A Fuzzy C-means Algorithm to Detect Learning Styles in Online Learning Environment

Bello Ahmad Muhammad[1], Wang Jianping[1], Guohong Gao[1], Chenping Zhao[1], Qian Li[1], Xiangang Zuo[1], and Yingying Lv[1]
[1]School of Computer Science and Technology, Henan Institute of Science and Technology, Xinxiang 453003, Henan, China

The ability to detect learners' learning styles based on their learning behaviors is of utmost importance for online educational systems, as it greatly enhances student engagement, motivation, and overall learning outcomes. Knowing the learning preferences of learners may significantly aid in creating personalized learning recommendations and empower learners to identify their own learning styles. However, learners exhibit diverse behaviors in an online setting, which poses significant difficulties in detecting their learning style. This paper proposes a novel approach for detecting learning styles using graph representation learning techniques and machine learning algorithms. While our approach is not reliant on a particular learning style model, our approach may be divided into two distinct parts. Initially, we represented the behavior of learners as a bipartite graph and then transformed this graph into a lower-dimensional representation using the graph embedding approach. This lower-dimensional representation was then utilized for machine learning tasks. Furthermore, we categorize the encoded learner's sequence using a clustering technique based on the selected learning style model. Our methodology uses the Felder-Silverman model as the learning style model and the Fuzzy C-means algorithm as a clustering technique. Our approach was evaluated using the 2015 KDD Cup dataset through a series of comprehensive experiments to showcase its effectiveness. The findings demonstrate that our approach surpasses the previous approach, achieving an average precision of 0.8737 and an accuracy of 0.9182.

Index Terms—**Learning style, online learning environment, Fuzzy C-means algorithm, Graph representation learning, learning behavior.**

I. Introduction

THE efficiency of online learning depends on numerous elements such as the quality of the platform, instructor engagement, student motivation, and the adaptability of the course content. Integrating interactive components, enabling collaboration, and giving proper assistance can dramatically enhance the online learning experience [1]. In recent years, online learning has gained immense popularity, revolutionizing the way education is delivered and accessed. With its flexibility and accessibility, online learning has become a vital tool for learners of all ages and backgrounds. However, despite its numerous advantages, one critical challenge in online learning is the lack of personalized instruction tailored to individual learning styles. Learning styles (LS), remain one of the most important features for delivering personalized learning and have become a core aspect in realizing personalized education [2]. LS can be defined as the description of learners' behavior and attitude, which indicate their preferred approach to learning [3], [4].

Understanding and accommodating diverse LS is crucial in achieving effective education outcomes and has so many advantages as it aids both learners and instructors in the learning process. By determining how a student learns first, educators can modify their teaching approaches to match these preferences based on the student's strengths and weaknesses, which can enhance comprehension and engagement. Furthermore, matching teaching styles to a student's learning preferences can increase information retention, stay engaged and interested in the materials, increase student confidence, provide efficient study strategies, and encourage students to engage with different styles that can make them more adaptive and versatile learners, which is beneficial in various educational settings and real-life scenarios [5], [6]. Therefore, there is a pressing need to develop intelligent algorithms capable of accessing and detecting LS in the online learning environment.

Most existing studies that looked at learning style detection (LSD) utilized either explicit or implicit approaches to determine the learner's preferred method of learning [7], [8], [9]. Finding out how learners learn best is the goal of the explicit approach, which uses tools like surveys, questionnaires, and observations. The approach has many benefits, such as being easy to implement and providing a good overview of students' learning preferences. It offers a quick assessment of student LS, so educators can adapt their method of instruction appropriately, offering valuable insight into students' broad learning preferences, enabling educators to create more diverse teaching strategies, and finally providing a structural framework for identifying LS, some explicit tools have been standardized and validated [7]. Researchers should not oversimplify the complex ways individuals learn by relying solely on a fixed classification of student learning styles; instead, they should supplement the explicit learning style detection approach with observations, feedback, and dynamic teaching strategies that accommodate various LS. This will help them understand learners' preferences. Another thing is that people's learning styles could change as they become older. It is possible that changes in student behavior will go unnoticed by an explicit approach. Additionally, validity and reliability may be lacking

Manuscript received February 29, 2024; revised May 31, 2024. Corresponding authors: Wang Jianping (email: wangjianping@hist.edu.cn), Bello Ahmad Muhammad (email: mbahmad@hist.edu.cn).

in certain surveys and questionnaires. Because of inherent biases, poorly worded questions, or a lack of alternatives, they may fail to capture students' preferred methods of instruction.

To address the limitations mentioned. In this paper, we propose an automatic learning style detection approach through graph representation learning techniques and machine learning algorithms. In our proposed approach, leaner behavior records were modeled as a bipartite graph representing learners' interactions with learning resources and the edge between them represents the frequency and duration of accessing each resource by the learner. With this, we encoded the constructed graph into a low-dimensional vector representation for various downstream machine-learning algorithms using graph representation learning techniques. Next, we provide a mapping between the learning resources and the selected learning style model in order to find the appropriate learning style of the concerned learner. The resulting features obtained were used as the input of the clustering algorithms to cluster learners with similar LSs. To sort the encoded learner's sequences by FSLSM dimension, this study makes use of the Felder-Silverman learning style model (FSLSM) in conjunction with the Fuzzy C-means (FCM) clustering technique.

The rest of the paper is organized as follows: In section II, the literature review of the existing automatic approaches for learning style detection was presented. In section III, we provide an extensive overview of the algorithms that were utilized. We lay out the description of the proposed approach in greater detail in section IV. Section V, we detail the experiment and analysis of the results, Finally, we give a general conclusion and outline potential avenues for future research.

