Interdisciplinary Journal of Information, Knowledge, and Management

An Official Publication of the Informing Science Institute
InformingScience.org

IJIKM.org

Volume 21, 2026

DEEP EMBEDDED CLUSTERING AND STATISTICAL VALIDATION FOR STUDENT MODALITY PROFILING

Elijah Ofori*	Department of ICT Education, University of Education, Winneba, Ghana	elijah.ofori@yahoo.com
Delali Kwasi Dake	Department of ICT Education, University of Education, Winneba, Ghana	dkdake@uew.edu.gh

* Corresponding author

ABSTRACT

Aim/Purpose	This study develops a theory-informed, data-driven framework for identifying student learning modalities using Deep Embedded Clustering (DEC), with the goal of improving scalable and interpretable learner profiling in secondary education.
Background	Traditional learning-style identification relies on qualitative or rule-based categorization that may not adequately capture complex learner behaviours. Although deep learning enables richer pattern discovery, many approaches lack interpretability for practical educational use. DEC offers a hybrid alternative by combining representation learning with unsupervised clustering.
Methodology	A 48-item questionnaire grounded in the Felder–Silverman Learning Style Model (FSLSM) was administered to senior high school students. After preprocessing and controlled, distribution-preserving augmentation, an autoencoder learned compact latent representations. Clustering was then applied to derive learner profiles, followed by statistical validation and comparative classification experiments to assess structural and predictive utility.
Contribution	This study proposes an integrated deep clustering and statistical validation pipeline that produces pedagogically interpretable learner profiles without relying on predefined labels. It extends prior work by demonstrating how unsupervised representation learning can reveal theory-consistent learning patterns.

Accepting Editor Natasha Boskic │ Received: December 26, 2025 │ Revised: February 9, February 11, 2026 │ Accepted: February 12, 2026.

Cite as: Ofori, E., & Dake, D. K. (2026). Deep embedded clustering and statistical validation for student modality profiling. *Interdisciplinary Journal of Information, Knowledge, and Management, 21,* Article 6. https://doi.org/10.28945/5724

Findings	Four interpretable student profiles emerged, aligning with FSLSM dimensions and differentiated primarily by multivariate response patterns rather than isolated mean differences. Although internal validity indices indicated moderate separation, models trained on latent representations consistently outperformed those trained on raw features, demonstrating the discriminative value of learned embeddings.
Recommendations for Practitioners	The identified profiles can support differentiated instruction, including structured visual materials, interactive learning tasks, blended strategies, and targeted scaffolding. Such profiling may enhance adaptive learning systems and early intervention strategies.
Recommendations for Researchers	Future studies should validate findings using multi-site, non-augmented datasets and examine cluster stability across cohorts. Lightweight DEC architectures and longitudinal evaluations of personalization strategies warrant further exploration.
Impact on Society	By enabling scalable, data-driven personalization, this framework may reduce reliance on subjective assessments and contribute to more equitable and responsive learning environments.
Future Research	Further work should explore real-time integration within learning management systems, adaptive feedback mechanisms, and ethical considerations surrounding large-scale learner profiling.

Keywords	deep embedded clustering (DEC), learning modalities, artificial intelligence in education, personalized learning, machine learning, classification, adaptive learning

INTRODUCTION

In contemporary education, understanding how students process and engage with learning materials is central to designing effective instructional strategies and fostering personalized learning experiences. Traditionally, attempts to categorize learners have relied on qualitative evaluations, rule-based classifications, or static survey instruments. While such approaches provide useful insights, they often fail to capture the complexity, variability, and multidimensional nature of student learning behaviours in modern digital environments (Lokare & Jadhav, 2024; Pardamean et al., 2022). As educational systems increasingly integrate online platforms and data-rich learning management systems, there is growing interest in leveraging Artificial Intelligence (AI) and Machine Learning (ML) to model student characteristics in a more adaptive and scalable manner (Abrar et al., 2025; Altamimi et al., 2022; Dake et al., 2023).

Within this landscape, clustering and classification techniques have been widely used to segment students based on cognitive and behavioural traits (Lailiyah et al., 2019; Pardamean et al., 2022). Methods such as K-means, hierarchical clustering, support vector machines (SVMs), and decision trees (DTs) have demonstrated utility in identifying patterns within predefined feature spaces (Dake et al., 2023; Tzenios, 2020). However, these approaches face notable limitations. Traditional clustering algorithms are sensitive to noise and struggle with high-dimensional data, often producing rigid partitions that inadequately reflect the dynamic nature of learning processes (Hasibuan et al., 2019; Pasina et al., 2019). Similarly, supervised classification models rely heavily on labelled datasets and predefined categories, limiting their ability to uncover latent or emergent structures in learner data (Hashmi et al., 2025).

Deep learning approaches, including convolutional neural networks (CNNs) and autoencoders, have shown promise in capturing complex nonlinear relationships in educational data (Xie et al., 2021;

Yang et al., 2014). Nevertheless, these models often require substantial computational resources and large datasets, and they are frequently criticized for limited interpretability in pedagogical contexts (Babić, 2017; Xie et al., 2021). As a result, a gap remains between powerful algorithmic modelling and educationally meaningful interpretation.

This study addresses that gap by proposing a hybrid framework that integrates Deep Embedded Clustering (DEC) with statistical validation techniques to profile student learning preferences in a theoretically informed manner. In this study, the term "student learning profiling" refers to empirically derived groupings of students based on patterns of responses aligned with the Felder-Silverman Learning Style Model (FSLSM). Rather than assuming rigid or deterministic "learning styles," this work conceptualizes modalities as probabilistic preference patterns that may inform adaptive instruction. By combining deep representation learning with statistical significance testing (ANOVA, Chi-square analysis, and cluster validation metrics), the proposed approach seeks to enhance both the technical robustness and interpretability of student profiling.

Specifically, this study is guided by the following research questions:

1. How can DEC be implemented to group students based on their learning preferences, and how do these clusters compare with traditional clustering methods?
2. What are the distinctive characteristics of each identified student cluster, and are these characteristics statistically significant across clusters?
3. How effective are the identified clusters based on Silhouette Score, Cluster Purity, and statistical validation metrics?
4. How does a classification model trained on DEC-derived representations compare with traditional machine learning classifiers trained on original statistical features?
5. What personalized instructional strategies can be recommended for different student clusters identified through deep and statistical analyses?

By integrating nonlinear representation learning with rigorous statistical validation, this research contributes to educational data mining and learning analytics in three ways. First, it demonstrates how deep clustering can uncover latent learner structures without predefined labels. Second, it strengthens interpretability by validating cluster distinctions using statistical tests and theoretical alignment with FSLSM dimensions. Third, it provides a scalable methodological pathway for adaptive learning systems, particularly in resource-constrained educational contexts. In doing so, this study advances ongoing discussions at the intersection of AI-driven analytics and pedagogical personalization, offering a framework that bridges computational modelling with actionable educational insights.

LITERATURE REVIEW

The integration of AI and ML has significantly transformed how researchers and educators model and understand student learning processes. Traditional approaches to identifying learning styles have largely relied on rule-based classifications or qualitative categorizations (Subagja & Rubini, 2023). While these approaches provide structured insights, they often lack scalability and adaptability in data-rich digital environments. More recent research has adopted clustering and classification algorithms to enable data-driven personalization in education (Dake et al., 2023; Risnasari et al., 2022; Sani Ibrahim, 2020). This section reviews existing machine learning approaches to learning style identification, highlights methodological limitations, and identifies research gaps that motivate the adoption of DEC.

The present study is grounded in the FSLSM (Felder & Silverman, 1988), which conceptualizes learner preferences across four bipolar dimensions: active–reflective, sensing–intuitive, visual–verbal, and sequential–global. The Index of Learning Styles (ILS), developed by Felder and Solomon (1991), operationalizes these dimensions and has demonstrated acceptable reliability and construct validity in higher education contexts (Felder & Spurlin, 2005). Although the broader learning styles literature has been criticized for rigid categorization and limited empirical support for instructional matching (Pashler et al., 2008), recent perspectives advocate probabilistic rather than deterministic modelling of learning preferences. In this study, FSLSM is used as a theoretically informed feature framework rather than a prescriptive instructional rule, enabling data-driven profiling grounded in educational theory.

CLUSTERING AND CLASSIFICATION OF LEARNING STYLES

Clustering algorithms such as K-means, DBSCAN, and hierarchical clustering have been widely used to group students based on learning behaviours (Lailiyah et al., 2019; Risnasari et al., 2022). K-means partitions learners into mutually exclusive groups, while DBSCAN can detect outliers and identify nonlinear patterns (Dhakal, 2025; Vives et al., 2024). Spectral clustering has also been explored to account for graph-based learning relationships (Tzenios, 2020).

However, these traditional clustering approaches often struggle with high-dimensional data and dynamic learner variability, leading to rigid partitions that inadequately capture evolving learning behaviours (Zhang & Dator, 2025). Moreover, clustering validity scores in educational datasets frequently remain moderate, suggesting limited separation between learner groups (Hasibuan et al., 2019).

Supervised classification models, such as SVMs, DTs, and ensemble techniques, have demonstrated strong predictive performance in identifying learning styles (Hmedna et al., 2020; Lokare & Jadhav, 2024). Feature engineering strategies have been shown to improve accuracy by up to 10% when irrelevant attributes are removed (Hasibuan et al., 2019). Nevertheless, supervised approaches depend heavily on pre-labelled datasets, which are costly to generate and may embed subjective bias (Kolekar et al., 2017). This dependence limits the discovery of latent or emergent learner structures.

