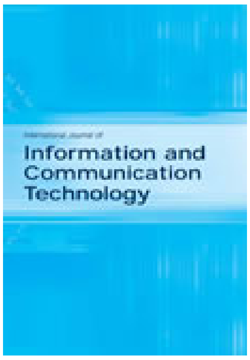


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Mining of tourism English learning mode based on temporal clustering and ensemble learning

Yan Jing

DOI: 10.1504/IJICT.2025.10073371

Article History:

Received:	16 May 2025
Last revised:	29 May 2025
Accepted:	29 May 2025
Published online:	26 September 2025

Mining of tourism English learning mode based on temporal clustering and ensemble learning

Yan Jing

School of Tourism Foreign Languages,
Zhengzhou Tourism College,
Zhengzhou, 451464, China
Email: happyfamilyyan@163.com

Abstract: With the increasing demand for tourism English learning, traditional learning mode analysis methods have limitations in capturing dynamic behaviour and personalised recommendations. This article proposes a tourism English learning mode mining framework that integrates temporal clustering and ensemble learning, aiming to extract multidimensional learning features from time series data and construct a high-precision prediction model. Firstly, the behaviour trajectory of learners is segmented using temporal clustering algorithm to identify their time distribution characteristics and knowledge mastery rhythm at different learning stages. Secondly, an ensemble learning model is used to fuse multi-dimensional features of clustering results, achieving learning effect prediction and pattern classification. In addition, the study revealed the nonlinear correlation between contextualised vocabulary memory and listening and speaking ability development in tourism English learning, providing data-driven decision support for the development of adaptive learning systems.

Keywords: temporal clustering; ensemble learning; attention mechanism; tourism English.

Reference to this paper should be made as follows: Jing, Y. (2025) 'Mining of tourism English learning mode based on temporal clustering and ensemble learning', *Int. J. Information and Communication Technology*, Vol. 26, No. 34, pp.45–59.

Biographical notes: Yan Jing received her Bachelor's degree from the Henan University in June 2005 and Master's degree from Henan University in June 2009. She is currently working at the Zhengzhou Tourism College. Her research interests include foreign linguistics, English teaching and applied linguistics.

1 Introduction

With the booming development of global tourism and the increasing frequency of cross-cultural communication, optimising the learning mode of tourism English as a language skill in specific scenarios has become an important issue in the field of educational technology (Wei, 2021). Traditional language learning analysis is often based on static features such as test scores and learning duration for linear modelling, which makes it difficult to capture the dynamic evolution of learner behaviour (Arifin et al., 2022). Especially in the context of tourism English, learners' vocabulary acquisition and listening and speaking ability development exhibit significant situational dependence and nonlinear temporal characteristics, which pose a dual challenge to existing analysis methods: on the one hand, it is necessary to effectively model the behavioural pattern drift in the time dimension, and on the other hand, it is necessary to solve the fusion problem of multi-source heterogeneous features (such as interaction frequency and knowledge mastery rhythm).

In recent years, significant progress has been made in the application of time-series data analysis technology in the field of education. Zheng et al. (2020) proposed a convolutional neural network model FWTS-CNN that integrates feature weighting and behavioural time series. It extracts continuous behavioural features from learners' learning activity logs, filters key features, and sorts them according to importance based on decision trees. Then, it weights the continuous behavioural features based on importance, and finally constructs a convolutional neural network model based on behavioural time series and weighted features. For modelling language learning behaviour, Huang and Samonte (2025) proposed a new cross-pattern adversarial learning framework. This framework combines multi-level feature extraction and transformer CNN-LSTM to more effectively process multimodal data and capture integrated models of complex relationships. Then extract low-level and high-level features from the original multimodal data. Meanwhile, transformer is utilised to mine long-range dependencies, CNN extracts local features, and long short-term memory (LSTM) is used to simulate time series. In terms of ensemble learning, Wu et al. (2024) proposed a multi-layer feature construction method that separates the mining of textual and numerical information, solving the problem of insufficient exploration of textual data in existing user profile processing methods. It is worth noting that Geng and Yamada (2023) used learning applications to understand the learning process and behavioural patterns of using augmented reality to acquire compound verbs. Perform lag sequence analysis using learning data to analyse changes in learning behaviour. In addition, frequent sequence mining is used to analyse frequent sequences and compare the learning behaviour patterns of learners with different learning performances.

Although the above research has achieved certain results, existing methods still have three limitations:

1. Most time series analysis frameworks adopt fixed time window segmentation strategies, which cannot adaptively identify the compound rhythm of 'situational modularisation' and 'skill progression' in tourism English learning.
2. The feature fusion mechanism relies heavily on artificial prior knowledge, making it difficult to capture implicit associations between cross-stage behavioural patterns.
3. The existing models lack specificity in identifying inefficient repetitive learners, resulting in weak targeted personalised intervention measures.

