Research Article

Construction of Flipped Classroom for College English Courses Using Big Data MOOCs and Information System

Mei Chen

Nanchong Vocational and Technical College, Nanchong 637000, China

Correspondence should be addressed to Mei Chen; chenmei163163@163.com

Received 8 June 2022; Revised 9 July 2022; Accepted 11 July 2022; Published 9 August 2022

Academic Editor: Liping Zhang

Copyright © 2022 Mei Chen. 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.

Big data refers to extremely large, sophisticated, and varied informational components. Better methods for controlling dynamic abilities, experiences, and interaction development are necessary. With the help of Massive Open Online Courses’ big data, this essay will examine how to research and analyse the flipped classroom in college English courses (MOOCs). Also covered is data mining. This paper addresses the data mining-based issue of big data in MOOCs. The idea of big data and associated algorithms is then further developed in this paper. In this essay, the flipped classroom for college English courses based on MOOCs is designed and examined. According to the experimental findings, 44.78 percent of students believe the MOOC learning platform is helpful after the implementation of the university English flipped classroom based on big data from MOOCs, and 26.96 percent of students find it extremely helpful. It is clear that the big data college English courses taught in a flipped classroom using MOOCs have had some educational impact.

1. Introduction

With the advancement of English majors in schools and colleges, more English majors and non-unknown dialect majors offer different courses. Educating English is progressively being attached great attention. Big data uses effective data mining techniques to change the way people work and acquire knowledge through visualization and predictive analytics, which in turn changes the way people communicate. With the increasing integration of educational concepts and information technology, the educational model of MOOC came into being. It encourages the digitization of education and fully utilises the technical benefits of informatics and big data to raise the standard and fairness of instruction. It actually makes it possible for the entire population to share excellent educational resources. It has gradually turned into an unavoidable trend in educational advancement.

With the advent of MOOCs in the web 2.0 era, English instruction in colleges and universities has been transformed since the turn of the century. MOOC stands for “massive open online course”. The characteristics of teaching languages are all present in real-time interactive practise between students, teachers, and students. The goal of MOOCs is to share the ways that Eastern and Western languages think. It presents a good chance to implement the integration of interdisciplinary knowledge in English teaching and support the enhancement of teaching standards.

The following is where this paper innovates: (1) this paper combines big data with MOOC flipped classroom and provides a detailed introduction to the methods associated with data mining and the MOOC theory. The information gain calculation, support vector machine, and BP neural network algorithm are the main introductions in this paper. (2) This paper designs the teaching classroom in response to the flipped college English classroom. This paper comes to the conclusion that the big data MOOCs-based flipped classroom for college English courses has had a particular impact through the evaluation of the experimental results.

2. Related Work

The rapid advancement of computer and data innovation, which has permeated all facets of human life, is bringing about the era of massive information. In order to reduce the
amount of data that the Internet of Things was gathering, Xue et al. [1] sped up big data processing. The proposed multi-objective molecule swarm improvement hereditary calculation (MOPSSGA), when compared to the standard GPSR-BB calculation, reduces the amount of emphasis by 51.6%, according to the application results. Its recreation power is 0.15 over the calculated value. Reconstructions are carried out more successfully. But its performance is worse [1]. Kuang et al. [2] proposed a single tensor model for the representation of unstructured, semistructured, and structured data. According to a fictitious analysis and test results, the proposed combined tensor model and IHOSVD technique are effective for large information portrayal and dimensionality reduction. However, he does not have very good accuracy [2]. In order to examine these innovations’ fundamental properties, Stergiou and Psannis [3] combined large information advancements with the Internet of Things (IoT) and Mobile Cloud Computing (MCC) innovations. It also points out the benefits of MCC and IoT that can aid in the creation of expansive informational applications. But his data are not sufficiently reliable [3]. Xu et al. [4] investigated different methods for protecting private data and looked at data mining privacy issues from a broader perspective. He briefly reviews the fundamentals of related research topics while reviewing cutting-edge methodologies. He also offered some preliminary ideas for potential lines of inquiry. However, his writing is incredibly subjective [4]. A patient-driven digital actual framework for clinical applications and administrations was proposed by Zhang et al. [5] in light of developments in cloud computing and large-scale information analysis. The results of this study show how large data and cloud innovations can be used to enhance clinical framework presentation and enable people to utilised a variety of top-notch clinical applications and services. But his approach is more time-consuming. Reference [5], Janssen et al. [6] identified the factors influencing decisions made with the help of case studies. The case study showed that his use of big data is a dynamic process. It is crucial to standardise procedures and to comprehend big data’s potential step by step. However, he lacks practicality [6]. The big data machine learning framework was introduced by Zhou et al. [7] to guide its opportunities and difficult conversations. Using the various ML stages and MLBiD components helps identify significant open doors and challenges. Future research in many neglected or understudied fields is now possible thanks to it. However, his introduction is biased. Reference [7], Zhang et al. [8] proposed a product life cycle architecture based on big data analysis. It facilitates access to information and data about the product life cycle. This architecture can benefit manufacturers, environments, product designs, and services as a theoretical basis for the execution of the product life cycle. However, its range of application is restricted [8].

