1,*Yuanyuan Qian

A Collaborative Study of English Teaching Based on Optimized Apriori Algorithm under the Integration of Curriculum Civics

Abstract: - The new period introduces new intellectual and political demands for college-level English courses and fosters favourable conditions as well as new intellectual and political opportunities. The knowledge and politics courses currently taught in our higher education institutions combine English instruction with cooperative research and knowledge and politics courses, producing some excellent results in China. However, there are still significant challenges to implementing the relatively new ideas for education reform. In order to increase the effectiveness of the mine, this paper will enhance the apriori calculation method in accordance with the features of a large amount of data, and it requests the English Education Graduate School's participation for a demonstration. To address the issue of inaccurate existing metrics, this document integrates the proposed measures with conventional measures. In the past, the new measuring technique and the enhanced algorithm were used to create a model of the mining system based on the performance of the pupils. In the end, we used the student successes to build the mine system model, integrate the new measurement technique with the improved algorithm, and draw significant data-driven conclusions. According on experimental findings, the proposed algorithm performs better than Mr. Apriori. The mixed metre tonnes serve as the foundation for the most precise selection as compared to the general metre tonnes. The system offers a range of data extraction services in accordance with the university's role, which serves to enhance college English education's collaboration with curriculum and governmental thinking.

Keywords: Curriculum Civics; Optimized Apriori Algorithm; English Teaching; Collaborative Research

I. INTRODUCTION

With greater teaching hours and a wider scope, collaborative study of English language instruction is a highly significant required public course in higher education and contributes to its relative weight in the field [1]. The secondary English language curriculum must be integrated with civics instruction in order for students to not only meet their basic English learning needs but also effectively improve their comprehensive quality and advance the overall development of their level. Secondary English language teaching is no longer able to meet the learning requirements of students [2].

Currently, in the information age, the speed of information dissemination is thousands of times faster than before, and whatever news can be spread to people through cell phones, and these news significantly affect the quality of students' thoughts, and are easy to mislead students whose minds are not yet developed and sound, so students' ideology is still in a very unstable state, which is also an unprecedented challenge to the ideological and political work[3]. A symposium on ideology and politics in schools was presided over by President Xi Jinping, and our nation proposed the transition education construction, making full use of the primary channel of classroom education to build ideology and politics well, creating a good learning atmosphere, and maintaining good contact with other courses, to create good conditions of responsibility. Abandon the outdated traditional teaching methods of the past, develop innovative secondary English teaching techniques, and integrate political and ideological construction with English education.

Civic education can further encourage English teaching in a collaborative manner [5]. The classroom lacked life and was unattractive to students in the past, and teaching English was merely a single mechanical transfer of English information and abilities [6]. In the English classroom, students merely acquired facts; their own habits and behaviour were unaffected. There is no harmonic teacher-student interaction since the relationship between professors and students is that of a professor and a pupil. The original repetitive and boring English classroom gets a new flavour with the addition of the thinking component. While teaching English, teachers can introduce students to the lives, historical events, and current affairs of Western nations, accurately analyse Western ideologies, fairly contrast Chinese and Western cultures, and support our traditional Chinese cultures [7-9]. Teaching for enjoyment

and teaching for improvement are genuinely realised as a result of this transformation, varied teaching methods, and friendly teacher-student relationships that add to the richness of the English teaching classroom.

The goal of classification is to group data into a certain category, such as identifying a student type's grades when that student type is classified as exceptional, good, passing, or failing. Artificial neural networks, decision trees, and Bayesian theory are all classification techniques [10]. For instance, the most popular approaches for predicting a student's English grade are logistic regression and linear regression [11]. According to certain rules, the technique of clustering involves placing data from related categories into the same category and data from unrelated categories into a different category. For instance, students with similar grades would be placed in the same class, and topics with a similar degree of difficulty would be placed in the same category[12]. While categorization is known, the categories to be classified by clustering are not known in advance. Analysis of association relationships between data is done through association rule mining. Examples include the correlation between student attendance at the library and grades, the correlation between English and math grades, etc. Common association algorithms include the Apriori algorithm, the FP-tree frequent set method, and transaction compression [13].

In conclusion, there are some theoretical and practical implications to the collaborative study of English instruction based on the optimised Apriori algorithm.

