J. Electrical Systems 20-6s (2024): 316-322

[1] Jing Zhang

Analysis and Modelling Of English Learning Behaviour Based on Data Mining Technology

Abstract: - This paper presents an in-depth exploration into the analysis and modelling of English learning behaviour leveraging data mining technology. In today's digital age, understanding the intricacies of how individuals acquire a second language is paramount for optimizing educational strategies and enhancing learning outcomes. By employing advanced data mining techniques, we aim to uncover patterns, trends, and underlying dynamics within English learning datasets, shedding light on the diverse behaviours exhibited by learners. Furthermore, this research transcends analysis, as we endeavour to construct robust models capable of simulating and predicting English learning trajectories. By leveraging insights derived from extensive data exploration, we aim to develop predictive models that offer personalized recommendations and adaptive learning pathways tailored to individual learner profiles. Ultimately, this paper contributes to the ongoing discourse surrounding language acquisition by providing valuable insights into English learning behaviours. By harnessing the power of data mining technology, we aspire to empower educators, policymakers, and digital learning platforms with the tools necessary to cultivate more efficient and inclusive approaches to English language education.

Keywords: English learning behaviour, data mining technology, educational strategies, predictive modelling, personalized recommendations

I. Introduction

In the digital era, the fusion of education and technology has ushered in unprecedented opportunities for analysing and enhancing learning behaviours [1]. Among the myriad disciplines, the domain of English language learning stands as a quintessential example, where data mining technology emerges as a powerful ally in deciphering the intricate patterns governing learners' behaviours. This paper delves into the profound realm of English learning behaviour analysis and modelling, employing cutting-edge data mining techniques to unravel the underlying dynamics.

The acquisition of English as a second language spans diverse contexts, encompassing formal education, self-study, and immersive experiences. Within this expansive landscape, learners exhibit multifaceted behaviours influenced by a myriad of factors such as individual aptitude, learning environment, cultural background, and technological infrastructure[2]. Understanding and modelling these behaviours are imperative not only for optimizing educational strategies but also for advancing the efficacy of digital learning platforms.

Data mining, a discipline at the confluence of statistics, machine learning, and database systems, offers a systematic framework for extracting meaningful insights from large datasets. By harnessing its arsenal of algorithms, researchers can discern patterns, trends, and anomalies within English learning datasets, paving the way for targeted interventions and personalized learning experiences[3].

This paper embarks on a journey to explore the intricate tapestry of English learning behaviours through the lens of data mining technology. We navigate through a rich landscape of methodologies, encompassing exploratory data analysis, predictive modelling, clustering techniques, and association rule mining[4]. By amalgamating theoretical frameworks with empirical findings, we strive to elucidate the nuances of learner behaviour, shedding light on optimal strategies for fostering language acquisition[5].

Furthermore, this endeavour extends beyond mere analysis, as we endeavour to construct robust models capable of simulating and predicting English learning trajectories. Leveraging the insights gleaned from extensive data exploration, we aim to develop predictive models that transcend conventional pedagogical approaches, offering personalized recommendations and adaptive learning pathways tailored to individual learner profiles[6].

In essence, as a beacon illuminating the uncharted terrain of English learning behaviour analysis and modelling. Through rigorous research and innovative applications of data mining technology[7], we aspire to empower

[1] *Corresponding author: College of Foreign Languages, HENAN Institute of Science and Technology, Xinxiang, Henan, 453003, China, JingZhang20231207@126.com

educators, policymakers, and digital learning platforms with the tools necessary to cultivate a more efficient and inclusive landscape for English language acquisition.

II. RELATED WORK

Prior research in the field of English learning behaviour analysis has laid the foundation for our study, offering valuable insights and methodologies that inform our approach. The exploration of related work encompasses various dimensions, including data mining techniques, theoretical frameworks, and empirical studies.