3. Integration Method of Big Data and MOOC

3.1. Related Concepts

3.1.1. MOOCs. MOOCs mean massive online open courses. The first letter M (Massive): one refers to a wide range of registered users and the other refers to a wide variety of course resources. The second letter O (Open) refers to open learning resources and learning spaces and students guided by interests. It chooses whether to register for the study. The third letter O (Online) represents teacher teaching, student learning, completing and submitting assignments, grading and testing homework assignments, teacher–student/student discussions, etc. The fourth letter C (Courses courses) refers to instructional videos, textbooks, homework and advanced exams, etc. Large-scale, open, unstructured, and autonomous are the characteristics of MOOCs [9].

3.1.2. Flipped Classroom. The knowledge transfer and internalisation processes are turned around in the “flipped classroom”. Its characteristics are as follows: First, the teaching process is reversed. Students must complete the classification and self-study of knowledge points in advance. Second, the teaching content is rich and colourful. In the flipped classroom, students no longer passively listen and take notes, but after internalizing their knowledge, they conduct research, discussions or answer questions with teachers or classmates. In addition, the roles of teachers and students have also changed [10]. In the flipped classroom, teachers are no longer the protagonists of the traditional classroom. They are no longer unified teaching and control but actively organize and plan teaching activities. It makes students no longer accept passively, but take the initiative to become real classroom thinkers.

3.1.3. Big Data

(1) Concept Value. Big data is often thought to have four components: volume (quantity), velocity (velocity), variety (variety), and value (value) (value). A valid and precise definition of big data is virtually unachievable at the current stage of development. A method is typically required when a new technical proposal is suggested [11, 12]. Big data’s conceptual meaning and application are depicted in Figure 1.

First, volume (number): The sheer volume of data, which is significant, is the feature that most distinguishes big data from traditional data. This trait is intrinsically tied to technological advancement. First off, the amount of data storage available in the past was extremely constrained. Large-scale data storage is now possible thanks to Moore’s Law, which describes how hardware performance has steadily increased and the price has gradually decreased. The proliferation of technologies like social networking, e-commerce, and the Internet of Things has also resulted in an oversupply of data.

Second, velocity: it is the quantity of data streams that are coming in at an ever-increasing rate and need to be processed in a reasonable amount of time. The main issue with big data is this. Traditional data processing and storage techniques simply cannot achieve the required efficiency.

Third, variety: Given the abundance of information sources, it alludes to a variety of informational subtypes. One of the main issues with huge amounts of information is the capacity and investigation of this information.
(2) Big Data Environment. Big data serves as a catalyst to advance educational digitization. MOOC will become a cutting edge in the field of education in the big information world. Through the use of information mining innovation, MOOC has gathered a significant amount of guiding data. It gathers essential data and evaluates the issues that students face during the learning process. It offers accurate learning and exam information results to pupils. Through the foundation, a lot of frameworks and stages have so far gathered a lot of data about the learning process. It filters the test results that are displayed and assesses the effects of using explicit models in the classroom [13, 14]. Figure 2 displays the study’s examination of the big data environment.

