Research Journal in Advanced Humanities

https://doi.org/10.58256/0a15m346

RJAH

Open Access

Check for updates

RESEARCH ARTICLE **Section:** *Literature, Linguistics & Criticism*

ETS-TDSVM for personalized French learning: An advanced framework for Vietnamese police cadets

Hue Thi. Nguyen[1]*

[1]Foreign language Department, Vietnam People's Police Academy, Hanoi, Vietnam

*Correspondence: huenguyen.vppa@gmail.com

ABSTRACT

In the context of globalized policing and international co-operation, the ability to communicate effectively in foreign languages has become increasingly essential for law enforcement officers. This research presents a machine learning (ML)-based frameworks to overcome the limitations of static models and provide real-time, personalized French language instruction at the People's Police Academy of Vietnam. The system leverages learner-specific data to dynamically adapt instructional content and progression, thereby enhancing language acquisition efficiency in a professional training environment. Data was collected over 16 weeks from 378 police cadets in Vietnam using a blended learning platform. Sources included weekly digital assessments, pronunciation accuracy metrics, user interaction logs, response times, and in-app feedback. The framework integrates the K-means clustering algorithm that initially segments learners based on behavioral and performance patterns. An Enriched Transient Search Tree-tuned Dynamic Support Vector Machine (ETS-TDSVM) model is applied to classify language proficiency levels, enabling adaptive content delivery. Implemented in Python, the ETS-TDSVM classifier achieved a high accuracy of 93.5% in proficiency level prediction. The integration of adaptive parameter optimization through the Enriched Transient Search Tree significantly enhances the convergence speed and predictive reliability of the Dynamic Support Vector Machine. Experimental evaluation demonstrates improved learner clustering consistency and reduced misclassification rates compared with baseline models. The results indicate that data-driven adaptive learning systems can substantially improve personalized instruction and learner engagement in specialized professional education environments. The high accuracy of the proposed framework demonstrates the potential to enhance language acquisition efficiency in specialized professional training institutions globally.

KEYWORDS: personalized learning, French language acquisition, machine learning (ML), adaptive instruction, police education

Research Journal in Advanced Humanities

Volume 7, Issue 2, 2026

ISSN: 2708-5945 (Print)

ISSN: 2708-5953 (Online)

ARTICLE HISTORY

Submitted: **04 April 2026**

Accepted: 30 **April 2026**

Published: 15 **May 2026**

HOW TO CITE

Nguyen Thi, H. (2026). ETS-TDSVM for personalized French learning: An advanced framework for Vietnamese police cadets. *Research Journal in Advanced Humanities*, 7(2). https://doi.org/10.58256/0a15m346

Published in Nairobi, Kenya by Royallite Global, an imprint of Royallite Publishers Limited

I. Introduction

A. Context and Importance of foreign language learning for police cadets

The necessity for foreign language acquisition in Vietnam has grown to be a requirement for an increasingly globalized society, and law enforcement is one of the fields where transnational business collaboration and communication across borders become more common (Xia et al., 2024). For police officers, knowledge of a second language is another example of a cognitive skill that is important for work involving international operations, diplomacy, and protection of national security (Chen et al., 2024).

B. Traditional teaching methods

Audio-visual materials, including tapes, dialogues, subtitled videos, translation-based instruction, and direct word-sentence mapping, is applied in language teaching (Tanweer & Ismail, 2024). Task-oriented learning using real-life communication and problem-solving enhances contextual vocabulary and grammar acquisition (Praveena & Anupama, 2025). Data-driven educational technologies further enable individualized learning, responsiveness to proficiency levels, and improved outcomes in professional language training (Dong et al., 2024). Figure 1 shows the Core Elements of Personalized Language Training.

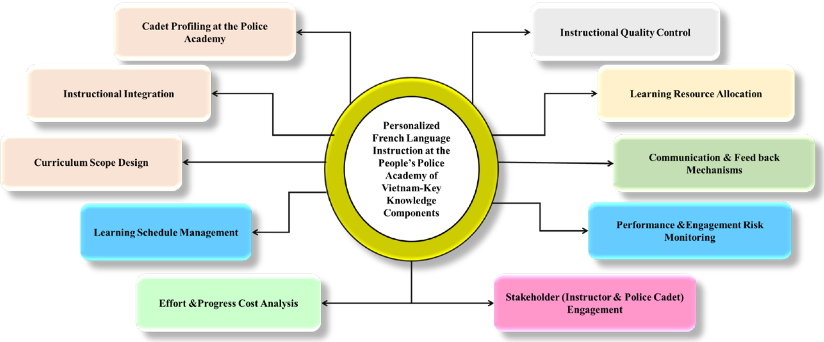

Figure 1 Key Components of Personalized French Instruction

C. Existing AI/ML approaches

The conventional AI and ML systems assist language instruction and enhance traditional education (Istanti et al., 2024). The first method uses NLP with the classical ML algorithms like SVM, Decision Trees, Random Forests, and Naive Bayes (Sajja et al., 2024). These models are efficient in assessing language by testing grammar, vocabulary and syntax. The intelligent tutoring systems provide customized feedback and personalized resources according to the performances of learners (Naseer et al., 2024). Real translation, speech recognition, and conversational practice, which are facilitated by NLP applications, improve learning results and communicative proficiency (Chen et al., 2024). Its applications in practice are automated grammars, scoring pronunciation, and vocabulary suggestions on online platforms (Wei, 2023).

D. Limitations of existing models

The key weakness of the traditional and AI-assisted approaches is their inability to capture intricate patterns with time and user feedback, leading to generic content (Chaipidech et al., 2022). Sequencing and fixed lesson difficulty does not fulfill the individualized and changing needs of police trainees during professional training (Gan et al., 2023).