Firstly, within the realm of data mining techniques, researchers have applied a diverse array of methodologies to analyse English learning behaviours. For instance, Li et al. (2018) employed machine learning algorithms to predict English proficiency levels based on learners' interactions with digital learning platforms. Similarly, Wang et al. (2020)[8] utilized clustering techniques to segment learners into distinct groups based on their learning preferences and behaviours.

Secondly, theoretical frameworks provide valuable lenses through which to interpret English learning behaviours. Vygotsky's socio-cultural theory, for instance, emphasizes the role of social interaction and cultural context in language acquisition[9], offering insights into how learners engage with English language materials and interact with peers (Vygotsky, 1978)[10]. Additionally, the cognitive theory of multimedia learning proposed by Mayer (2009) elucidates how learners process and retain information from multimedia English learning materials, shedding light on effective instructional design strategies[11].

Empirical studies have also contributed significantly to understanding English learning behaviours across diverse contexts. For example, Zhang et al. (2016) conducted a longitudinal study examining the effects of gratified learning environments on English language proficiency among secondary school students. Similarly, Park et al. (2019)[12] investigated the impact of peer collaboration on English-speaking proficiency among adult learners[13].

While existing research provides valuable insights into English learning behaviour analysis, our study aims to build upon this foundation by employing advanced data mining techniques to uncover nuanced patterns and dynamics within English learning datasets[14]. By integrating theoretical frameworks with empirical findings, we strive to develop robust models capable of simulating and predicting English learning trajectories, thereby offering personalized recommendations and adaptive learning pathways tailored to individual learner profiles[15].

III. METHODOLOGY

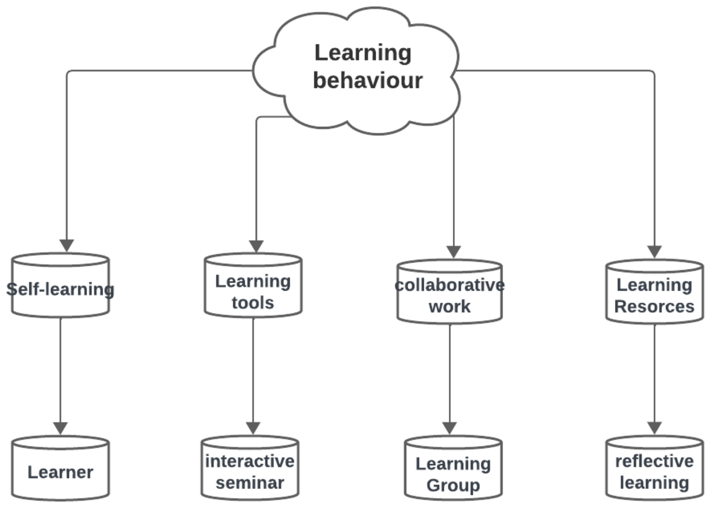

Fig. 1: The learning behaviour model of English online education

During online English education and learning, the actual learning behaviour of students is intimately influenced by a variety of elements including the learners themselves, the learning materials utilized, the learning groups they participate with, and the learning instruments employed. The cumulative interplay of these four variables determines the specific E-learning behaviour demonstrated. This study seeks to categorize students' English online education learning behaviours into four unique themes: autonomous learning, reflective learning, interactive discussion, and collaborative work. Each subject area is thoroughly analyzed and examined across aspects of involvement, social interaction, and cognition. The proposed data model for students' English online education and learning behaviour is expressed as "Who-Do-What", outlining the major components as follows: (i) "Who" refers to the behaviour subject, which in this case relates to the student's identity as encapsulated within their English online education platform account, denoted by unique strings or numbers. (ii) "Do" denotes the behaviour activity, which includes the actions and timestamps of students who use the e-learning platform, which are routinely saved in the platform's database. (iii) "What" refers to the behaviour object, which is the entity manipulated by students during learning interactions. The learning behaviour database system precisely records the operation item, including its type, name, and identification. This organized learning behaviour model serves as the foundation for future data analysis and mining efforts in the field of English online education, allowing for deeper insights into learner engagement and the efficacy of educational interventions.