The classroom paradigm, which is supported by cutting-edge technology, displays educational traits like dataization, intelligence, function, and efficiency. Big data-related classroom activities are conducted in a big data MOOC environment. The information produced by the teachers and students during the interactive teaching activities is mined and accumulated during the entire teaching process. In addition, it sends large data to the cloud for integrated analysis, and the educational data produced by mining operations is a crucial component of measuring the effects of the classroom [15].

3.2. Related Technologies of Big Data

3.2.1. Big Data Storage. Information was initially only kept in documents when it first started to be saved. Document stockpiling has disadvantages, as is obvious. The only demanding tasks that can support it are inquiry, embed, and delete. It requires designed data. Information can now be stored thanks to frameworks for distributed documents. These are the advantages of distribution: It can accommodate a lot of data. With the help of its distributed file system, represented by HDFS, a sizable file can be stored across numerous computers. Each device can store a piece of the file, relieving demand on a single machine. Backup for redundant data is supported. HDFS saves three copies of each piece of data and distributes them among several machines to ensure that no data is lost even if a machine in the cluster is taken offline or destroyed totally. It has strong scalability because as the amount of data progressively increases, adding more cluster machines will enhance the data storage capacity. Three components are crucial to HDFS: DataNode, NameNode, and Cie. The NameNode receives the information about the data block from the DataNode [16, 17]. The organisational structure of the distributed file system is shown in Figure 3.

3.2.2. Data Mining. Data mining is a multi-stage process. It is not only a modelling process but also a step-by-step data mining process. A complete data mining process includes six steps of problem formulation and understanding, data understanding, data preparation and preprocessing, model establishment, model evaluation and optimization, and program implementation [18], as shown in Figure 4.

3.3. Data Mining Algorithms. Data mining incorporates an assortment of examination techniques to mine and investigate informational indexes, which gets designs and apply them. Among them, grouping possesses a spot, and arrangement strategies are additionally notable. The most effective method to appropriately group the information will straightforwardly influence the exactness and standard proficiency of mining results [19].

3.3.1. Information Gain Calculation. There was only a blank decision tree in the early days of machine learning; there was no idea of how to split existence depending on features. The recently learned decision tree model is used to classify the full feature space. It states that class X is designated as L in the training set and is represented by samples from class a, |L| represents the total number of cases L in the training set, and the probability \( P(Z_a) \) of unknown instances belonging to class a is defined as:
Figure 3: HDFS architecture.

Figure 4: Data mining flowchart.
Among them, the penalty parameter $C > 0$, and the most $E$ solution $\alpha^* = (\alpha^*_1, \alpha^*_2, \ldots, \alpha^*_N)^T$.

The inner product between instances can be transformed into a kernel function for nonlinear classification. After a nonlinear transformation, the inner product between two instances is the kernel function $Z(x, k)$. It denotes the existence of a mapping between input space and feature space $\phi(x)$. For $x, k$ in any input space, we have

$$Z(x, k) = \phi(x) \cdot \phi(k).$$

Because the technique is good at handling high-dimensional data, support vector machines are frequently utilised in neuroscience and bioinformatics. Its primary drawbacks are the extensive calculation and simple overfitting.

3.3.3. BP Neural Network Algorithm. Information forward propagation and error backpropagation are two processes that the BP neural network algorithm uses. Weight initialization: Each node in the neural network has an associated bias, and the connection weights between those nodes are initially initialised to a small random number. The Forward Propagation. The neural network's input layer receives the training samples, and the value of this layer is kept constant. For the input node $j$, its output value $S_j$ is equal to the input value $E_j$, as shown in Figure 6.

The outlet of the node in the layer above connects all of the inputs from each node, which each has a number of inputs. Weights are attached to each connection. The $b$-net node's input $E_b$ is the following if it is in the exit layer or hidden layer:

$$E_b = \sum_a w_{ab} S_a + \theta_b.$$  

This function depicts the activity of the neuron that this node symbolises using logistic or sigmoid functions. Given the net input $E_b$ of node $b$, the output $S_b$ of node $b$ is:

$$S_b = \frac{1}{1 + e^{-\Delta_b}}.$$  

Backpropagation error: with the continuous updating of the burden and bias representing the network prediction error, the error will propagate backward. For the node of the output layer $b$, the error $\text{Err}_b$ is calculated as follows:

$$\text{Err}_b = S_b (1 - S_b) (G_b - S_b).$$

$S_b$ is the actual output of node $b$, and $G_b$ is the known target value of node $b$ based on a given training sample. In fact, $S_b (1 - S_b)$ is actually the derivative of the logistic function.

