Proceedings of

2025 3rd International Conference on Information Education and Artificial Intelligence (ICIEAI 2025)

December 12-14, 2025
Guangzhou, China

ISBN: 979-8-4007-2071-0

The Association for Computing Machinery
1601 Broadway, 10th Floor
New York, New York 10019, USA

ACM ISBN: 979-8-4007-2071-0

2025 3rd International Conference on Information Education and Artificial Intelligence (ICIEAI 2025)

Table of Contents

Preface

The 2025 3rd International Conference on Information Education and Artificial Intelligence (ICIEAI 2025) was successfully held in Guangzhou, China, from December 12 to 14, 2025. The conference brought together scholars, educators, researchers, and industry practitioners from around the world, reflecting the rapid development and growing importance of digital education, artificial intelligence, and intelligent learning technologies in the era of educational transformation.

ICIEAI 2025 focused on the deep integration of informatization and artificial intelligence with modern education. Key themes included AI-empowered teaching innovation, smart education platforms, data-driven learning analytics, human–computer collaborative learning, intelligent educational tools, and the digital transformation of basic, vocational, and lifelong education. The conference provided a high-level international platform for presenting cutting-edge research, sharing practical experiences, and fostering interdisciplinary collaboration among academia, industry, and educational institutions.

The conference featured five keynote speeches delivered by distinguished experts. Prof. Haoran Xie discussed emerging advances in AI-driven educational technologies and intelligent learning systems. Prof. Xinguo Yu presented core breakthroughs and ecosystem development in digital intelligence technologies for mathematics education. Prof. Dong Huang introduced recent progress in large-scale graph learning and its applications in intelligent education. Prof. Kai Liu examined the logic and technological pathways through which artificial intelligence is reshaping the rule-of-law ecosystem. Prof. Xiaoli Li explored the connection between novice programming feedback and industrial time-series intelligence. These keynote presentations offered valuable insights into both theoretical advancements and practical innovations across the fields of informatization education and artificial intelligence.

In addition to keynote speeches, ICIEAI 2025 organized extensive oral and poster presentation sessions covering a wide range of topics, including AI-supported e-learning, educational big data, intelligent curriculum design, digital literacy development, AI-enhanced teaching platforms, and interdisciplinary applications of machine learning in education. The diversity and depth of the presented work reflect the accelerating integration of AI technologies into educational research and practice.

A total of 407 manuscripts were submitted to ICIEAI 2025, of which 214 were accepted after rigorous peer review. The papers included in this proceedings volume represent high-quality research contributions from scholars and practitioners worldwide. These works span theoretical frameworks, empirical studies, algorithmic innovations, system design, and practical educational applications, offering valuable references for researchers and professionals in information education, artificial intelligence, educational technology, and related fields.

This proceedings volume captures the latest developments and emerging trends in intelligent education and AI-driven learning technologies. We believe the collected papers will inspire further research, promote interdisciplinary collaboration, and support the continued advancement of digital and intelligent education ecosystems.

We sincerely thank all authors, reviewers, keynote speakers, session chairs, and organizing committee members for their dedication and contributions, which made ICIEAI 2025 a successful, influential, and inspiring academic event.

ICIEAI 2025 Conference Committee

Committee Member

Conference Chairmen
Prof. Benjamin W.Wah, The Chinese University of Hong Kong, China
Prof. Ishfaq Ahmad, The University of Texas at Arlington, USA
Prof. Haoran Xie, Lingnan University, China

Technical Program Committee Chairs
Prof. Mohammed El-Abd, American University of Kuwait, Kuwait
Prof. Shadiev Rustam, Zhejiang University, China
Prof. Mohammed Akour, Prince Sultan University, Saudi Arabia

Technical Program Committee Co-Chairs
Prof. Fahriye Altinay, Near East University, Northern Cyprus
Assoc. Prof. Sam Goundar, RMIT University, Vietnam

Publication Chairs
Prof. Zhengxi Yang, South China Agricultural University, China
Assoc. Prof. Dong Huang, South China Agricultural University, China
Prof. Zehra Altınay, Near East University, Northern Cyprus

Local Organizing Committee Chairs
Prof. Ying Wang, South China Agricultural University, China
Prof. Haiyan Zhong, South China Agricultural University, China
Prof. Quandian Wang, South China Agricultural University, China
Prof. Lijuan Wang, South China Agricultural University, China