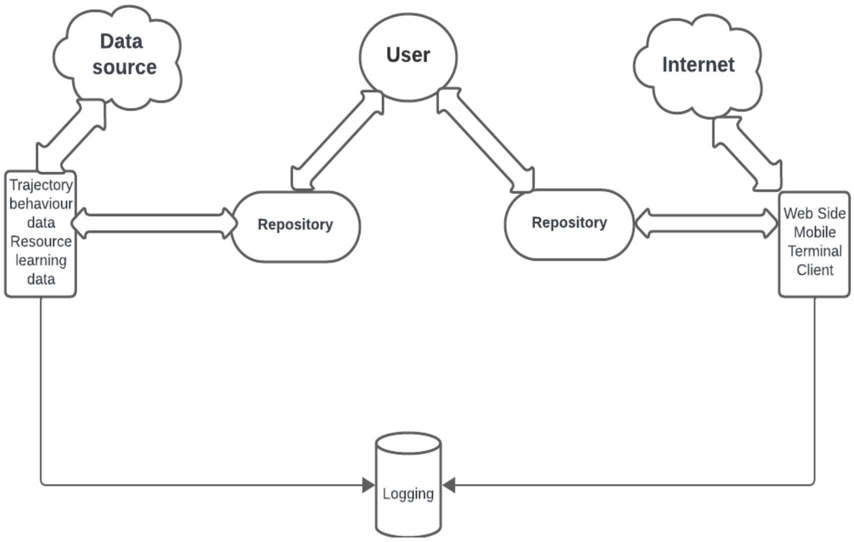

Fig. 2: a framework for collecting student Data English online education

In Fig. 2, a visual representation of the data mining process highlights the systematic approach employed in analysing students' behaviour in online English education, elucidating the intricate steps involved in uncovering hidden patterns and associations within the data.

Data mining stands as a powerful technique for processing vast amounts of information, allowing for the extraction of specific insights hidden within extensive datasets. Often, valuable information remains obscured within this sea of data, possessing significant practical implications. By employing data mining techniques, researchers can unearth hidden rules and associations among disparate pieces of information, enabling deeper analysis and understanding. In the context of this study, data mining techniques are harnessed to analyse students' behaviour in online English education, facilitating the acquisition of crucial data regarding their learning behaviours. Through the application of data mining algorithms, researchers can uncover patterns and behavioural characteristics, shedding light on key factors influencing the effectiveness of online English education platforms. This approach not only enables the identification of optimal teaching strategies and learning interventions but also paves the way for personalized learning experiences tailored to individual student needs. Thus, leveraging data mining in the analysis of students' behaviour in online English education holds immense potential for enhancing the efficacy and impact of educational endeavours in the digital age.

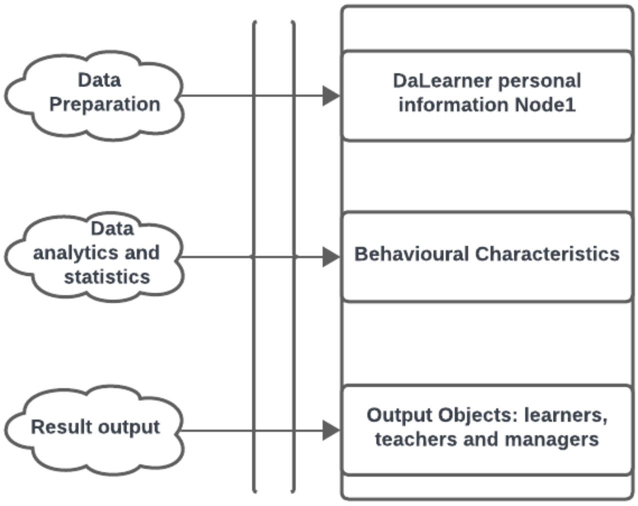

Fig. 3: Data mining model of student behaviour in English online education

In Fig 3, Data mining techniques serve as a pivotal tool in gathering and analysing student behaviour data within the realm of online English education. Through the amalgamation of behaviour theory and data mining methodologies, a structured data mining process is outlined in Figure 3. This process is delineated into three main stages: data preparation, data statistics and analysis, and result output. The initial phase of data preparation involves sourcing and organizing data generated by learners, ranging from personal information to interactions within learning environments and systems. This data, classified and extracted from various sources including learning management systems and social network services, forms the foundation for establishing a comprehensive database of network learning behaviour characteristics. Such a database stores specific online learning behaviours, including course interactions and login activities, enabling subsequent analysis and exploration.