It is necessary to balance the mistakes of nodes connected to node $b$ on the next layer in order to calculate the error of node $b$ on the hidden layer. Node $b$'s hidden layer mistake is as follows:

$$\text{Err}_b = S_b (1 - S_b) \sum_k \text{Err}_k w_{kb},$$

$$G(Z_a) = \left| \frac{Z_a}{L} \right|.$$  

At this point, the partitioned $Z$ uncertainty measure is:

$$H(L, Z) = - \sum_i F(Z_i) \log F(Z_i).$$  

If the test attribute $p$ is used for testing, when $p = p_b$, the samples belonging to class $a$ can be regarded as $Z_{ab}$, then there are:

$$F(Z_a; p = p_b) = \left| \frac{Z_{ab}}{L} \right|.$$  

$$H(L_b) = \sum_a F(Z_a | p_b) \log F(Z_a | p = p_b).$$

When all $p = p_b$, protruding after test attribute $P$ is selected, the information entropy of branch $I$ for classification information is:

$$H(L, p) = \sum_b F(p = p_b) H(I_b).$$

The information gain $I(L; p)$ provided by the attribute $p$ for classification is:

$$I(L; p) = H(L) - H(L | p).$$

3.3.2. Support Vector Machine. The support vector machine method is proposed according to the optimal classification surface under the condition of linear separability. Based on such a classification hyperplane, SVMs can not only correctly classify all training samples, but also make the points in the training samples closer to the grading surface to the longest distance from the grading surface [20], as shown in Figure 5.

Solving the optimal separating hyperplane is the basis of a support vector machine; this paper shows that the training data can be divided correctly and the geometric spacing is maximized. For linearly separable data sets, the basic idea and the formalized convex quadratic programming problem are shown in formulas (6) and (7).

$$\max_{\alpha} \chi,$$

s.t.$\chi \left( \frac{w}{||w||} \cdot x_i + b + \frac{b}{||w||} \right) \geq \chi, \quad i = 1, 2, \ldots, N.$$

$$\chi; x_i, y_i \left( (w/||w||) \cdot x_i + (b/||w||) \right)$$ defines the geometric space of the hyperplane relative to the sampling point $(x_i, y_i)$, and $\chi = \min_{i=1,2,\ldots,N} \chi_i$ is the minimum value of the geometric space of the hyperplane at all sampling points.

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i x_i \cdot x_j - \sum_{i=1}^{N} \alpha_i,$$

s.t.$\sum_{i=1}^{N} \alpha_i y_i = 0 \quad (0 \leq \alpha_i \leq C, i = 1, 2, \ldots, N).$
wbk is the connection weight from node k to node b in the next higher layer, and Errk is the error of node k.

The formula for weight update is as follows, \( \Delta w_{ab} \) is the change amount of weight \( w_{ab} \).

\[
\Delta w_{ab} = (1)Err_bS_a, \\
w_{ab} = w_{ab} + \Delta w_{ab}.
\] (14)

The offset is updated by the following formula, \( \Delta \theta_b \) is the amount of change in the offset \( \theta_b \).

\[
\Delta \theta_b = (1)Err_b, \\
\theta_b = \theta_b + \Delta \theta_b.
\] (15)

By processing tuples, the aforementioned method updates existing biases and weights. In actuality, however, bias and weight increases can accumulate in variables and, after all tuples have been processed, bias and weight updates can affect training samples. The period represents a training sample iteration, and the method is a continuously updated method.

4. Flipped Classroom Experiment and Big Data College English Course Based on MOOC

4.1. Design of MOOC Flipped Classroom Teaching Mode.

Flipped study hall is a method of instruction that is shaped by varying the amount of time spent inside and outside of the classroom. It transfers dynamic power from teachers to students. It mixes traditional teaching techniques with online learning platforms to enhance each other. It contains a blend of various learning environments, learning methodologies, and learning theories. Whatever
the mixing method, the end goal is to realise the natural blending of teachers and pupils. As depicted in Figure 7, this work develops a MOOC to create a flipped classroom teaching strategy based on its design concepts. This was done in accordance with the pertinent theories and the analysis of the characteristics of subversion based on MOOC.