This article proposes an innovative framework that integrates temporal clustering and ensemble learning to address the aforementioned issues. Firstly, an improved DBSCAN algorithm based on knowledge unit perception is designed to dynamically adjust clustering density parameters in typical scenarios of tourism English, such as hotel check-in and scenic spot guidance; secondly, a multi-granularity feature cross-over mechanism is constructed, utilising attention weighted fusion of stage labels and original behaviour sequences output by temporal clustering; finally, develop a Shapley value analysis module for pattern interpretation to reveal the differential impact paths of key behavioural features on three typical learning patterns (efficient/intermittent/inefficient). Experimental results have shown that this method not only significantly improves the accuracy of pattern classification, but also provides insights into the generation of adaptive teaching strategies through visualised temporal evolution graphs.

2 Relevant technologies

2.1 Temporal clustering

Temporal clustering is a key technology at the intersection of time series data analysis and unsupervised learning, aimed at identifying sets of objects with similar evolutionary patterns from dynamic and ordered observational data (Chen et al., 2022). Compared to traditional static clustering methods, temporal clustering not only focuses on the distribution pattern of data in the feature space, but also emphasises the continuity of behaviour, stage transition rules, and pattern drift characteristics in the time dimension.

The core task of time series clustering is to segment and classify time series data (Melser et al., 2024). Its input is a multidimensional set of observation sequences indexed by timestamps, and its output is a cluster structure with internal consistency (intra cluster homogeneity) and external differentiation (inter cluster heterogeneity) (Moosavi et al., 2024). Its particularity is reflected in three aspects: the observed values at adjacent time points usually have autocorrelation, and the traditional assumption of independence no longer holds (Long et al., 2023). For example, the daily practice duration of tourism English learners may be positively influenced by the previous day's learning outcomes; the clustering pattern may undergo non-stationary changes over time and requires detection of phase transitions. In the process of language acquisition, the accumulation of vocabulary often presents an S-shaped growth curve with alternating 'plateau period explosive period'; behavioural patterns may exhibit different clustering characteristics at different time granularities (such as hours, weeks and months), and it is necessary to choose an analysis scale that matches the task objectives (Park et al., 2024).

According to the representation of time series and similarity measurement strategies, mainstream temporal clustering algorithms can be divided into the following four categories:

1. Clustering based on the original sequence. Directly calculating the similarity of the original time series, representative methods include:
   - Dynamic time warping (DTW): By aligning the timeline elastically, it solves the problems of inconsistent sequence length and local time offset, and is suitable for analysing behaviour logs with irregular intervals in tourism English learning (Han and Lee, 2023).
   - K-means++extension: By improving the initial centroid selection strategy, the convergence stability of traditional K-means algorithm in temporal data is enhanced, commonly used for learning coarse-grained segmentation of duration distributions.

2 Clustering based on feature extraction. Extract statistical features (such as mean and variance), time-domain features (such as autocorrelation coefficients), or frequency-domain features (such as wavelet coefficients) from the original sequence, and then perform clustering in the feature space. For example, by extracting the sliding window variance of learners' weekly test scores, the fluctuation characteristics of their knowledge mastery can be quantified (Chen et al., 2021).

3 Clustering based on model parameters. Assuming that the time series follows a specific generative model (such as ARIMA, hidden Markov model), the sequence dynamics are characterised by clustering model parameters (Benevento et al., 2024). For example, using the state transition probability matrix of the hidden Markov model, the pattern switching patterns in tourism English listening and speaking exercises can be identified.

4 Time series extension based on density clustering. By improving the temporal awareness of density clustering algorithms such as DBSCAN, dense regions with spatiotemporal proximity can be identified (Oyewole and Thopil, 2023).

Temporal clustering has shown extensive potential in educational data analysis, especially in the field of language learning behaviour modelling. Its typical application scenarios include three directions: learner behaviour grouping, teaching stage division, and abnormal pattern detection. For example, by analysing the temporal characteristics of learners' daily practice duration, test score fluctuations, etc., three typical groups can be identified: efficient (continuous and stable investment), intermittent (periodic strong weak alternation), and inefficient (random scattered learning), which can provide a basis for personalised resource recommendation. In the division of teaching stages, temporal clustering can detect the turning point of knowledge mastery. For example, when the vocabulary accumulation of tourism English learners exceeds a certain critical threshold, the rate of improvement in their listening and speaking abilities may undergo a significant transition. At this time, the clustering results can provide decision signals for the stage switching strategy of the adaptive learning system. In addition, an improved algorithm based on density clustering has been successfully applied to abnormal behaviour detection in programming learning. This method identifies learning fragments in low-density areas (such as long-term stagnation or impulsive learning) and can be applied to tourism English scenarios to discover inefficient repetitive learning patterns. However, the practical application of temporal clustering still faces multiple challenges: firstly, the problem of high-dimensional noise interference is prominent, and tourism English learning data often contains non-semantic behaviour records such as device misoperation and network delay, which need to be preprocessed through wavelet denoising or density based outlier filtering; secondly, it is difficult to align variable length sequences, and the differences in learner participation periods lead to varying lengths of raw data. Traditional Euclidean distance measures are ineffective, and elastic alignment methods such as DTW or longest common subsequence (LCSS) need to be introduced; thirdly, concept drift is a common phenomenon, and learning patterns may gradually or abruptly change with course progress or external environment (such as exam pressure). It is necessary to design sliding window mechanisms or online clustering algorithms to achieve dynamic pattern updates.