II. Related Work

The explicit approach to identifying a learner's LS is the utilization of specialized questionnaires. While the implementation of this approach is straightforward and provides a foundation for comprehending learners' preferred LSs, it is inherently unchanging, certain questionnaires may lack credibility and accuracy, oversimplifying the intricate ways in which individuals learn, and restricting personalization [6], [9]. The implicit technique is suggested as an alternate approach that relies on analyzing and interpreting the student's behavior, participation, and motivations during the learning process [6]. As stated in [7], [10], the implicit method for detecting learning styles (LSD) can be implemented using either a literature-based or a data-driven approach. The literature-based strategy utilizes the behaviors and inclinations of learners to get an understanding of their learning preferences, and then adopts a straightforward rule-based method to evaluate the learning style by evaluating the quantity of relevant indicators [11], [12]. On the other hand, the data-driven approach relies on the observation of learners' behavior during the learning process and utilizes data mining, machine learning, or deep learning algorithms to analyze these behavior patterns [13]]. The proposed approach in this paper focuses on utilizing a data-driven approach. Therefore, we provide a summary of the current data-driven approaches used for LSD.

Most existing data-driven approaches based their studies on different learning style models and use either clustering, classification, or hybrid approaches to identify and predict the learner's LS [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24]. For instance, [15] uses a K-means clustering algorithm for learning style prediction. The authors mapped the extracted learner's sequences into the eight categories of FSLSM and the resulting features were used as input to the K-means clustering algorithm. [23] uses a graph representation learning techniques to capture the behavior of learners as a bipartite graph and learn the structural and weight information of the graph into a low dimensional representation. The authors employ the use of the K-means algorithm to cluster learners with similar learning styles according to FSLSM categories.

[20] employs the use of a Fuzzy c-means algorithm to predict the learner's preferred method of learning. First, learners' behavior data was collected and analyzed using web usage mining to discover the learner's sequences, then these sequences were mapped into FSLSM and the resulting features were used as input to the Fuzzy c means algorithms to cluster learners according to the eight FSLSM categories. [19] utilizes web mining techniques to extract learners' sequences from their log files. The Fuzzy C means algorithm is then employed to cluster learners who have similar learning styles based on the 16 combinations of FSLSM. [17] presented a three-step method for LS prediction based on a clustering algorithm. First, multidimensional space is constructed using learners' pre-tested LS preference patterns. Secondly, Euclidean distance was used to calculate their similarities. Third, the authors compared the mutual similarity pattern to the pre-tested pattern to get the target LS dimension. After the dimension was determined, a cluster core construction algorithm was used to group learners based on their behavior along this dimension.

[16] suggested a technique that integrates clustering and classification algorithms to discern and forecast individuals' learning styles. The authors employ a K-mean clustering approach to group learners based on their choice for each FSLSM category. The categories were subsequently assigned a quantitative measure of their level of dominance (i.e., very weak, weak, moderate, strong), resulting in the creation of a labeled dataset. Subsequently, four classification algorithms, namely Neural Network (NN), Random Forest (RF), K-Nearest Neighbors (KNN), and Decision Tree (DT) were utilized to forecast Learning Styles (LSs) using traces produced by learners during various course activities. DT demonstrated the highest accuracy among the three dimensions of FSLSM, as indicated by the result. [24] use an FCM algorithm to cluster learners according to the eight categories of FSLSM. Then, the Gravitational Search back propagation Neural network algorithm (GSBPNN) was used to predict the learner's learning style in the real world. The use of graph neural networks to forecast the actual learning style of the learner was suggested by [18]. The authors cluster the encoded sequence of learner behavior according to the FSLSM categories using the k-means clustering algorithm. The GNN algorithm was fed the acquired data in order to forecast the learning style of the incoming or current sequence of learners.

The decision tree is the most popular classification method for learning style prediction. It simplifies a complex decision-making process into a series of smaller decisions and outputs an often easy-to-understand solution [6]. For instance, [25] used a Decision tree as a classification method for learning style prediction to adaptivity adjust interfaces for them. The DT aims to simplify a complex decision-making process into a series of smaller decisions and outputs an often easy-to-understand solution. [26] achieved better prediction results with the J48 algorithm (a DT prune with significant rules) than with the Bayesian Network (BN) classifier.

[27] combines an ant colony system and an artificial neural network to predict learners' learning styles. The innovative solution makes it a valuable contribution to the field of educational technology and personalized learning. [7] compares the performance of four classifiers (Neural network (NN), Genetic Algorithm (GA), Ant Colony, and Particle swarm Optimization) to improve the precision of learning style prediction using a dataset collected from 75 learners. The NN achieves the most promising result. [3] suggested an automated method that utilizes diverse machine-learning algorithms to detect learners' learning styles by analyzing the patterns of their learning behavior in relation to FSLSM. The Support Vector Machine (SVM) has shown superior performance in terms of accuracy compared to other classification algorithms. [5] used a range of machine-learning algorithms to identify the learning styles of learners. The study aims to find characteristics that can be used to detect various bits of intelligence and learning styles, using the Felder-Silverman model as a basis. The results indicated that Support Vector Machine and Logistic Regression attained the highest level of accuracy. The majority of these methodologies evaluate their performance by comparing them to various machine learning algorithms, rather than comparing them to existing methods.