Publicity Chair
Prof. Luan Trong Nguyen, FPT University, Vietnam

International Technical Program Committee
Prof. Fuchang Wang, South China Agricultural University, China
Prof. Guoming Du, South China Agricultural University, China
Prof. Andrew Csizmadia, Newman University, UK
Prof. Zehui Zhan, South China Normal University, China
Prof. Dariusz Jakóbczak, Koszalin University of Technology, Poland
Prof. Yoshifumi Manabe, Kogakuin University, Japan
Prof. Hui-Yu Zhou, University of Leicester, UK
Prof. Yudong Zhang, University of Leicester, UK
Prof. Addisson Salazar, Universidad Politécnica de València, Spain
Prof. Manuel Padilla, Universidad Nacional de San Martin, Peru
Prof. Shakil Akhtar, Clayton State University, USA
Prof. Tomokazu Nakayama, Jissen Women's University, Japan
Prof. Zarinah Hamid, International Islamic University Malaysia, Malaysia
Prof. Yongfa Kong, Nankai University, China
Prof. Marcin Paprzycki, Polish Academy of Sciences, Poland
Prof. Vijayakumar Varadarajan, The University of New South Wales, Australia
Prof. Yongjun Feng, Shaanxi Normal University, China
Prof. Bode Li, Lanzhou University of Finance and Economics, China
Prof. Yu-Mei Wang, University of Alabama at Birmingham, USA
Prof. Zhenghong Xiao, Guangdong Polytechnic Normal University, China
Prof. Huimin Zhao, Guangdong Polytechnic Normal University, China
Prof. Ya Chen, Guangdong Polytechnic Normal University, China
Assoc. Prof. Gege Yu, South China Agricultural University, China
Assoc. Prof. Yan Li, South China Agricultural University, China

Assoc. Prof. Bing Xiao, Guangdong Polytechnic Normal University, China
Assoc. Prof. Goh Wei Wei, Taylor's University, Malaysia
Assoc. Prof. William Yeoh, Deakin University, Australia
Assoc. Prof. Yu-Mei Wang, University of Alabama at Birmingham, UK
Assoc. Prof. Huey Lei, Caritas Institute of Higher Education, Hong Kong, China
Assoc. Prof. Qiong Li, Lanzhou University of Finance and Economics, China
Assoc. Prof. Apicha Dangchamroon, Ramkhamheang University, Thailand
Assoc. Prof. Apostolos Tsagaris, International Hellenic University, Greece
Assoc. Prof. Aryusmar Kartadilaga, Bina Nusantara University, Indonesia
Assoc. Prof. Yasushi Akiyama, Saint Mary's University, Canada
Assoc. Prof. Xie Ming, Nanyang Technological University, Singapore
Asst. Prof. El Hadji Yaya Koné, University of Ottawa, Canada
Asst. Prof. Sho Ooi, Ritsumeikan University, Japan
Asst. Prof. Janchai Yingprayoon, Suan Sunandha Rajabhat University, Thailand
Dr. Xiuli Liu, South China Agricultural University, China
Dr. Fei Li, South China Agricultural University, China
Dr. Dingyong Deng, South China Agricultural University, China
Dr. Lvkun Zhang, South China Agricultural University, China
Dr. Zebing Lai, South China Agricultural University, China
Dr. Xiaochun Cheng, Middlesex University, UK
Dr. Li Zhao, Shandong Technology and Business University, China
Dr. Arvind Dhaka, Manipal University Jaipur, India
Dr. Bin Hu, Henan University, China
Dr. Ann Rosnida Md Deni, Sunway University, Malaysia
Dr. Edmond W. M. Lam, The Hong Kong Polytechnic University, China
Dr. Liviu Octavian Mafteiu-Scai, West University of Timisoara, Romania
Dr. Nasir Abdul Jalil, Sunway University, Malaysia
Dr. Shin Ling Wu, Sunway University, Malaysia

Enhancing Language Service Education with Literary Narratives: An XGBoost Regression Approach

Jia Liu*
School of English Language and Literature
Xi'an Fanyi University
Xi'an, Shannxi, China
15399461266a@gmail.com

Hamoud Yahya Ahmed Mohsen
Department of English Language and Literature,
AbdulHamid AbuSulayman Kulliyyah of Islamic Revealed Knowledge and Human Sciences
International Islamic University Malaysia
Kuala Lumpur, Kuala Lumpur, Malaysia
hamoud@iium.edu.my