E. Dynamic and Adaptive Machine Learning Models for Language Instruction

The modern globalized world requires integration and communication in English. Sun (2025) created Gated Recurrent Neural Network (GRNN) model to offer various instruction based on the needs of learners. GRNN is a dynamic module that adapt to a person, leading to a better fluency, lexical richness and contextual language use. On the same note, Song et al. (2024) presented the Dynamic Feedback-Driven Learning Optimization Framework (DFDLOF), which uses machine learning to provide personalized content and real-time feedback.

Both methods foster motivation and individualized learning and provide teachers with useful tools to meet the varied needs of learners by technology.

F. Natural Language Processing and Adaptive Learning Frameworks

Personalized learning is adjusted to the level of the learner, where NLP models Hussain et al. (2024) apply adaptive algorithms and machine learning to evaluate the level and provide individualized content to enhance the engagement, performance, and accessibility. AI-based applications demonstrate potential in promoting language learning, but ethical aspects suggest that there should be responsible integration. The hybrid English learning model, utilizes LSTM networks, is Ezhilmathi et al. (2024) personalizing assignments and material according to the profiles of the learners, which are gathered on Kaggle datasets. Outcomes show improved participation and skills, suggesting the dynamism of LSTM in the personalized learning of language via utilization of the advanced NLP technologies.

G. Traditional Machine Learning Models in Personalized Education

Liu and Yang (2024) argue that Support Vector Machine (SVM) can be used to promote personalized instructions in simulative courses with high accuracy using augmented questionnaire data. They facilitated AI-based personalization in the form of large language models, AI text generation, and knowledge graphs in educational applications. Personalized resources enhanced learning possibilities, and this was proved in LMS-based datasets of SVM, CNN, and XGBoost, which had a high level of accuracy (Kanchon et al., 2024). NLP generators, such as spaCy, GPT-3 or T5, could be used to further customize the content to increase the interest of visual, auditory, and kinesthetic learners.

H. AI-Driven Tools for Interactive and Contextual Language Learning

The globalized society is based on the necessity of learning the language to communicate effectively. One AI-based system that uses Natural Language Processing (NLP) to assist vocational English learners is the Bag of Words (BoW), TF-IDF, and traditional classifiers, including SVM, NB, and ensemble models to predict tense (Peng and Wang, 2025), where SVM and Bagging have high precision. Cai and Li (2024) designed an intelligent chatbot, which is based on NLP and Deep Learning, to facilitate active language learning, providing precise and correct and easy instructions. This method shows considerable prospects of revolutionizing the foreign language education by utilizing the high-tech incorporation.

I. Hybrid and Intelligent Systems for Personalized Language Teaching

The Fuzzy-associated Frequent Pattern-Growth-PELT model of Personalized English Language Teaching is an algorithm that uses fuzzy-association rule mining (combined with frequent pattern-growth) to find meaningful patterns in learners (Chen, 2024). It is responsive to personal needs, facilitating interaction and communication, and experiments demonstrated better learning results and teacher-student interaction. Features are converted and linked with time using speech recognition with MLP-LSTM (Orosoo et al., 2025), which obtain high accuracy and low Word Error Rate. SSA-LSTM (Sun, 2025) integrates Sparrow Search Algorithm with LSTM to refine hyper-parameters and use behavioral data to provide custom learning routes. Nevertheless, the traditional AI methods such as GRNN, SVM, and LSTM are mostly fixed and offer generalized content and reduced flexibility. To overcome it, the ETS-TDSVM model of police cadets in Vietnam combines adaptive classification and dynamic optimization, which allow real-time and personalized training in the French language. ETS-TDSVM leverages behavioral data to measure proficiency, provide personalized tutorials, and dynamically adapt content in professional classes to overcome the drawbacks of classic AI-based language teaching.

J. Research Gaps

The current AI use in education is more engaging, personalized, and efficient, yet the majority of AI tools cannot adequately adapt in real-time and use scarce resources. More than 60 percent of adaptive systems rely on large datasets (Liu and Yang, 2024), limiting their application to target applications such as police academies. GRNN-ELL model (Sun, 2025) involves a lot of data and computation, which is tailored to low-tech application. To deal with this, Vietnamese police cadets ETS-TDSVM system enhances flexibility, less computational overhead, and its operations are reliable with sparse input. Decision Tree-based adaptive sequencing balances performer with content, and is a versatile, responsive solution to professional training.

II. METHODOLOGY

The proposed methodology uses a Personalized French Language Learning Dataset containing performance, behavioral, and interaction data. Preprocessing includes min-max normalization and one-hot encoding of categorical variables. K-means clustering groups learners, and the ETS-TDSVM model predicts proficiency levels. A Decision Tree forecasts topic mastery and guides adaptive content sequencing. The system adjusts difficulty based on learner feedback and task performance. Implemented on a web-based platform, it ensures real-time customization, high accuracy, and scalability for institutional use in Vietnam. Figure 2 illustrates the Pipeline of Data-Driven Personalized Language Learning with ETS-TDSVM.