II. RELATED WORKS

Academic conferences and journals on educational data mining have exploded in recent years, and many of these have discussed the topic. Educational data mining techniques have also been thoroughly studied and applied, making it one of the hottest areas of research in today's academic world.

A mature intelligent tutoring system has emerged, and foreign data mining technology for collaborative English education is more advanced, with a certain breadth and depth. There is also a certain breadth and depth of research related to data mining technology, focusing primarily on the improvement and optimisation of data mining algorithms, and has produced very positive results [14]. Representative outcomes include the classification algorithm developed by Professor J.R. Quinlan of Australia, the association algorithm developed by R. Agrawal of IBM, etc. Additionally, many top-notch data mining programmes have been developed [15–16]. In the early stages of data mining research, questionnaire surveys and management platforms were the primary sources of data. The validity and quantity of the data from these sources were not assured, and the major techniques employed were association rule algorithms and straightforward statistics. Since the turn of the century, educational data mining has advanced quickly. At this point, the majority of the data comes from online educational learning platforms, and the data mining techniques are generally varied.

Domestic research on collaborative English education data mining technology began somewhat later, focusing mostly on applications, literature [17–18] used correlated data to solve the problems of learning disorientation and cognitive overload; literature [19–21] proposed a research method for educational decision making, emphasising the importance of educational data; literature [22] used improved clustering and association rule algorithms in student performance analysis was . From 2008 to the present, while having developed quickly, it has lagged behind studies from other countries. The major goals of the study are to forecast students' performance, examine how courses relate to one another, and provide learning tools to students. Grey system theory was applied in the literature [23] to forecast student grades, with some success. In the literature [24], student grades were examined using association rule algorithms to determine how well students performed across various courses. In order to attain the goal of teaching to each student's ability, the literature [25] employed the research of personalised recommendations based on knowledge points to recommend various knowledge points to various students. Prior to this, student performance predictions simply took into account past performance, which was biassed and did not produce the desired outcomes. The correlation study of student performance ignores the impact of students' everyday behaviours on courses and simply examines the association between courses. In this study, we forecast student performance using artificial neural networks and the Apriori algorithm.

A database, artificial intelligence, and quantitative statistics are the three main pillars of collaborative research on English language education in the framework of curriculum thinking, which is to say.

III. EXPERT SYSTEM AND APRIORI ALGORITHM OPTIMIZATION

A. *Expert System for Collaborative English Teaching in the Context of Curriculum Civics*

When human experts are not available, expert systems are used to imitate their decision-making processes using the knowledge base that is provided by human experts. An expert system can be thought of as a computer intelligence system that derives its capabilities from the expert knowledge provided by human domain specialists and whose knowledge base incorporates expertise and experience. The simulation of complicated problems is necessary for an expert system for collaborative English language education in the context of curriculum thinking, and this simulation largely draws on the idea of knowledge representation and knowledge acquisition. An expert system may perform reasoning and judgement at a level that is comparable to that of a human expert at that level. Expert systems often have knowledge provided by subject matter experts, which is mostly utilised for processing knowledge-based information. Expert systems excel at solving complex issues that cannot be handled by current computers, but conventional artificial intelligence systems are limited to dealing with straightforward numerical issues, etc. Additionally, expert systems have the ability to resolve broad issues; this ability results from the use of knowledge and reasoning. The expert system can use the knowledge in the knowledge base and the offered reasoning approach to reason effectively about the user-proposed problem based on the data provided by the user and arrive at a solution. General data processing systems, on the other hand, are limited to handling mathematical operations and statistical functions and are unable to address general issues. The expert system can describe how it handled the problem and came to its ultimate output decision during operation. It may be claimed that the expert system has an explanation function because the problem-explanation process is transparent. This function and experience are not found in conventional computer applications. Similar to human experts, expert systems are limited in their knowledge to a narrow field, and their experience is very specific. As a result, the problems they solve only pertain to this narrow field, which increases the effectiveness of the system and gives the work they do more significance for the users.