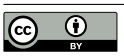
Abstract

In the context of globalization and the collaborative development of regional economies, the cultivation of high-quality language service professionals has become a critical requirement for advancing the construction of the Hainan Free Trade Port in China. Based on the theoretical perspective of language-culture symbiosis, this study intends to build a two-track training track including "language transformation" and "cultural mediation" to compensate for the deficiencies of language service education, cutuespecially low cultural literacy and weak practice orientation. In terms of methodology, this research constructs a modular teaching program structured around scenario-based literary translation workshops, interdisciplinary coursework, and internships with regional cultural enterprises. To validate this model, a stratified survey was administered to 378 stakeholders-comprising students, educators, and industry practitioners recruited via institutional networks-and analyzed using the XGBoost Regression model to explore the influencing factors of students' cultural awareness and professional capability. Finally, the XGBoost Regression model is applied to analyze the collected data in order to explore the influencing factors of students' cultural awareness, intercultural communication ability, and high-level language service ability. The empirical results show that the inclusion of literary narratives embedded in AI-based digital tools such as natural language processing and virtual/augmented reality can improve students' cultural sensitivity, communicative flexibility, and professional performance. These results indicate that a narrative-based, data-driven model of language service education meets the strategic needs of the Hainan Free Trade Port, and this model also shows high replicability and scalability in other multilingual and multicultural educational practices.

CCS Concepts

• **General and reference** → Cross-computing tools and techniques; Empirical studies.

*Corresponding author



ICIEAI 2025, Guangzhou, China

ACM ISBN 979-8-4007-2071-0/2025/12
https://doi.org/10.1145/3799457.3799550

Keywords

Literary Narratives, Language Services, Language Services Education, Dual Language-Culture Pathway, Language Transformation, Cultural Translation, XGBoost Regression

ACM Reference Format:
Jia Liu and Hamoud Yahya Ahmed Mohsen. 2025. Enhancing Language Service Education with Literary Narratives: An XGBoost Regression Approach. In *2025 3rd International Conference on Information Education and Artificial Intelligence (ICIEAI 2025), December 12–14, 2025, Guangzhou, China.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3799457.3799550

1 INTRODUCTION

The building of Hainan Free Trade Port [1] is a major project in the new round of reform and opening-up in China. Hainan will function as a highly liberalized platform for trade, investment, and cross-border people flow. A corresponding high-level language services system is vital for international trade and tourism, foreign investment and engagement, and for people to be not only good transmitters of accurate meanings across cultural boundaries but also intermediaries of culture. However, the development of language service talent in China still has a long way to go. Multilingual education is mainly focused on widely used languages like English, French, and Spanish [2], while the enrollment of non-dominant languages, particularly those of Belt and Road countries, is very limited. There is also a shortage of practice-oriented instructors, limiting the supply of qualified language service professionals [3]. Surveys conducted in enterprises have repeatedly pointed out the gaps in intercultural communication, industry-specific knowledge, and technical proficiency in CAT and AI enabled tools.

International experience can provide instructive examples for the development of language services in China. Comparative studies from established multilingual hubs, such as the European Union and Singapore, demonstrate the efficacy of holistic models that integrate linguistic and cultural literacy. However, unlike these regions which often focus on dominant global languages, the Hainan Free Trade Port faces the unique challenge of servicing under-represented 'Belt and Road' languages, requiring a more adaptable pedagogical model. By contrasting these established approaches with Hainan's specific needs, this study proposes a narrative-based framework that addresses the documented gaps in cultural mediation often found in purely vocational training [4]. Universities in Europe, Australia, and other regions have incorporated literary narratives and cultural studies into translation and language service education. Empirical studies have shown that students exposed to literary

narrative-based instruction have improved in terms of cultural sensitivity and cross-cultural adaptability.

In this context, the Hainan Free Trade Port, which is a highly liberalized trade platform, offers an excellent opportunity to conduct empirical studies. Based on the theory of linguistic and cultural symbiosis, language services are redefined as a dual process of "language transformation" and "cultural mediation." Literary narratives are redefined as abundant symbolic resources, carriers of historical memory and value references, and key resources for cultural literacy and interpretive competence.

Meanwhile, rapid development in educational data mining and learning analytics provides methodological opportunities to evaluate innovations in language service education. Machine learning methods, especially tree-based ensemble methods, are increasingly used to predict students' academic performance and identify the determinants of learning outcomes. The XGBoost Regression model [5] is particularly suitable because of its high prediction accuracy, ability to capture nonlinear relationships, and interpretability of feature importance. The application of XGBoost on questionnaire data from language service education is possible and makes it possible to quantitatively model how different curricular and contextual factors, such as literary narrative-based instruction, regional practicum opportunities, and digital tools [6], jointly affect students' cultural awareness, intercultural competence, and advanced language service skills, thereby filling the gap in the rigorous and quantitative evaluation of narrative-based instruction.

