity of the two sample data with the postreplacement evaluation value. The example is shown as follows:

$$\cos(x, y) = \frac{(-1 \times 2) + (-1 \times 2) + (-1 \times 2) + (-1 \times 2)}{\sqrt{(-1)^2 + (-1)^2 + (-1)^2 + (-1)^2} \times \sqrt{2^2 + 2^2 + 2^2 + 2^2}} = -1. \tag{4}$$

It can be seen that after the improvement, the cosine distance similarity model has a good consistency with the real situation.

The improved dissimilarity calculation formula is shown in Formula (5).

$$Sim(X, Y) = \cos(X, Y) = \frac{\sum_{i=1}^{m}\left((x_i - p_x) \times (y_i - p_y)\right)}{\sqrt{0.0001^2} \times \sqrt{\sum_{i=1}^{m}(y_i - p_y)^2}}$$
$$= \frac{X - Y}{\|X\| \times \|Y\|}. \tag{5}$$

Solve the problem of who to compare the similarity with. The cosine distance similarity comparison is for the comparison of two multidimensional sample data, so it is necessary to construct a target sample for comparison (hereinafter referred to as the target sample). For the partial center distance, the study takes such sample one corresponding to the average data as the target sample, denoted as T, namely,

$$T = \left(\frac{1}{N}\sum_{i=1}^{N}x_{i1} - p_1, \frac{1}{N}\sum_{i=1}^{N}x_{i2} - p_2, \cdots, \frac{1}{N}\sum_{i=1}^{N}x_{im} - p_m\right). \tag{6}$$

M will be dimensions, x_{ij} is the evaluation value of the jth column of the ith sample data, and p_j is the average value of the value range of the jth dimension of the sample data.

Abnormal student evaluation results and analysis. Applying the above method, 237,924 student evaluation records in R College's 1,326 classified sample data files were eliminated:

The data is scaled and uniformly mapped to the [0, 1] interval, so that all evaluation data on the same attribute column are standardized to the same equivalent, so that the evaluations are comparable. Its mapping method is as follows:

$$x_{ij}' = \frac{x_{ij} - \min_{i \leq j \leq m} x_{ij}}{\max_{i \leq j \leq m} x_{ij} - \min_{i \leq j \leq m} x_{ij}}. \tag{7}$$

The standardized flow chart is obtained according to Formula (7), as shown in Figure 1.

After eliminating it, based on the teaching evaluation situation for teachers in a certain term of R College, through the discussion on the K-modes algorithm in Chapter 2, it can be seen that each attribute value will be the one, wide range of applications. But the algorithm also has three serious deficiencies when it solves them. The first or the second is the determination of the initial K cluster centers; the third is the measurement and the following is the evaluation of
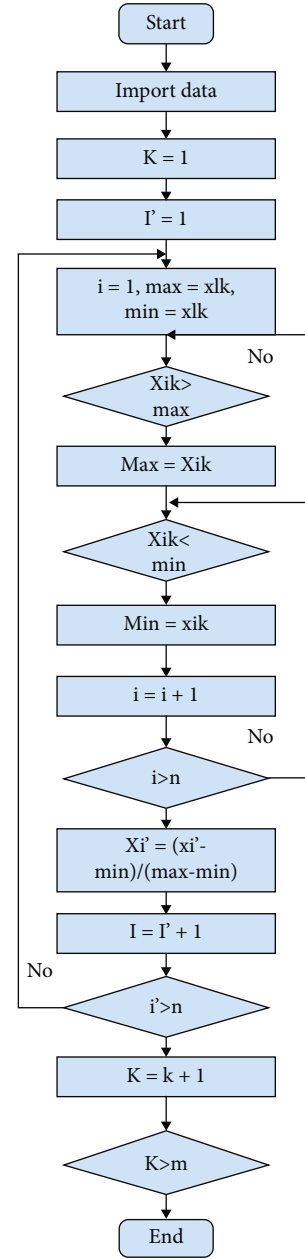

FIGURE 1: Standardization flow chart of teaching evaluation data files.

teachers' teaching ability by improving one in these three aspects.