In the context of curriculum thinking, the organisation and creation process of each component of the expert system are referred to as the "structure" of the collaborative English teaching expert system. Because different expert systems have varied application domains and application aims, as well as different knowledge bases depending on the features of the subjects they are used to analyse, different expert systems have different functions and architectures. Figure 1 depicts the abstract expert system architecture.
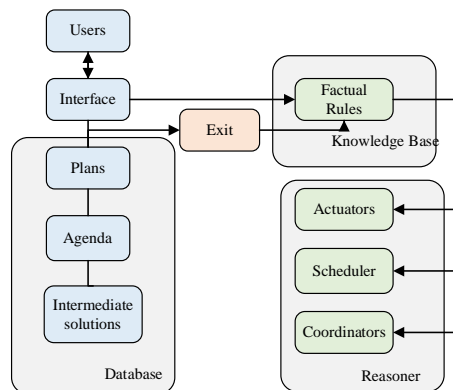

Figure 1 Structure of the collaborative English teaching expert system in the context of curriculum thinking and politics

Figure 2 depicts the collaborative English teaching expert system's three-layer structure in relation to political and curriculum-related thinking. In principle, the three-layer structure is clearly divided, but in actual use, there is no set standard. In this structure, only the aforementioned browser version will allow the user to access and use the system.
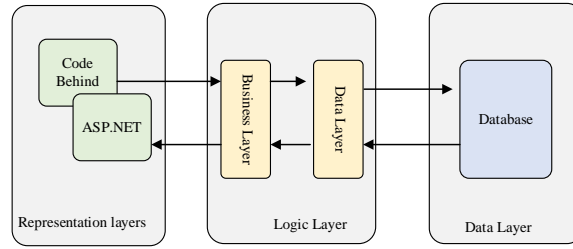
Figure 2 Three-tier structure model of collaborative English teaching in the context of curriculum thinking and administration

This three-layer structure has a rigid business hierarchy; no layer can access a layer below it without first accessing the layer above it. The data and instructions are processed and translated by the business logic layer, which then connects to the data access layer to finish the data's operation. The results are subsequently uploaded via the opposite channel to the representation layer. The outcome is then communicated back to the user after cascading to the representation layer via the opposing path.

B. Knowledge Representation and Forward and Backward Reasoning

The uncertainty reasoning is modelled and put into practise using deterministic theory in accordance with the actual circumstances of this system. The degree of certainty is separated into four categories, ranging from very certain to very uncertain, and the particular defined values are provided individually by domain experts, who are unrelated to and unaffected by one another. The inference technique is one of them, and it serves as the functional foundation upon which the expert system's primary functionality rests. Expert systems that primarily use generative rules can reason in one of two ways: forward reasoning or backward reasoning.

The multiple-input multiple-output model, which may be broken down into numerous single-input multiple-output variants, is the general model of forward rule inference.

$$R(k) : If \ x \ then \ y \ is \ (y_k, c_k) \ (k = (I...M) : CF(y_k) \ (1)$$

In the system, several different inputs may get the same y_k, if this happens, it means that the same knowledge is examined several times, at this time the same y_k to be superimposed to increase the trust, so it is necessary to modify the child corresponding to y_k, indicating that the inference of this knowledge should be more credible. Its reasoning diagram is shown in Figure 3.
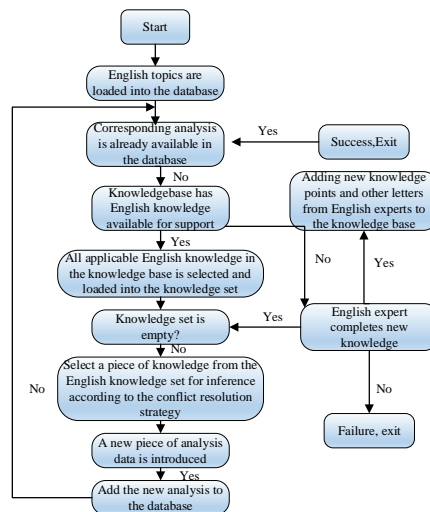

Figure 3 Forward reasoning diagram of collaborative English teaching in the context of curriculum thinking

Here, the backward reasoning does not exactly entail the forward reasoning. The uncertainty factor is an additional consideration for this system in the reverse reasoning. The given basis, which is uncertain in and of itself, the knowledge in the knowledge base, which is likewise uncertain, and the conclusion with uncertainty property, which is ultimately produced by the uncertainty reasoning process, are all layers through which this factor is traversed.