DEEP LEARNING AND HYBRID MODELS

Deep learning models, including convolutional neural networks (CNNs), autoencoders, and reinforcement learning frameworks, have enhanced pattern recognition in learning analytics (Yang et al., 2014; Zhang & Dator, 2025). These models can analyse engagement logs, assessment performance, and interaction metrics to predict evolving learning demands (Dhakal, 2025; Palakurti, 2025).

However, deep learning models are frequently criticized for limited interpretability, making it difficult for educators to translate predictions into actionable strategies (Hasib et al., 2022). Additionally, they often require substantial computational resources, restricting adoption in resource-constrained institutions (Huang & Chen, 2024)

Hybrid approaches have attempted to combine clustering with predictive modelling to improve adaptability (Li & Liu, 2021; Ramírez-Correa et al., 2021). While these models improve predictive performance, such as enhancing dropout prediction accuracy by 15% (Awedh & Mueen, 2025), they rarely integrate statistical validation with deep representation learning in a unified framework.

LEARNING ANALYTICS AND VALIDATION APPROACHES

Learning analytics has become central to adaptive education systems, enabling data-driven monitoring of student engagement and achievement (Daoud et al., 2025; Pasina et al., 2019). Transfer learning and reinforcement learning have further enhanced student modelling by leveraging pre-trained architectures (Abrar et al., 2025; Zhang & Dator, 2025).

Despite these advances, many studies focus primarily on performance metrics such as accuracy or engagement improvements, with limited attention to statistical validation of cluster stability or theoretical alignment with established learning frameworks. As a result, computational performance often outpaces pedagogical interpretability.

GAPS IN EXISTING RESEARCH

Although significant progress has been made in applying machine learning to learning style identification, several gaps remain. First, many studies rely exclusively on either clustering or supervised classification without integrating both approaches within a statistically validated framework (Tzenios, 2020; Xie et al., 2021). Second, traditional clustering methods struggle with high-dimensional nonlinear educational data, often yielding moderate validity indices (Hasib et al., 2022). Third, while deep learning improves predictive performance, interpretability and theoretical grounding remain limited (Dissanayake et al., 2018).

DEC represents a methodological advancement by jointly optimizing feature representation learning and cluster assignment through iterative refinement (Yang et al., 2014). Unlike traditional clustering methods that operate on fixed feature spaces, DEC integrates nonlinear feature extraction with clustering objectives, improving cluster compactness and separation in high-dimensional data. While DEC has demonstrated effectiveness in pattern recognition domains, its application in educational data mining remains limited, particularly in connection with statistically validated pedagogical interpretation. Table 1 summarizes the reviewed studies, including datasets, algorithms, metrics, and identified gaps.

Overall, prior research demonstrates that machine learning enhances predictive modelling in education; however, three persistent gaps remain. First, few studies integrate deep representation learning with statistical validation to strengthen interpretability. Second, high-dimensional learner data is often handled using shallow clustering methods that struggle with nonlinear structure. Third, theoretical grounding in established learning models is frequently underdeveloped. These limitations motivate the present study's adoption of DEC combined with statistical testing to bridge computational robustness and educational meaning.

Table 1. Summary of reviewed literature

Authors	Population and sample	Algorithm used	Clustering or classification	Metrics and performance	Key findings	Gaps
Pardamean et al. (2022)	322 students (final sample: 269), 11 teachers	Matrix Factorization-Based Collaborative Filtering	Classification	RMSE (0.9035), paired t-test ($p < 0.05$)	AI and teacher assessments differ (35.13% agreement)	Limited to primary school students; lacks generalizability to higher education and different learning environments.
Abrar et al. (2025)	200 students, split into experimental and control groups	Supervised ML models, Reinforcement Learning, NLP	Classification	Post-test improvement (25% vs 14%), engagement up 15%	AI-driven adaptive learning improves learning efficiency	Short duration (6 weeks) and small sample size limit findings. AI model lacks robustness due to absence of deep learning approaches and unsupervised clustering.
Lokare and Jadhav (2024)	52 college students	Decision Tree, Random Forest, Ensemble	Classification	Random Forest (87.5% accu-	EEG-based classification is more accurate	Relies on EEG data, which may not be scalable or practical for large-scale adoption.

Authors	Population and sample	Algorithm used	Clustering or classification	Metrics and performance	Key findings	Gaps
		Bagging Classifier		racy), Decision Tree (62.5%)	than self-reported methods	No exploration of alternative feature selection methods or clustering approaches.
Subagja and Rubini (2023)	100 junior high school students	None (qualitative analysis)	Classification (manual)	Descriptive statistics, frequency analysis	Kinesthetic learners were the majority (35%)	Focuses on qualitative classification rather than AI-driven learning prediction. Misses integration with modern machine learning techniques for improved accuracy.
Lailiyah et al. (2019)	100 senior high school students	K-means, Fuzzy C-Means	Clustering	Silhouette Score, Calinski-Harabasz Index	Fuzzy C-Means allows mixed learning styles, K-means provides strict classification	Fuzzy C-Means clustering allows mixed learning styles, but lacks interpretability and validation on larger datasets.
Risnasari et al. (2022)	108 students in Software Engineering class	K-means	Clustering	Silhouette Coefficient (0.302)	Identified six clusters, but moderate separation.	Hierarchical clustering approach leads to rigid classification that may not adapt to dynamic learning behaviors. Study does not explore deep learning alternatives.
Pasina et al. (2019)	72 engineering students	Complete Linkage Clustering Algorithm	Clustering	Similarity coefficient matrix, dendrogram analysis	Identified clear learning style clusters, but some students exhibited mixed traits	K-means clustering effectively categorizes students but struggles with high-dimensional data. Study does not evaluate performance across diverse student populations.
Tzenios (2020)	310 undergraduate students	K-means, Hierarchical Clustering	Clustering	Inter-cluster distance, cohesion measures	Effective grouping into Sensory, Intuitive, and Visual/Verbal clusters	No investigation into transfer learning or deep feature extraction for clustering.

Authors	Population and sample	Algorithm used	Clustering or classification	Metrics and performance	Key findings	Gaps
El Aissaoui et al. (2019)	1,235 learner interactions in an e-learning platform	K-Modes Clustering, Naïve Bayes Classifier	Both	High classification accuracy with Naïve Bayes	Hybrid model effectively predicts learning styles dynamically	Comparative study lacks empirical validation with real students, making conclusions more theoretical. No investigation of adaptive clustering techniques
Palakurti (2025)	Analysis of 5 major adaptive learning platforms	Neural Networks, Reinforcement Learning, K-means, Decision Trees, Gradient Boosted Models	Both	Accuracy, scalability, engagement, adaptability	Hybrid AI models outperform rule-based adaptive systems	Pattern recognition approach is promising but lacks scalability due to manual tuning. Absence of self-learning mechanisms for dynamic student classification.
Yang et al. (2014)	50 undergraduate computer science students	Pattern Recognition Algorithm, 2-Means Clustering, Simulated Annealing	Both	Accuracy improvement from 63% to 85% over iterations	Adaptive learning style prediction improves personalization	DBSCAN clustering is more robust to noise, but struggles with high-dimensional educational data. No experimentation with deep clustering techniques to improve clustering validity.
Awedh and Mueen (2025)	Systematic review of 45 studies	RF, SVM, ANN	Both	Varied (Accuracy, F1-Score, RMSE, Clustering Validity)	ML-based education systems improve learning outcomes	Study does not validate clusters using predictive modeling.
Yang et al. (2014)	450 university students	CNN, Autoencoders	Both	Accuracy (CNN - 91.2%), Precision (89.3%)	Deep learning models outperform traditional ML in personalizing education	Study does not compare clustering results to classification outcomes.
Hmedna et al. (2020)	600 students from MOOC courses	Self-Organizing Maps (SOM), Hierarchical Clustering	Clustering	Silhouette Score (0.42), Davies-Bouldin Index (0.35)	Unsupervised methods uncover hidden patterns in stu-	Computationally expensive for large-scale student populations. Study does not explore alternative clustering solutions

Authors	Population and sample	Algorithm used	Clustering or classification	Metrics and performance	Key findings	Gaps
					dent learning behavior	
Dake et al. (2023)	395 Level 100 and 200 university students	Support Vector Machine (SVM), Random Forest, Decision Tree, K-means Clustering	Both	AdaBoost (RF) 61.77 % accuracy	Machine learning models effectively classify students into different learning patterns. Identified clusters provide insight into study habits and engagement levels.	Research should integrate more diverse datasets across multiple institutions.
Daoud et al. (2025)	500 online learners	Deep Reinforcement Learning, Neural Networks	Classification	Engagement improvement (20%), learning retention boost (15%)	AI-based engagement tracking enhances learning effectiveness	Study does not investigate clustering to identify at-risk students proactively.
(Hasib et al., 2022)	600 students in collaborative learning environments	Clustering with NLP, Spectral Clustering	Clustering	NLP clustering coherence (0.52)	AI detects patterns in student collaboration that impact learning outcomes	Lacks interpretability for educators. Study does not compare different classification approaches for curriculum optimization.
Zhang and Dator (2025)	700 students across various disciplines	Deep Q-Networks (DQN), K-means Clustering	Both	Recommendation accuracy (91%), student satisfaction (87%)	AI-powered curriculum customization enhances student engagement	Lacks interpretability for educators. Study does not compare different clustering approaches for curriculum optimization.

METHODOLOGY

As illustrated in Figure 1, the methodological procedure consists of data collection and processing, dimensionality reduction, the proposed DEC model, and cluster-based classification.