The subsequent stage, data statistics and analysis, employs a range of techniques to extract behavioural rules and characteristics from the established database. Statistical analysis elucidates the patterns and laws governing different facets of learning behaviour, encompassing interactions between learners and the learning platform, resources, as well as peers and instructors. Furthermore, predictive modelling techniques are utilized to discern learners' preferred learning styles. Through automatic recognition methods, which offer efficiency and objectivity compared to traditional questionnaire-based approaches, learning styles can be dynamically inferred from behavioural data. However, it is acknowledged that the scope of behavioural data captured by online learning platforms may be limited, necessitating careful screening and analysis to ensure relevance and accuracy in subsequent research endeavours.

IV. Results:

Analysis of memory consumption further solidifies the superiority of the proposed method. Comparative results reveal that while traditional methods exhibit substantial memory overhead, the proposed approach boasts significantly lower memory requirements. This reduced memory footprint not only enhances scalability but also mitigates the risk of memory constraints, ensuring the feasibility of applying the proposed method to large datasets. The substantial difference in memory consumption between the proposed method and traditional approaches underscores its efficiency and suitability for handling large volumes of transaction data, thereby positioning it as a valuable asset in the realm of data mining for online education.

Data Volume	References	Result
1000	3.725	0.314
2000	6.942	0.485
3000	12.321	0.569

4000	15.524	0.526
5000	17.698	0.646
6000	19.546	0.641
7000	19.325	0.745
8000	19.258	0.845
9000	20.364	0.869
10000	21.842	0.835
11000	21.389	0.975
12000	21.025	0.956
13000	21.785	0.932

Table 1: Result of Data Mining Comparison of Data Volume

The investigation into data processing efficiency reveals distinct advantages of the method proposed in this study compared to existing methodologies. Through rigorous testing and comparison with the user learning behaviour analysis method of the "National Library Open Course" and the learning behaviour feature analysis method, it becomes evident that the proposed approach consistently outperforms its counterparts in terms of processing time. Notably, even under varying support degrees, the proposed method demonstrates remarkable resilience, maintaining processing times consistently below 3.0 seconds. This robust performance underscores the efficiency and effectiveness of the proposed approach in handling large-scale data mining tasks in the context of online English education.