Faced with these opportunities and challenges, this study is conducted in response to the Hainan Free Trade Port in terms of a literary narratives' pathway for culture through the insertion of literary narratives into language service education via modular curriculum design, interdisciplinary training, regionally rooted practice, and AI digitalization. We analyze a large-scale questionnaire survey of language service majors and employ XGBoost Regression to identify and weight the important factors stakeholders consider when supporting this model and their perceived learning gains. The paper offers three main contributions. First, we construct a narrative-driven and culture-oriented framework to respond to the Hainan Free Trade Port's strategic needs and the international trend of multilingual education. Second, we provide empirical evidence that the insertion of literary narratives and digital technologies can contribute to students' cultural literacy, intercultural adaptability, and professional performance. Third, we show how machine learning can be applied to extract significant educational dimensions from questionnaire data, providing a data-driven reference for the curriculum construction and policymaking of language service education.

2 RESEARCH METHOD

2.1 Data Collection

This study distributed a designed questionnaire to collect quantitative information for machine-learning analysis of the effectiveness of literary narratives in language service education. The targeted population were stakeholders in the field of language service education, including students, teachers in universities, practitioners in language services, and other relevant professionals. The questionnaire was circulated online through institutional mailing lists, course websites, and social media platforms for professionals. After removal of incomplete questionnaires and obviously unreasonable answers, 378 valid cases were retained. The final sample consisted of 378 valid responses, including undergraduate students, graduate students, and language-service practitioners from institutions in Hainan and other provinces. Participants were recruited through purposive sampling to ensure inclusion of individuals with prior exposure to translation or intercultural communication courses, and demographic variables such as age, gender, and academic background were recorded to support subgroup analysis.

The questionnaire, named "Questionnaire on the Effectiveness of Literary Narratives in Language Service Education," consisted of 20 single-choice questions. Question 1-4 collected information about the participants' identity (student, teacher, practitioner in language service, or other), highest level of education attained (from junior college and below to doctorate), participation in language service education (e.g., in a program of translation or intercultural communication; yes/no), and participation in courses considered literary-narrative-related (e.g., selected readings in English and American literature or histories of English literature; or courses on the society and culture of English-speaking countries; yes/no). Question 5 collected the degree to which the respondents agreed or disagreed with the statement: "In general, I support the inclusion of literary narratives in the cultivation of talent in language services." Questions 6-20 were statements (from 1 = strongly disagree to 5 = strongly agree, coded as five-point Likert scales) about the following aspects: the contribution of literary narratives to the understanding of target-language cultures, improvement of intercultural communication competence, assistance in dealing with cultural differences in translation, impact on learning motivation and sense of professional identity, and views on the modular curriculum, industry-university collaboration, and the future of language service education.

For XGBoost Regression modeling, question 5 was defined as the main continuous output variable representing overall support for the inclusion of literary narratives in language service talent cultivation. All other questions (1-4 and 6-20) were used as candidate explanatory variables reflecting the demographic characteristics, educational experience, and perceived educational outcomes associated with language service education based on literary narratives.

2.2 Data Processing

All of the questionnaire data were downloaded from the survey website in CSV format and cleaned in a Python environment before modeling. In step 1, we screened the raw dataset and eliminated observations with significant missing values or obvious patterned responses (i.e., marking the same response for all items without variation). The cleaned dataset consisted of 378 valid observations.

Role, level of education, prior education on language service and literary narrative courses, and prior exposure to courses on literary narrative were categorical variables that were one-hot encoded. The responses to items 5-20 on the Likert scale were recoded into 1, 2, 3, 4, or 5 depending on whether the person had, respectively, the weakest, second weakest, third weakest, fourth weakest, or strongest attitude toward the idea that literary narratives can play

a role in language service education (i.e., the higher the number, the more positive the attitude).

We visualized the distribution of each variable and dealt with the few missing values by imputing the mode for categorical variables and the median for Likert scale items. XGBoost is a tree-based ensemble model that is not sensitive to feature scaling; therefore, we kept the original 1-5 coding for the Likert items and only standardized the descriptive statistics/correlation matrices.