It is based on the known clustering number K in advance. In many practical applications, K is not known. Even if K is known, the effect of clustering according to K may be very poor. No valuable information can be mined, and the effect of clustering cannot be achieved. Therefore, there are many software packages currently used to determine K, such as the Mclust package; the user can input the upper limit of the desired clustering family; the system performs a large number of calculations by distance, density, and other methods, and finally determines an optimal clustering. The number of classes, but for some problems cannot be calculated, and the efficiency is very low. Another

example is the Nbclust package. Its idea is similar to the Mclust package. By defining multiple evaluation indicators, various traversals are performed, and finally the cluster with the largest number of indicators supported is selected. Definition of error sum of squares is shown as follow:

$$SSE = \sum_{i=1}^{k} \sum_{x \in L_I} Dist(x, Z_l)^2. \tag{8}$$

The form is as follows:

$$f(x) = \begin{cases} C_1, x_{ij} = 5, \\ C_2, x_{ij} = 4, \\ C_3, x_{ij} = 3, \\ C_4, x_{ij} = 2, \\ C_5, x_{ij} = 1. \end{cases} \tag{9}$$

The frequency calculation formula for calculating the similarity of sample data based on frequency (AVF) is as follows:

$$AVF(x_i) = \frac{1}{m} \sum_{j=1}^{m} f(x_{ij}). \tag{10}$$

Calculate the value of SSE when each sample data point is the cluster center, and select the one when $l = 1$; in each sample except the data point that has been used as the assumed +1st initial cluster center, the SSE with $l + 1$ cluster centers is shown as follows:

$$\begin{cases} Z = \{Z_l | Z_l \in X, \text{and } Z_l \text{ and } Z_1, Z_2, \cdots, Z_{l-1} \text{ different}, l \leq k\}, \\ Z_l = \max_{x \in X} DXJSSE(x), \\ DXJSSE(x) = \left(\sum_{j=1, x \in L_j}^{l-1} Dist(x, x_j)^2 + \min_{x \in X} \sum_{i=1}^{n} Dist(x, x_i)^2 \right). \end{cases} \tag{11}$$

The calculation formula of the improved K-modes will in line with definition of the distance metric between two different values of the two sample data under a certain attribute and another attribute given above. The definition of the distance metric between the two sample data is as follows:

$$d(x, y) = \sum_{i=1}^{m} \sum_{j=1, \cdots m, j \neq i} d_{ij}(x_{Ai}, y_{Ai}). \tag{12}$$

3.2. Application Research of Neural Network Model in College Students' Learning Prediction. This paper combs, cleans, transforms, analyzes, and mines the data of students' teaching evaluation in a university. The improved K-Modes clustering algorithm is used to model and analyze the operation of teaching. The evaluation model of teachers' teaching status is established. On the basis of improving the similarity

calculation model in three aspects, the improved model is applied to deal with the abnormal teaching evaluation data. The results were normalized by normalization method. As the main channel of personnel training, the positive and negative excitation of neurons > 5859. We take the Kth neuron as an example to design, as shown in Figure 2.

We discussed a general artificial neural network structure, at the same time, we can also construct a neural network with another structure (the structure here refers to the connection of two neurons), that is, a neural network with multiple hidden layers. For example, there is a neural network with nl layers; then, the first layer is the input layer; the nth layer is the output layer, and each layer l in the middle is closely connected with the H + 1 layer. In this configuration, it is easy to calculate the output value of the neural network, we can follow the formula derived earlier, step by step forward propagation, calculate each activation value of the L2 layer cell by cell, and so on, then is the activation value of the L3th layer until the last Lnth layer. This connection graph has no loops or closed loops, so this kind of neural network is called a feed forward network.

The mathematical expression of Figure 2 is as follows:

$$\begin{cases} u_k = \sum_{i=1}^{n} w_{kj} x_{ij}, \\ v_k = u_k + b_k, \\ y_k = \phi(v_k). \end{cases} \tag{13}$$

The input item xij in Formula (13) is the jth feature of the ith training sample. In this problem, $i = 1, 2, 3, \cdots, n$, where $n = 150$, $j = 1, 2, 3, .., d$, and $d = 6$; wkj is the weight of the kth neuron on the jth feature of the sample; uk is the linear combination of each input item and the corresponding weight; bk is the kth neuron. The threshold (vk) of the unit is the excitation function of the kth neuron; yk is the output item of the kth neuron. The topology diagram of the neural network for this problem is shown in Figure 3.

For the feedforward neural network, there are only two types of neurons, one is the output unit, and the other is the computing unit. For the computing unit, it can accept multiple different inputs, but since there is no feedback information, there can only be one output. But this only one output can be coupled to any other unit as input, so the input of other layers except the input house is only related to the previous layer; the input layer and the output layer are connected to the peripheral, and the other layers in the middle both are hidden layers.