Firstly, a fuzzy uncertainty rule relation is established.

$$\mu_z = \{\mu_y \cdot W\} \quad (2)$$

The inference model is obtained as follows.

$$\mu_y \to \mu_{CF} \to \mu_z \quad (3)$$

μ_z is a weighted ensemble, which is a logical expression of the operation, then the following equation is obtained.

$$\mu_p = \sum_{t=1}^{n} W_j \cdot \mu_y \quad (4)$$

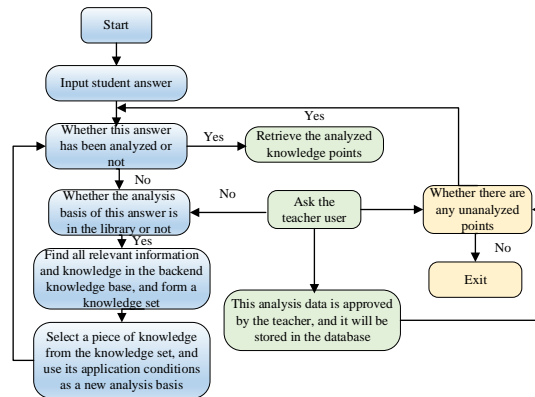Figure 4 explains the process of reverse uncertainty inference for this system.

Figure 4 Reverse reasoning diagram of collaborative English teaching in the context of curriculum thinking

In practical applications, challenges are typically complicated, making it risky to approach or draw conclusions solely from one angle. Reverse reasoning relies on assumptions that may not apply to the actual situation, which can waste system resources and diminish operational efficiency. Forward reasoning may provide sub-results that the system does not need during operation. Therefore, combining forward and backward thinking is a great way to solve this issue because it addresses both of their advantages and disadvantages. Hybrid reasoning is the name for this type of reasoning.

In addition, each item in the set represents a knowledge point, and the knowledge points corresponding to each question are shown in Table 1.

Table 1 Set knowledge points correspondence table

Title	Knowledge points
J1	I4, I1, I2
J2	I1, I2
J3	I2, I4
J4	I1, I3
J5	I0, I2, I3, I5
J6	I3, I4, I5
Q7	I2, I3

Q8	I2, I4
Q9	I1, I2, I5

At this moment, the frequent item set is the outcome of the Apriori algorithm's mining operation. Figure 5 displays the candidate item set generation diagram.
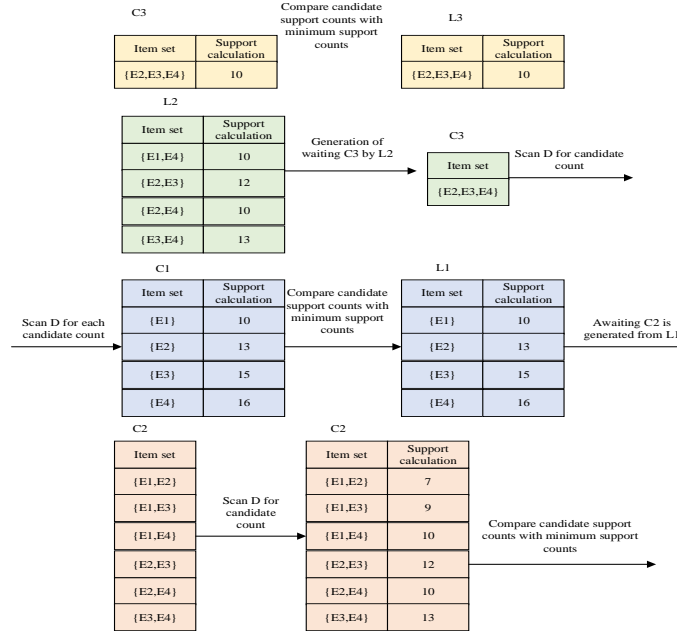

Figure 5 Candidate item set generation diagram

The connection between the knowledge points is what is sought in the collaborative system of English teaching in the context of civics, and there is essentially no causal relationship between the knowledge points, more of an equal and mutually exclusive relationship. The system's necessary function is to extract the knowledge points that are easily confused, and according to the extracted knowledge points, the system's existing questions are analysed and aggregated.