DATA COLLECTION AND PREPROCESSING

This study employed purposive sampling to select participants from Assin Nsuta Senior High School (ANSHS) in the Central Region of Ghana. The rationale for selecting this population was based on their developmental stage, which represents a critical phase in the Ghanaian education system. Prior to data collection, approval was obtained from the Department of ICT Education, University of Education, Winneba. A formal letter of permission was also issued to the Assin Nsuta Senior High School administration. Following ethical clearance from ANSHS, participants were briefed on the purpose of the study. They were informed about the objectives of the research, the voluntary nature of their participation, and the measures taken to ensure data privacy. Informed consent was subsequently obtained from all participants.

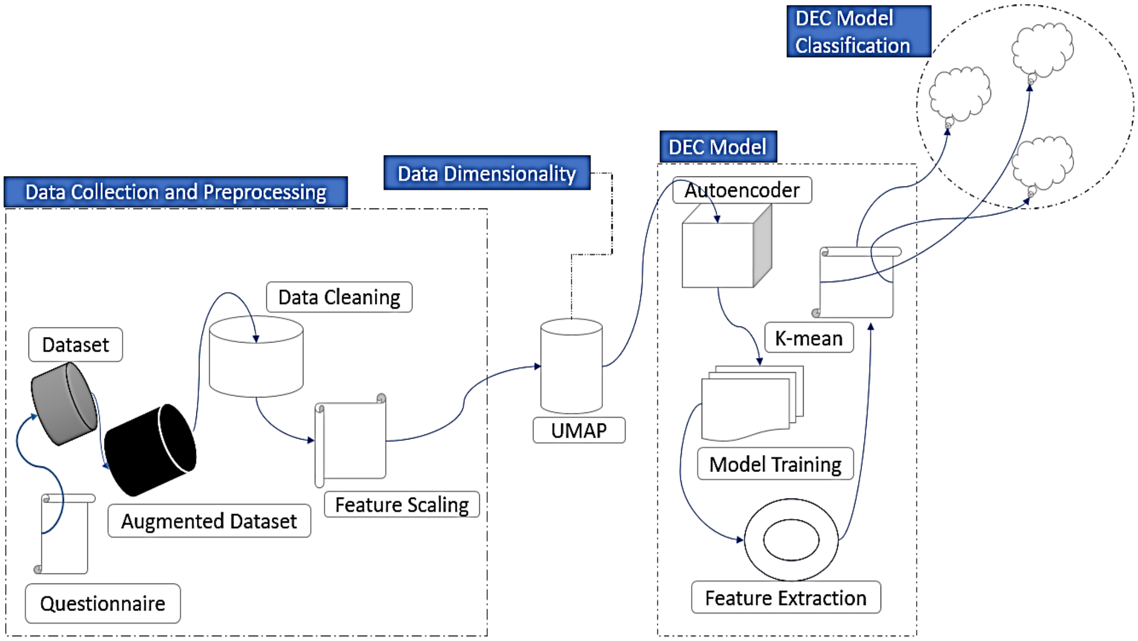


Figure 1. The study methodology workflow

The dataset was collected through a structured questionnaire based on the FSLSM. The questionnaire items were adapted from the Index of Learning Styles (ILS) developed by Felder and Solomon (1991), which operationalizes the four dimensions of the FSLSM (Felder & Silverman, 1988). The ILS has demonstrated acceptable reliability and construct validity across higher education populations (Felder & Spurlin, 2005). In this study, the instrument was used to capture learning preference tendencies rather than to assign rigid learning-type labels.

The questionnaire contained 48 categorical items designed to assess students' learning preferences. Additionally, four demographic attributes (gender, age, course, and form) were included. The questionnaire was created using Microsoft Forms for accessibility. A total of 412 responses were collected. No missing values were observed, and the dataset was stored in CSV and Excel formats for further processing and analysis.

A sample of the dataset is presented in Table 2. Due to the insufficiency of 412 instances for deep learning, we employed data augmentation techniques to expand the dataset. This increased the dataset to 5,928 records. This was achieved through synthetic oversampling using a duplication and controlled perturbation strategy, where categorical values were programmatically altered within domain-consistent limits to mimic realistic diversity without distorting class distribution.

We verified that the frequency distribution of each categorical feature in the augmented dataset closely mirrored that of the original dataset. Implausible combinations were filtered out during post-augmentation review. Following augmentation, duplicate records (380) were removed, resulting in 5,548 unique instances. To ensure uniformity in scale, all responses within each category are coded numerically and normalized to a range of 0 to 1.

Table 2. Sample attributes and options respondents

Attribute	Options
I like to learn new things through	a. visual aids like pictures, diagrams, graphs, or maps b. written instructions or spoken explanations
I prefer to study	a. in a study group b. alone
When studying with a group on hard topics, I'm more likely to	a. actively join in and share ideas b. sit quietly and listen to others
If I were a teacher, I would prefer to teach a course that is	a. focuses on real-life facts and situations b. explores ideas and theories
When I work on math problems	a. I usually solve problems step by step b. I often see the answers, but then struggle to understand how I got there
In my class	a. I've usually made friends with many of the students. b. I haven't really gotten to know many of the students
I like teachers who	a. use many diagrams on the board. b. take their time to explain things

DIMENSIONALITY REDUCTION USING UMAP

To enhance the performance of classification and clustering models, a dimensionality reduction technique was employed to transform the high-dimensional dataset into a representation with fewer dimensions. Reducing the 48-dimensional feature space helps eliminate redundancy while preserving meaningful structural information. The reason for selecting the Uniform Manifold Approximation and Projection (UMAP) was its ability to preserve both global and local structures of data at low computational complexity.

UMAP was utilized to project the 48-dimensional dataset to a 10-dimensional latent space. This transformation retained essential structural information while reducing noise. DEC used the compact and informative feature set of the resulting latent space representation as its input.

The combination of UMAP and an autoencoder was intentionally designed to enhance clustering performance. UMAP preserves both local and global manifold structures in high-dimensional data through nonlinear projection. However, UMAP alone does not optimize representations for clustering objectives. The subsequent autoencoder further refines the representation by learning task-specific latent features via reconstruction loss minimization. This two-stage nonlinear transformation reduces noise, improves separability, and produces compact feature representations before DEC optimization.

DEEP EMBEDDED CLUSTERING FOR STUDENT LEARNING STYLES

To enhance reproducibility, all preprocessing steps, dimensionality transformations, model architectures, hyperparameters, and validation splits are explicitly reported. The experimental pipeline was implemented using standard deep learning libraries with deterministic preprocessing operations. After obtaining the 10-dimensional latent space from UMAP, the next step was to use DEC to put students into groups based on how they learn best. DEC integrates deep neural network-based representation learning with clustering objectives to identify meaningful latent group structures.

The encoder, as shown in Table 3, receives the 10-dimensional UMAP representation and progressively reduces its dimensionality through a hidden layer with 8 neurons (ReLU activation) before reaching a bottleneck layer with 5 neurons to build compressed representations. The decoder follows this stepwise expansion of information: symmetrically expanding from 5 neurons through an intermediate 8-neuron layer to a final 10-neuron output, at which point the original input was recovered.

To ensure that essential information is not lost, the reconstruction loss was optimized using the Mean Squared Error (MSE) loss function.

Table 3. DEC hyperparameters and architecture (UMAP-reduced inputs)

Component	Setting	Justification
Input Dimension	48 (encoded & normalized features)	Uses full signal; AE learns task-specific latent structure
Encoder Hidden Layers	48 → 32 → 16 (ReLU)	Gradual compression avoids information bottlenecks; stable reconstruction
Bottleneck (Latent)	10 (ReLU)	Compact manifold enabling separability while retaining structure
Decoder Hidden Layers	16 → 32 (ReLU)	Symmetric expansion aids faithful reconstruction
Output Layer	48 (Sigmoid)	Matches inputs normalized to [0,1]
Loss/Optimizer	MSE / Adam	Standard and robust for tabular AEs
Learning Rate	0.001	Stable training; avoids divergence
Batch Size	32	Stable updates / hardware-fit
Epochs (pretrain AE)	50	Converged at validation loss ≈ **0.2189**
Train/Validation Split	80% / 20%	Overfitting control
k (clusters)	4	WCSS elbow and interpretability
DEC target refresh	Every 10 epochs	Standard DEC practice

For the autoencoder component of the DEC framework, the Adam optimizer (learning rate = 0.001) was used, and the model was trained for 50 epochs with a batch size of 32. For classification models, an 80%–20% training–testing split was used, whereas the autoencoder employed an 80%–20% training–validation split. During training, the reconstruction loss was minimized to learn latent representations that capture students' learning preference patterns. Following autoencoder training, latent representations from the bottleneck layer were used for clustering. The optimal number of clusters (k) was identified using the Elbow Method based on within-cluster sum of squares (WCSS). Subsequently, K-means clustering was applied to the latent space to derive distinct learning-style groupings.

The following pseudocode outlines the key steps involved in the DEC model:

```
BEGIN
  Define INPUT_LAYER with 10 neurons (from UMAP)
  ENCODER_LAYER_1 = Dense (8 neurons, activation=ReLU)
  BOTTLENECK_LAYER = Dense (5 neurons, activation=ReLU)
  DECODER_LAYER_1 = Dense (8 neurons, activation=ReLU)
  OUTPUT_LAYER = Dense (10 neurons, activation=Sigmoid)
  Compile AUTOENCODER using:
    - Loss function: Mean Squared Error (MSE)
    - Optimizer: Adam (learning rate=0.001)
  Train AUTOENCODER with:
    - Batch Size = 32
```

```
        - Epochs = 50
        - Train-Validation split (80%-20%)
    Extract encoded feature representations from the BOTTLENECK_LAYER
    Store the latent space representations for clustering
END
```

CLUSTERING AND EVALUATION OF THE DEC MODEL

Once the latent representations were obtained from the bottleneck layer, K-means clustering was conducted using the trained autoencoder. We used the Elbow Method to determine the optimal number of clusters by calculating the within-cluster sum of squares (WCSS) for different values of k. The best number of clusters is the one with the highest bending point of the WCSS curve.