4. Result Analysis and Discussio.

4.1. Result Analysis of Improved K-Modes Algorithm. In the collected e-commerce live video course data, there will be outliers in the students' teaching evaluation data. The k-modes algorithm can automatically correct the outliers, estimate the fitting value according to the similar distance, and automatically store the samples into the abnormal value. In the sample data file library, after correction, the abnormal
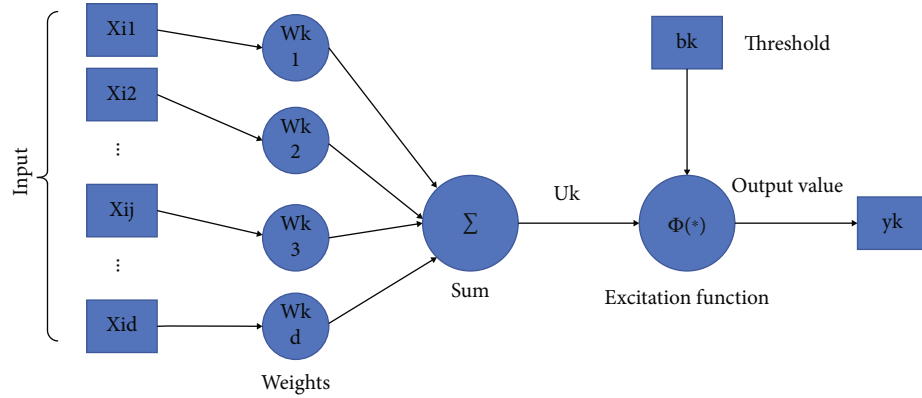
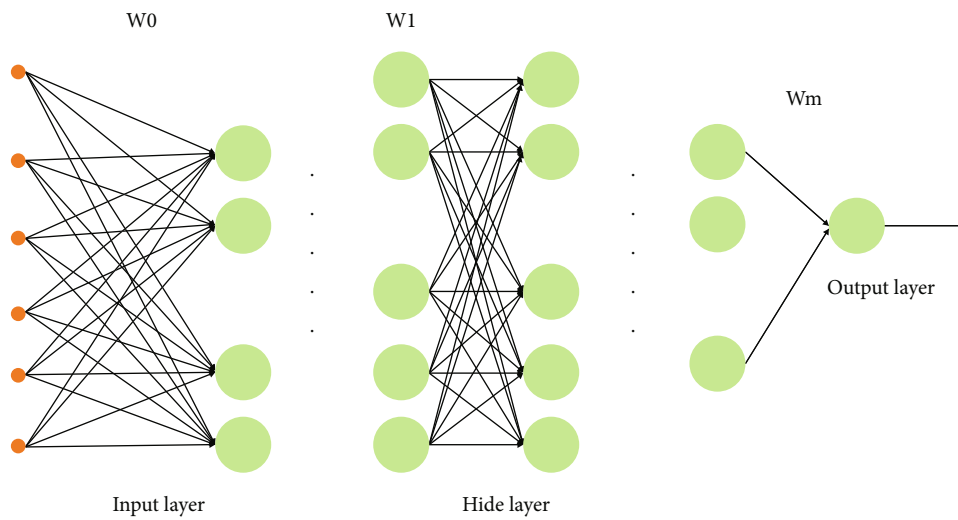FIGURE 2: Structure diagram of the kth neuron.

FIGURE 3: Topological structure diagram of m-layer of neural network.

sample is deleted from the classification file. As shown in Figure 4.

In the cluster analysis, the most important thing will be the one with K-Modes, and obtain the wrong ones in Figure 5. Diagram of the relationship of K.

However, it should be noted that the ones are carried out by the same method as above. The preclustering result found that it was increased when it is doubled; the preclustering results find that the inflection point of the image appears.

After removing outliers and determining the optimal clustering K value, the best fitting evaluation value is obtained, which can be used for comprehensive evaluation of the course teaching in the innovative education model of e-commerce video live broadcast. The effect of the modes algorithm on the clustering of a certain semester of Y school is shown in Figure 6.

It will be in poor clustering effect. Cooccurrence is used as a distance measure to improve the traditional ways. According to the clustering results, ones in the whole school and the teaching situation are analyzed by semester. In this way, it provides a scientific basis for correct decision-making and targeted policy implementation; on the other hand, it enables teachers to understand their own and other teachers' teaching conditions in a timely manner, and take targeted measures to enhance the internal driving force for continuous improvement of teaching.