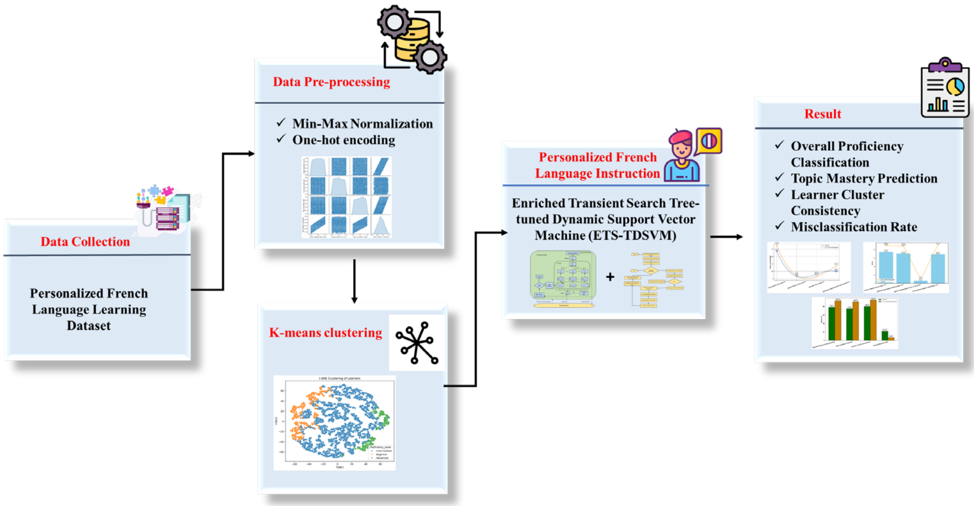

Figure 2 ETS-TDSVM-Based Framework for Personalized French Language Instruction

A. Dataset

This Dataset of Personalized French Language Learning helps to develop personalized instructions of the French language for police cadets at the People's Police Academy of Vietnam. It has performance, interaction, and behavioral data of 378 learners who participated in a 16-week blended learning program. Source: https://www.kaggle.com/datasets/ziya07/personalized-french-language-learning-dataset/data.

B. Data Preprocessing using One-Hot Encoding

One-hot encoding transforms categorical features into binary vectors, preventing ordinal bias. It encodes cadet proficiency, interaction type, and feedback, improving pattern detection and enabling adaptive, individualized content delivery. Equation 1 illustrates this encoding.

With classes labeled 0 to –1, one-hot encoding ensures equal separation (Hamming = 2, Euclidean = √2) and enhances task-based classification performance using cross-entropy loss and Softmax activation.

C. Min-max Normalization

Normalization enhances model performance by removing biases and aligning cadet learning records on the same scale. Scaling features to [0,1] and zero-centering captures individual learning habits and performance, improving proficiency prediction and enabling data-driven personalization of instructional content and pacing at the People's Police Academy of Vietnam. Equation 2 shows min-max normalization, and Figure 3 presents a Pairplot of normalized features.

Where, is the original value, is the normalized value, and and are the minimum and maximum values of the feature, respectively.

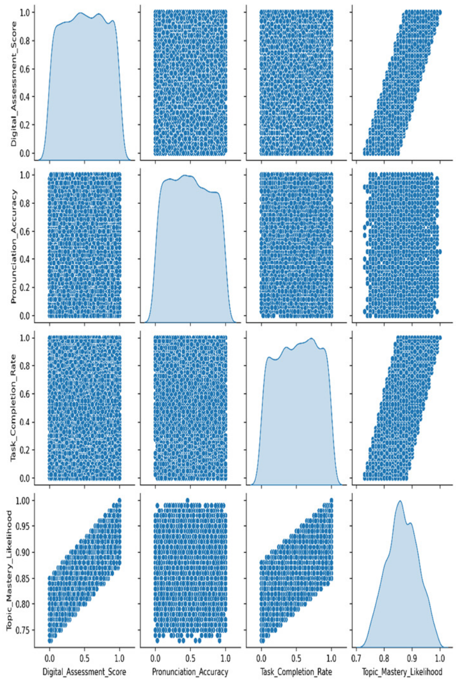

Figure 3 Normalized Correlation Patterns in Cadet Language Performance

D. K-Means clustering

The most significant step is the choice of centroids in each cluster and their location to construct significant learner groups. Assigning cadet data points to the closest centroids forms groups of similar performance patterns. Recalculation of centroids is repeated till they converge to consistent and well-defined clusters. This process makes the model pick local temporal and spatial dynamics, which can be used to categorize language proficiency levels more accurately to facilitate adaptive content delivery. As a result, police trainees get more effective and individualized learning. The K-Means clustering is shown in Equation 3.

In the M-dimensional feature space, is a K-cluster partition of the entity set represented by vectors . It is made up of non-empty, non-overlapping clusters, , each of which has a centroid . Figure 4 shows Proficiency-Level Segmentation of Learners Using K-means Dimensionality Reduction.

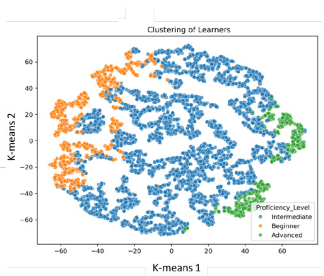

Figure 4 K-Means Clustering of Learner Proficiency Levels

E. ETS-TDSVM

ETS-TDSVM is a sophisticated classification system, which combines DSVM and ETST to automatically optimize hyper-parameters. The former batches the data stream sequentially whereas ETST finds the candidate parameters through chaotic mapping, opposition-based learning, adaptive inertia weights and neighborhood dimensional learning. These parameters update or train the DSVM model. Performance is assessed at the end of every iteration, and ETST concentrates on successful areas of hyper-parameters and then creates new candidates. The process will be repeated until convergence, which allows the adaptive classification of the proficiency levels in learning the French language among the cadets. The stepwise pseudocode is given in Appendix A.

(a). TDSVM

DSVM improves upon SVM by updating parameters according to changing data patterns, enhancing the accuracy of cadet proficiency classification in the customized French learning system in Vietnam. It enables real-time adaptation of instructional content. SVM offers flexibility, global optimization, and strong generalization, improving classification performance by leveraging prior knowledge without depending on feature dimensionality. By eliminating low error sample points from the original variable space, the value-containing function supports sparse planning. Equation 4 provides the most often used -insensitive loss function.

The output of SVM is a linear collection of intermediate nodes, and it looks like a neural network. Consequently, Equation 5 might be used to display the function's form.

Non-parametric kernel functions are commonly represented using SVM. When estimating the value of in Equation 6 for regression, a parametric straight basis function is used to create a Dynamic SVM hybrid, which increases the operational risk of SVM.

A non-linear map denoted by transforms inputs into a high-dimensional space. The non-parametric SVM model's weight is denoted by , given in Equation 7.