IV. METHODS

We can see from the programme's implementation that the Apriori algorithm needs to traverse the database multiple times and has a lot of access to it. To obtain the set of frequency sets L, two MapReduce operations calculate Figure 6. Figure 6 illustrates how the algorithm mines all frequency sets. The mining procedure in the figure reduces the database scan to twice.
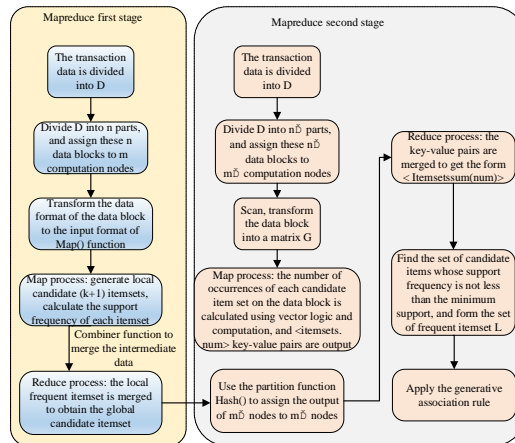

Figure 6 Flow chart of the improved Apriori algorithm

Among them, the transaction database D has 6 transactions S1~S6, and the specific data are shown in Table 2.

Table 2 Matter database simulation table

Transaction ID(Sid)	Items
W1	I1, I2, I3, I4, I5
W2	I0, I3, I5
W3	I1, I6
W4	I3, I5,
W5	I1, I4
W6	I0, I4, I5

The algorithm gets the corresponding L1 after the first scan of the database D. Since sup(I6) =2<min_sup, I6 is removed and L1= {I1, I2, I3, I4, I5}, after L1 is obtained, the two parts of the database are matrixed and I6 is removed, and the resulting matrix is shown in Table 3 and Table 4.

Table 3 First matrix

Transaction ID(Sid)	I1	I2	I3	I4	I5
S1	0	0	1	1	0
S2	1	1	0	0	1
S3	1	1	0	1	1

Table 4 Second matrix

Transaction ID(Sid)	I1	I2	I3	I4	I5
S4	1	0	1	0	0
S5	1	1	1	1	0
S6	1	0	0	1	1

Since $L5 = \phi$, all the mining process is finished, all the frequency sets are merged and the duplicate parts of the frequency sets are removed, and the final set of frequency sets L={I1I2I3, I1I2I4, I1I2I5, I3I4, I3I5, I4I5} is obtained, and the set L is regularized by association to find the rules required by the user and return them to the user.

A.   Experimental Environment

With six parallel computing nodes, one management node, and each node having a consistent configuration of 4GB RAM and a 1T SATA hard drive, Hadoop is used as the cloud computing platform architecture in this experiment. Java programming is done with Eclipse. The experiments' data is mined using the suggested algorithm, and the experiments' rules are produced using the Apriori_MR algorithm. Figure 7 illustrates the division phases in accordance with the system's actual requirements.
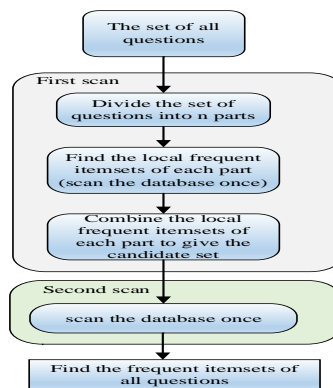

Figure 7 The division of optimization algorithm steps

B. Experimental Data

The experiment employs association rule mining on three data sets to illustrate the algorithm's generalizability. Retail, accidents, and graduates' grades in a university academic system are chosen as the experimental test sets, with retail being the transaction database of a shopping mall with a size of 1 GB and the number of transactions being over 3 million. The experimental data are collected from websites. accidents is a city's collected database of traffic accidents, with a size of 372 MB and more than 500,000 transactions; the database of graduates' grades has a size of 87 MB and more than 60,000 transactions. The data on accidents and graduating grades were duplicated and expanded to 1 GB before to the experiment in order to create the same amount of data for the test. The data were then pre-processed, the files were backed up multiple times, and the files were divided into files of the same size. The files are then posted to HDFS after processing. In Figure 8, the business procedure is displayed.
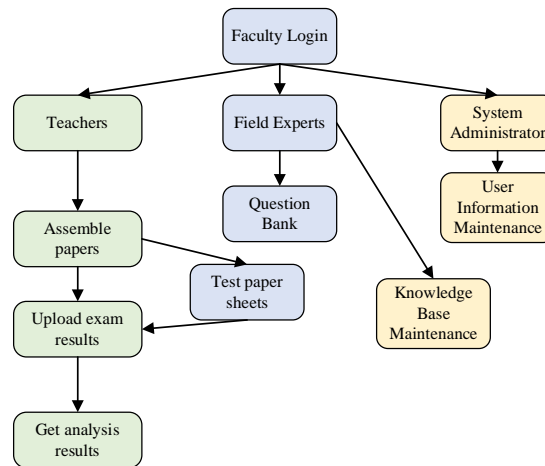