How the teaching situation can better academic year semester as the time period, the teaching situation of the school has been clustered in three houses, including the teachers of the whole school in the past five years. The analysis is as follows, although it is slightly different in each one, the evaluation results are relatively stable. Figure 7 shows the statistical results of student teaching evaluation in a semester of school Y.

Figure 8 shows the change chart of the cluster center, the distribution of various types of people, the proportion of various types, and the proportion of teachers in the three categories.

The proportion is basically within the range of (43% and 45%); The second category (basically, 2-3 of the 4 indicators are rated as "good", and at least one of them is rated as "moderate"), and the proportion is basically the category (the evaluation of the four indicators is below "moderate"), and the proportion is basically as a whole.

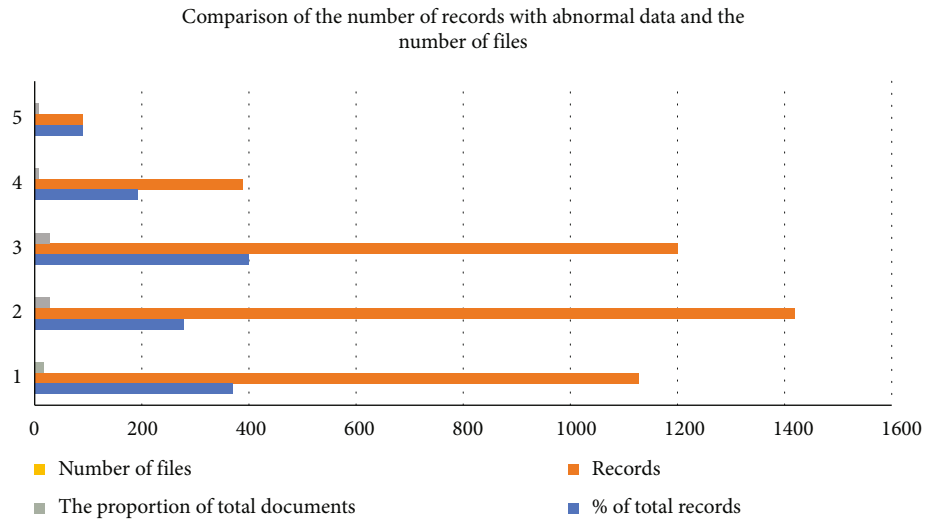About 55% are rated as "excellent" and "good", and about 45% are rated as "moderate" and below. Although

Comparison of the number of records with abnormal data and the
number of files

- Number of files
- The proportion of total documents
- Records
- % of total records

FIGURE 4: Comparison of the number of records containing abnormal data, the number of files, and the number of records.

Pre-clustering minimized error sum of squares vs. K

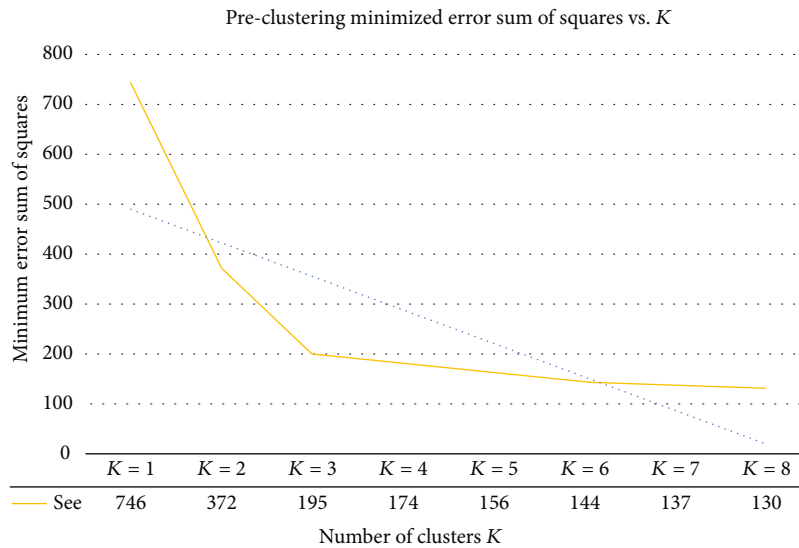|	$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$	$K = 6$	$K = 7$	$K = 8$
— See	746	372	195	174	156	144	137	130