Where, the parametric model's basis function is denoted by . is the basis function's weight. Relaxing variables are added based on the structural risk minimization concept to fit the problem easier. Equations 8 and 9, confine the optimization problem, are used to represent the regression issue while lowering the computing complexity and estimation errors.

Using the Radial Base Function (RBF) kernel, which is specified in Equation 10, the inner product in the high-dimensional feature space is implicitly calculated in the low-dimensional input space

The Lagrange multipliers are defined, and the Lagrange multiplier approach is used to build the following dual problem as expressed in Equations 11 and 12.

The Karush-Kuhn-Tucker (KKT) condition is derived and given as shown in Equations 13 and 14. Partial derivative of the problem concerning the variable and The weight vector for the parameter model is given by Equation 15.

KKT conditions provide optimal assistance in selecting the support vectors in Dynamic SVM, relating Lagrange multipliers to predict errors. This improves the sparsity and accuracy of models. Dynamic SVM achieves a better classification of proficiency in intelligent language learning systems, integrating adaptive parameterization and value estimation. Figure 5 illustrates the Architecture of the TDSVM-Based Learning Sy

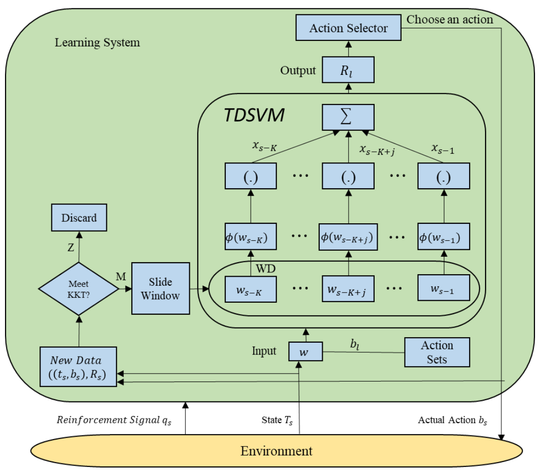

Figure 5 Flowchart of TDSVM

1. *Enriched Transient Search Tree*

The simple Transient Search Tree may converge slowly due to poor random initialization. Using logistic chaos mapping generates high-quality populations, enhancing ETS-TDSVM's accuracy in classifying cadet proficiencies for individualized French instruction. Equation 16 defines logistic chaos mapping.

In the range [0, 1], is a randomly distributed number with a broad distribution. is the chaos regulation of the parameter, and is the preset maximum number of chaotic iterations. According to Equation 17, the location vectors of each d-dimensional searching agent are then utilized to map the upper and lower boundaries of the search space using the chaotic series and the created set of chaotic elements

Where, the coordinate of the search agent's dimension is , Additionally, is the coordinate of dimension after internal random ordering. Opposition-based learning looks at the other side of the search space to find better options, which produces better candidates. Whereas chaotic sequences present varied populations, the creation of conflicting solutions enhance coverage and the search efficiency by enhancing the likelihood of finding optimal agents, as given in equation 18 is the opposite positioning for the agent, is the greatest boundary's position, is the smallest boundary's position, and is the individual's location. These are then joined to create 2N, a different population of search agents. Following an assessment of each population's fitness, the categories with the greatest fitness are selected as the starting populations.

(1). Adaptive Inertia Weights

Particle swarm-inspired adaptive inertia weights improve merit-seeking and convergence. Unlike fixed weights in standard TSO, they enhance global search initially and refine local search later. Equation 19 defines the method for adaptive inertia weight adjustment per iteration.

Where, the optional parameters , , and are used. The initial value of inertia weight is large, and then declines very quickly to amplify global search, and then decreases slowly with a small value to make local fine search. Enrich-TSO weighs exploration and exploitation as specified in Equations 20 and 21 with adaptive inertia weights.

(2). Neighborhood Dimensional Learning

Global and local search is enhanced by Dimension Learning-based Hunting (DLH) through agent information sharing. Neighbor Dimension Learning (NDL) identifies local agents for dimension exchange, preserving diversity and avoiding local optima, with candidate populations updated iteratively as shown in Equation 22.

Where, determines the position vector space of the candidate population, indicates the vector indicating the searching agent's position, and represents the random probability. Equation 23 shows how to calculate a radius in the neighborhood, utilizing the Euclidean distances between the candidate position , and the present location iteratively filters the population according to the Euclidean distance lower than the searching agent, before displaying the individual's neighbors. Lastly, the suitability of the NDL and candidate populations is computed using Equation 24 and the most recent data on the population members.

Enriched Transient Search Tree (ETST) uses chaos mapping, antagonistic learning, dynamically changing inertia weights and Neighborhood Dimensional Learning to augment standard search methods in terms of convergence, diversity, and global local search. It is used when complex classification is needed, e.g., when it comes to predicting personalized skills during advanced training in the French language. The flowchart of the enriched Transient Search Tree is shown in figure 6. Table 1 is a representation of the hyper-parameters of the ETS-TDSVM.