Figure 8 Experimental business flow chart

Grammar points are primarily responsible for this, as is the question base, which also includes more specific information like the question's explanation. The system's function is to be used by the teacher, who may either utilise the system's prompts to construct the papers or edit them by directly entering the questions. The system offers a database of questions and separates the questions' levels of difficulty. The system has two attributes: certainty and reliability, which recommend changes for the teacher's lesson plan once the teacher uploads the test results. These functions can be used by the class or by each individual student.

V. CASE STUDY

The experiment compares the improved algorithm with the Apriori_MR algorithm to show the performance improvement in the mining process because the algorithm is based on the context of curriculum thinking and collaborative English teaching, and it is improved on the basis of the Apriori_MR algorithm using matrixing.

A. Good performance with the same dataset

The performance of the two algorithms is displayed in Figure 9 as the number of nodes rises, and the improved algorithm does not establish the target items to assure the fairness of the comparison. The algorithm changes slowly, especially after the upgrade. This is due to the fact that the data block that has to be processed already has a Map task assigned to it, thus there is no need to wait for the execution of an existing Map task before beginning a new one. This also shows that the current system is capable of processing all data blocks simultaneously in parallel. The outcomes also demonstrate that the enhanced algorithm uses less time.
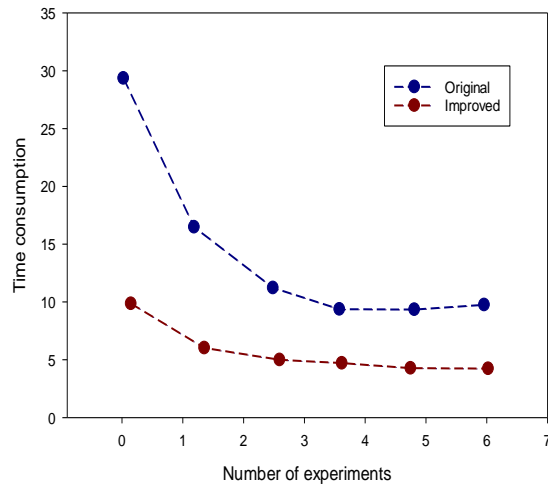
Figure 9 Time consumption ratio of the improved algorithm

B. *Good performance of the optimization algorithm for different data sizes*

To obtain data of the order of 1 million, the collaborative data of ELT were duplicated and replicated. These results are then applied successively to the experiments. Since it is known from the prior experiment that the current experimental data volume can produce the best results by using only 4 nodes, 4 nodes are used in this experiment, and the experimental results are displayed in Figure 10 in order to make the experimental comparison more clear. The improved algorithm does not set the target term in the experiments. From the graphic, we can see that as the data amount increases, the ratio of the time required by both algorithms reduces, suggesting that the greater the data volume. Because the system requires some necessary I/O operations, such as storage, communication, and coordination management, it does not function properly when the data amount is tiny.
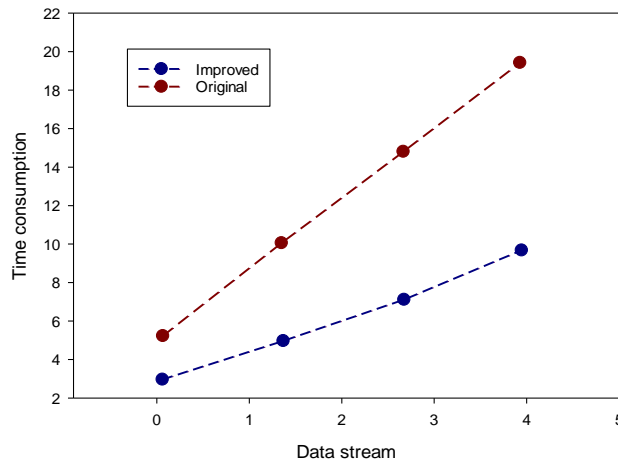