Number of clusters K

FIGURE 5: Minimum squared error, the relationship between the sum and the number of clusters K.
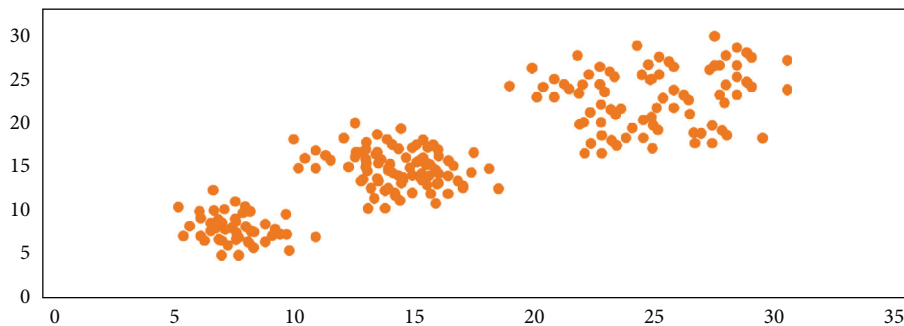

FIGURE 6: Clustering results of teachers' teaching status after removing outliers in a semester in Y school.

there have been slight fluctuations in each semester in the past five years, the first category and the second category show a downward trend from semester to semester, while the third category shows an obvious upward trend from semester to semester, which indicates that the overall teaching situation of the whole school presents a clear downward trend, and the teaching management department should do it.
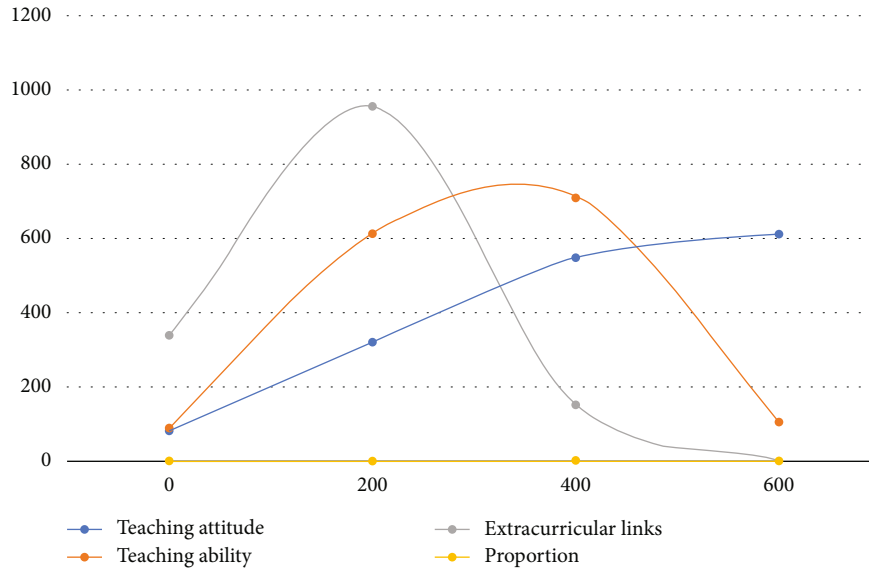
FIGURE 7: Statistical table of students' teaching evaluation in a certain semester of Y school.
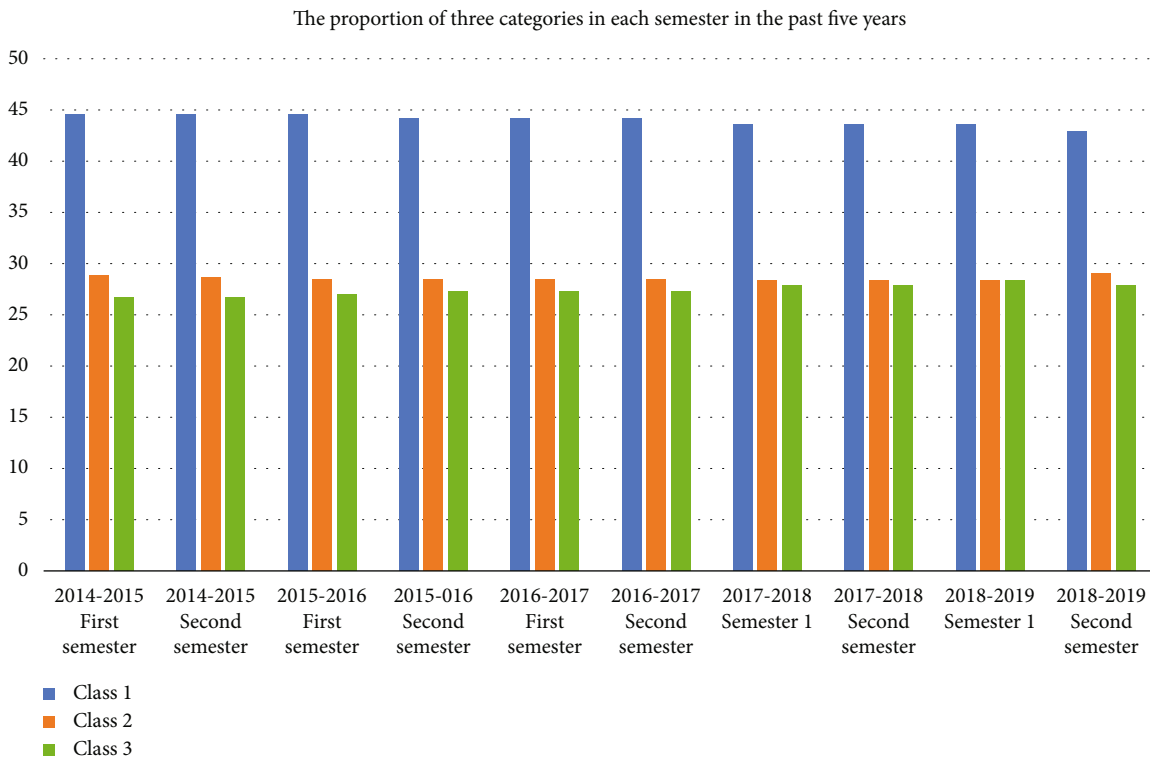