2.3 Feature Extraction

Considering that enabling effective training and enhancing interpretability could benefit from not directly using each questionnaire item as an independent input variable and conceptually related items were aggregated into a set of higher-level features by averaging scores , this study's feature engineering process reduced dimensionality, mitigated the multicollinearity problem that exists between highly correlated Likert-scale items, and enhanced the robustness of the XGBoost Regression model. Six higher-level features were extracted as model inputs according to the questionnaire structure and theoretical basis of language-culture integration.

1. Participant's Profile. Respondents' background characteristics were summarized in this feature, which included three aspects: role identity, highest attained education level, prior education related to language-service and literary-narrative courses. This feature reflected structural differences in respondents' educational experience and professional identity that might systematically affect their evaluation of literary-narrative integration.

2. Understanding of Target-Language Culture and Cultural Connotation. This feature was extracted from items measuring the degree to which literary narratives would deepen understanding of target-language culture and cultural connotation. This feature reflected the cultural foundation of language service education that was highly emphasized in the language-culture symbiosis model.

3. Intercultural Competence. This feature was extracted from items measuring the degree to which literary narratives would handle cultural difference, understand implicit meaning, and adapt expression strategy in cross-cultural context. This feature reflected the practical intercultural mediation skill that was a key component of professional language service work.

4. Translation Mediation. This feature was extracted from items measuring translation performance and language-culture transfer (e.g. handle cultural difference in translation, literary translation training). This feature reflected advanced language service skill beyond linguistic conversion.

5. Identification of Learning Motivation. This feature was extracted from items measuring the degree to which literary narratives would stimulate learning interest, participation, sense of belonging, and adaptability in language service field. This feature reflected the affective impact of literary narrative on learner's long-term development.

6. Integrated Training. This feature was extracted from items measuring the modular curriculum design, scenario-based teaching, industry-university collaboration, technological support, and future development direction. This feature reflected the integrated training evaluation of literary narrative based language service education's systemic feasibility and sustainability.

2.4 Model Specification

To quantitatively assess how literary narratives influence language service education, we use the regression algorithm Extreme Gradient Boosting (XGBoost) [7]. XGBoost is a gradient-boosted ensemble of regression trees that iteratively update their predictions by fitting additional trees to the residual errors of the current model.

In the diagram of conceptual model, the inputs to the questionnaire flow through a series of decision trees, and the outputs of all trees are aggregated to produce the predicted support score for integrating literary narratives.

For each respondent i with feature vector x_i and target value y_i, the prediction at boosting iteration t is updated as

$$\widehat{y_i^{(t)}} = \widehat{y_i^{(t-1)}} + f_t(x_i), \tag{1}$$

where f_t denotes the regression tree added at iteration t. XGBoost optimizes a regularized objective function that combines the empirical loss and a penalty on model complexity:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{N} \ell\left(y_i, \widehat{y_i^{(t-1)}} + f_t(x_i)\right) + \Omega(f_t), \tag{2}$$

In this study, $\ell(\cdot)$is the squared error loss, which is appropriate for the continuous outcome variable representing respondents' support levels. The regularization term takes the form

$$\Omega(f_t) = \gamma T_t + \frac{1}{2}\lambda \sum_{j=1}^{T_t} w_j^2, \tag{3}$$

where T_t is the number of leaves in tree f_t, $\mathrm{w_j}$ is the weight of leaf j, γ and λ are regularization hyperparameters controlling tree complexity and the magnitude of leaf weights. This design allows XGBoost to balance data fit and model simplicity, thereby mitigating overfitting.

The questionnaire includes both categorical variables (e.g., person-characteristics such as role, educational level, prior exposure to language service programs and literary narrative courses) and items measuring cultural understanding and intercultural competence, motivation, and views on curriculum development on a Likert scale. XGBoost is well-suited to these data because it can naturally model nonlinear relationships and high-order interactions between features without strong assumptions about their distribution. Furthermore, the score importance values produced by the model support interpretable analysis of which aspects of literary narrative-based education most influence respondents' level of support for integrating literary narratives into language service training.

2.5 Model Training and Validation

After cleaning, we had 378 usable responses available for construction of models. To rigorously evaluate the generalization performance of the XGBoost regression model, we randomly split the data into a training set and held-out test set. We fit the model and tune the hyperparameters using 80% (302 cases) of the samples and evaluate the final model on the remaining 20% (76 cases). We split the data such that the overall dependent variable distribution was preserved so that there was some of the dependent variable in both the training and test sets (i.e., both sets reflect the full range of support scores).