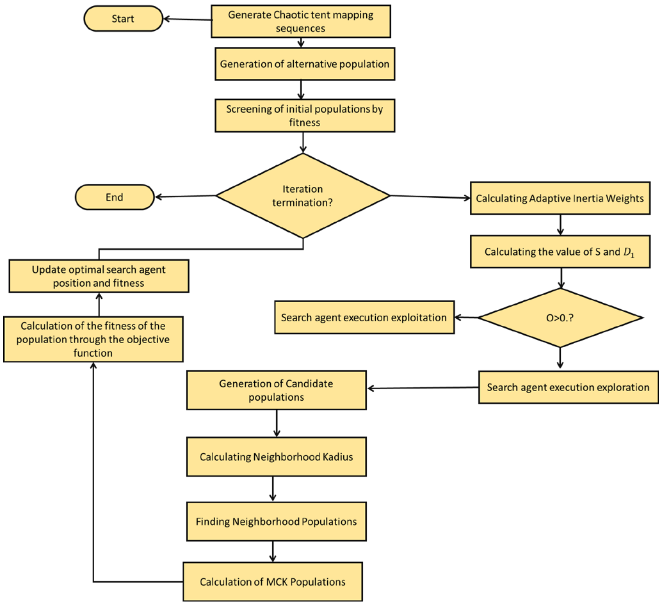

Figure 6 Schematic diagram of the ETS

Table 1: Hyperparameters of the ETS-TDSVM

Category	Hyperparameter	Example Range / Value
SVM (Core)	CC	[0.1, 10]
	γ\gamma	[0.01, 1]
	Kernel Type	RBF (fixed)
	ϵ\epsilon (optional)	[0.001, 0.1]
ETS (Search)	Tree Depth DD	3–6
	Branching Factor BB	2–5
	Enrichment Factor EE	0.2–0.8
	Transience Decay TT	0.1–0.5
	Iterations II	5–20
Dynamic SVM	Batch Size	50–200
	Update Strategy	LASVM
	Budget Size (optional)	500–2000
Stopping Criteria	Convergence Threshold	0.001–0.01 (1%)
	Max Iterations	10–30

III. RESULTS AND DISCUSSION

The ETS-TDSV method was implemented on Python 3.9 to compare the performance of the method with a baseline model on Pronunciation & Oral Skills, Adaptive Content Delivery, and Prediction Accuracy. The Personalized French Language Learning Dataset was the training set of both methods. Findings reveal that the suggested model is superior to the baseline in that it is more consistent and efficient in meeting the needs of various language competencies.

The plot illustrated trends in the four major features of learning as Digital Assessment Score, Pronunciation Accuracy, Task Completion Rate, and Topic Mastery Likelihood categorized by the level of proficiency. There are high values in advanced learners as compared to their beginners who have lower values, which clearly demonstrate trends in language acquisition. Figure 7 presents the trend visualization of four major learning features in relation to the levels of language proficiency.

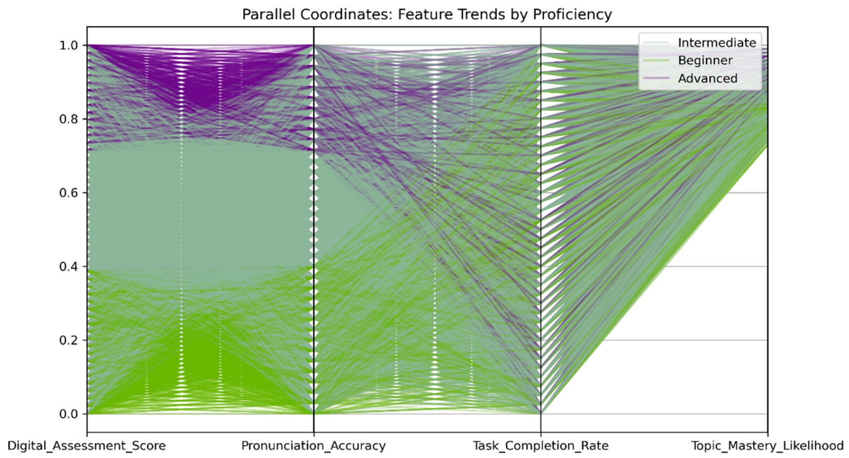

Figure 7 Trends in key learning features grouped by language proficiency level (Beginner, Intermediate, Advanced)

The heatmap depicts the patterns of pronunciation accuracy in language learning. The color yellow tells the existence of more density, which is a typical cadet performance. This visualization ssassists in determining the progress, stagnation or regression and customizing the learning strategies accordingly. Figure 8 shows the faceted histograms of the digital assessment scores and shows the distinct patterns based on the level of proficiency.

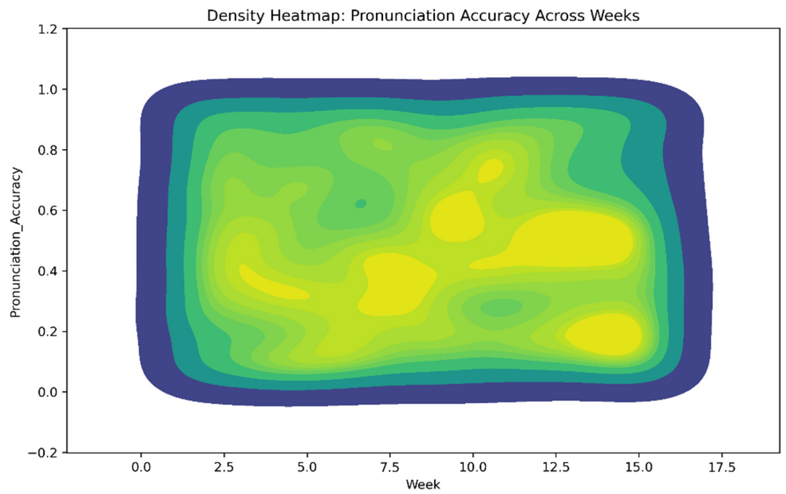

Figure 8 Density histograms of digital assessment scores, highlighting distinct patterns by proficiency level

The faceted histogram demonstrates the scores of digital assessments at the levels of Beginner, Intermediate, and Advanced. Novices have mass at low scores, the high scores are always in advanced learners and the intermediate learners tend to be very variable. These trends show specific performance footprints, which inform customized digital training measures. Figure 9 shows the distribution of scores by level of proficiency.