FIGURE 8: Histogram of the proportion of teachers in three categories in each semester of Y school in the past five years.

In the evaluation of the "extracurricular link" evaluation indicators, less than 0.7% were rated as "excellent", less than 10% were rated as "good", and less than 10% rated as "average" The ratio is about 64%; about 20% is rated as "pass", and about 5% is rated as "fail". It can be seen that students' evaluation of this indicator is the lowest. It also reflects that students are very dissatisfied with this part. The problem may be in two aspects, one is that students have higher requirements for teachers to participate in extracurricular guidance; the other is that there is a problem in the management of this teaching link; therefore, the teaching management department urgently needs to make great efforts in this link.

With the increase of the number of users, the difference between users is gradually expanding. In this paper, we select the classical user behavior impact indicators, and then use k-means analysis method to cluster the existing historical data. Applying K-means clustering analysis one by one, on the

The distribution of the absolute error between the actual value and the predicted value of the two groups of forecast data
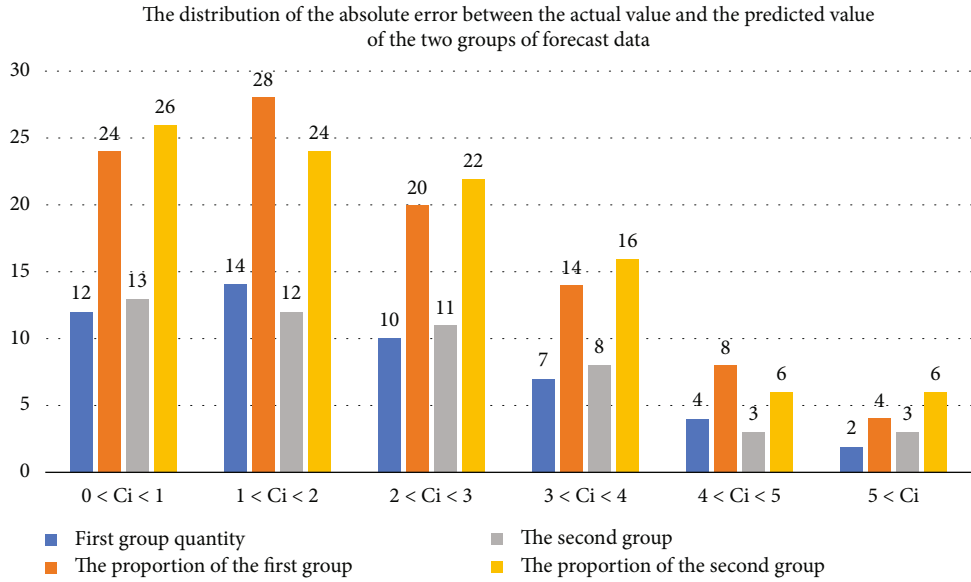