III. Background on the Algorithm

This research proposes to use a graph representation learning technique and a Fuzzy C-means (FCM) algorithm to improve the accuracy and precision of learning style detection approaches. A literature review was conducted in order to select the algorithms for this paper as described in the previous section. Existing research shows that the best accuracy and precision were obtained using graph autoencoder and k-mean clustering algorithm [23]. Hence, in this paper, we use the same procedure to capture and learn the behavior of the learner into low-dimensional vector representations. Then, we employ the use of a FCM algorithm to cluster learners according to the FSLSM dimension. In what follows, the description of learning style, graph representation learning techniques, and Fuzzy C-mean algorithm is provided.

A. *Learning Style*

Learning styles simply refer to a preferred or dominant way by which an individual gets, processes, and retains information [4]. It primarily is an innate approach that an individual has towards understanding how things work and how new material is processed. Learning styles can be anything from sensory modalities (viz., visual, aural, kinesthetic), social aspects (group or individual learning), environmental factors (the setting where one prefers to learn), and cognitive strategies (differences in analytical vs. intuitive approaches to solving problems) [6].

Individuals and educators, therefore, can be of the ability to tailor design teaching methods as well as materials based on understanding what learning style one identifies with in order to impact the effort maximally. But at the same time, one should keep in mind the fact that the idea of learning styles has been controversial in educational psychology and some research reveals that the effects of explicit instruction of teaching by individual learning styles may be limited upon enhancing learning outcomes. However, recognizing and accommodating different learning style preferences can provide a more varied and inclusive educational environment [2].

In the learning style detection and prediction process, a learner model is one of the most important components in an online educational system, because of its ability to represent the behavior of learners according to which online educational system provides personalized learning resource recommendations [6].

Various learning style models have been utilized to understand and detect learning styles. These include the VARK model, Kolb's experimental learning theory, Honey and Munford's learning styles, Gardner's theory of multiple intelligence, and Felder and Silverman's learning style (FSLSM) among others [10], [11]. In this paper, we utilized the use of Felder and Silverman learning style models for several reasons. First, the model has gained recognition and applicability in online education systems because it offers insight into the diverse learning preferences among learners. The model doesn't prescribe a one-size-fits-all strategy. Instead, it recognizes that individuals may have preferences across a spectrum in each dimension. This flexibility allows educators to adapt teaching methods to accommodate a range of learning styles. Third, the model aims to enhance learning engagement, understanding, and retention of information. Fourth, the model encourages both learners and instructors to be aware of their own learning preferences. This self-awareness can prompt individuals to reflect on their learning approaches and potentially adapt strategies that suit their learning styles. Finally, the model takes into account multiple dimensions of learning style including perception (Sensory/Intuitive), input (visual/verbal), processing (active/reflective), and understanding (sequential/global). This multidimensional approach provides a more nuanced understanding of how individuals prefer to learn and suggests ways to cater to these varied styles for more effective learning experiences [4], [6], [16], [23], [28]. Below is a detailed description of the FSLSM dimensions.

Perception Dimension: This dimension of FSLSM involves how learners prefer to take the information. Sensory Learners are the type of learners that prefer concrete examples, practical applications, and real-life case studies. Intuitive learners prefer conceptual frameworks, Creative problem-solving exercises, and theoretical information.

Input Dimension: The dimension here focuses on how learners process information. The dimension deals with visual and verbal learners. Visual learners prefer visual aids like videos, diagrams, infographics, and charts and Verbal learners prefer text-based resources, written and spoken explanations, forum discussions, and lectures.

Processing Dimension: This dimension relates to how learners internalize and understand information. The dimension takes into account active and reflective learners. Active learners learn better through experiments that engage learners in hands-on activities, group discussions, quizzes, and problem-solving. Reflective learners prefer to think quietly and take their time to process information, reflective journals, or opportunities for self-paced learning.

Understanding Dimension: This dimension deals with how learners make sense of information. Learners can either be sequential learners or global learners. Sequential learners may provide structured courses with clear step-by-step instructions, and logical progression and build upon previous knowledge. Global learners might incorporate big-picture concepts, interdisciplinary connections, and non-linear learning paths to understand the overall concepts before delving into details.

FSLSM suggests that individuals possess preferences along these dimensions and are typically stronger in one preference over the other in each dimension. Most educational systems and instructors can use this model in order to create diverse teaching strategies that cater to various learning styles, thereby enhancing the learning experience of learners with different preferences.

B. Graph Representation Learning

Graph representation learning (GRL) techniques have gained significant attention in recent years due to their ability to capture complex relationships and dependencies in data. People generally find it easier to interpret information when it is presented graphically rather than in tabular form. These will allow the systems to automatically identify the necessary representation from large network data for automatic feature engineering and classification and it has been widely applied in various domains including social networks, recommender systems, and knowledge graphs [29].

The GRL is a subfield of machine learning that specifically deals with the task of extracting meaningful and valuable representations of data that are structured in the form of graphs [30]. This entails converting the complex and interrelated connection within a graph into a structure that is appropriate for diverse machine-learning tasks. Graphs are composed of nodes and edges, which make them a suitable form for describing intricate linkages and interactions. The objective of the GRL is to enhance machine learning models with the capability to understand and explore the abundant structural information inherent in graph-structured data for various applications [30], [31]. It aims at converting nodes and edges into a low-dimensional, uninterrupted vector representation commonly referred to as embeddings. These embeddings efficiently capture the graph's structural and relational information, enhancing the performance of downstream machine-learning tasks like clustering, classification, and link prediction [32], [33].

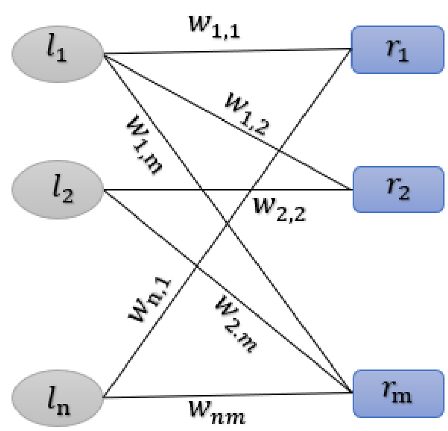

Figure 1. An example of an L-R bipartite graph.