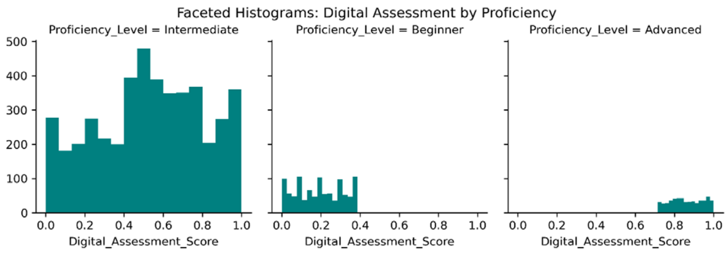

Figure 9 Digital Assessment Scores by proficiency level, showing high variation in Intermediate learners and consistently high scores in advanced learners

A. Comparative analysis of ETS-TDSVM

Pronunciation & Oral Skills assesses the capacity of the model to identify speech errors and give personal feedback. Adaptive Content Delivery was indicative of its ability to deliver learning contents to the expertise and progress of each cadet to increase the engagement and speed of learning the language. Prediction Accuracy reflects the expertise of the model in predicting the language levels of the cadets. Table 2 provides a comparison of these metrics.

TABLE 2: PRONUNCIATION AND ORAL SKILLS COMPARISON

	Pronunciation & Oral Skills	TDSVM	ETS-TDSVM [Proposed]
Result Parameter	Pronunciation Accuracy (%)	74.5	89.8
	Speech Fluency Score (0-10)	6.1	8.4
	Phoneme Error Rate (%)	11.7	5.6
	Listening Comprehension Improvement (%)	19.3	33.7

Table 2 compares the baseline method with the proposed machine learning structure based on the key factors of pronunciation and oral skills measurement. The proposed model has significantly better pronunciation rate (89.8 % vs 74.5%), speech flow (8.4/10 vs 6.1/10), phoneme error (5.6 % vs 11.7 %), and listening comprehension (33.7 % vs 19.3 %), compared to the current model, which indicates the excellent facilitation of language learning. Figure 10 presents the three main properties of the model to drive personalized language acquisition and enhance instructional planning among cadets using adaptive, data-driven approaches. Table 3

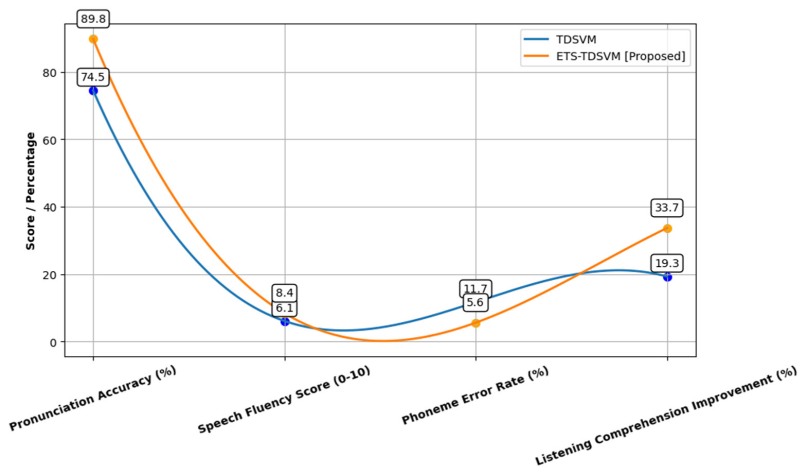

Figure 10 Key model capabilities: Pronunciation & Oral Skills analysis, Adaptive Content Delivery, and accurate Proficiency Prediction for personalized language instruction

Table 3: Comparative Evaluation of Adaptive Content Delivery Efficiency

Adaptive Content Delivery	TDSVM	ETS-TDSVM [Proposed]
Lesson Sequencing Accuracy (%)	74.2	89.7
Difficulty Adjustment Effectiveness (%)	70.5	88.4
Learner Engagement Score (1–10 scale)	6.2	8.9
Personalized Content Satisfaction (%)	68.9	91.3

A comparative training of the Adaptive Content Delivery with the baseline approach and the proposed framework is presented in Table 3. The metrics compared are Lesson Sequencing Accuracy (74.2% vs. 89.7%), Difficulty Adjustment Effectiveness (70.5% vs. 88.4%), and the scores of Learner Engagement (6.2 vs. 8.9) and Personalized Content Satisfaction (68.9% vs. 91.3%). The proposed method indicates remarkable improvements in all parameters, particularly in terms of being student-centered and engaging teaching. The graph in Figure 11 represents the effectiveness of personalized content delivery strategies, comparing a baseline method to a proposed method across four key metrics, and Table 4 provides the prediction performance comparison: Baseline vs ETS-TDSVM.

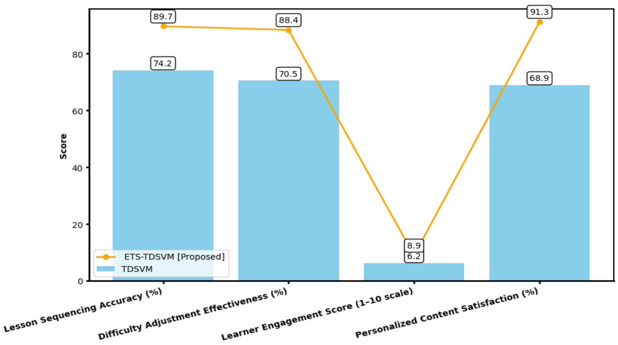

Figure 11 Comparison of personalized learning metrics, showing the proposed method's significant improvements in content satisfaction and engagement over the baseline

Table 4: Comparative Evaluation of Prediction Accuracy Metrics

Prediction Accuracy (%)	TDSVM	ETS-TDSVM [Proposed]
Overall Proficiency Classification	78.2	93.5
Topic Mastery Prediction	74.6	91.2
Learner Cluster Consistency	80.1	95.0
Misclassification Rate	21.8	6.5

Table 4 illustrates how the accuracy of the baseline method's prediction was measured against the newly proposed ETS-TDSVM technique through the set of evaluation measures which include Overall Proficiency Classification (78.2% vs. 93.5%), Topic Mastery Prediction (74.6% vs. 91.2%), Learner Cluster Consistency (80.1% vs. 95.0%), and Misclassification Rate (21.8% vs. 6.5%). The new set-up showed that the baseline was not up to in all metrics including consistency, which is indicative of long-term and solid predictive skills. Comparison of model performance between the baseline and proposed ETS-TDSVM method on evaluation metrics is shown in Figure 12.