FIGURE 9: Distribution of the absolute error between the actual value and the predicted value of the two groups of forecast data.
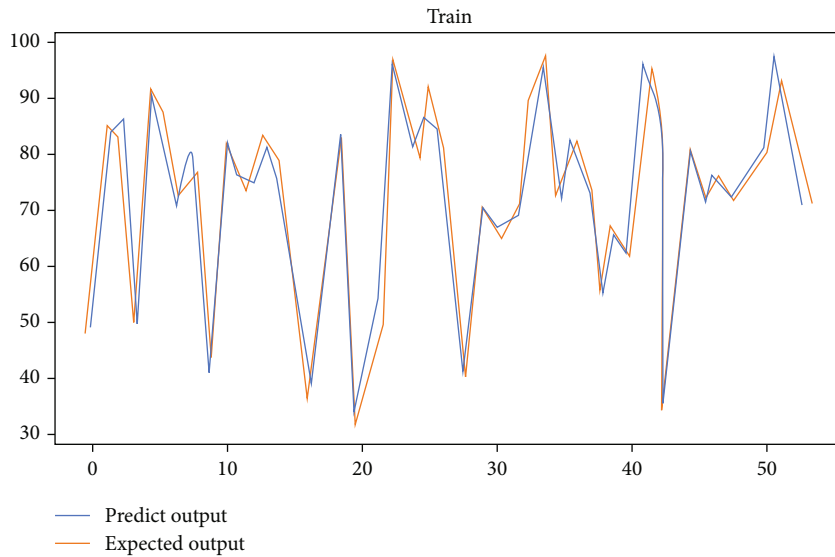

FIGURE 10: Comparison of training results with the ground truth.

one hand, we can identify singular points in the data, but in fact, we attach great importance to users. On the other hand, it makes the degree of user classification controllable and the level of user classification clear. The attributes and behaviors of users of the same kind are relatively consistent. It is convenient for enterprises to reasonably classify users, so as to provide accurate services for users, so as to achieve a win-win situation for enterprises and users.

4.2. Application Research of Neural Network Model in College Students' Learning Prediction.

Classroom teaching is the main channel for talent training for the quality of talent training. E-commerce live video courses are a new education model. In this section, I will discuss the behaviors of donkey students during online learning through this model. Data, using neural network model to predict the learning effect, so as to provide support for students to adjust their learning behavior in a timely manner and teachers to provide targeted teaching.

Based on the above neural network model, by training 150 training samples for 15,000 times, the output of the training result is shown in Figure 9. The training output value basically coincides with the real image, indicating that the model is effective. Figure 10 shows the comparison between the training results and the ground situation.

It can be seen that the average error is within 1.73; about two-thirds of the output value is smaller than the actual value and about one-third of the output value.

For the students' course learning, this section analyzes the model's response to the training data "same class with the same teacher" and "different class with different teachers" in two groups of 50 students each. Based on the

prediction of learning effect, the comparison of the prediction data with the model established by traditional regression analysis further shows that the model is more powerful in predicting the learning effect of students. Its histogram is shown in Figure 9.

5. Conclusion

This paper studies the evaluation method of e-commerce video live innovation education mode and draws the following conclusions: according to the k-mode clustering results, for the whole school, the overall teaching situation of unit teachers can provide policy implementation. Teachers can timely learn about their own and other teachers' teaching conditions and classroom teaching for talent training. Based on the behavior data generated for students, the neural network model predicts the online learning effect of students, and provides targeted basis for adjusting their learning behavior and teachers. Today, with the rapid development of online education, online learning is an indispensable part of talent training in the new era, especially for the training of various professionals in the new engineering field. Guide teaching by improving the construction of the evaluation system and formulate corresponding teaching strategies according to the reality of each student. Truly create a diversified, multilevel and multiangle teaching evaluation model based on big data. Truly let the teaching evaluation objectively show the actual situation of students, and really let the teaching evaluation promotes the all-round development of students.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J. Fu, Y. Zhang, Y. Wang et al., "Optimization of metabolomic data processing using NOREVA," *Nature Protocols*, vol. 17, no. 1, pp. 129–151, 2022.