We trained and validated our models in Python using the official XGBoost implementation. On the training set, we performed a grid-search procedure using five-fold cross-validation to select reasonable hyperparameters. We searched over learning rate $\eta \in \{0.05,0.1,0.2\}$, maximum tree depth max_depth $\in \{4,5,6\}$, number of boosting rounds n_estimators $\in \{100,150,200\}$, subsample ratio of training instances subsample $\in \{0.7,0.8,0.9\}$, and column subsample ratio of training fields per tree colsample_bytree $\in \{0.7,0.8,0.9\}$. We also searched over $\alpha \in \{0,0.5,1.0\}$ and $\lambda \in \{1.0,1.5,2.0\}$ for L1 and L2 regularization parameters to control complexity in addition to the learning rate.

Based on cross-validation, the best set of parameters were lr = 0.1, max_depth = 5, n_estimators = 150, subsample = 0.8, colsample_bytree = 0.8, $\alpha = 0.5$, $\lambda = 1.5$. We also implemented an early-stopping procedure based on performance on the validation set: if the loss on the validation set did not decrease over the last 10 boosting iterations, training was stopped to prevent overfitting. Using these parameters, we re-fit the XGBoost model on the entire training set and evaluated it on the test set. We report mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination as measures of model performance and report the results in the Results section. We also investigated the relative importance of each feature as estimated by the trained model. That is, which aspects of demographic characteristics, educational experiences, and perception dimensions do respondents consider most highly correlated with their support for including literary narratives in language service education? The rankings of importance provide a data-driven guide to iteratively improve curriculum design and to identify elements of literary narrative-based pedagogy to prioritize in future language service training.

2.6 Hardware and Software Environment

All experiments in this study were conducted on a standalone computational workstation configured to support training and evaluation of machine learning models. The hardware platform featured an Intel Core i7-13700 processor (16 cores, up to 5.2 GHz), 32 GB system memory, and an NVIDIA GeForce RTX 3060 graphics card with 12 GB dedicated VRAM. Although we considered a moderate-sized dataset, we nevertheless enabled GPU support during model training to leverage the computational power offered by GPU acceleration and to allow for reproducibility if we were to explore larger feature sets and more complex models in future work. We conducted all data analysis, modeling, and evaluation experiments in a 64-bit Ubuntu Linux operating system (version 22.04 LTS). Data processing and feature encoding were implemented using pandas (version 2.0.3) and NumPy (version 1.24.4), and we used the XGBoost library (version 1.7.6) to construct and train a regression model with GPU support. We used scikit-learn (version 1.3.2) to split datasets, perform cross-validation, and provide evaluation metrics for each model. This configuration provided us with a stable, efficient, and scalable hardware-software environment to conduct the XGBoost analysis of questionnaire data in language service education.

3 RESULT

3.1 Sample Characteristics and Data Distribution

The final dataset comprised 378 valid responses. In terms of stakeholder identity, students formed the largest group (56.08%), followed by teachers (29.63%), language service practitioners (9.26%), and other respondents (5.03%). Regarding educational background, 63.23% of participants reported a bachelor's degree, 24.34% a master's degree, and 10.85% a doctoral degree, while only 1.59% had junior-college education or below. Most respondents had direct experience with language service education: 82.01% indicated that they had studied in language-service-related programs (such as translation or intercultural communication), and 73.02% reported having taken courses explicitly focused on literary narratives or related cultural content, including selected readings in English and American literature, literary history courses, or courses on the society and culture of English-speaking countries.

Items 6-20, which captured perceived educational benefits and views on future development directions using five-point Likert scales, showed similarly favorable distributions. Across these items, approximately 71%-76% of respondents chose to agree or strongly agree. Higher levels of agreement were particularly evident for statements indicating that literary narratives deepen understanding of target-language cultures, enhance intercultural communication and problem-solving in translation, increase learning motivation and professional identity, and support the integration of modular curriculum design, scenario-based teaching, industry-university cooperation, and technology-enhanced resource development. The share of clearly negative responses (options 1-2) remained relatively low for all items, indicating that skepticism towards literary narrative-based approaches is confined to a small minority.

For subsequent machine learning analysis, the 378 observations were randomly partitioned into a training set and a held-out test set using an 80/20 split, resulting in 302 and 76 cases, respectively. The two subsets were compared on key descriptive indicators (role composition, educational level, prior language service education, exposure to literary narratives, and support scores), and no substantial differences were observed. This indicates that the random split preserved the overall structure and distributional characteristics of the full sample, thereby providing a reliable foundation for training and evaluating the XGBoost regression model described in the following sections.