In the context of LSD in online education, GRL techniques present a potential approach for extracting significant features from learner behavior data. By portraying learners and their interactions with learning resources as nodes and edges in a graph, these techniques can capture the data's underlying structure and patterns, allowing the detection of LS. For instance, [23] proposed the use of bipartite graphs to simulate the interactions between learner (L) and Learning resources (R) in online education platforms. The GRL module leverages this bipartite graph to extract the latent representation of learners and resources, which is subsequently utilized to detect LS as seen in Fig. 1. The latent representations obtained from graph representation learning can then be used to identify and group learners with similar learning methods using clustering algorithms. This clustering procedure groups nodes (learners) according to their strong connections throughout the graph, providing insights into various LSs.

C. Fuzzy C- means Algorithm

The Fuzzy C-means clustering algorithm is an effective unsupervised learning algorithm developed by Dunn that can be used in data mining, machine learning, pattern recognition, etc [34]. Various Scholars have applied the FCM algorithm to several fields, such as image processing, Natural language processing, scheme ranking, data mining, and information retrieval among others. As the FCM has been widely applied to address the limitations of the existing learning style identification, some scholars have utilized the use of FCM algorithm to deal with conflicts. [19] utilizes the use of the FCM algorithm to detect learning styles based on the 16 combinations of the FSLSM model. [20] uses web mining techniques to capture the behavior of learners, then uses the FCM algorithm to cluster the captured learning behavior according to the FSLSM categories. While the majority of current methods employing the FCM algorithm are effective in detecting LS, they rely on either web-based mining or manual feature engineering on unlabeled behavioral datasets to extract features, leaving the potential for enhancement.

In this paper, we utilize the use of the FCM algorithm on the encoded learner's sequence for effective learning style detection. This is due to the following reasons: 1) FCMs have the ability to perform soft clustering, unlike the traditional K-means clustering by allowing nodes in the graph to have partial memberships in multiple clusters, reflecting the inherent fuzziness in community boundaries within graphs. 2) Its soft assignment nature enables nodes to belong to multiple clusters, accommodating the overlapping nature of communities in the graph. 3) FCMs has the ability to accommodate weight graphs where edges have different strengths or weights. This flexibility allows the algorithm to consider edge weight in determining node relationships and cluster assignment, capturing the importance of connections between nodes. 4) FCMs can reveal the underlying structure or communities within the graph by assigning nodes to clusters based on their connection patterns. 5) it has the advantage of providing a flexible approach for partitioning graph data into meaningful groups, and finally, it can help to identify groups of learners with similar interests or behavior, facilitating targeted recommendations or community-based analysis [34], [35].

IV. The Proposed Approach

This paper presents a novel strategy that utilizes graph representation learning techniques and machine learning algorithms to construct a model capable of detecting learners' learning styles in online education. The initial stage of the graph representation learning approach procedure involves gathering the recorded actions of the learner from the online platform's weblog. then, the behaviors undergo preprocessing and are then represented as a bipartite graph. The structural information and weight relations of the graph are then transformed into a fixed-length vector representation, also known as an embedded vector using graph autoencoder. This embedded vector is utilized as input to a clustering algorithm. The embedded vector consists of a low-dimensional representation of the group of learners, the group of learning resources, and their weighted relationships, which are unlabeled [23].

The focus of this paper is to explore the use of the embedded vector for accurately identifying and grouping learners who share similar learning styles. In order to identify the LS of a student, we must first discover the LS preferences associated with each learning resource, as the embedded vectors are comprised of the set of learning resources used by learners during the learning process. The learning resources were correlated with learning style using the FSLSM method. Subsequently, a clustering algorithm was employed on the correlation-resulting features to discover and group learners who have similar learning styles.

In the following part, we will demonstrate the process of associating each learning resource with its corresponding LS model and subsequently describe how the clustering algorithm was utilized in our approach to detect LS. The proposed approach is depicted in Fig. 2.

A. Mapping Learning resources (LR) onto FSLSM dimension

Online learning environments consist of multiple learning subjects and activities that make up the courses. Each subject consists of learning content that is classified as learning resources. In order to categorize the learning resources based on FSLSM, it is assumed that each learner possesses a preferred learning style, which may be measured across four dimensions. Learning resources (LR) consist of diverse digital resource nodes, including text, image, video, audio, discussion forums, wiki, problems, etc. These resources are distinguished by their capacity for reusability, interoperability, durability, and accessibility. Here, the learning resources must be categorized based on the individual preferences of each learner. Each learning resource is associated with a corresponding FSLSM dimension, as described in the literature by [8], in order to identify the learning style of learners based on the encoded elements of their behavior. The features obtained from each FSLSM dimension were subsequently utilized as input for the FCM method in order to detect and classify the LSs of the learners.

B. Learning style detection using Fuzzy-C-mean Algorithm

FCM algorithm was used in this paper to group the encoded learning sequence according to the FSLSM categories and obtained labels [35]. The grouping is achieved by considering the encoded learning resources and the feature value for each FSLSM category as shown in Fig. 2. The algorithm considers the set of all learners in the encoded space as well as the embedded feature related to each FSLSM dimension as input.