After clustering, key metrics were computed to evaluate the quality. Silhouette Score measured how separated the clusters were; higher values indicated that groups were more distinct. In addition, the homogeneity of clusters was evaluated in terms of cluster purity, which made it useful for understanding how good our clustering method was at grouping learners according to their learning styles.

INTERPRETATION OF DISTINCT CLUSTERS

The distinct groups were evaluated to identify their defining characteristics and determine the types of learners represented in the separate categories.

To do this, we investigated the characteristics of the cluster centres in the latent space. In addition, analysing the original dataset features enables identification of important attributes that are significantly different across clusters. The differences in how the student groups like to learn were interpreted by looking at the correlations between the clusters and the most important features that set them apart.

We assigned descriptive labels to each cluster based on its most obvious characteristics. This made it easier for us to understand how students in each cluster approached learning, as well as their study preferences and their response to different educational strategies. These insights led to the development of personalized learning recommendations tailored to each cluster.

STATISTICAL VALIDATION OF CLUSTERS

To assess whether the clusters identified by the DEC model represent statistically distinct groups of students, additional statistical analyses were performed. After assigning each student to a cluster, one-way Analysis of Variance (ANOVA) was conducted for each continuous or ordinal feature to determine whether mean values differed significantly across clusters. For variables that did not meet the assumptions required for ANOVA, the Kruskal-Wallis H-test was applied as a non-parametric alternative. In the case of categorical variables, including gender and course enrolment, Chi-square tests of independence were used to evaluate whether the distribution of categories varied significantly between clusters.

Classification using cluster representations

Following the discovery and explanation of distinct clusters, a classification task was carried out to see how well standard machine learning models could predict cluster assignments. The purpose of this phase was to evaluate whether the formed features would be used to appropriately identify student learning styles. Three classification approaches were implemented:

Initially, traditional machine learning models were trained using the original dataset attributes. The dataset was split into 80% and 20% subsets for training and testing, respectively. The selected models included Random Forest (RF), SVM, and Logistic Regression (LR). These models were used to predict cluster labels, and performance was evaluated using accuracy, precision, recall, and F1-score.

Second, classification performance on the DEC latent-space representations was evaluated. The same machine learning models were retrained using the latent features that were lower-dimensional and extracted from the bottleneck layer of the trained autoencoder. Therefore, we compared the classification performance of the models trained using the original dataset to these latent representations.

A different correlation-based feature selection method was used to find the most important first features in the original data set that were used in the clustering. These attributes were then used to train traditional machine learning models, leading to another comparative study of classification performance. By evaluating classification models trained on the original dataset, the DEC latent space, and selected features, these steps allowed us to understand how valuable a variety of attributes can be for predicting student learning style categories.

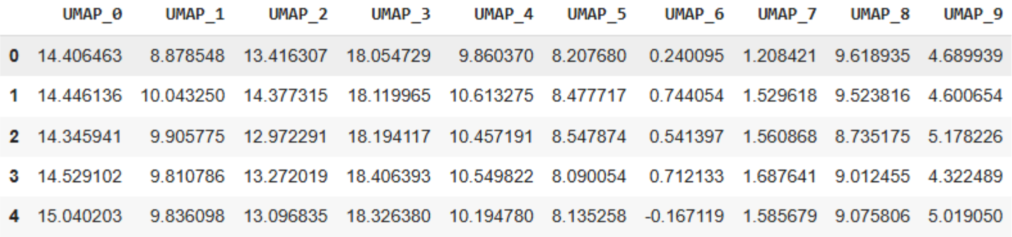
EXPERIMENTAL RESULTS

CLUSTERING OUTCOMES USING DEEP EMBEDDED CLUSTERING

The DEC model was implemented through a structured pipeline consisting of autoencoder training, dimensionality reduction, and clustering. The autoencoder architecture comprised an encoder, a bottleneck layer, and a decoder. The encoder transformed the input features into a lower-dimensional latent space, while the decoder reconstructed the original input from these compressed representations. The bottleneck layer served as the feature extraction component. Mean Squared Error (MSE) was used as the reconstruction loss function. The Adam optimizer with a learning rate of 0.001 was applied, and the model was trained for 50 epochs. An 80%–20% training–validation split was used. The final validation loss was 0.2189, indicating stable reconstruction performance. Each of the 44 questionnaire items was numerically encoded, with binary encoding applied to dichotomous (“a/b”) responses and ordinal encoding applied to Likert-scale items. Together with four demographic attributes (gender, age, form, and course), this resulted in a 48-dimensional feature space. All features were normalized to the range [0,1].

UMAP was applied for dimensionality reduction. The dataset was projected into a 10-dimensional space for clustering and into 2D/3D for visualization, while preserving pairwise relationships among samples (Figure 2). The elbow method was used to determine the optimal number of clusters. The WCSS was computed for k values ranging from 2 to 10. A pronounced elbow was observed at k = 4, as shown in Figure 3. The 10-dimensional latent representations extracted from the autoencoder bottleneck layer were clustered using the K-means algorithm with k = 4. This resulted in four clusters with the following counts in the augmented dataset (n = 5,548): 2,116; 1,546; 1,560; and 326 samples. The relative distribution of samples across clusters is shown in Figure 4. A two-dimensional UMAP visualization illustrating the spatial separation of the four clusters is shown in Figure 5.

Baseline clustering methods were evaluated for comparison, including K-means applied directly to the original feature space, Gaussian Mixture Models (GMM), and DBSCAN. Due to computational constraints, baseline models were evaluated on a representative subsample (n = 600) of the standardized 48-dimensional dataset. Performance was assessed using the Calinski–Harabasz index and the Davies–Bouldin index. The results are summarized in Table 4. Under the tested parameter settings, DBSCAN produced a single cluster, indicating limited density-based separability in the dataset.



	UMAP_0	UMAP_1	UMAP_2	UMAP_3	UMAP_4	UMAP_5	UMAP_6	UMAP_7	UMAP_8	UMAP_9
0	14.406463	8.878548	13.416307	18.054729	9.860370	8.207680	0.240095	1.208421	9.618935	4.689939
1	14.446136	10.043250	14.377315	18.119965	10.613275	8.477717	0.744054	1.529618	9.523816	4.600654
2	14.345941	9.905775	12.972291	18.194117	10.457191	8.547874	0.541397	1.560868	8.735175	5.178226
3	14.529102	9.810786	13.272019	18.406393	10.549822	8.090054	0.712133	1.687641	9.012455	4.322489
4	15.040203	9.836098	13.096835	18.326380	10.194780	8.135258	-0.167119	1.585679	9.075806	5.019050