3.2 XGBoost Results Evaluation

Figure 1(a) shows the measures indicating the relative importance of features that were produced by the method using the approach to analysis. The method examined six factors that were included as variables for making predictions. The factors labeled as understanding relating to culture and competence in situations involving interaction across cultures received the measures indicating the highest importance. This indicates that gains that participants perceived in areas related to understanding of culture and abilities for communication across cultures represent factors that are central in determining the extent to which individuals in the study support the approach that integrates narratives from literature into

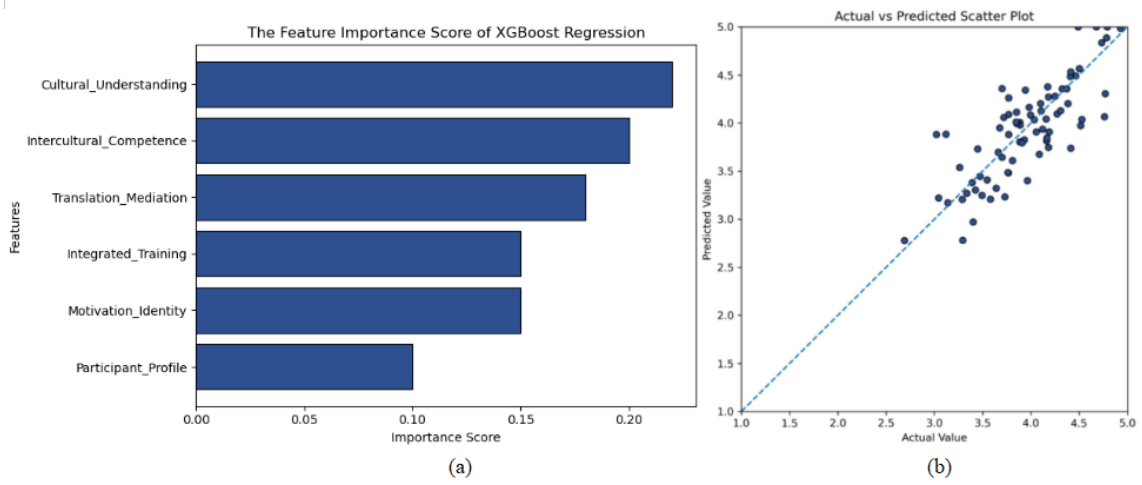


Figure 1: The XGBoost Results of Evaluation: (a) The Feature Importance Score of XGBoost Regression Model; (b) Actual vs Predicted Scatter Plot

education for language services. The factors labeled as activities involving translation and connection between elements and training that follows an integrated design showed importance at a level that is moderate. This suggests that abilities relating to translation at an advanced level and activities that involve connection between elements, and the way that individuals in the study assess approaches to instruction that use modules and scenarios and collaboration between institutions for education and settings for professional work, represent influences that are substantial. The factors labeled as motivation relating to identity of participants and profile features of participants showed measures of importance that were lower in a relative sense but were not equal to zero. This suggests that factors relating to emotions, which include motivation for activities involving learning and identity relating to professional work, and differences relating to features that describe individuals, which include role and level of education and previous training, represent contributions that are meaningful for the process of making predictions. The results that the method produced indicate overall that gains relating to understanding and culture and abilities for practical work involving interaction across cultures represent the factors that are strongest in determining the degree to which individuals who are involved in this context support the approach to education for language services that is based on narratives from literature. Factors relating to structure and factors relating to emotions provide additional capacity for providing explanations.

Figure 1(b) shows the relationship between the actual support measures and the values that the model using XGBoost Regression indicates on the separate test data. The line that appears as dashed diagonal represents the reference at forty-five degrees, where values that the model predicts would match the observed measures in a complete manner. The points that appear as scatter are located around this line in large part and show alignment that is positive and clear, indicating that the model provides the means to represent the overall pattern of support that stakeholders demonstrate for integrating narratives from literature into education for language service. A degree of dispersion that is certain can be observed, however, in particular in the range that is medium between approximately three point zero and four point five, where individual values that the model predicts deviate from the diagonal. In addition, some measures that are high in actual terms are underestimated to a slight degree and some measures that are relatively low are overestimated to a slight degree, suggesting that the model shows performance that is less accurate at the points that are extreme on the measure. Overall, however, the distribution of points around the line at forty-five degrees suggests that the model using XGBoost Regression provides approximation that is reasonably good of the actual support measures, while leaving space for improvement that is further through construction of features that is more refined and adjustment of parameters.