The algorithm starts by initializing random membership for each sequence. It then converts the encoded graph into an adjacency matrix for each FSLSM dimension, which is used as input for the algorithm. The algorithm was implemented on the adjacency matrix. It iteratively updates the membership matrix and cluster centers until convergence or reaches the maximum number of iterations using the following equations.

$$d_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\|X_i - C_j\|}{\|X_i - C_k\|} \right)^{\frac{2}{m-1}}} \tag{1}$$

$$c_j = \frac{\sum_{i=1}^{n} d_{ij}^{m} \times X_i}{\sum_{i=1}^{n} d_{ij}^{m}} \tag{2}$$

where d_{ij} is the degree membership of X_i in the cluster C_j, X_i is the i^{th} of the d-dimensional measure data and C_j is the dimension center of the cluster. The algorithm is run minimizing the following objective function.

$$O_m = \sum_{i=1}^{n} \sum_{k=1}^{c} d_{ij}^{m} X_i - C_i^2, \quad 1 \leq m \leq \infty \tag{3}$$

The FCM algorithm categorizes learners' learning styles by selecting center values. The center values are determined by computing the feature values and assigning each value to a cluster. We utilize tools such as NetworkX for graph manipulation and implementation in the FCM method.

C. Evaluation Metrics

Evaluation metrics are essential to assess the performance and quality of clustering algorithms, including K-means and

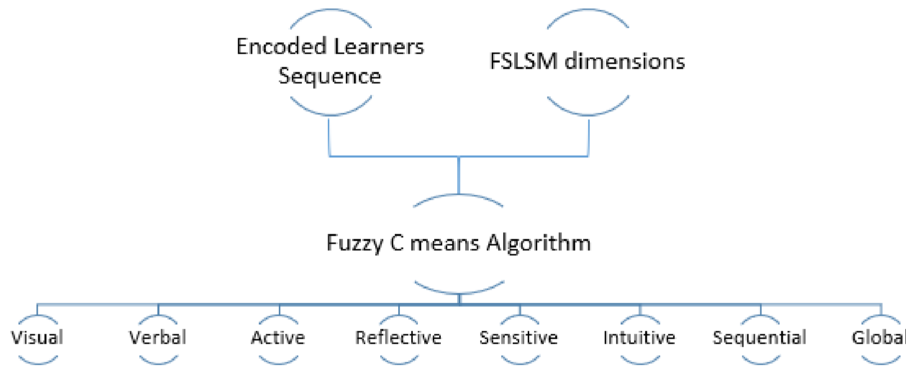

Figure 2. The proposed model for learning style detection w.r.t FSLSM categories.

fuzzy C-means (FCM), especially when applied to graph data. The proposed approach was evaluated using the 2015 KDD cup data set. We used accuracy, precision, recall, and F1 measures as evaluation matrices to evaluate the effectiveness of the proposed approach. These matrices are chosen as the most commonly used confusion matrix techniques to assess the quality of the classifier. Detailed descriptions of the metrics are discuss below.

Precision (P): It denotes the ratio of correctly identified learner's sequences with positive predictions to the total number of projected learner's sequences with positive class values. The Precision is determined using Eq. 4.

$$P = \frac{TP}{TP + FP} \tag{4}$$

Recall (R): The accuracy rate is the ratio of correctly detected positive predictions to the total number of positive class values correctly classed as positive. The recall is computed using the Eq. 5.

$$R = \frac{TP}{TP + FN} \tag{5}$$

Accuracy (A): The performance metric is highly intuitive and represents the accuracy of the approach in accurately classifying the projected learner's sequences. The accuracy is determined by applying Eq. 6 as shown below.

$$A = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

F1-Meassure (F1_score): The F1 score, commonly known as the harmonic mean of precision and recall, quantifies the accuracy of a model. The F1 score is computed using Eq.7.

$$F1score = \frac{2 \times (P \times R)}{P + R} \tag{7}$$

V. Evaluation of the Proposed Approach

The subsequent section outlines the materials and methodologies employed to assess the suggested approach. First, we outline the data utilized and the preparation phase, taking into account the dataset's reliability in assessing the suggested approach. Subsequently, the experimental findings from the evaluation are presented, together with a thorough analysis and discussion of the results and observations encountered throughout the development of the proposed approach.

A. Data Description

The data set used to evaluate the effectiveness of our proposed approach is the same dataset used for evaluating GRL-LS [23]. The learner's behavior data was collected from the KDD-Cup 2015 dataset. The KDD-Cup is the first and among the most prestigious competitions in data science. The KDD Cup is an annual data mining and knowledge discovery competition organized by the ACM Special Interest Group on Knowledge Discovery and Data Mining.

The dataset contains a record of events documenting user activities and interactions with learning resources. A total of 5069 students from 40 courses were included in this study. Initially, we utilize the learner's enrollment information to convert the dataset into a bipartite graph that depicts the collection of learners, the collection of learning materials, and their weighted connections. Subsequently, a graph autoencoder was employed to extract and encode the built graph into a low-dimensional representation for further machine-learning tasks.

B. Experiments and results analysis

In order to assess the efficacy of our proposed methodology, we conducted experiments utilizing behavioral data extracted from the 2015 KDD Cup dataset. The paper utilizes the embedding outcome acquired from [23], which encompasses the set of learners and learning resources in the low-dimensional space (latent representation) with their weight relations. The latent representation identifies and organizes the learner's sequences based on GRL approaches. The learner's sequences lack labels and contain information regarding 5069 sets of learners and 7 sets of learning resources.