Figure 2. Sample records of UMAP-reduced dimensionality

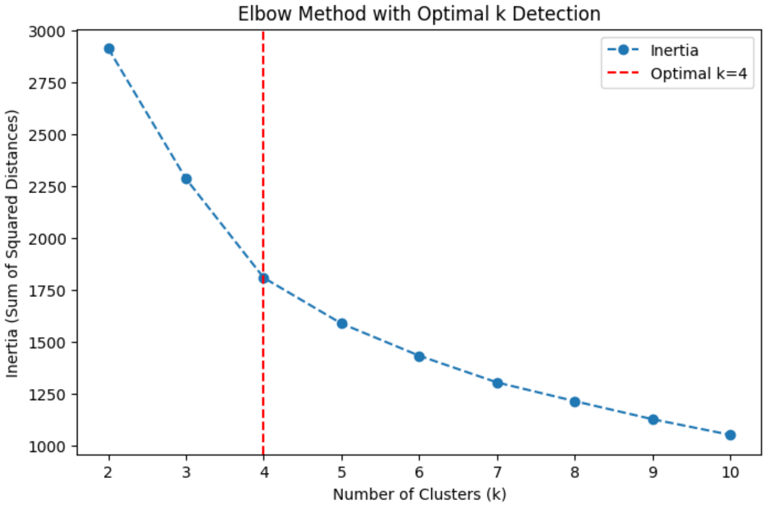


Figure 3. Elbow method for determining the optimal number of clusters

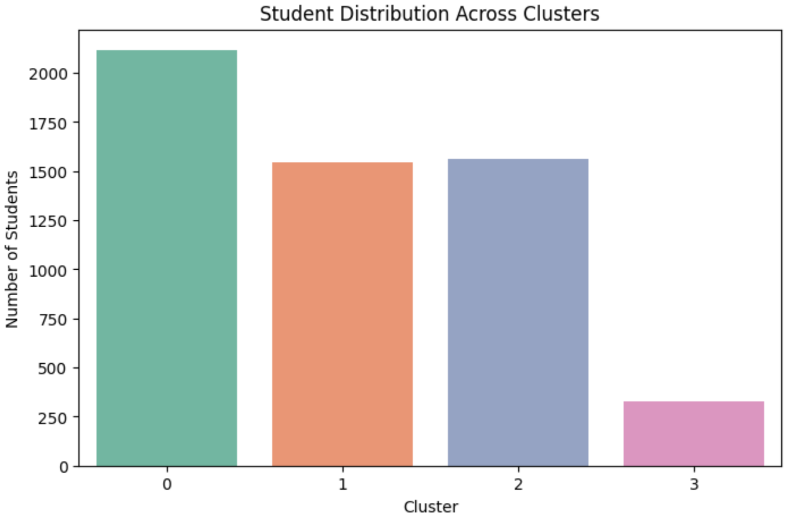


Figure 4. Cluster distribution

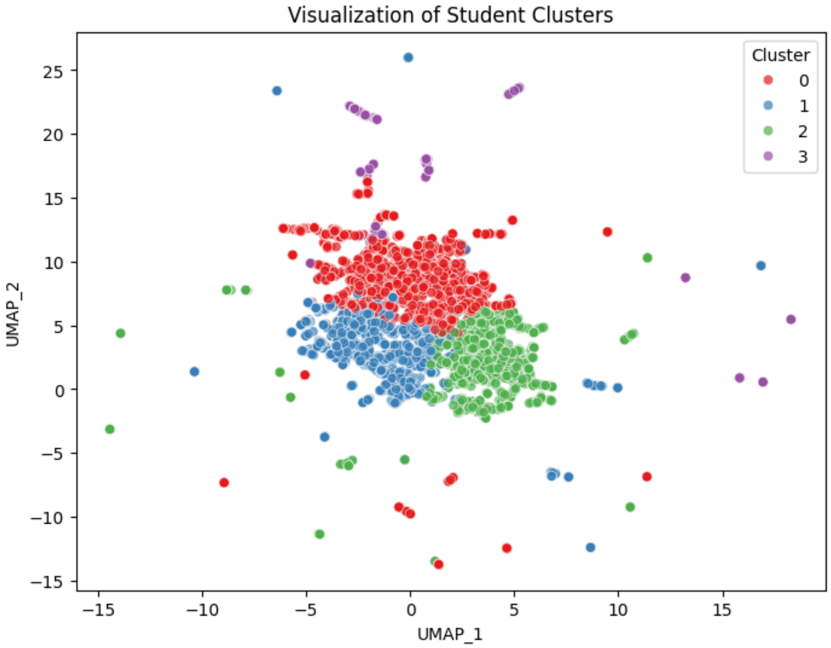


Figure 5. Visualization of clusters

Table 4. Baseline clustering performance on 10-dimensional latent space

Model	Calinski–Harabasz	Davies–Bouldin	Clusters
K-means	10.5694	5.3771	4
GMM	11.5344	5.2730	4
DBSCAN			1

STATISTICAL ANALYSIS OF CLUSTER DIFFERENCES

To examine differences across clusters, all 48 learning attributes were evaluated using one-way ANOVA for ordinal or continuous variables and χ^2 tests (with Cramér's V) for binary variables. P-values were adjusted using the Benjamini–Hochberg false discovery rate (FDR) procedure with q = 0.05. Results are reported as F(3, 5544), p, FDR-adjusted q, and η^2 for ANOVA, or $\chi^2(3, N = 5{,}548)$, p, FDR-adjusted q, and Cramér's V for χ^2 tests.

The ANOVA results indicated that, although a small number of attributes exhibited nominal significance prior to correction, none of the learning attributes remained statistically significant after FDR adjustment (all $q > 0.05$). In addition, effect sizes were consistently small ($\eta^2 \approx 0.001$–0.002), suggesting minimal between-cluster separation in terms of mean values. This outcome is consistent with the ordinal nature of learning-style questionnaire items and indicates that mean-based comparisons alone do not adequately capture cluster differences.

In contrast, χ^2 analyses revealed statistically significant differences in response distributions across clusters. Chi-square tests indicated significant associations between cluster membership and Gender ($\chi^2 = 11.02$, $p = 0.0116$) as well as Course ($\chi^2 = 33.66$, $p = 0.0139$), while no significant association was observed for Form ($\chi^2 = 9.24$, $p = 0.161$). Expected cell counts satisfied χ^2 assumptions. The most discriminative attributes based on χ^2 statistics, all of which remained significant after FDR correction, are reported in Table 5. Item-level significance patterns across clusters are further summarized in Figure 6.

Table 5. Top discriminative attributes by Chi-square

Attribute	Statistic (χ^2)	p	FDR q	Effect (Cramér's V)	Highest cluster
The_idea_of_doing_homework_in_groups_and_getting_one_grade_for_the_whole_group	805.8225	<0.0001	<0.0001	0.3687	2
I_like_to_learn_new_things_through	625.5966	<0.0001	<0.0001	0.3249	0
When_reading_a_book_with_many_pictures_and_charts__I_m_likely_to	572.9135	<0.0001	<0.0001	0.3109	3
I_like_topics_that_focus_on	569.7640	<0.0001	<0.0001	0.3100	1
I_prefer_to_study	457.4879	<0.0001	<0.0001	0.2778	2
When_I_m_reading_for_fun__I_prefer_writers_who	451.7046	<0.0001	<0.0001	0.2760	1
I_prefer_to_begin_by	425.7163	<0.0001	<0.0001	0.2680	2
When_I_see_a_diagram_or_sketch_in_class__I_usually_remember	423.7224	<0.0001	<0.0001	0.2674	2
I_understand_new_things_	422.2641	<0.0001	<0.0001	0.2669	1
I_remember_things_more_easily_when	410.1711	<0.0001	<0.0001	0.2630	1
For_entertainment__I_would_rather	400.1611	<0.0001	<0.0001	0.2598	0
I_think_it_s_a_bigger_compliment_to_say_someone_is	363.7882	<0.0001	<0.0001	0.2477	1
When_I_have_a_task_to_do__I_like_to	311.8409	<0.0001	<0.0001	0.2294	1
When_reading_nonfiction__I_like	297.3094	<0.0001	<0.0001	0.2239	1
People_are_more_likely_to_see_me_as	293.1698	<0.0001	<0.0001	0.2224	2

CLUSTERING PERFORMANCE EVALUATION

Clustering quality was evaluated using the Silhouette Score and Cluster Purity metrics. The Silhouette Score was computed for each sample and averaged across clusters. The DEC model achieved a Silhouette Score of 0.2863. Cluster purity was computed by assigning each cluster to its most frequent class label and calculating the proportion of correctly assigned samples. The DEC model obtained a cluster purity score of 0.2092.

Figure 6. Item-level significance by cluster

CLASSIFICATION RESULTS AND COMPARISONS

Three classification models, RF, SVM, and LR were trained and evaluated using three different feature representations: the original dataset, the DEC latent space, and a subset of 20 features selected using correlation-based feature selection. Classification performance was assessed using accuracy, recall, precision, and F1-score. Results for each feature representation are presented in Tables 6–9.

Models trained on the original dataset achieved accuracy values ranging from 34.8% to 38.35%. Training on the DEC latent space substantially improved performance, with accuracy values of 97.57% RF, 98.56% SVM, and 95.95% LR. Models trained on the selected feature subset showed marginal improvements over the original dataset but did not approach the performance obtained using the latent representations.

Table 6. Performance of classification models on original dataset

Model	Accuracy	Best recall	Best precision	Best F1-score	Remarks
RF	34.80%	68% (Cluster 0)	39% (Cluster 0)	49% (Cluster 0)	Poor recall for clusters 1, 2, and 3
SVM	38.35%	100% (Cluster 0)	38% (Cluster 0)	55% (Cluster 0)	Failed to classify any samples in Clusters 1, 2, or 3
LR	35.86%	78% (Cluster 0)	38% (Cluster 0)	51% (Cluster 0)	Failed to classify Cluster 3 completely

Table 7. Performance of classification models on DEC representation

Model	Accuracy	Best recall	Best precision	Best F1-score	Remarks
RF	97.57%	97-99%	97-99%	97-98%	Small misclassifications between clusters 0 & 1
SVM	98.56%	100% (Cluster 0)	99% (Cluster 2 & 3)	99% (Cluster 0 & 3)	Minor misclassifications in Cluster 1
LR	95.95%	99% (Cluster 2), 97% (Cluster 0)	100% (Cluster 3)	99% (Cluster 2)	Lower recall for Cluster 3 (78%)

Table 8. Performance of classification models on feature selection

Model	Accuracy	Best recall	Best precision	Best F1-Score	Remarks
RF	34.80%	57% (Cluster 0)	40% (Cluster 0)	49% (Cluster 0)	Cluster 3 still underperforms
SVM	38.35%	100% (Cluster 0)	38% (Cluster 0)	55% (Cluster 0)	Fails to classify Clusters 1, 2 & 3
LR	36.63%	86% (Cluster 0)	38% (Cluster 0)	53% (Cluster 0)	Cluster 3 not classified

Table 9. Overall classification model performances

Model		Accuracy	Recall	Precision	F1 Score
RF	Original Dataset	0.35	0.26	0.25	0.23
	Latent Space	**0.98**	**0.98**	**0.97**	**0.98**
	Selected Feature	0.35	0.33	0.35	0.33
SVM	Original Dataset	0.38	0.10	0.25	0.14
	Latent Space	**0.99**	**0.99**	**0.99**	**0.99**
	Selected Feature	0.38	0.15	0.38	0.21
LR	Original Dataset	0.36	0.24	0.25	0.20
	Latent Space	**0.96**	**0.97**	**0.92**	**0.94**
	Selected Feature	0.37	0.28	0.37	0.26

FINDINGS

This section synthesizes the experimental results and interprets them in relation to the study's research questions, emphasizing both technical outcomes and educational meaning.

RQ1: How can Deep Embedded Clustering be implemented to group students based on their learning preferences?

The findings indicate that DEC provides an effective mechanism for grouping students into latent learner profiles without reliance on predefined labels. These profiles reflect patterned preferences in the FSLSM-aligned questionnaire responses, rather than fixed or deterministic "learning styles". The four-cluster structure reflects coherent and interpretable learning patterns rather than arbitrary partitions. Examination of the standardized cluster profiles (Table 10) shows consistent differentiation across multiple learning attributes, suggesting that the DEC framework successfully captured underlying structure in students' learning preference patterns. The heatmap visualization (Figure 7) further illustrates clear contrast patterns across clusters, reinforcing the interpretability of the learned representations.