[2] N. P. Jayasri and R. Aruna, "Big data analytics in health care by data mining and classification techniques," *ICT Express*, vol. 8, no. 2, pp. 250–257, 2022.

[3] A. Rosário and R. Raimundo, "Consumer marketing strategy and E-commerce in the last decade: a literature review," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 16, no. 7, pp. 3003–3024, 2021.

[4] H. Yang and W. Zhang, "Data mining in college student education management information system," *International Journal of Embedded Systems*, vol. 15, no. 3, pp. 279–287, 2022.

[5] C. M. Montalcini, B. Voelkl, Y. Gómez, M. Gantner, and M. J. Toscano, "Evaluation of an active LF tracking system and data processing methods for livestock precision farming in the poultry sector," *Sensors*, vol. 22, no. 2, p. 659, 2022.

[6] V. Grover, R. H. L. Chiang, T. P. Liang, and D. Zhang, "Creating strategic business value from big data analytics: a research framework," *Journal of Management Information Systems*, vol. 35, no. 2, pp. 388–423, 2018.

[7] D. Wang, J. Wu, L. Deng, Z. Li, and Y. Wang, "A real-time optimization control method for coagulation process during drinking water treatment," *Nonlinear Dynamics*, vol. 105, no. 4, pp. 3271–3283, 2021.

[8] K. O. Iwuozor, "Prospects and challenges of using coagulation-flocculation method in the treatment of effluents," *Advanced Journal of Chemistry-Section A*, vol. 2, no. 2, pp. 105–127, 2019.

[9] P. Garikapati, K. Balamurugan, T. P. Latchoumi, and R. Malkapuram, "A cluster-profile comparative study on machining AlSi7/63% of SiC hybrid composite using agglomerative hierarchical clustering and K-means," *Silicon*, vol. 13, no. 4, pp. 961–972, 2021.

[10] I. G. N. M. Jaya and H. Folmer, "Identifying spatiotemporal clusters by means of agglomerative hierarchical clustering and Bayesian regression analysis with spatiotemporally varying coefficients: methodology and application to dengue disease in Bandung, Indonesia," *Geographical Analysis*, vol. 53, no. 4, pp. 767–817, 2021.

[11] M. Y. C. Yim, S. C. Chu, and P. L. Sauer, "Is augmented reality technology an effective tool for e-commerce? An interactivity and vividness perspective," *Journal of Interactive Marketing*, vol. 39, no. 1, pp. 89–103, 2017.

[12] T. Tang, Y. Wu, Y. Wu, L. Yu, and Y. Li, "Videomoderator: a risk-aware framework for multimodal video moderation in E-commerce," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, pp. 846–856, 2021.

[13] J. Shuja, M. A. Humayun, W. Alasmary, H. Sinky, E. Alanazi, and M. K. Khan, "Resource efficient geo-textual hierarchical clustering framework for social IoT applications," *IEEE Sensors Journal*, vol. 21, no. 22, pp. 25114–25122, 2021.

[14] Y. Deng, H. Levine, X. Mao, and L. M. Sander, "Collective motility and mechanical waves in cell clusters," *The European Physical Journal E*, vol. 44, no. 11, pp. 1–15, 2021.

[15] Q. Hou, M. Han, and Z. Cai, "Survey on data analysis in social media: a practical application aspect," *Big Data Mining and Analytics*, vol. 3, no. 4, pp. 259–279, 2020.

[16] Y. Bae, J. Choi, M. Gantumur, and N. Kim, "Technology-based strategies for online secondhand platforms promoting sustainable retailing," *Sustainability*, vol. 14, no. 6, p. 3259, 2022.

[17] T. Chen, L. Peng, B. Jing, C. Wu, J. Yang, and G. Cong, "The impact of the COVID-19 pandemic on user experience with online education platforms in China," *Sustainability*, vol. 12, no. 18, p. 7329, 2020.

[18] M. Sandanayake, R. Kumanayake, and A. Peiris, "Environmental impact assessments during construction stage at different geographic levels–a cradle-to-gate analysis of using sustainable concrete materials," *Engineering, Construction and Architectural Management*, vol. 29, no. 4, pp. 1731–1752, 2021.

[19] S. Jung, J. Jeoung, H. Kang, and T. Hong, "3D convolutional neural network-based one-stage model for real-time action detection in video of construction equipment," *Computer-Aided Civil and Infrastructure Engineering*, vol. 37, no. 1, pp. 126–142, 2022.