Initially, the learning resources were aligned with the FSLSM dimensions. The resulting characteristics were then utilized as the input for the FCM algorithm, which generated labels based on the FSLSM categories. This allowed for the clustering of learners who possess comparable learning styles. For each dimension, we ascertain the number of clusters $(K = 2)$ and calculate the cluster value for each cluster

Table I. Accuracy and precision results of the clustering algorithm

Max. Iteration	Accuracy				Precision			
	Vis./Ver	Act./Ref.	Sen./Int.	Seq./Glo.	Vis./Ver	Act./Ref.	Sen./Int.	Seq./Glo.
100	0.9100	0.9480	0.9080	0.9071	0.9056	0.9475	0.8100	0.8290
200	0.8600	0.9339	0.9140	0.9041	0.8700	0.8703	0.8600	0.9041

based on the features' values. The membership matrix was established, with each element denoting the data point's degree of membership to each cluster. The learner's sequences are categorized into eight groups of FSLSM, which are defined as Active, Reflective, Visual, Verbal, Sensory, Intuitive, Sequential, and Global. We demonstrate and examine the result of the proposed approach using the 2015 KDD Cup dataset on the clustering outcome using two widely used evaluation metrics: precision and accuracy. The ultimate clusters are derived after 100 and 200 iterations and Table 1 displays the accuracy and precision results. The results demonstrate that the suggested method achieves a high degree of accuracy and precision for every dimension of the FSLSM.

The experiment is conducted with the aim of achieving the following objectives. The main goal of the experiment is to showcase the effectiveness of our suggested approach in leveraging graphs to represent learners' activity and utilize this data to determine their learning styles. Moreover, we establish that the proposed method outperforms the current method in terms of accuracy and precision.

Table II. Accuracy and precision performance comparison with existing approach

FSLSM dimension	Algorithm	Accuracy	Precision
Visual/Verbal	K-means	0.9100	0.8300
	FCM	0.9100	0.9056
Active/Reflective	K-means	0.9300	0.8700
	FCM	0.9480	0.9475
Sensitive/Intuitive	K-means	0.8500	0.7300
	FCM	0.9080	0.8100
Sequential/Global	K-means	0.8400	0.7100
	FCM	0.9071	0.8290

Table 2 displays the performance comparison between the results generated using the k-means method. Based on the results, it is evident that the utilization of the Fuzzy C-means method outperforms K-means clustering in every FSLSM dimension. Among the existing approaches, [36] demonstrates superior precision performance for sensitive/Intuitive with a rate of 0.7730. Additionally, [26] achieves better performance for sequential/global with a precision rate of 0.8000. However, our approach surpasses both of these existing approaches, with a precision rate of 0.8100 for sensitive/intuitive and a precision rate of 0.8290 for sequential/global. The utilization of the fuzzy c-means method for clustering learners' learning styles based on the embedding findings yields superior performance, with an average accuracy and precision of 0.9182 and 0.8737, respectively.

The exceptional performance of the acquired result in this research demonstrates the efficacy of the suggested technique for detecting learning styles. The utilization of the FCM algorithm has significantly improved the performance of this technique by efficiently identifying and clustering the encoded learner's behavior gained through the usage of a graph autoencoder. The precision and recall achieved by the technique showed that it is feasible and effective in detecting learning styles, in comparison to the k-means clustering algorithm. The findings of the research show an emphasis on tailoring educational strategies to meet the needs of learners aligned with the current trends in educational research and practices. Moreover, the proposed approach aims to aid teachers and institutions in designing and delivering personalized learning resource suggestions that cater to diverse learning styles, ultimately enhancing the effectiveness of online education. This will enable learners to gain a deeper understanding of their LS to improve their learning outcomes. Lastly, the proposed approach is designed to be adaptable to multiple educational systems and may be implemented across various LS models.

VI. Conclusion

This work presents a refined method for detecting LSs by utilizing the Graph representation learning techniques and FCM algorithm. The objective is to boost the precision of the current methodology. Initially, the activity of learners was recorded using a graph-based method to acquire a low-dimensional representation of the set of learners, the set of learning resources, and the weight relations between them. Next, the research aims to categorize the encoded learner's sequences based on the FSLSM categories using the FCM method. The proposed approach is versatile and can be used in different educational systems and theories of learning styles. The experimental findings demonstrate that the proposed methodology surpasses the conventional K-means algorithm, resulting in enhanced precision and accuracy. In future research, we will explore other methods of graph representation learning in order to enhance the precision and accuracy of learning style recognition. Additionally, we will evaluate the proposed strategy using a substantial dataset including a greater number of learning resources.

Acknowledgment

This work was partly supported by the Key Scientific and Technological Project of Henan Province (232102111128, 222102320181, 212102310087), in part by the Major Special Project of Xinxiang City(21ZD003), in part by the Key Scientific Research Projects of Colleges and Universities in Henan Province (23B520003, 21A520001), in part by the Henan Province Postdoctoral Support Program(HN2022165). The authors approved the version of the manuscript to be published. They agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

REFERENCES

[1] S. G. Essa, T. Celik, and N. E. Human-Hendricks, *"Personalized Adaptive Learning Technologies Based on Machine Learning Techniques to Identify Learning Styles: A Systematic Literature Review,"* IEEE Access, vol. 11, pp. 48392–48409, 2023, doi: 10.1109/ACCESS.2023.3276439.

[2] B. A. Muhammad, C. Qi, Z. Wu, and H. K. Ahmad, "An evolving learning style detection approach for online education using bipartite graph embedding," Appl Soft Comput, vol. 152, Feb. 2024, doi: 10.1016/j.asoc.2024.111230.

[3] Z. Mehenaoui, Y. Lafifi, and L. Zemmouri, "Learning Behavior Analysis to Identify Learner's Learning Style based on Machine Learning Techniques," Journal of Universal Computer Science, vol. 28, no. 11, pp. 1193–1220, 2022, doi: 10.3897/jucs.81518.