Table 10. Sample of standardized cluster profiles

Feature	Cluster 0	Cluster 1	Cluster 2	Cluster 3
Course	-0.63	0.8	-1.3	1.13
I like the idea of	0.71	-0.14	1	-1.57
I remember things more easily when	-0.19	-0.56	-0.93	1.67
I usually	-0.7	-0.3	-0.7	1.71
When looking at a lot of information, I'm more likely to	-0.65	-1.09	0.22	1.53
When writing a paper, I'm more likely to	-0.94	-0.94	0.47	1.41
Form (Academic Level)	-0.35	-1.27	1.5	0.12
When I'm studying a new subject, I like to	-0.24	-0.73	-0.73	1.7
When reading a book with many pictures/charts, I'm likely to	1.26	-0.98	0.7	-0.98
When I receive directions to a new place, I like them to be	-0.53	0	-1.07	1.6
I prefer to begin by	1.6	0	-1.07	-0.53
When I think about what I did yesterday, I usually feel	0.63	-0.63	1.26	-1.26
The idea of doing homework in groups and getting one grade for the whole group	0.33	0.33	-1.67	1
When I'm learning something new, it's useful for me to	0.17	1.52	-0.51	-1.18
When someone is sharing data with me, I like it when they use	-0.91	-0.39	-0.39	1.69
When I work on math problems	0	-1.41	0	1.41
I remember things best when	0.58	-0.96	-0.96	1.35
In my class	-0.58	0.96	-1.35	0.96
Age	1.52	-0.51	0.17	-1.18
When working on problems with a group, I'm more likely to	0.33	-1.67	0.33	1

RQ2: What are the unique characteristics of each identified student cluster?

Distinct learning preference profiles were observed across the four clusters, as evidenced by the standardized feature values (Table 10) and the cluster-wise characteristic patterns in the heatmap (Figure 7).

Cluster 0 (Structured Visual Learners) exhibited strong alignment with visually organized and sequential learning approaches. High standardized scores on visually oriented items and structured study preferences (Table 10) suggest greater reliance on diagrams, charts, and clearly organized materials. The heatmap in Figure 7, which summarizes cluster-wise patterns in the learned latent feature space, shows consistent over-representation associated with visual and sequential learning tendencies in this cluster. Cluster 1 (Active Experimental Learners) was characterized by elevated engagement in interactive and experiential learning. Correlation analysis (Figure 8) shows strong associations between this cluster and features related to discussion, collaboration, and practical application. These learners showed stronger preferences for experimentation and group-based problem-solving.

Cluster 2 (Balanced Adaptive Learners) displayed moderate and evenly distributed feature values across learning attributes. As shown in Table 10 and Figure 7, this cluster does not strongly dominate any single learning modality, indicating adaptive switching between structured and interactive learning strategies depending on context. Cluster 3 (Support-Need Learners) demonstrated lower standardized scores across several learning preference dimensions. Figure 7 highlights under-representation in both visual and active engagement features, suggesting difficulty in sustaining effective learning strategies. Feature–cluster correlations (Figure 8) further indicate weaker alignment with dominant learning behaviours observed in other clusters. Together, these profiles suggest that the identified clusters represent qualitatively distinct learning patterns.

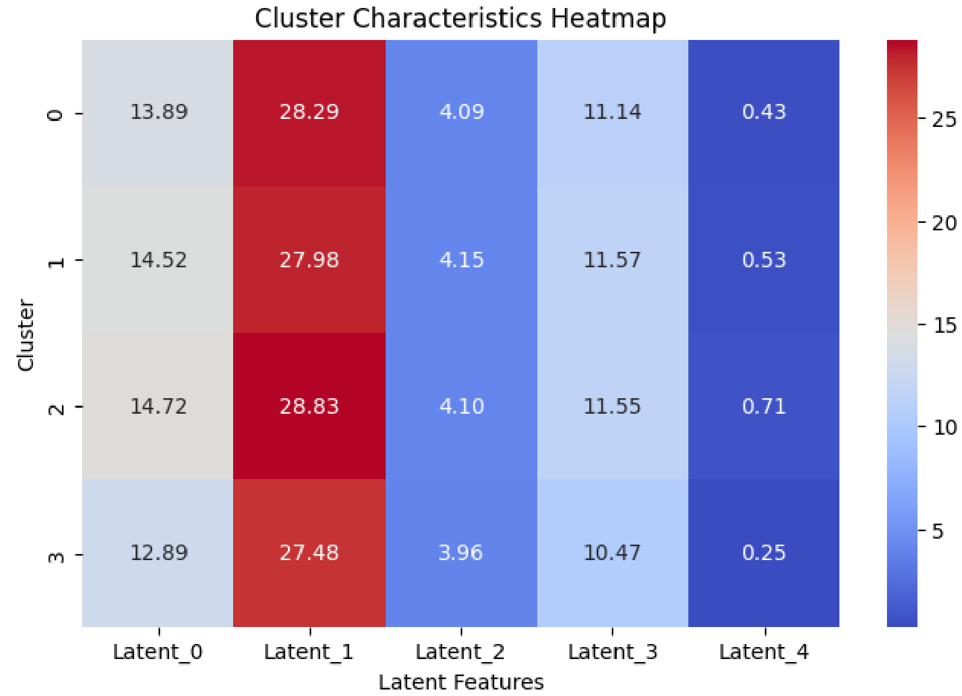


Figure 7. Cluster characteristics heatmap

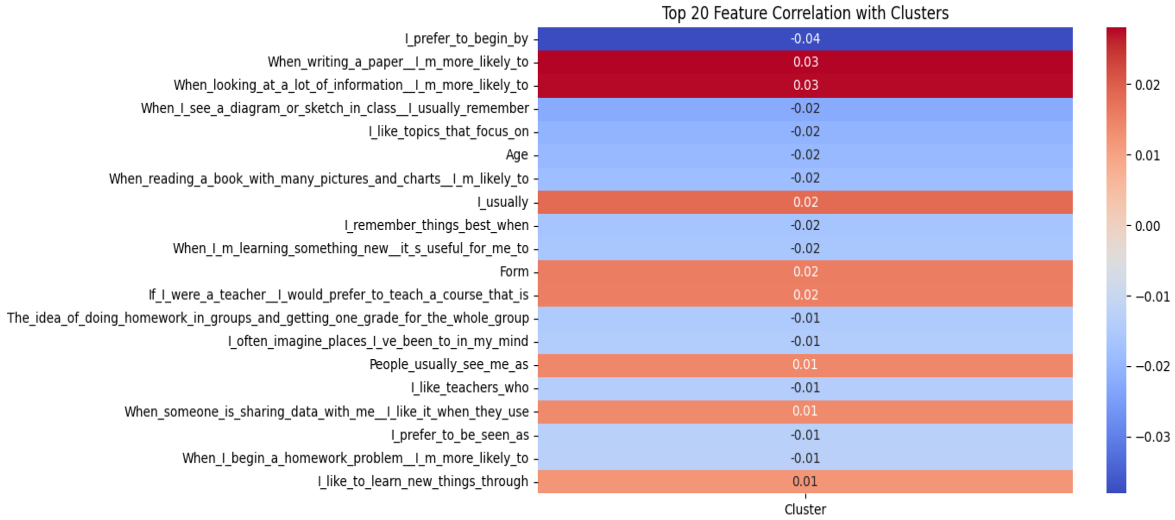


Figure 8. Feature correlation with cluster

RQ3: How effective are the identified clusters based on Silhouette Score and Cluster Purity metrics?

The evaluation findings suggest that the clusters possess moderate but meaningful structural validity (Silhouette = 0.286; Purity = 0.209). Although classification performance using DEC-derived embeddings was high, the classifiers predicted cluster membership learned from the unsupervised pipeline rather than externally assigned labels. To reduce leakage risk, cluster identifiers were not used as input features, and train–test splits were applied consistently during classification. Taken together, the moderate internal indices and the strong latent-space separability suggest structured differentiation without implying perfectly separated learner categories.

Although the clusters are not entirely separable, the observed overlap reflects the inherently fluid nature of learning preferences. The standardized cluster profiles (Table 10) show that while dominant tendencies exist within each cluster, no cluster is defined by a single exclusive attribute. This reinforces the view that learning styles function along continua rather than as rigid categories. The DEC-based grouping, therefore, captures nuanced distinctions that traditional clustering approaches tend to overlook.

RQ4: How does a classification model trained on Deep Embedded Clustering representations compare with traditional machine learning classifiers?

The findings show that classifiers trained on DEC-derived representations achieve substantially higher predictive performance than those trained on raw or feature-selected data. Feature–cluster correlations shown in Figure 8 indicate that DEC embeddings concentrate discriminative information into a compact latent space, enabling clearer class boundaries. In contrast, standardized profiles in Table 10 illustrate the dispersion and overlap present in the original feature space, which contributes to weaker classification performance when deep representations are not used. These results highlight the value of deep clustering as a feature learning mechanism for downstream predictive tasks. From an educational perspective, the improved classification performance indicates that latent behavioural patterns among students are internally coherent. Rather than merely increasing statistical accuracy, the DEC framework reveals structured preference tendencies that may not be directly observable through surface-level questionnaire responses. This supports the pedagogical interpretability of the identified profiles by showing that the latent representations encode consistent response-pattern structure.

RQ5: What personalized teaching strategies can be recommended for different student clusters?

The cluster-specific patterns observed in Table 10 and Figure 7 provide a foundation for differentiated instructional strategies. To further ground these interpretations theoretically, learning attributes were mapped to the FSLSM. The resulting dimension-level profiles (Table 11) reveal consistent alignment between cluster behaviours and established learning-style dimensions.