Figure 10 Variation of algorithm analysis time with the amount of mined data

C. *Support and confidence enhancement*

The collaborative English teaching criteria are the most traditional rule measure in association rule mining, which necessitates pre-set minimum support and minimum confidence for the rules. The database is searched to count the itemsets in the frequency set mining stage, and the itemset counts are then compared with min_sup to filter the frequency sets. Following the mining of all frequency sets, association rules are formed from the frequency sets; a strong association rule is one in which the transaction counts of the previous and following parts of the rule are divided by the transaction counts of the preceding pieces. Generally speaking, support and confidence stand for the

rule's universal applicability and reliability, respectively. The majority of association rule mining algorithms base their evaluation of the results on this criterion, but if the rules are evaluated solely by this criterion, it may produce some meaningless or even deceptive rules. Using the information in Table 5 as a starting point, the following example illustrates the inadequacies of this criterion.

Table 5 Case transaction database

Transaction ID(Sid)	Item
T1	BCDE
T2	ABE
T3	BCDE
T4	ABF

The relative probability that the term set X does not occur is denoted as XP, and the relative confidence level is specified as follows, provided that X is a frequency set.

$$(P(X) > \min\_\sup) \& (P(X) < 1) \Rightarrow \text{Relative confidence level=}P(Y|X)\text{-}P(Y|\overline{X}) \quad (5)$$

$$(P(X) = 1) \& (P(Y) = 1) \Rightarrow \text{Relative confidence level=1} \quad (6)$$

$$(P(X) = 1) \& (P(Y) < 1) \Rightarrow \text{Relative confidence level=0} \quad (7)$$
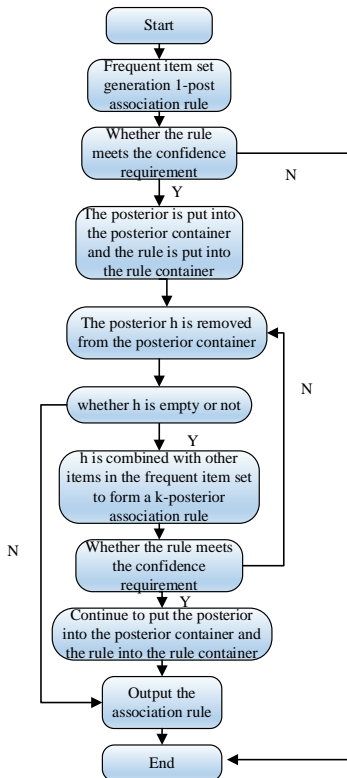
The association rules are shown in Figure 11.

Figure 11 Relative confidence judgment rule

In conclusion, this chapter first introduces two algorithms that are similar to the proposed algorithm: Apriori_MR algorithm (a MapReduce-based Apriori algorithm) and Apriori_M algorithm (a matrix-based Apriori algorithm). The two algorithms are then combined to create Apriori_PM algorithm (a parallel matrix-based Apriori algorithm), which incorporates the improvements of the two algorithms. Then, to further enhance the algorithm's efficiency, it is optimised for multi-stage strategic mining in accordance with its rules during the process of frequency set mining.

VI. CONCLUSION

In order to further promote the implementation of curriculum thinking in English classrooms in higher education institutions, this paper first examines the development and issues of collaborative English teaching in the context of curriculum thinking. It then investigates the current state of curriculum thinking promotion in English classrooms in higher education institutions. Finally, it proposes some superficial suggestions and countermeasures. Second, the apriori calculation method of mining association rules is enhanced, which accelerates mine foundation, enhances fundamental measurement, and enhances mining rules' effectiveness. In order to comprehend the applicability of the calculating techniques suggested in this paper, the Cooperative English training system is realised. First, the amount of data on the Internet has increased quickly. Second, the parallel Apriori matrix algorithm has been improved. Third, the algorithm has been better integrated with the hard strip distribution system. Fourth, the mining algorithm has been made more effective. Second, mining is now a multi-stage decision-making process with the addition of many candidates produced by the algorithm, increasing the system's efficiency. Third, we offer numerous examples, such as usefulness, effectiveness, and usefulness, to help overcome the drawbacks of supporting confidence-building methods.