[4] A. Khamparia and B. Pandey, "Association of learning styles with different e-learning problems: a systematic review and classification," Educ Inf Technol (Dordr), vol. 25, no. 2, pp. 1303–1331, 2020, doi: 10.1007/s10639-019-10028-y.

[5] F. Rasheed and A. Wahid, "Learning style detection in E-learning systems using machine learning techniques," Expert Syst Appl, vol. 174, no. February, p. 114774, 2021, doi: 10.1016/j.eswa.2021.114774.

[6] M. Raleiras, A. H. Nabizadeh, and F. A. Costa, "Automatic learning styles prediction: a survey of the State-of-the-Art (2006–2021)," Journal of Computers in Education, vol. 9, no. 4, pp. 587–679, Dec. 2022, doi: 10.1007/s40692-021-00215-7.

[7] J. Bernard, T. W. Chang, E. Popescu, and S. Graf, "Learning style Identifier: Improving the precision of learning style identification through computational intelligence algorithms," Expert Syst Appl, vol. 75, pp. 94–108, 2017, doi: 10.1016/j.eswa.2017.01.021.

[8] B. A. Muhammad, Z. Wu, and H. K. Ahmad, "A Conceptual Framework for Detecting Learning Style in an Online Education Using Graph Representation Learning," in Proceedings - 2020 International Conference on Networking and Network Applications, NaNA 2020, 2020, pp. 136–140. doi: 10.1109/NaNA51271.2020.00031.

[9] J. Feldman, A. Monteserin, and A. Amandi, "Automatic detection of learning styles: state of the art," Artif Intell Rev, vol. 44, no. 2, pp. 157–186, 2015, doi: 10.1007/s10462-014-9422-6.

[10] S. Graf, Kinshuk, and T. C. Liu, "Supporting Teachers in Identifying Students' Learning Styles in Learning Management Systems: An Automatic Student Modelling Approach," Educational Technology and Society, vol. 12, no. 4, pp. 3–14, 2009.

[11] I. Katsaris and N. Vidakis, "Adaptive e-learning systems through learning styles: A review of the literature," Advances in Mobile Learning Educational Research, vol. 1, no. 2, pp. 124–145, 2021, doi: 10.25082/amler.2021.02.007.

[12] E. Okewu, P. Adewole, S. Misra, R. Maskeliunas, and R. Damasevicius, "Artificial Neural Networks for Educational Data Mining in Higher Education: A Systematic Literature Review," Applied Artificial Intelligence, vol. 35, no. 13, pp. 983–1021, 2021, doi: 10.1080/08839514.2021.1922847.

[13] N. R. Ngatirin and Z. Zainol, "Automatic Detection of Learning Styles: A Decade Review on Data-driven Approaches," in Journal of Physics: Conference Series, IOP Publishing Ltd, Aug. 2021. doi: 10.1088/1742-6596/1997/1/012001.

[14] M. Raleiras et al., "A robust classification to predict learning styles in adaptive E-learning systems," Educ Inf Technol (Dordr), vol. 25, no. 2, pp. 437–448, 2020, doi: 10.1007/s10639-019-09956-6.

[15] N. Hidayat, R. Wardoyo, A. Sn, and H. D. Surjono, "Enhanced performance of the automatic learning style detection model using a combination of modified K-means algorithm and Naive Bayesian," International Journal of Advanced Computer Science and Applications, vol. 11, no. 3, pp. 638–648, 2020, doi: 10.14569/ijacsa.2020.0110380.

[16] B. Hmedna, A. El Mezouary, and O. Baz, "A predictive model for the identification of learning styles in MOOC environments," Cluster Comput, vol. 23, no. 2, pp. 1303–1328, 2020, doi: 10.1007/s10586-019-02992-4.

[17] J. Yang, Z. X. Huang, Y. X. Gao, and H. T. Liu, "Dynamic learning style prediction method based on a pattern recognition technique," IEEE Transactions on Learning Technologies, vol. 7, no. 2, pp. 165–177, 2014, doi: 10.1109/TLT.2014.2307858.

[18] B. A. Muhammad, B. Liu, H. K. Ahmad, M. Umar, and K. Qin, *"GNN-LS: A Learning Style Prediction in Online Environments using Graph Neural Network,"* Journal of Networking and Network Applications, vol. 2, no. 4, pp. 172–182, 2022, doi:10.33969/J-NaNA.2022.020405.

[19] O. El Aissaoui, Y. El, A. El, L. Oughdir, and Y. El Allioui, "A fuzzy classification approach for learning style prediction based on web mining technique in e-learning environments," Journal of Education and Information Technologies, vol. 24, pp. 1943–1959, 2019.

[20] I. Azzi, A. Jeghal, A. Radouane, A. Yahyaouy, and H. Tairi, "A robust classification to predict learning styles in adaptive E-learning systems," Educ Inf Technol (Dordr), vol. 25, no. 1, pp. 437–448, 2020, doi: 10.1007/s10639-019-09956-6.

[21] O. El Aissaoui, Y. E. A. El Madani, L. Oughdir, and Y. El Allioui, "Integrating web usage mining for an automatic learner profile detection: A learning styles-based approach," in 2018 international conference on intelligent systems and computer vision (ISCV), 2018, pp. 1–6.