Cluster 0 aligns strongly with the Visual–Sequential dimension, supporting the use of structured visual instruction. Cluster 1 exhibits Active learning tendencies, favouring collaborative and experiential approaches. Cluster 2 demonstrates balanced positioning across FSLSM dimensions, indicating suitability for blended instructional designs. Cluster 3 shows weak or mixed alignment across dimensions, suggesting the need for personalized support, scaffolding, and adaptive intervention.

These findings indicate that DEC-based clustering can inform adaptive learning systems capable of aligning pedagogical strategies with students' latent learning preferences. The alignment between cluster characteristics and FSLSM dimensions enhances interpretive stability. Rather than generating arbitrary machine-defined groupings, the clusters demonstrate theoretical coherence with FSLSM dimensions used as a guiding framework. This correspondence supports interpretability, while acknowledging that cross-institutional validation would strengthen generalizability.

Table 11. FSLSM dimension profiles by cluster

FSLSM dimension	Cluster 0	Cluster 1	Cluster 2	Cluster 3
Active – Reflective	Reflective learning	Active learning	Balanced	Mixed/weak
Sensing – Intuitive	Sensing learning	Sensing/ Practical learning	Balanced	Mixed/weak
Visual – Verbal	Visual dominant	Neutral	Neutral	Verbal learning
Sequential – Global	Sequential learning	Neutral	Neutral	Less sequential/ Global-like

DISCUSSION

The outcomes of this study demonstrate that deep embedded clustering provides an effective way to identify distinct groups of students with similar learning preferences. In addition, the clustering process revealed coherent response-pattern groupings that extend beyond traditional rule-based classification methods. Classifiers based on representations generated by DEC showed significantly better performance compared to traditional machine learning classifiers, indicating that deep representation learning captures latent behavioural structure. Furthermore, the statistical validation confirmed that the clusters were not only computationally distinct but also statistically robust. Although one-way ANOVA was conducted across all learning attributes, none remained statistically significant after false discovery rate correction, and effect sizes were small.

In contrast, χ^2 analyses revealed significant distributional differences across clusters for multiple learning preference items. This indicates that cluster separation is better characterized by response-pattern differences rather than mean-level differences. These findings provide additional evidence that the clusters reflect structured differences in student learning profiles. The absence of significant ANOVA results after false discovery rate correction, combined with significant χ^2 distributional differences, suggests that learning preferences in this dataset are not characterized by large mean shifts in isolated attributes but rather by patterned response configurations across multiple dimensions. This indicates that learning preferences operate as multivariate behavioural tendencies rather than single-variable effects. The DEC framework is therefore particularly suited to capturing such distributed response structures, which traditional mean-based statistical comparisons may fail to detect.

While GMM and K-means provided reasonable partitions on the standardized 48-dimensional feature space, DBSCAN collapsed to a single cluster, suggesting a lack of density separability (see Table 4). In contrast, the learned DEC latent space produced more stable and interpretable clusters, reinforcing its suitability for profiling learning behaviours. The superior stability of DEC likely stems from its ability to learn task-specific nonlinear embeddings that preserve manifold geometry while compressing noise. Unlike centroid-based clustering, which assumes spherical separability, DEC iteratively refines cluster assignments within a learned latent space. This mechanism allows subtle behavioural distinctions to become linearly separable, thereby explaining its improved discriminative performance.

This study stands out from others due to its significant development of methodologies, datasets, and classification performances. Previous studies in the classification of learning styles have exhibited noticeable differences in the size of datasets used as well as the demographics of students involved in those datasets. For instance, a study by Pardamean et al. (2022) involved 269 primary school students, whereas Lokare and Jadhav (2024) investigated only 52 college students, and Subagja and Rubini (2023) concentrated on 100 junior high school students. In contrast, this study utilized an expanded dataset, which increases internal modelling stability; however, generalizability remains bounded by the single-site sampling design. By improving on the dataset quality, the model demonstrated stronger internal structural coherence across learner preference patterns.

Previous research primarily focuses on the use of supervised learning strategies, which are inherently dependent on large amounts of labelled data. In their study, Pardamean et al. (2022) employed matrix factorization-based collaborative filtering in anticipating learning styles. Abrar et al. (2025) utilized supervised machine learning models along with reinforcement study concepts. Referencing previous work by Lailiyah et al. (2019), they employed K-means and fuzzy C-means clustering methods based on fixed cluster numbers, leading to low adaptability. This research did not rely on labelled data. It dynamically learned feature representations by adapting cluster structures depending on students' responses through the DEC, thereby enabling adaptive latent representation learning rather than predefined classification.

In relation to the classification performance, previous studies showed some success in classification, yet without appropriate validations. In handling diverse student learning styles, many basic machine learning algorithms performed poorly. Lokare and Jadhav (2024) reported a classification accuracy of 87.5% with a Random Forest trained on EEG signals. Significantly higher classification accuracy was achieved in this study when Random Forest was trained on DEC-generated representations, which reached 97.57%. In addition, high precision and recall scores were maintained by DEC-based classification between 97% and 99%, while traditional models performed poorly in certain cases of misclassification, notably when trying to differentiate similar learning styles. The results suggest that deep representation learning frameworks such as DEC provide stronger internal separability because they allow for automatic extraction of intricate relationships from responses made by students. It is important to note that classification models predicted unsupervised cluster assignments rather than externally labelled outcomes. Therefore, the high performance reflects internal structural coherence within the latent space rather than external validation accuracy. This distinction underscores that the contribution of DEC lies in representation learning and behavioural profiling rather than supervised prediction of predefined categories.

Yet another significant drawback in previous literature was using human-made feature selection for examples like test results, ECG signals, or self-reported learning modalities. Works written by Subagja and Rubini (2023) were based on predefined qualities by professionals. This method assumes expert knowledge in specific fields, thus resulting in human subjectivity in data collection and interpretation. On the other hand, the study deploys a more scalable learning style classification approach, as DEC is used to automatically extract significant aspects from student interactions without the necessity of human intervention. Because DEC learns feature representations directly from response patterns, it reduces reliance on manual feature engineering and associated subjectivity.

Theoretically, these findings contribute to the ongoing learning-style debate by demonstrating that machine-derived clusters can align with established FSLSM dimensions without enforcing rigid categorical assignments. Rather than validating fixed typologies, the results support a continuum-based interpretation of learning preferences, where behavioural tendencies emerge from distributed feature interactions. This challenges deterministic interpretations of learning styles while supporting adaptive, probabilistic modelling approaches. From an instructional standpoint, structured learner profiling may support more targeted differentiation strategies. Such insights from DEC clustering assist educators in customizing instructional strategies, maximizing learning environments, and producing adaptive learning tools. An AI-based Learning Management System (LMS) might use DEC for prompt recommendation of study and learning materials to align instructional resources with observed response-pattern tendencies. The findings suggest that DEC-based models offer a data-driven alternative to traditional self-reported learning style assessments, which are often subjective and inconsistent. By using AI-driven clustering, educational institutions can implement scalable, real-time solutions that dynamically adapt to student learning behaviours and provide personalized learning experiences.

Even though this study has shown strong results, future research needs to address certain limitations. The dataset that was used should be validated across different educational environments and contexts in order to improve its generalizability. Real-time implementation of DEC may be difficult in resource-constrained settings, mainly because of the computational complexity. However, it is possible to address this using lightweight or hybrid approaches, which combine accuracy with efficiency in order to scale well. Moreover, though this particular research is mainly on unsupervised clustering, other future studies may include the use of reinforcement learning methods to further improve the classification of learning styles. Integrating adaptive feedback mechanisms could allow continuous model improvement and better adaptation to student learning patterns over time.

LIMITATION

While the study demonstrates methodological rigor, several limitations warrant critical consideration. The use of purposive sampling from a single senior high school provides contextual depth but limits generalizability to broader educational populations. Future studies should incorporate multi-institutional and cross-cultural datasets to enhance external validity.

Although augmentation procedures were distribution-preserving, synthetic data expansion may introduce structural regularities not present in naturally occurring datasets. Additionally, model architecture and hyperparameter choices may influence cluster formation. While statistical validation was employed to assess cluster distinctiveness, residual algorithmic bias and representational distortion inherent in nonlinear embedding techniques cannot be entirely ruled out and should be examined in future cross-validation studies.

CONCLUSION

This study demonstrates that a theory-informed, data-driven pipeline can recover meaningful learner profiles in secondary education. Using an FSLSM–based questionnaire (48 items), a nonlinear representation learning framework combining dimensionality reduction and deep embedding enabled the identification of four interpretable clusters aligned with established FSLSM dimensions.

Unlike prior studies that relied on supervised prediction of predefined learning-style labels or fixed centroid-based clustering, this study demonstrates that unsupervised deep embedding can recover multivariate behavioural configurations consistent with theory while preserving flexibility beyond rigid typologies. The combination of per-item statistical validation with false discovery rate control, dimensional mapping to FSLSM composites, and downstream predictive evaluation provides a comprehensive validation framework that is not commonly applied in educational clustering research.

After encoding/normalization and synthetic oversampling with controlled perturbations plus de-duplication, the analysis set comprised 5,548 records. Internal indices indicated moderate separation (Silhouette ≈ 0.286, Purity ≈ 0.209). Statistical testing revealed that cluster differentiation was driven by distributional response patterns rather than isolated mean shifts, reinforcing the interpretation of learning preferences as multivariate continua rather than discrete categories. Cluster membership was independent of academic form, suggesting behavioural profiling beyond grade-level stratification.

In downstream tasks, classifiers trained on DEC latent representations substantially outperformed models trained on raw features (≈96–99% vs. ≈35–38% accuracy), indicating that representation learning enhances structural separability and pedagogical interpretability. Importantly, this performance reflects internal structural coherence of the learned latent space rather than the prediction of externally validated labels.