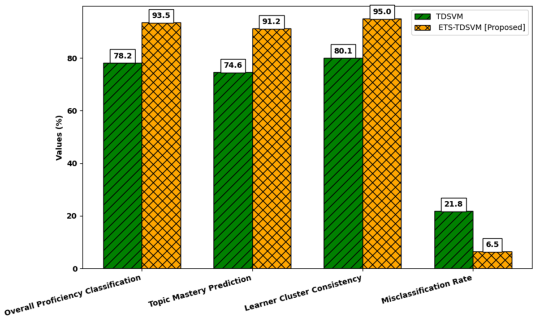

Figure 12 Comparative performance of the proposed ETS-TDSVM msethod versus the baseline across key evaluation metrics, demonstrating superior accuracy and consistency.

B. Statistical Performance Validation of the ETS-TDSVM Model

The validation assesses the effectiveness of the ETS-TDSVM model in individualized training of the French language by comparing the accuracy of the topics mastered, the accuracy of the classification of the learners, the consistency of the prediction through the confidence intervals and standard deviations, and the inaccuracy, which should be statistically significant ($p< 0.01$). Table 5 is a summary of these results.

Table 5: Statistical Analysis Of Ets-Tdsvm Model Outcomes

Metric (%)	Mean	Standard Deviation (±)	95% Confidence Interval	*p*-value
Overall Proficiency Classification	93.5	1.7	[91.8-95.2]	< 0.01
Topic Mastery Prediction	91.2	1.9	[89.3-93.1]	< 0.01
Learner Cluster Consistency	95.0	1.4	[93.6-96.4]	< 0.01
Misclassification Rate	6.5	0.8	[5.7-7.3]	< 0.01

The ETS-TDSVM model showed a low rate of misclassification along with good accuracy and learner cluster consistency. Low standard deviations and narrow confidence intervals show steady performance, and -values less than validate statistical significance, confirming its efficacy for individualized, real-time French language training in professional training settings.

IV. DISCUSSION

The ETS-TDSVM system which it was based on the ML and developed in this study improved proficiency classification, topic mastery, and adaptive learning by offering police students customized and real time French training. The language learning models that are accurate, computationally efficient, and flexible were needed to accomplish the task of learning language effectively within the specialized setting, like law enforcement academies. Earlier methods such as GRNN (Sun, 2025) enhanced the contextual relevance and fluency but required a high level of computing and a huge number of labeled data. In the same way, customized instruction was improved by DFDLOF (Song et al., 2024) that lacked scalability whereas hybrid LSTM-NLP techniques (Ezhilmathi et al., 2024) required large amounts of training data and did not always explain transient learner behavior. Conventional ML models, such as SVM (Liu and Yang, 2024), were very accurate, but lacked sequence of lessons and real-time adaptation. The drawbacks of ETST-TDSVM is managed through combining the ETST-based classification with dynamic SVM and DT-based adaptive sequencing to provide accurate predictions of proficiency, low misclassifications as well as deliver lessons with scale and in a contextual manner. It also uses less data to assess skills as compared to previous models, adapts dynamically to the performance of learners, has less computational load, and is designed to fit structured professional settings. The mutual dependency of the ML parts enables accurate prediction of proficiency and dynamic delivery of content, which makes it easier to learn the language. Conversely, the baseline model has worse performance and the overall proficiency classification accuracy is 78.2%, topic mastery prediction is 74.6%, learner cluster consistency is 80.1% and high misclassification (21.8%) is limiting personalized instruction. ETS-TDSVM addresses these loop holes with dynamic classification and improved parameter optimization to provide more accurate, stable and responsive French language training to cadets in Vietnam.

V. CONCLUSION

The foreign language learning in Vietnam allows the enforcers in Vietnam to improve cross-border communication, and adaptive training features add efficiency to the learning environment in the workplace. The suggested ETS-TDSVM model helps the individualized approach to the French language learning with the help of a dataset that has been gathered during 16 weeks and includes 378 police cadets. Cleaning, min-max normalization, one-hot encoding, and K-means clustering are used as data preprocessing to cluster learners. Hyper-parameter optimization by means of ETS, together with the DSVM model, results in a better adaptive framework with corrective instructional feedback. The performance of ETS-TDSVM in all the evaluation measures was high making it possible to increase the predictiveness of the instructional method and learner involvement in professional training settings. The inclusion of this model in the Learning Management Systems (LMS) enables academies, universities, and specialized organizations to enjoy the convenience of real-time progress monitoring, automated curriculum modification, and data-driven planning in instruction. Although these are encouraging findings, the method is currently restricted to French language instruction and makes use of static clustering which is not responsive to changing behavior of the learners. To enhance the generalizability, future studies must support the approach using larger and more diverse datasets, dynamic clustering and Deep Reinforcement Learning (DRL), and run the model across multi-lingual learning environments, which allow realizing a greater generalizability and training at a personal level across languages and professional settings.

Author Biodata

Hue Ng. Thi is a Senior Lecturer in French Language at the People's Police Academy. She holds a Bachelor's degree in French and English from the University of Languages and International Studies, Vietnam National University, Hanoi.

Ms. Hue Ng. Thi completed her Master's degree in French and PhD in Teaching Methodology at the University of Languages and International Studies, Vietnam National University, Hanoi. Her research areas include Teaching Methodology Theory, Educational Technology, and Foreign Language Testing and Assessment.

She has authored numerous articles with ISBN indexes and participated in national and international conferences on French language teaching and learning.