The preclass stage is mainly for students to conduct autonomous learning on the MOOC platform. It belongs to the stage of knowledge imparting. At this stage, this paper designs different learning activities for teachers and students, as shown in Figure 8(a).

Face-to-face classroom teaching is an essential and basic link in blended teaching. In the classroom learning stage, not only students must systematically acquire curriculum knowledge, but also develop their multi-faceted learning ability for specific mid-term learning activities, as shown in Figure 8(b).

The testing phase of the student’s entire learning process occurs after school. The teacher’s assigned homework should be carefully completed by the students, along with a focused review. At the same time, they can share learning experiences and broaden their learning concepts by conversing with teachers and students online. Students should also summarise their achievements and feedback problems to teachers or team members in time, so as to solve problems in time, reflect, and correct their learning.

4.2. Teaching Implementation. This paper conducts a preliminary questionnaire survey of students’ online learning before class for the case implementation of this study in order to understand what is happening with students’ Internet learning. Students from a variety of academic majors at a university make up the survey’s subjects. There are 230 individuals total among them, with 102 boys making up 44.35 percent of the group and 128 girls making up 55.65 percent. They have a range of academic backgrounds, such as those in education, computing, design, and other fields. Making the data obtained authentic and universal is the goal. 230 valid questionnaires were ultimately located, and the recovery rate was 100%.

An overview of the basic situation of students’ Internet implementation is a basic review of students’ requirements for web-based learning. Thirst for knowledge is the internal main thrust of advanced demand. Only the thirst for knowledge can stimulate students’ advanced requirements, which are the basic conditions of the era of learning inspiration. The specific analysis is shown in Table 1:

According to the table, for the overview of students explaining web-based learning, boys’ normal understanding of Internet learning is 3.96, and the average understanding of online learning for girls is 3.98. The average value shows that the online learning situation of the learners in the case-implemented class is average, and the average value is 3.97. As to study on “individuals think Internet educating in

![Figure 7: The flipped classroom teaching mode of college English courses based on MOOCs.](https://example.com/figure7)
schools and colleges is fundamental”, it very well may be seen from Table 1 that there is not much difference between male and female students. Boys are 4.02, girls are 4.00, and the average is 4.01. Since the gap between the two is not large, we will no longer analyse it from the perspective of gender.

Regarding the teaching methods that learners prefer, students prefer the mixed teaching method of “classroom teaching combined with network teaching”. The second is the classroom teaching method, and the last is the simple network teaching method; in general, 195 students prefer the blended teaching method of “classroom teaching combined with online teaching”, accounting for 84.78%. The number of people who like the classroom teaching method is 27, accounting for 11.74%. There are 8 people who simply use the online teaching method, accounting for 3.49%, as shown in Figure 9(a).

Regarding “If people choose an online course, which of the following criteria would people choose?”, a multiple-choice question is used, as shown in Figure 9(b). According to the analysis of statistical results, this paper finds that if students choose online courses, the primary criterion is what they are interested in, accounting for 90.43%. The second is courses taught by well-known professors, accounting for 75.65%, and the third is courses related to professional knowledge, accounting for 70.43%. The proportion of the number of people who choose these three options is the highest. This also shows that, in the selection of MOOC course content, educators should focus on the instructing of popular instructors and the interests of undergraduates.

As shown in Table 2, the design and implementation of MOOC blended learning activities primarily include three learning stages: prior to class, during class, and after class.

### 4.3 Effect Evaluation

The questionnaire is aimed at learners who participate in the blended learning of MOOCs. From the perspective of students, they can understand the effect of this blended learning practice based on MOOCs. Since the basic information of the students was already available in the questionnaire survey in the early stage of the course, this questionnaire will no longer be included in the basic information survey of learners.

The result of the question “What do people think is the main role of the MOOC learning platform in this MOOC-based blended learning?” results are shown in Figure 10(a). The topic “People’s satisfaction with the learning effect of MOOC flipped classroom?” is shown in Figure 10(b).