When applying temporal clustering to mining tourism English learning patterns, special attention should be paid to its two core characteristics of contextual correlation and nonlinear cumulative effects. Firstly, the formation of tourism English proficiency is highly bound to specific scenarios such as hotel check-in and scenic spot guidance, and learners' vocabulary memory, listening and speaking training, and other behaviours often exhibit contextual clustering characteristics. For example, learning behaviours related to airport customs clearance scenarios may focus on short-term high-intensity exercises, while learning related to cultural differences and communication may manifest as long-term low-frequency interactions. This requires clustering algorithms to recognise knowledge unit boundaries driven by scenarios, rather than simply relying on temporal proximity. Secondly, there is a significant nonlinear threshold effect in the improvement of language ability, which means that learners' knowledge mastery may experience a sudden qualitative transition after long-term quantitative accumulation (such as the ability transition from mechanical memory to natural application). This characteristic makes it difficult for traditional linear interpolation methods (such as uniform time segmentation) to accurately capture the transition points during the learning stage, and clustering strategies based on density changes or hidden state transitions need to be adopted. For example, by analysing the second derivative of test scores (representing the acceleration of progress), the critical state of the 'plateau burst period' can be more sensitively detected. These two special characteristics jointly determine that the temporal clustering of tourism English requires deep integration of domain knowledge: on the one hand, it is necessary to design a feature encoding strategy for scene perception, embedding tourism context labels (such as scene type, task complexity) into similarity measurement functions; on the other hand, it is necessary to establish a dynamic threshold adjustment mechanism to enable clustering algorithms to adaptively identify pattern jump features in nonlinear evolution, thereby providing theoretical support for accurate judgement of educational intervention timing.

2.2 Ensemble learning

Ensemble learning significantly improves the generalisation ability and robustness of models by collaborating the prediction results of multiple base learners, and is one of the core paradigms in the field of machine learning (Zhang et al., 2022). Its core idea stems from the theory of group intelligence, that is, by reasonably combining the output of multiple weak learners, the deviation or variance defects of a single model can be remedied, and the '1 + 1 > 2' decision optimisation effect can be achieved.

The essence of ensemble learning is to construct high-performance models through a two-stage strategy of diversity generation and result fusion. Its effectiveness depends on two basic conditions: first, the base learner needs to have a certain level of accuracy (at least better than random guessing), and second, the prediction errors between different learners should be as uncorrelated as possible (Mian et al., 2024). Compared with a single model, the core advantages of integrated learning are reflected in three aspects: first, improving generalisation performance by reducing model variance (such as bagging) or bias (such as boosting); secondly, enhance tolerance for noisy data and outliers to avoid overfitting risks; thirdly, it supports joint modelling of multimodal feature spaces, suitable for heterogeneous fusion scenarios of temporal behaviour data and static knowledge graphs in tourism English learning (Mienye and Sun, 2022).

According to the generation method and combination strategy of base learners, ensemble learning can be divided into the following three categories:

1 Parallelisation method: Multiple training subsets are generated through bootstrap sampling, and independent training base learners are trained using voting or averaging strategies to aggregate the results. It performs well in dealing with class imbalance problems and can effectively identify minority class patterns of inefficient repetitive learners.

2 Serialisation method: Iteratively adjust sample weights or model weights, so that subsequent base learners focus on correcting the prediction errors of the preceding model. AdaBoost weights misclassified samples through an exponential loss function, while XGBoost and LightGBM become the preferred tools for processing large-scale time-series data through gradient optimisation and efficient feature binning techniques. This type of method has strong modelling ability for the non-stationary distribution of 'intermittent reinforcement' behaviour in tourism English learning.

3 Heterogeneous model integration: By integrating the prediction results of different types of base models (such as decision trees and neural networks) through meta learners, we can fully utilise the complementarity of the models.

The temporal, multi-source heterogeneous, and pattern implicit characteristics of tourism English learning data make ensemble learning an ideal choice for modelling in this field (Yang et al., 2023). Firstly, the requirement for temporal dynamic modelling requires algorithms to capture the long-term evolution patterns and short-term fluctuation characteristics of learner behaviour. By integrating basic models such as LSTM network and Prophet (time series prediction model), it is possible to collaboratively analyse the gradual accumulation trend of vocabulary (such as monthly growth curve) and sudden behaviour (such as peak value of concentrated review before exams). Secondly, the ability to fuse multimodal features is the core advantage in processing multi-source data of tourism English. For example, learners' text practice logs (structured data), speech pronunciation scores (temporal signals), and interface interaction heatmaps (spatial features) require collaborative processing of heterogeneous models: random forests excel at mining statistical patterns in structured logs, convolutional networks can extract local patterns in speech spectrograms, and attention mechanisms can focus on key areas in interaction heatmaps. Finally, pattern implicitness requires models to have strong nonlinear representation capabilities to deconstruct complex behavioural associations. The integration of XGBoost and LightGBM can jointly identify 'high-frequency but low-quality' exercise features among 'inefficient repetitive' learners (such as repeating the same questions multiple times a day without improving accuracy), which are easily masked by noise in a single linear model (Ngo et al., 2022).