ACKNOWLEDGMENT

REFERENCES

[1] Yuanhong, X. U. (2021). Research on the reform of humanities English teaching based on ideological and political theories teaching in all courses. Canadian Social Science, 17(2), 52-56.

[2] Wang, Y. (2020). Analysis on the construction of ideological and political education system for college students based on mobile artificial intelligence terminal. Soft Computing, 24(11), 8365-8375.

[3] Yu, Y. (2022). On the ideological and political education of college students in the new media era. Open Journal of Social Sciences, 10(1), 1-14.

[4] Liu, G. (2021). The ways and methods of ideological and political education for postgraduates. Advances in Educational Technology and Psychology, 5(3), 80-87.

[5] Zhao, X., & Zhang, J. (2021). The analysis of integration of ideological political education with innovation entrepreneurship education for college students. Frontiers in Psychology, 12, 610409.

[6] Liu, X., Xiantong, Z., & Starkey, H. (2021). Ideological and political education in Chinese Universities: structures and practices. Asia Pacific Journal of Education, 1-13.

[7] Yin, Y. (2021). Research on ideological and political evaluation model of university students based on data mining artificial intelligence technology. Journal of Intelligent & Fuzzy Systems, 40(2), 3689-3698.

[8] Di Wu, Yin Lei, Maoen He, Chunjiong Zhang, Li Ji, "Deep Reinforcement Learning-Based Path Control and Optimization for Unmanned Ships", Wireless Communications and Mobile Computing, vol. 2022, Article ID 7135043, 8 pages, 2022. https://doi.org/10.1155/2022/7135043

[9] Xiaoyang, H., Junzhi, Z., Jingyuan, F., & Xiuxia, Z. (2021). Effectiveness of ideological and political education reform in universities based on data mining artificial intelligence technology. Journal of Intelligent & Fuzzy Systems, 40(2), 3743-3754.

[10] Xia, Y. (2021). Big data based research on the management system framework of ideological and political education in colleges and universities. Journal of Intelligent & Fuzzy Systems, 40(2), 3777-3786.

[11] Rong, Z., & Gang, Z. (2021). An artificial intelligence data mining technology based evaluation model of education on political and ideological strategy of students. Journal of Intelligent & Fuzzy Systems, 40(2), 3669-3680.

[12] Su, L., Xiao, L., & Wang, J. (2021). A case study of the ideological and political education of college English translation course driven by words. Creative Education, 12(2), 317-328.

[13] Zhou, Y., & Zhou, H. (2022). Research on the quality evaluation of innovation and entrepreneurship education of college students based on extenics. Procedia Computer Science, 199, 605-612.

[14] Huda, M., Anshari, M., Almunawar, M. N., Shahrill, M., Tan, A., Jaidin, J. H., ... & Masri, M. (2016). Innovative teaching in higher education: The big data approach. TOJET, 1210-1216.

[15] Qasem, Y. A., Abdullah, R., Jusoh, Y. Y., Atan, R., & Asadi, S. (2019). Cloud computing adoption in higher education institutions: A systematic review. IEEE Access, 7, 63722-63744.

[16] Zeide, E. (2017). The structural consequences of big data-driven education. Big Data, 5(2), 164-172.

[17] Murumba, J., & Micheni, E. (2017). Big data analytics in higher education: a review. The International Journal of Engineering and Science, 6(06), 14-21.

[18] Williamson, B. (2018). The hidden architecture of higher education: building a big data infrastructure for the 'smarter university'. International Journal of Educational Technology in Higher Education, 15(1), 1-26.

[19] Zhu, Z. T., Yu, M. H., & Riezebos, P. (2016). A research framework of smart education. Smart learning environments, 3(1), 1-17.

[20] Samerkhanova, E., Krupoderova, E., Krupoderova, K., Bakhtiyarova, L., Ponachugin, A., & Kanyanina, T. (2017). Developing an information educational environment based on cloud technologies. Journal of Entrepreneurship Education, 20(3), 1-9.

[21] An, P., Wang, Z., & Zhang, C. (2022). Ensemble unsupervised autoencoders and Gaussian mixture model for cyberattack detection. Information Processing & Management, 59(2), 102844.