As can be seen from Figure 10(a), the MOOC stage is basically utilised for undergraduates to help homeroom
Figure 9: Early part of the survey. (a) Which teaching method do people prefer? (b) Criteria for selecting online courses.

Table 2: Design of MOOC flipped classroom learning activities.

| Learning phase activities | Guidance strategies | Cognitive tools | Evaluation tool | Technical means |
|---------------------------|---------------------|-----------------|-----------------|-----------------|
| Before class              | Distribute study guides and study materials, introduce the MOOC platform, and formulate study plans | Draw knowledge maps, communicate online, and answer questions | Prelearning test, MOOC platform test | MOOC platform, WeChat, QQ group, community forum |
| In class                  | Difficult and difficult explanations, personalized guidance | Micro-class mind map drawing, team learning | Class performance, teacher–student communication, feedback, peer evaluation | Multimedia teacher, PPT |
| After class               | Summarize           | Micro-lesson production | Results report, project work | WeChat, QQ group, community forum |
learning and lead to free learning, while less understudies decide to download learning materials and speak with peers. It shows that understudies’ familiarity with Internet learning is not high, and their energy for online connection is low. As can be seen from Figure 10(b), 103 people felt that the MOOC learning platform was helpful, accounting for 44.78%, and 62 people felt it was very helpful, accounting for 26.96%. It tends to be seen that the flipped classroom of big data college English courses based on MOOCs has played a certain effect.

5. Discussion

Prior to analysing how to use big data from MOOCs to perform research on the flipped classroom in college English courses, this paper first acquires the necessary foundational knowledge through the examination of essential knowledge points in literary works. This paper explains the data mining principles and algorithms. This study examines information gain calculation, support vector machines, and BP neural network techniques. This study investigates the use of big
data from MOOCs in flipped classrooms for undergraduate English courses.

The development of information technology and the rise of the Internet have had an impact on how people learn. MOOCs are rapidly expanding abroad for advanced education. In the first MOOC, there were over 2000 enrolled students, and there are now countless students in the current massive MOOC stage. MOOCs have reached an “unstable stage” of development in just over five years. Since their founding in 2008, international MOOCs have become a significant component of academic instruction. In China, with the progression of online courses and information development, e-learning has turned into a wave, and this is only the start, an ever-increasing number of colleges are joining or making their own MOOC field. The high-level requirements of substitute understudies are more customized, empowering substitute understudies to assimilate new data. This approach is presently not just through school instruction plans or manuals. For current understudies, the electronic learning stage is not just a significant wellspring of their learning, it is one of the significant apparatuses for independent learning.

This paper designs the English flipped classroom in the MOOC big data environment in detail, as demonstrated by the experimental analysis. It uses the model to implement particular classroom instruction. The results demonstrate that the model is effectively applied through the gathered lessons, insightful opinions, and research reviews. This function has sparked changes in teaching strategies and the growth of educational technology, but there is still room for improvement. Such a teaching method can also be improved and optimised after use. It offers a more efficient and clever method of instructing students in classrooms for learning and teaching.

6. Conclusion

College English has been impacted by big data, which has also presented opportunities. Applying new teaching methodologies to track content-based English teaching is becoming more and more important in the age of big data. This is not like the regular general English classes. Additionally, the development of an online learning environment results in students having more varied learning needs. The approach that combines the benefits of an online classroom with a traditional classroom is the one that students are most receptive to and expect. This paper uses a hybrid teaching approach that combines “MOOC” and “flipped classroom” to achieve the ideal teaching effect when dealing with the teaching conundrum in the English improvement stage and the actual needs of students in order to complete the English monitoring course. It has a lot of practical significance and meets the expectations that teachers and students share.

**Data Availability**

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

**Conflicts of Interest**

The author declares that there are no conflicts of interest.

**References**

[1] J. W. Xue, X. K. Xu, and F. Zhang, “Big data dynamic compressive sensing system architecture and optimization algorithm for internet of things,” *Discrete and Continuous Dynamical Systems - Series S*, vol. 8, no. 6, pp. 1401–1414, 2017.

[2] L. Kuang, F. Hao, L. T. Yang, M. Lin, C. Luo, and G. Min, “A tensor-based approach for big data representation and dimensionality reduction,” *IEEE Transactions on Emerging Topics in Computing*, vol. 2, no. 3, pp. 280–291, 2014.