Although ensemble learning has significant advantages in tourism English analysis, its practical application still needs to overcome three bottlenecks. Firstly, the contradiction between computational overhead and real-time performance is prominent, and the training and inference costs of large-scale integration (such as hundred model level stacking) are high, making it difficult to deploy directly to resource limited educational terminal devices. Redundant base learners can be removed through model pruning, or lightweight inference can be achieved through hardware acceleration techniques such as GPU parallelisation (Matloob et al., 2021). Secondly, the problem of concept drift adaptation urgently needs to be addressed, as learners' behavioural patterns may undergo sudden changes due to course schedule adjustments, external interventions (such as teacher feedback), or environmental changes (such as tourism policy updates). The traditional batch training mode cannot respond to such dynamic changes in a timely manner, and an online integration framework (such as incremental boosting) needs to be designed to dynamically update the model weights, for example, by using a sliding window mechanism to only retain the behaviour data from the past three months for training. Thirdly, insufficient embedding of domain knowledge may lead to models deviating from educational laws, and purely data-driven integration may overlook cognitive science theories (such as the exponential decay characteristics of forgetting curves). Improvement directions include: introducing a memory decay penalty term in the loss function, forcing the model to focus on the impact of recent learning behaviour; alternatively, a hybrid base learner can be constructed by combining cognitive diagnostic models to ensure that the integrated results meet both data fitting and educational psychology constraints. Future research needs to further explore integrated architectures that are lightweight, adaptive, and enhance domain knowledge to support the real-time and precise needs of tourism English learning analysis.

3 Temporal clustering and ensemble learning framework

3.1 Sorting target feature extraction

This section proposes a temporal clustering and ensemble learning framework (TCELF) that integrates temporal clustering and ensemble learning, aiming to achieve accurate recognition and interpretability analysis of learning patterns through dynamic behaviour segmentation and multimodal feature fusion. The overall process of the method is shown in Figure 1.

In the data preprocessing stage, the first step is to address the issue of non-random missing values in tourism English learning data. Based on the assumption of time proximity, linear interpolation is performed on the missing values within time window $t \in (t_k, t_{k+1})$:

$$\hat{x}_t^{(i)} = \frac{(t_{k+1} - t)\, x_{t_k}^{(i)} + (t - t_k)\, x_{t_{k+1}}^{(i)}}{t_{k+1} - t_k} \tag{1}$$

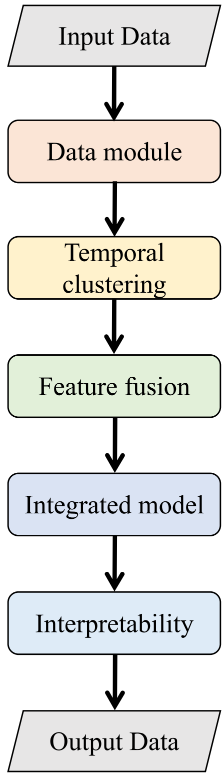
where $x_{t_k}^{(i)}$ and $x_{t_{k+1}}^{(i)}$ are the effective observation values before and after the missing window, respectively. This strategy ensures temporal continuity while avoiding the introduction of false fluctuations. To further enhance the semantic information of the scene, standardise the features according to the scene s:

$$z_{t,s}^{(i)} = \frac{x_{t,s}^{(i)} - \mu_s}{\sigma_s} \tag{2}$$

$$\mu_s = \frac{1}{|T_s|} \sum_{t \in T_s} x_{t,s}^{(i)} \tag{3}$$

where T_s represents the timestamp set of scene s, and standardised features can eliminate dimensional differences and preserve semantic boundaries between scenes.

Figure 1 Method flowchart (see online version for colours)



The temporal behaviour modelling module divides the learning stages through an improved density clustering algorithm. The algorithm first dynamically adjusts the neighbourhood radius based on the scene semantics:

$$\varepsilon(s_t) = \alpha \cdot \bar{d}_{intra}(s_t) + (1-\alpha) \cdot \bar{d}_{inter} \tag{4}$$

where $\bar{d}_{intra}$ and $\bar{d}_{inter}$ are the average distances within and across

scenes, respectively, and $\alpha \in [0, 1]$ controls the weight of scene specificity. Further define the spatiotemporal joint distance metric function:

$$D_{SADC}(\tilde{x}_t, \tilde{x}_{t'}) = \left\|\tilde{x}_t - \tilde{x}_{t'}\right\|_2 \cdot e^{-\lambda|t-t'|} \tag{5}$$

where $\lambda = 0.05$ is the time decay factor, which strengthens the behavioural correlation of time adjacent points. The clustering results are merged with adjacent similar stages through post-processing to generate C behavioural stages $\{P_1, P_2, \ldots, P_C\}$.