The findings suggest that DEC-based profiling can support scalable personalization strategies within AI-enabled learning environments. Rather than replacing teacher judgment, such models may augment instructional planning by identifying latent behavioural tendencies (e.g., Visual/Sequential, Active) that inform differentiated material design. Theoretically, the results contribute to the learning-style discourse by demonstrating that probabilistic, continuum-based modelling approaches may reconcile structure and flexibility within established frameworks such as FSLSM.

While the study demonstrates methodological rigor, it is limited by reliance on synthetic oversampling and sampling from a single institution. These design constraints restrict external generalizability and warrant cautious interpretation of cluster stability beyond the present context. Replication across multi-site, non-augmented datasets and longitudinal validation of adaptive interventions will be essential to strengthen generalizability.

Overall, the integration of deep embedded clustering, statistical validation, and theory alignment demonstrates that data-driven profiling can inform targeted, scalable personalization in secondary education while maintaining interpretive coherence.

REFERENCES

Abrar, M., Aboraya, W., Khaliq, R. A., Subramanian, K. P., Husaini, Y. A., & Hussaini, M. A. (2025). AI-powered learning pathways: Personalized learning and dynamic assessments. *International Journal of Advanced Computer Science and Applications, 16*(1), 454–462. https://doi.org/10.14569/IJACSA.2025.0160145

Altamimi, A. M., Azzeh, M., & Albashayreh, M. (2022). Predicting students' learning styles using regression techniques. *Indonesian Journal of Electrical Engineering and Computer Science, 25*(2), 1177–1185. https://doi.org/10.11591/ijeecs.v25.i2.pp1177-1185

Awedh, M. H., & Mueen, A. (2025). Early identification of vulnerable students with machine learning algorithms. *WSEAS Transactions on Information Science and Applications*, *22*, 166–188. https://doi.org/10.37394/23209.2025.22.16

Babić, I. D. (2017). Machine learning methods in predicting the student academic motivation. *Croatian Operational Research Review*, *8*(2), 443–461. https://doi.org/10.17535/crorr.2017.0028

Dake, D. K., Bada, G. K., & Techie-Menson, H. (2023). Using machine learning to cluster and predict the learning pattern of university students. *Telematique*, *22*, 77–84.

Daoud, R. A., Amine, A., Abouelmehdi, K., & Razouk, A. (2025). Student engagement in e-learning during crisis: An unsupervised machine learning and exploratory data analysis approach. *Journal of Applied Data Sciences*, *6*(1), 508–525. https://doi.org/10.47738/jads.v6i1.458

Dhakal, S. D. (2025). Impact of algorithms and big data on educational field. *OCEM Journal of Management, Technology & Social Sciences*, *4*(1), 125–131. https://doi.org/10.3126/ocemjmtss.v4i1.74754

Dissanayake, D., Perera, T., Elladeniya, C., Dissanayake, K., Herath, S., & Perera, I. (2018). Identifying the learning style of students in MOOCs using video interactions. *International Journal of Information and Education Technology*, *8*(3), 171–177. https://doi.org/10.18178/ijiet.2018.8.3.1029

El Aissaoui, O., El Alami El Madani, Y., Oughdir, L., & El Allioui, Y. (2019). Combining supervised and unsupervised machine learning algorithms to predict the learners' learning styles. *Procedia Computer Science*, *148*, 87–96. https://doi.org/10.1016/j.procs.2019.01.012

Felder, R. M., & Silverman, L. K. (1988). Learning and teaching styles in engineering education. *Journal of Engineering Education, 78*(7), 674–681.

Felder, R. M., & Solomon, B. A. (1991). *Learning styles and index of learning styles*. North Carolina State University. https://engr.ncsu.edu/stem-resources/legacy-site/learning-styles/

Felder, R. M., & Spurlin, J. E. (2005). Applications, reliability, and validity of the index of learning styles. *International Journal of Engineering Education, 21*(1), 103–112. https://doi.org/10.1037/t43782-000

Hashmi, S. A., Singh, Y., & Bhardwaj, H. (2025). *Implementing a novel approach using PSO-enhanced SVM with custom kernel for classification of student and recommendation.* Authorea. https://doi.org/10.22541/au.173882389.92058235/v1

Hasib, K. M., Rahman, F., Hasnat, R., & Alam, M. G. R. (2022, January). A machine learning and explainable AI approach for predicting secondary school student performance. *Proceedings of the IEEE 12th Annual*

Computing and Communication Workshop and Conference, Las Vegas, NV, USA, 399–405. https://doi.org/10.1109/CCWC54503.2022.9720806

Hasibuan, M. S., Nugroho, L. E., & Santosa, P. I. (2019). Model detecting learning styles with artificial neural network. *Journal of Technology and Science Education*, *9*(1), 85–95. https://doi.org/10.3926/jotse.540

Hmedna, B., El Mezouary, A., & Baz, O. (2020). A predictive model for the identification of learning styles in MOOC environments. *Cluster Computing*, *23*, 1303–1328. https://doi.org/10.1007/s10586-019-02992-4

Huang, Q., & Chen, J. (2024). Enhancing academic performance prediction with temporal graph networks for massive open online courses. *Journal of Big Data*, *11*, Article 52. https://doi.org/10.1186/s40537-024-00918-5

Kolekar, S. V., Pai, R. M., & Manohara Pai, M. M. (2017). Prediction of learner's profile based on learning styles in adaptive e-learning system. *International Journal of Emerging Technologies in Learning*, *12*(6), 31–51. https://doi.org/10.3991/ijet.v12i06.6579

Lailiyah, S., Yulsilviana, E., & Andrea, R. (2019). Clustering analysis of learning style on Anggana high school student. *Telkomnika*, *17*(3), 1409–1416. https://doi.org/10.12928/telkomnika.v17i3.9101

Li, S., & Liu, T. (2021). Performance prediction for higher education students using deep learning. *Complexity*, *2021*(1), Article 9958203. https://doi.org/10.1155/2021/9958203

Lokare, V. T., & Jadhav, P. M. (2024). An AI-based learning style prediction model for personalized and effective learning. *Thinking Skills and Creativity*, *51*, 101421. https://doi.org/10.1016/j.tsc.2023.101421

Palakurti, N. R. (2025). Comparative study of adaptive learning platforms: A machine learning perspective. *International Journal of Advanced Research in Education and Technology, 12*(1), 17–26. https://doi.org/10.15680/IJARETY.2025.1201002

Pardamean, B., Suparyanto, T., Cenggoro, T. W., Sudigyo, D., & Anugrahana, A. (2022). AI-based learning style prediction in online learning for primary education. *IEEE Access*, *10*, 35725–35735. https://doi.org/10.1109/ACCESS.2022.3160177

Pashler, H., McDaniel, M. A., Rohrer, D., & Bjork, R. A. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest, 9*(3), 105–119. https://doi.org/10.1111/j.1539-6053.2009.01038.x

Pasina, I., Bayram, G., Labib, W., Abdelhadi, A., & Nurunnabi, M. (2019). Clustering students into groups according to their learning style. *MethodsX*, *6*, 2189–2197. https://doi.org/10.1016/j.mex.2019.09.026

Ramírez-Correa, P., Alfaro-Pérez, J., Gallardo, M., & Rodrigues, F. (2021). Identifying engineering undergraduates' learning style profiles using machine learning techniques. *Applied Sciences*, *11*(22), 10505. https://doi.org/10.3390/app112210505

Risnasari, M., Aulia, N., & Cahyani, L. (2022). Clustering of student learning styles in the industry 4.0 using KMeans algorithm. *Jurnal Teknologi Pendidikan, 24*(2), 246–257. https://doi.org/10.21009/jtp.v24i2.28029

Sani Ibrahim, M. (2020). Learning style detection using K-means clustering. *FUDMA Journal of Sciences*, *4*(3), 375–381. https://doi.org/10.33003/fjs-2020-0403-351

Subagja, S., & Rubini, B. (2023). Analysis of student learning styles using Fleming's VARK model in science subject. *Jurnal Pembelajaran dan Biologi Nukleus, 9*(1), 31–39. https://doi.org/10.36987/jpbn.v9i1.3752

Tzenios, N. (2020). Clustering students for personalized health education based on learning styles. *Sage Science Review of Educational Technology (SSRET)*, 22–36.

Vives, L., Cabezas, I., Vives, J. C., Reyes, N. G., Aquino, J., Condor, J. B., & Altamirano, S. F. S. (2024). Prediction of students' academic performance in the programming fundamentals course using long short-term memory neural networks. *IEEE Access*, *12*, 5882–5898. https://doi.org/10.1109/access.2024.3350169

Xie, S.-t., Chen, Q., Liu, K.-h., Kong, Q.-z., & Cao, X.-j. (2021). Learning behavior analysis using clustering and evolutionary error correcting output code algorithms in small private online courses. *Scientific Programming*, *2021*(1). https://doi.org/10.1155/2021/9977977

Yang, J., Huang, Z. X., Gao, Y. X., & Liu, H. T. (2014). Dynamic learning style prediction method based on a pattern recognition technique. *IEEE Transactions on Learning Technologies*, *7*(2), 165–177. https://doi.org/10.1109/TLT.2014.2307858

Zhang, H., & Dator, W. L. (2025). Establishment and validation of a prediction model for compassion fatigue in nursing students. *BMC Nursing*, *24*, Article 193. https://doi.org/10.1186/s12912-025-02834-2

AUTHORS

Elijah Ofori completed his MPhil in ICT Education at the University of Education, Winneba, Ghana. His research interests include deep learning, educational data mining, ICTs, and software development.

Delali Kwasi Dake, PhD is an Associate Professor in the Department of ICT Education at the University of Education, Winneba, Ghana. His research interests include educational data mining, artificial intelligence, sentiment analysis, software-defined networking, and Internet of Things.