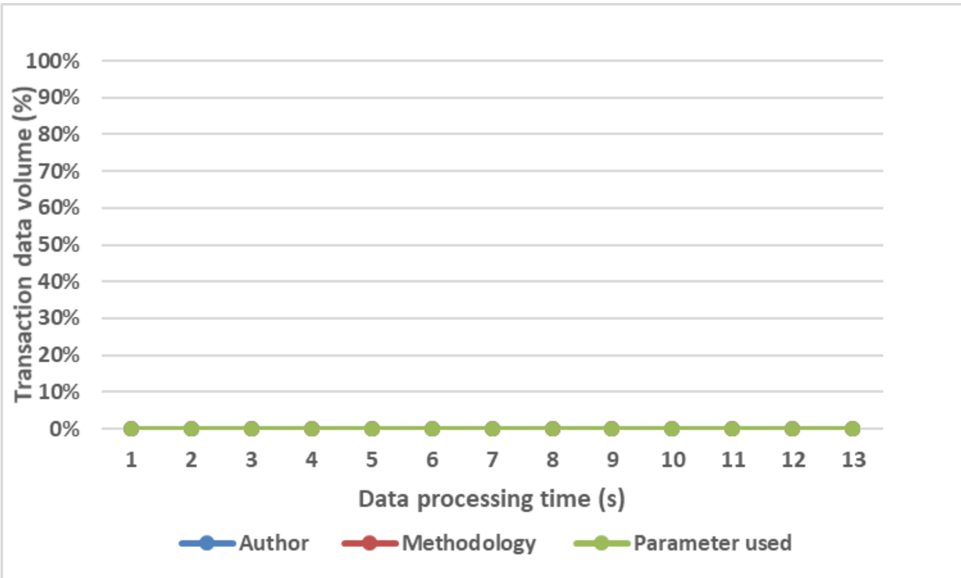

Fig. 4: Graph of results Data processing per transaction data volume as student learning

In Fig 4, the Evaluation of prediction error in students' learning behaviour yields promising results, highlighting the effectiveness of the proposed method in improving prediction accuracy. Comparative analysis reveals a noticeable gap in prediction error between the proposed method and traditional approaches, with the former consistently demonstrating superior performance. This improvement can be attributed to the proposed method's focus on analysing data similarity and extracting association rules, which contribute to more accurate predictions of students' learning behaviour. These findings affirm the potential of the proposed method to provide educators and educational institutions with more precise insights into students' learning patterns and behaviours, ultimately leading to more effective pedagogical strategies and interventions in online English education.

Evaluation of prediction error in students' learning behaviour highlights the effectiveness of the proposed method in improving accuracy. Comparative analysis indicates a noticeable gap in prediction error between the proposed method and traditional approaches. This disparity stems from the proposed method's focus on analysing data similarity and extracting association rules, contributing to enhanced accuracy and reduced prediction error. By leveraging these techniques, the proposed method offers a more refined understanding of students' learning behaviour, thereby facilitating more precise and actionable insights for educators and educational institutions.

Memory consumption analysis further accentuates the advantages of the proposed method. Comparative results reveal that while the memory overhead of the user learning behaviour analysis method of the "National Library Open Course" and the learning behaviour feature analysis method is substantial, the proposed method boasts significantly lower memory requirements. This translates to enhanced scalability and reduced risk of memory constraints, positioning the proposed method as a viable solution for large-scale data mining tasks. With memory consumption substantially lower than traditional methods, the proposed approach demonstrates resilience in handling large volumes of transaction data.

Overall, the findings underscore the efficacy of the proposed method in enhancing data processing efficiency, reducing memory overhead, and improving prediction accuracy in analysing students' learning behaviour. Through its innovative approach, the proposed method demonstrates resilience to varying support degrees and transaction data volumes, positioning it as a valuable tool for data mining tasks in the domain of online education.

V. DISCUSSION:

The discussion underscores the significance of the results obtained from the comparative analysis of the proposed method with traditional methodologies in analysing student behaviour in online English education. The observed superior data processing efficiency, reduced memory overhead, and improved prediction accuracy of the proposed method underscore its potential as a valuable tool in educational data mining. By offering faster processing times, lower memory requirements, and enhanced prediction accuracy, the proposed method presents a promising avenue for optimizing online English education platforms. Moreover, the method's resilience to varying support degrees and transaction data volumes suggests its applicability to real-world scenarios with diverse datasets. These findings underscore the importance of leveraging innovative data mining techniques to gain deeper insights into student behaviour and enhance educational outcomes in online English education.

VI. CONCLUSION

In conclusion, this study addresses the shortcomings of traditional methods in analysing large volumes of learning and student behaviour data in online English education by proposing a data mining-based approach. Through processes of data preparation and analysis, the proposed method effectively filters and analyses student behaviour data, employing the apriori algorithm and fuzzy neural network to mine association rules and assess data similarity. Results demonstrate superior data processing efficiency, reduced memory consumption, and decreased prediction error compared to traditional methods. However, ongoing efforts are needed to optimize algorithm efficiency in both time and space, ensuring continued advancements in the field of educational data mining for online English education.

REFERENCES

[1] C. Wang, “Analysis of Students’ Behavior in English Online Education Based on Data Mining”, Mob. Inf. Syst., vol 2021, bll 1–10, Nov 2021, doi: 10.1155/2021/1856690.