The feature fusion module extracts multi-granularity features from the clustering results. For each stage P_C, calculate the mean and variance features:

$$\bar{x}_c = \frac{1}{|P_c|}\sum_{t \in P_c} \tilde{x}_t \tag{6}$$

$$\sigma_c^2 = \frac{1}{|P_c|}\sum_{t\in P_c}(\tilde{x}_t - \tilde{x}_c)^2 \tag{7}$$

The central tendency and fluctuation characteristics of behaviour during the quantification phase. Simultaneously calculate the transition differences between adjacent stages:

$$f_{trans}^{(c)} = f_{stat}^{(c)} - f_{stat}^{(c-1)} \tag{8}$$

Reflect the sudden changes in learning pace. The original time-series data is dimensionality reduced through principal component analysis:

$$f_{raw}^{(i)} = W_{PCA}^T X^{(i)} \in \mathbb{R}^d \tag{9}$$

The final concatenation is a multimodal feature vector $F^{(i)} \in \mathbb{R}^{8M+d}$.

Integrated modelling adopts a hierarchical architecture. LSTM units update modelling temporal dependencies through hidden states:

$$h_t = \sigma\left(W_h\left[h_{t+1}; F_t^{(i)}\right] + b_h\right) \tag{10}$$

where σ is the Sigmoid function, and W_h and b_h are learnable parameters.

The pattern interpretation module quantifies feature contributions based on Shapley values:

$$\phi_j^{(i)} = \sum_{S\in D\setminus\{j\}} \frac{|S|!(D-|S|-1)!}{D!}\left[\hat{y}(S\cup\{j\} - \hat{y}(S))\right] \tag{11}$$

Analyse the key driving factors of efficient, intermittent, and inefficient modes through feature masking experiments.

This framework provides a systematic methodology for the dynamic analysis of tourism English learning behaviour. Firstly, data is preprocessed through temporal interpolation and scene standardisation, and behaviour stages are divided using dynamic density clustering. After extracting statistical and transfer features, an integrated model combining LSTM, LightGBM, and gating is constructed. Finally, the decision logic of the model is analysed through Shapley value parsing.

4 Experiment

To verify the effectiveness of TCELF, this section conducted multidimensional experiments based on a real tourism English learning dataset, covering pattern recognition accuracy, clustering quality, generalisation ability, and interpretability analysis, and compared them with mainstream baseline methods. The experimental data is sourced from an online education platform, covering 1,254,790 behavioural records of 2,318 learners, including temporal features such as learning duration, test accuracy, scene switching frequency, as well as static features such as initial language proficiency (CEFR level). The data was divided into training set, validation set, and testing set in a ratio of 7:2:1. All experiments were repeated five times and the mean was taken to reduce the impact of randomness.

4.1 Experimental setup and baseline method

The baseline method selects four representative models:

1 K-means + RF (traditional time-series clustering combined with random forest)

2 dynamic time warping alignment hidden Markov model (DTW-HMM)

3 LSTM only (single LSTM time series classification model)

4 DeepCluster (self-supervised deep clustering).

The key parameters of TCELF are set as follows: the neighbourhood radius adjustment factor of the temporal clustering module is 0.7, the time decay factor is 0.05, the LSTM hidden layer dimension in ensemble learning is 128, and the number of LightGBM trees is 200. The evaluation indicators include pattern classification accuracy, F1 score (balancing the recognition ability of efficient and inefficient types), silhouette score to measure clustering quality, and Shapley consistency (SC) to quantify the Spearman correlation between feature contributions and educational theory.

4.2 Comparison of model classification performance

To evaluate the performance advantages of TCELF in three types of learning modes (efficient, intermittent and inefficient) classification tasks, this experiment compared four mainstream baseline methods: K-means + RF, DTW-HMM, LSTM only and DeepCluster. The experimental dataset contains temporal behaviour records of 2,318 learners, divided into training, validation, and testing sets in a 7:2:1 ratio. All models use the same data preprocessing process to ensure fairness. As shown in Table 1 and Figure 1, the classification accuracy of TCELF on the test set reached 89.7%, which is 7.2 percentage points higher than the suboptimal DeepCluster (82.5%); the F1 score is 87.3%, which is 8.5 percentage points higher than DeepCluster (78.8%). Especially in efficient (minority class) recognition, its F1 score is 12.6% higher than LSTM only. Figure 2 visually compares the accuracy and F1 score of each model through a bar chart. It can be seen that TCELF significantly alleviates the problem of class imbalance through the feature fusion and heterogeneous integration strategy guided by temporal clustering. The results indicate that TCELF can effectively capture the complex correlation between dynamic behaviour patterns and static ability profiles in tourism English learning, providing reliable evidence for subsequent educational interventions.