References

Cai, J. and Li, Y., (2024). Fuzzy association rule mining for Personalized English Language Teaching in higher education. *Journal of Computational Methods in Sciences and Engineering*, *24*(6), 3617-3631.https://doi.org/10.1177/14727978241296748

Chaipidech, P., Srisawasdi, N., Kajornmanee, T., and Chaipah, K. (2022). A personalized learning system-supported professional training model for teachers' TPACK development. *Computers and Education: Artificial Intelligence*, *3*,100064.https://doi.org/10.1016/j.caeai.2022.100064

Chen, J., Liu, Z., Huang, X., Wu, C., Liu, Q., Jiang, G., Pu, Y., Lei, Y., Chen, X., Wang, X. and Zheng, K., (2024). When large language models meet personalization: Perspectives of challenges and opportunities. *World Wide Web*, *27*(4), .42.https://doi.org/10.1007/s11280-024-01276-1

Chen, Y., (2024), May. Enhancing language acquisition: The role of AI in facilitating effective language learning. In *2024, the 3rd International Conference on Humanities, Wisdom Education and Service Management (HWESM 2024)* (593-600). Atlantis Press https://doi.org/10.2991/978-2-38476-253-8_71

Dong, W., Pan, D., and Kim, S. (2024). Exploring the integration of IoT and Generative AI in English language education: Smart tools for personalized learning experiences. *Journal of Computational Science*, *82*, 102397.https://doi.org/10.1016/j.jocs.2024.102397

Ezhilmathi, K., Durairaj, M., Vazhangal, M., Mohammad, R., Pathak, P., and Karthik, M. (2024), July. Tailored English Language Learning Support: Leveraging Long Short-Term Memory Networks for Personalized Assistance. *In 2024, the Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, 1-6. IEEE.https://doi.org/10.1109/ICEEICT61591.2024.10718506

Gan, W., Qi, Z., Wu, J., and Lin, J.C.W. (2023), December. Large language models in education: Vision and opportunities. In *2023 IEEE international conference on big data (BigData)*,4776-4785. IEEE.https://doi.org/10.1109/BigData59044.2023.10386291

Hussain, S.M., Satti, S.M.J., and Khan, Z. (2024). AI-Powered Personalized Learning: Advancing Language Education in the Digital Era. *Journal of Social Signs Review*, 2(4), 730-740.

Istanti, W., Pratiwi, S., and Saddhono, K. (2024), November. AI-Driven Personalized Learning: Revolutionizing Language Education. In *2024 International Conference on IoT, Communication and Automation Technology (ICICAT)*,329-334. IEEE.10.1109/ICICAT62666.2024.10923274

Kanchon, M.K.H., Sadman, M., Nabila, K.F., Tarannum, R., and Khan, R. (2024). Enhancing personalized learning: AI-driven identification of learning styles and content modification strategies. International Journal of Cognitive Computing in Engineering, 5,269-278. https://doi.org/10.1016/j.ijcce.2024.06.002

Liu, C. and Yang, S., (2024). Personalized learning ability classification using SVM for enhanced education in system modeling and simulation courses. *Frontiers of Digital Education*, 1(4), 295-307. https://doi.org/10.1007/s44366-024-0035-6

Naseer, F., Khan, M.N., Tahir, M., Addas, A., and Aejaz, S.H. (2024). Integrating deep learning techniques for personalized learning pathways in higher education. *Heliyon*, *10*(11).

Orosoo, M., Raash, N., Treve, M., Lahza, H.F.M., Alshammry, N., Ramesh, J.V.N., and Rengarajan, M. (2025). Transforming English language learning: Advanced speech recognition with MLP-LSTM for personalized education. *Alexandria Engineering Journal*, *111*,21-32.https://doi.org/10.1016/j.aej.2024.10.065

Peng, X. and Wang, Y. (2025). An AI-Driven Approach for Advancing English Learning in Educational Information Systems Using Machine Learning. International Journal of Advanced Computer Science & Applications, 16(1).https://doi.org/10.14569/ijacsa.2025.0160116

Praveena, T. and Anupama, K. (2025). Machine Learning Meets Language Learning: The Transformative Potential of Artificial Intelligence in English Language Instruction. *Human Research in Rehabilitation*, *15*(1). https://doi.org/10.21554/hrr.04251

Sajja, R., Sermet, Y., Cikmaz, M., Cwiertny, D., and Demir, I. (2024). Artificial intelligence-enabled intelligent assistant for personalized and adaptive learning in higher education. *Information*, *15*(10), 596.https://doi.org/10.3390/info15100596

Song, C., Shin, S.Y., and Shin, K.S. (2024). Implementing the dynamic feedback-driven learning optimization framework: a machine learning approach to personalize educational pathways. Applied Sciences, 14(2),

916. https://doi.org/10.3390/app14020916

Sun, B. (2025). Gated recurrent deep learning approaches to revolutionizing English language learning for personalized instruction and effective instruction. Scientific Reports, 15(1), 13028. https://doi.org/10.1038/s41598-025-96351-6

Sun, Y., (2025). Construction and optimization of personalized learning paths for English learners based on the SSA-LSTM model. *Systems and Soft Computing*, *7*, 200218.https://doi.org/10.1016/j.sasc.2025.200218

Tanweer, M. and Ismail, A. (2024). Generative AI in curriculum development: A framework for adaptive, customized, and personalized learning. In *Impacts of generative AI on creativity in higher education*,193-226. IGI Global. https://doi.org/10.4018/979-8-3693-2418-9.ch008

Wei, L. (2023). Artificial intelligence in language instruction: impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in psychology*, *14*, 1261955.https://doi.org/10.3389/fpsyg.2023.1261955

Xia, Y., Shin, S.Y. and Shin, K.S., (2024). Designing personalized learning paths for foreign language acquisition using big data: Theoretical and empirical analysis. Applied Sciences, 14(20), 9506. https://doi.org/10.3390/app14209506