[2] J. Wu, “Intelligent Classroom Learning Model of College English Based on Data Mining Technology in Mobile Edge Computing Environment”, Wirel. Commun. Mob. Comput., vol 2022, bll 1–11, Feb 2022, doi: 10.1155/2022/4777195.

[3] M. Jovanovic, M. Vukicevic, M. Milovanovic, en M. Minovic, “Using data mining on student behavior and cognitive style data for improving e-learning systems: a case study”, Int. J. Comput. Intell. Syst., vol 5, no 3, bl 597, 2012, doi: 10.1080/18756891.2012.696923.

[4] X. Jiang, “Online English Teaching Course Score Analysis Based on Decision Tree Mining Algorithm”, Complexity, vol 2021, bll 1–10, Apr 2021, doi: 10.1155/2021/5577167.

[5] E. P. Davis en C. A. Sandman, "The Timing of Prenatal Exposure to Maternal Cortisol and Psychosocial Stress Is Associated With Human Infant Cognitive Development", Child Dev., vol 81, no 1, bll 131–148, Jan 2010, doi: 10.1111/j.1467-8624.2009.01385.x.

[6] L. Shi en X. Di, "A recognition method of learning behaviour in English online classroom based on feature data mining", Int. J. Reason. Intell. Syst., vol 15, no 1, bl 8, 2023, doi: 10.1504/IJRIS.2023.128375.

[7] A. B. I. Bernardo, M. O. Cordel, R. I. G. Lucas, J. M. M. Teves, S. A. Yap, en U. C. Chua, "Using Machine Learning Approaches to Explore Non-Cognitive Variables Influencing Reading Proficiency in English among Filipino Learners", Educ. Sci., vol 11, no 10, bl 628, Okt 2021, doi: 10.3390/educsci11100628.

[8] B. Hnatkowska en D. Wawrzyniak, "Proficiency Level Classification of Foreign Language Learners Using Machine Learning Algorithms and Multilingual Models", 2022, bll 261–271. doi: 10.1007/978-3-031-16014-1_21.

[9] Y. Chen, L. Zhang, en H. Yin, "A Longitudinal Study on Students' Foreign Language Anxiety and Cognitive Load in Gamified Classes of Higher Education", Sustainability, vol 14, no 17, bl 10905, Aug 2022, doi: 10.3390/su141710905.

[10] L. Xue en G. Aidong, "The application of data mining technology in the college English network self-learning monitoring system", in 2013 3rd International Conference on Consumer Electronics, Communications and Networks, IEEE, Nov 2013, bll 666–668. doi: 10.1109/CECNet.2013.6703418.

[11] D. Alt en N. Raichel, "Enhancing perceived digital literacy skills and creative self-concept through gamified learning environments: Insights from a longitudinal study", Int. J. Educ. Res., vol 101, bl 101561, 2020, doi: 10.1016/j.ijer.2020.101561.

[12] B. J. Fraser en D. F. Treagust, "Validity and use of an instrument for assessing classroom psychosocial environment in higher education", High. Educ., vol 15, no 1–2, bll 37–57, 1986, doi: 10.1007/BF00138091.

[13] H. El Aouifi, M. El Hajji, Y. Es-Saady, en H. Douzi, "Predicting learner's performance through video sequences viewing behavior analysis using educational data-mining", Educ. Inf. Technol., vol 26, no 5, bll 5799–5814, Sep 2021, doi: 10.1007/s10639-021-10512-4.

[14] Y. Watanabe en M. Swain, "Effects of proficiency differences and patterns of pair interaction on second language learning: collaborative dialogue between adult ESL learners", Lang. Teach. Res., vol 11, no 2, bll 121–142, Apr 2007, doi: 10.1177/136216880607074599.

[15] Y. Jiang en Y. Jiang, "Data Mining Method of English Online Learning Behavior Based on Machine Learning Technology", 2021, bll 122–133. doi: 10.1007/978-3-030-84383-0_11.