Table 1 Comparison of model classification performance

Method	*Accuracy (%)*	*F1-score (%)*
K-means + RF	76.3	72.1
DTW-HMM	81.5	78.4
LSTM-only	83.2	79.8
DeepCluster	82.5	78.8
TCELF	89.7	87.3

Figure 2 Comparison chart of model classification performance (see online version for colours)

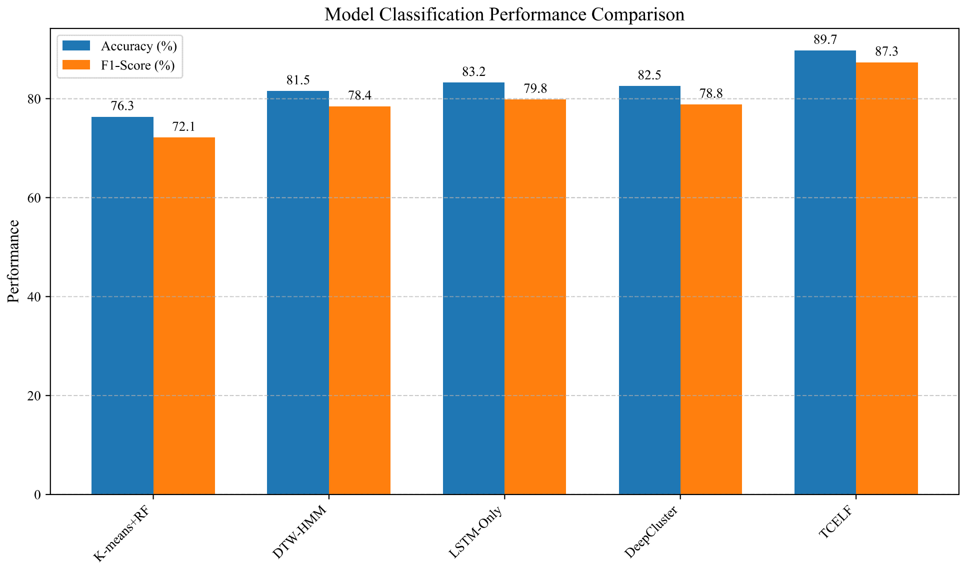
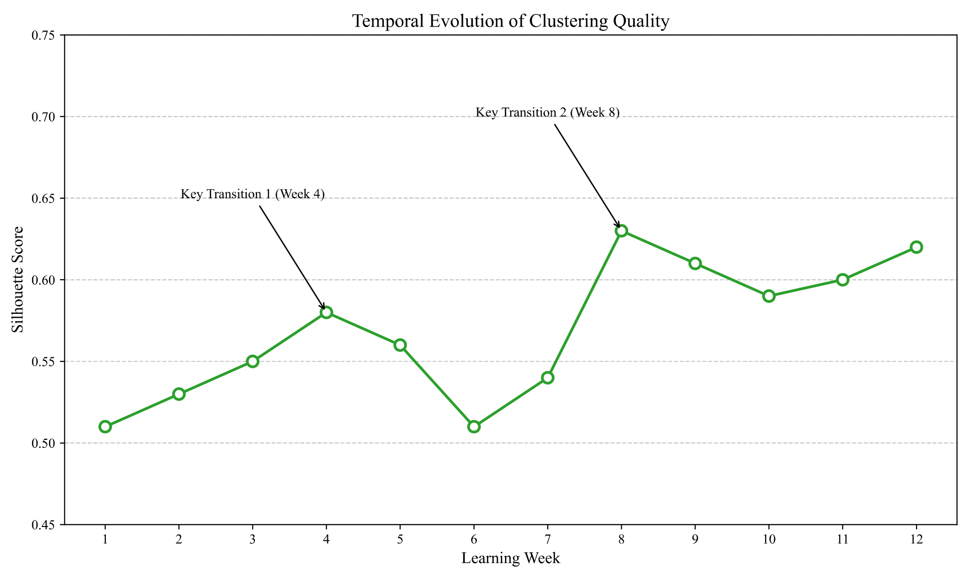


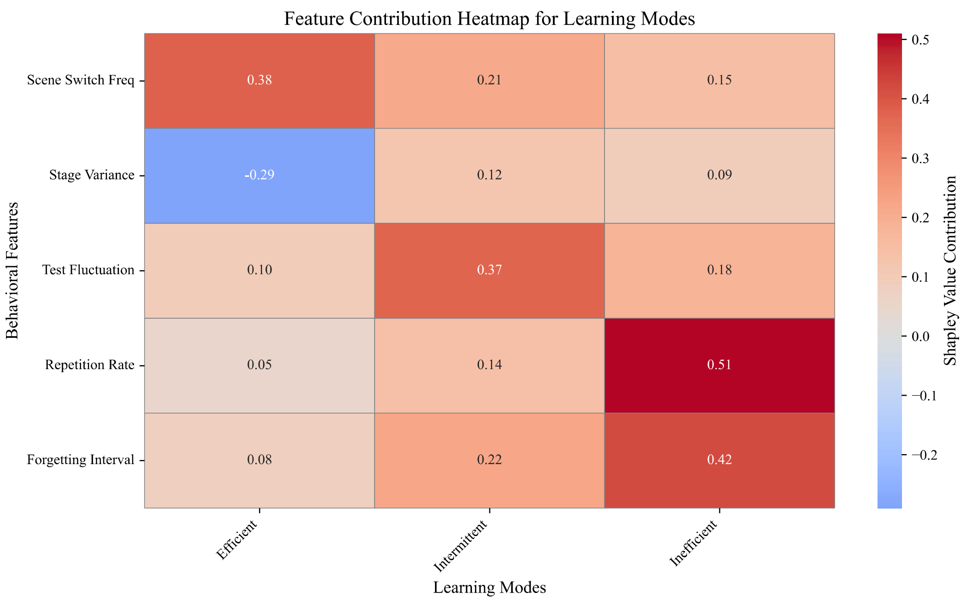
Figure 3 Analysis of cluster quality evolution over time (see online version for colours)