[22] A. S. Aziz, R. A. El-Khoribi, and S. A. Taie, "AFCM MODEL TO PREDICT THE LEARNER STYLE BASED ON QUESTIONNAIRE AND FUZZY C MEAN ALGORITHM," J Theor Appl Inf Technol, vol. 31, no. 2, 2021, [Online]. Available: www.jatit.org

[23] B. Ahmad Muhammad, C. Qi, Z. Wu, and H. Kabir Ahmad, "GRL-LS: A learning style detection in online education using graph representation learning," Expert Syst Appl, vol. 201, no. June 2021, p. 117138, 2022, doi: 10.1016/j.eswa.2022.117138.

[24] S. V. Kolekar, R. M. Pai, and M. M. Manohara Pai, "Prediction of Learner's Profile Based on Learning Styles in Adaptive E-learning System," International Journal of Emerging Technologies in Learning, vol. 12, no. 6, pp. 31–51, 2017, doi: 10.3991/ijet.v12i06.6579.

[25] H. J. Cha, Y. S. Kim, S. H. Park, T. B. Yoon, Y. M. Jung, and J. H. Lee, *"Learning styles diagnosis based on user interface behaviors for the customization of learning interfaces in an intelligent tutoring system,"* Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 4053 LNCS, pp. 513–524, 2006, doi: 10.1007/11774303-51.

[26] I. Karagiannis and M. Satratzemi, "An adaptive mechanism for Moodle based on automatic detection of learning styles," Educ Inf Technol (Dordr), vol. 23, no. 3, pp. 1331–1357, 2018, doi: 10.1007/s10639-017-9663-5.

[27] J. Bernard, E. Popescu, and S. Graf, "Improving online education through automatic learning style identification using a multi-step architecture with ant colony system and artificial neural networks," Appl Soft Comput, vol. 131, p. 109779, 2022, doi: 10.1016/j.asoc.2022.109779.

[28] E. Özpolat et al., *"Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities,"* Comput Educ, vol. 9, no. 3, pp. 1–16, 2020, doi: 10.1016/j.tsc.2015.04.003.

[29] I. Makarov, D. Kiselev, N. Nikitinsky, and L. Subelj, *"Survey on graph embeddings and their applications to machine learning problems on graphs,"* PeerJ Comput Sci, vol. 7, pp. 1–62, 2021, doi: 10.7717/peerj-cs.357.

[30] W. L. Hamilton, R. Ying, and J. Leskovec, *"Representation Learning on Graphs: Methods and Applications,"* in Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, 2017, pp. 1–23. [Online]. Available: http://arxiv.org/abs/1709.05584

[31] F. Chen, Y. C. Wang, B. Wang, and C. C. J. Kuo, *"Graph representation learning: A survey,"* APSIPA Trans Signal Inf Process, vol. abs/1909.0, pp. 1–17, 2020, doi: 10.1017/ATSIP.2020.13.

[32] M. Seo and K. Y. Lee, *"A Graph Embedding Technique for Weighted Graphs Based on LSTM Autoencoders,"* Journal of Information Processing Systems, vol. 16, no. 6, pp. 1407–1423, 2020, doi: 10.3745/JIPS.04.0197.

[33] E. Giamphy, J. L. Guillaume, A. Doucet, and K. Sanchis, *"A survey on bipartite graphs embedding,"* Social Network Analysis and Mining, vol. 13, no. 1. Springer, Dec. 01, 2023. doi: 10.1007/s13278-023-01058-z.

[34] C. Troussas, A. Krouska, C. Sgouropoulou, and I. Voyiatzis, *"Ensemble learning using fuzzy weights to improve learning style identification for adapted instructional routines,"* Entropy, vol. 22, no. 7, 2020, doi: 10.3390/E22070735.

[35] S. E. Hashemi, F. Gholian-Jouybari, and M. Hajiaghaei-Keshteli, *"A fuzzy C-means algorithm for optimizing data clustering,"* Expert Syst Appl, vol. 227, Oct. 2023, doi: 10.1016/j.eswa.2023.120377.

[36] S. Graf, T. Liu, S. Graf, and T. Liu, *"Supporting Teachers in Identifying Students' Learning Styles in Learning Management Systems: An Automatic Student Modelling Approach,"* International Forum of Educational Technology & Society, vol. 12, no. 4, 2016.

Bello Ahmad Muhammad received his B.Sc. and M.Sc. degrees in computer science from Bayero University, Kano, Nigeria, in 2010 and 2015, respectively, and received his Ph.D. degree in Computer Software Theory from Shaanxi Normal University, China. He is currently a postdoctoral researcher in the School of Computer Science and Technology, Henan Institute of Science and Technology, China. His research interests include Learning style detection, Graph representation learning, Learning resources recommendation, and Underwater Object Detection.

Wang Jianping is currently a Professor with the School of Computer Science and Technology, Henan University of Science and Technology. His research interests include underwater acoustic sensor networks, software-defined networks, and wireless sensor networks.

Guohong Gao is currently a Professor with the School of Computer Science and Technology, Henan University of Science and Technology. His research interest includes wireless sensor networks.

Chenping Zhao received the B.S. and M.S. degrees from the School of Mathematics and Statistics, Henan University, Kaifeng, in 2003 and 2006, respectively, and the Ph.D. degree in applied mathematics from the School of Mathematics and Statistics, Xidian University, Xi'an, in 2018. She is currently a Teacher at the School of Computer Science and Technology, Henan Institute of Science and Technology. Her research interests include variation methods and optimization algorithms for image processing.

Qian Li is currently a Teacher with the School of Information Engineering, Henan Institute of Science and Technology. Her main research interest includes big data.

Xiangang Zuo is currently a Teacher with the School of Computer Science and Technology, Henan Institute of Science and Technology.

Yingying Lv is currently a Teacher with the School of Computer Science and Technology, Henan Institute of Science and Technology. Her research interest includes computer science and technology.