[3] C. Stergiou and K. E. Psannis, “Recent advances delivered by mobile cloud computing and internet of things for big data applications: a survey,” *International Journal of Network Management*, vol. 27, no. 3, 2017.

[4] L. Xu, C. Jiang, J. Wang, J. Yuan, and Y. Ren, “Information security in big data: privacy and data mining,” *IEEE Access*, vol. 2, no. 2, pp. 1149–1176, 2017.

[5] Y. Zhang, M. Qiu, C. W. Tsai, M. M. Hassan, and A. Alamri, “Health-CPS: healthcare cyber-physical system Assisted by cloud and big data,” *IEEE Systems Journal*, vol. 11, no. 1, pp. 88–95, 2017.

[6] M. Janssen, H. van der Voort, and A. Wahyudi, “Factors influencing big data decision-making quality,” *Journal of Business Research*, vol. 70, no. JAN, pp. 338–345, 2017.

[7] L. Zhou, S. Fan, J. Wang, and A. V. Vasilakos, “Machine learning on big data: opportunities and challenges,” *Neurocomputing*, vol. 237, pp. 350–361, 2017.

[8] Y. Zhang, S. Ren, Y. Liu, and S. Si, “A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products,” *Journal of Cleaner Production*, vol. 142, pp. 626–641, 2017.

[9] W. Yanxia and W. Qian, “Analysis of the fine management strategy for the cultivation of vocational college students’ core quality based on big data,” *Journal of Physics: Conference Series*, vol. 1881, no. 4, Article ID 042056, 2021.

[10] Y. Wang, L. A. Kung, and T. A. Byrd, “Big data analytics: understanding its capabilities and potential benefits for healthcare organizations,” *Technological Forecasting and Social Change*, vol. 126, pp. 3–13, 2018.

[11] J. Li and P. Wei, “Three-dimensional landscape rendering and landscape spatial distribution of traditional villages based on big data information system,” *Mobile Information Systems*, vol. 1, Article ID 4945918, 2022.

[12] M. M. U. Rathore, A. Paul, A. Ahmad, B. W. Chen, B. Huang, and W. Ji, “Real-time big data analytical architecture for remote sensing application,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 10, pp. 4610–4621, 2015.

[13] W. Xu, H. Zhou, N. Cheng et al., “Internet of vehicles in big data era,” *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 1, pp. 19–35, 2018.

[14] H. Cai, B. Xu, L. Jiang, and A. V. Vasilakos, “IoT-based big data storage systems in cloud computing: perspectives and challenges,” *IEEE Internet of Things Journal*, vol. 4, no. 1, pp. 75–87, 2017.

[15] A. Barbu, Y. She, L. Ding, and G. Gramajo, “Feature selection with annealing for computer vision and big data learning,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 2, pp. 272–286, 2017.
[16] I. H. Chung, T. N. Sainath, B. Ramabhadran et al., “Parallel deep neural network training for big data on blue gene/Q,” IEEE Transactions on Parallel and Distributed Systems, vol. 28, no. 6, pp. 1703–1714, 2017.

[17] G. Zhuo, Q. Jia, L. Guo, M. Li, and P. Li, “Privacy-preserving verifiable set operation in big data for cloud-assisted mobile crowdsourcing,” IEEE Internet of Things Journal, vol. 4, no. 2, pp. 572–582, 2017.

[18] W. A. Guenther, M. Mehrizi, M. Huysman, and F. Feldberg, “Debating big data: a literature review on realizing value from big data,” The Journal of Strategic Information Systems, vol. 26, no. 3, pp. 191–209, 2017.

[19] L. A. Tawalbeh, R. Mehmood, E. Benkhifa, and H. Song, “Mobile cloud computing model and big data analysis for healthcare applications,” IEEE Access, vol. 4, no. 99, pp. 6171–6180, 2016.

[20] Z. Lv, H. Song, P. Basanta-Val, A. Steed, and M. Jo, “Next-generation big data analytics: state of the art, challenges, and future research topics,” IEEE Transactions on Industrial Informatics, vol. 13, no. 4, pp. 1891–1899, 2017.