4.3 *Dynamic evolution analysis of cluster quality*

To further explore the ability of TCELF temporal clustering module to capture learning stage transitions, this experiment quantifies the clustering quality of different learning cycles through silhouette score and analyses its correlation with educational psychology theory. As shown in Figure 3, the silhouette coefficient exhibits significant fluctuations during the learning cycle: it reaches its first peak (0.58) in the fourth week, corresponding to the stage of 'basic scene vocabulary accumulation completion' in tourism English learning; the second peak (0.63) appeared in the 8th week, reflecting the transition of learners from 'modular training' to 'comprehensive application transfer'. It is worth noting that in the sixth week, due to the introduction of high complexity scenes such as cultural differences, the contour coefficient briefly decreased to 0.51, indicating that the algorithm is sensitive to dynamic changes in the scene. This result is highly consistent with the 'platform explosion' theory, verifying the educational rationality of TCEFL in dividing learning stages. By displaying the temporal evolution of contour coefficients through a line chart, decision signals for stage switching can be provided for adaptive teaching systems, such as recommending cross-scenario exercises in week 4 to consolidate knowledge, or adding comprehensive simulation training in week 8 to enhance application abilities.

Figure 4 Feature contribution heatmap (see online version for colours)



4.4 *Explanatory analysis of feature contribution*

To reveal the key behavioural features that affect the classification of learning patterns, this experiment quantifies the contribution of features based on Shapley values and visualises their cross-pattern differences through heat maps. As shown in Figure 4, the core driving characteristics of efficient learners are 'scene switching frequency' (contribution of 0.38) and 'intra stage variance' (–0.29), indicating a positive synergistic effect between multi-scene alternating practice and stable learning rhythm; the low efficiency mode is strongly related to the 'repetition rate' (0.51) and 'forgetting interval' (0.42), reflecting the negative impact of mechanical repetition and long-term interruption on learning efficiency; the intermittent mode is dominated by 'test score fluctuations' (0.37), which is consistent with its alternating behaviour characteristics of 'assault stagnation'. In addition, the contribution of 'initial language level' to the three types of patterns in the heatmap is less than 0.1, indicating that dynamic behavioural features are more discriminative than static attributes. This result provides a direct basis for teachers to develop personalised intervention strategies, such as designing interval repetition algorithms for inefficient learners to shorten forgetting cycles, or providing stability reinforcement training for intermittent learners.

The experimental results show that TCELF has significant advantages in mining tourism English learning patterns: through scene time joint modelling, it accurately identifies three types of groups: 'efficient intermittent inefficient'; the clustering results reveal key transition points in the learning stage, providing stage switching signals for adaptive teaching systems; interpretability analysis identifies the core characteristics of inefficient repetitive behaviour, such as long forgetting intervals, and guides teachers in designing targeted reinforcement strategies.

5 Conclusions

This article proposes an innovative framework that integrates temporal clustering and ensemble learning to address the core issues of difficulty in capturing dynamic behaviour patterns and insufficient personalised recommendations in tourism English learning scenarios. By designing a scenario adaptive density clustering algorithm, the time boundaries of learner behaviour stages are dynamically identified, effectively addressing the limitations of traditional methods in modelling contextual dependencies and nonlinear learning patterns; the multimodal feature fusion mechanism constructed on this basis, combined with the complementary advantages of temporal modelling and statistical learning, achieves accurate discrimination of efficient, intermittent and inefficient learning modes. Further introduction of the interpretability analysis module based on game theory reveals the differential impact path of key behavioural characteristics such as scene switching frequency and stage stability on pattern classification, providing theoretical support for educational intervention strategies. Experiments have shown that this framework not only significantly improves pattern recognition accuracy, but also guides the teaching system to trigger personalised resource recommendations at appropriate times through stage transition detection and feature contribution visualisation. The research results provide a new paradigm for the synergy between temporal perception and educational cognition in language learning behaviour analysis, demonstrating direct application value in adaptive education platforms. In the future, it can be further expanded to multi-language skill transfer and real-time interaction scenario optimisation.

Declarations

All authors declare that they have no conflicts of interest.

References

Arifin, A., Norain, S. and Ridwan, R. (2022) 'ESP course design: the need analysis on English for tourism book for travel business department of eleventh grade students at SMKN 1 Tarakan', *Technium Education and Humanities*, Vol. 2, No. 2, pp.1–7.

Benevento, A., Durante, F. and Pappadà, R. (2024) 'Tail-dependence clustering of time series with spatial constraints', *Environmental and Ecological Statistics*, Vol. 31, No. 3, pp.801–817.

Chen, J., Du, C., Zhang, Y., Han, P. and Wei, W. (2021) 'A clustering-based coverage path planning method for autonomous heterogeneous UAVs', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 23, No. 12, pp.25546–25556.

Chen, Y., Shu, T., Zhou, X., Zheng, X., Kawai, A., Fueda, K., Yan, Z., Liang, W. and Wang, K.I-K. (2022) 'Graph attention network with spatial-temporal clustering for traffic flow forecasting in intelligent transportation system', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 24, No. 8, pp.8727–8737.

Geng, X. and Yamada, M. (2023) 'Using learning analytics to investigate learning processes and behavioural patterns in an augmented reality language learning environment', *Journal of Computer Assisted Learning*, Vol. 39, No. 2, pp.532–546.

Han, M. and Lee, J-S. (2023) 'Temporal peak-management and clustering system for sustainable learning with application to the excimer laser annealing process', *IEEE Access*, Vol. 11, No. 3, pp.24535–24544.

Huang, A. and Samonte, M.J.C. (2025) 'Dynamic analysis of classroom engagement sentiment based on multilevel feature extraction and transformer CNN-LSTM integrated model: personalized behavior prediction for non-English learners in a cross-modal adversarial learning framework', *Journal of Combinatorial Mathematics and Combinatorial Computing*, Vol. 127, No. 5, pp.7153–7177.

Long, Y., Ouyang, X., Yan, C., Wu, Z., Huang, X., Pu, W., Cao, H., Liu, Z. and Palaniyappan, L. (2023) 'Evaluating test-retest reliability and sex-/age-related effects on temporal clustering coefficient of dynamic functional brain networks', *Human Brain Mapping*, Vol. 44, No. 6, pp.2191–2208.

Matloob, F., Ghazal, T.M., Taleb, N., Aftab, S., Ahmad, M., Khan, M.A., Abbas, S. and Soomro, T.R. (2021) 'Software defect prediction using ensemble learning: a systematic literature review', *IEEE Access*, Vol. 9, No. 2, pp.98754–98771.

Melser, R., Coops, N.C. and Derksen, C. (2024) 'Characterizing satellite-derived freeze/thaw regimes through spatial and temporal clustering for the identification of growing season constraints on vegetation productivity', *Remote Sensing of Environment*, Vol. 309, No. 21, p.114210.

Mian, Z., Deng, X., Dong, X., Tian, Y., Cao, T., Chen, K. and Al Jaber, T. (2024) 'A literature review of fault diagnosis based on ensemble learning', *Engineering Applications of Artificial Intelligence*, Vol. 127, No. 12, p.107357.

Mienye, I.D. and Sun, Y. (2022) 'A survey of ensemble learning: concepts, algorithms, applications, and prospects', *IEEE Access*, Vol. 10, No. 4, pp.99129–99149.

Moosavi, V., Mostafaei, S. and Berndtsson, R. (2024) 'Temporal cluster-based local deep learning or signal processing-temporal convolutional transformer for daily runoff prediction?', *Applied Soft Computing*, Vol. 155, No. 5, p.111425.

Ngo, G., Beard, R. and Chandra, R. (2022) 'Evolutionary bagging for ensemble learning', *Neurocomputing*, Vol. 510, No. 2, pp.1–14.

Oyewole, G.J. and Thopil, G.A. (2023) 'Data clustering: application and trends', *Artificial Intelligence Review*, Vol. 56, No. 7, pp.6439–6475.

Park, T.K., Liew, H-P. and Lee, H. (2024) 'The potential influences of COVID-19 and social vulnerability measures on child maltreatment in the United States: spatial temporal cluster analysis and negative binomial regression', *Child Abuse & Neglect*, Vol. 157, No. 22, p.107082.

Wei, X. (2021) 'A classification method of tourism English talents based on feature mining and information fusion technology', *Mobile Information Systems*, Vol. 2021, No. 1, p.5520079.

Wu, D., Du, X. and Peng, F. (2024) 'Multi-layer and multi-source features stacking ensemble learning for user profile', *Electric Power Systems Research*, Vol. 229, No. 8, p.110128.

Yang, Y., Lv, H. and Chen, N. (2023) 'A survey on ensemble learning under the era of deep learning', *Artificial Intelligence Review*, Vol. 56, No. 6, pp.5545–5589.

Zhang, Y., Liu, J. and Shen, W. (2022) 'A review of ensemble learning algorithms used in remote sensing applications', *Applied Sciences*, Vol. 12, No. 17, p.8654.

Zheng, Y., Gao, Z., Wang, Y. and Fu, Q. (2020) 'MOOC dropout prediction using FWTS-CNN model based on fused feature weighting and time series', *IEEE Access*, Vol. 8, No. 6, pp.225324–225335.