Table 1: Evaluation Metrics of XGBoost

Model	R^2	Mean Square Error	Root Mean Square Error	Mean Absolute Error
XGBoost	0.566	0.098	0.313	0.242

Figure 2 presents the residuals of the model using XGBoost Regression that are plotted against the support measures that the model predicts. The line that is horizontal and dashed marks residual at zero; in ideal terms, points should vary in a manner that is random around this line if the model represents the structure that is systematic in the data and the remaining differences are mostly random. In the case that we examine, residuals are distributed both above and below zero across the full range of values that the model predicts, and no pattern that is clear as funnel shape or curved form is observed, suggesting that there is no problem that is severe with variation that changes or misspecification that is obvious and nonlinear. Overall, however, the group of residuals remains compact to

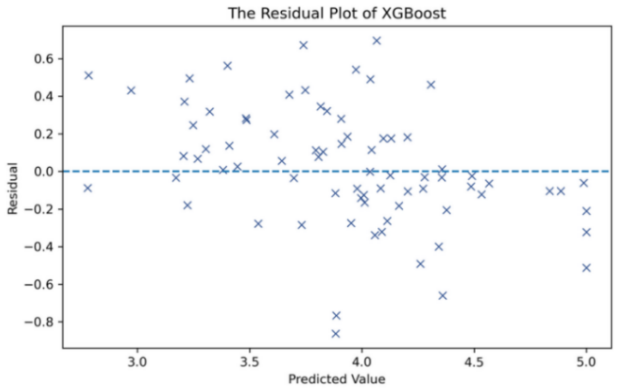


Figure 2: The Residual Plot of XGBoost Regression Model

a reasonable degree, and the majority of deviations remain within a band that is narrow around the line at zero. The results provide evidence that is further that the model using XGBoost Regression provides a fit that is satisfactory for the data from the study using questionnaire, while leaving some space for improvement through construction of features that is refined or adjustment of parameters that is more focused.

Table 1 shows the measures that assess the model on the test data. The measure of determination, $R^2 = 0.566$, indicates that the model accounts for around 56.6% of the variation in the support that individuals show for including narratives from literature in training for language work. The sample in the study is relatively small and the features that the analysis includes are limited. This degree of explanation appears acceptable and suggests that the features selected represent a substantial part of the structure that the data contain. The measure showing error in squared form is 0.098. The measure that presents the square root of this error is 0.313. These results indicate that the typical difference between the values that the model predicts and the values that the data provide is roughly 0.31 points on a scale that uses five points for rating. The measure showing the absolute value of error is 0.242. This finding reveals that the predictions that the model produces differ from the observed values by less than a quarter of one unit on the scale. The measures suggest that the model provides a reasonably good balance. It shows fit to the data while also maintaining a level of simplicity. The model represents the main patterns that the data from the study contain. However, the results also indicate that improvement is possible. Future work that develops richer features, uses larger samples, or applies more refined procedures for selecting parameters may enhance the model.

4 DISCUSSION

This study presents an approach to training in language work that includes narratives from literature and uses data to inform the design of the training. The results from the study indicate that individuals show a high level of support for including literary narratives in training for language work. By prioritizing XGBoost over conventional linear regression, this analysis effectively unveiled the non-linear relationships between cultural literacy and professional skills, offering a level of granular insight that simpler statistical methods often obscure. Although this machine learning approach demands more rigorous validation to mitigate overfitting compared to standard models, its superior ability to rank feature importance explicitly quantifies how 'Cultural Understanding' drives stakeholder support, providing a robust empirical basis for the proposed modular curriculum. The analysis examines six features that combine multiple aspects. The feature relating to understanding of culture and the feature relating to communication across cultures show the largest influence on the attitudes that individuals report. The feature that represents ability in translation and mediation also contributes to explanation. The feature that relates to design of integrated training, the feature that concerns motivation and identity, and the feature that describes the profile of individuals in the study provide additional explanation. The measures that evaluate the model and the plots that provide diagnosis indicate that the model represents the main structure in the perceptions that individuals hold. These findings suggest that the model provides a reasonable representation of the support that individuals show for including narratives in training for language work.

The focus on features relating to culture and interaction between cultures in the model appears consistent with the theory [8] that language and culture show a relationship of this form. Individuals in the study indicate a clear preference for a form of teaching that moves beyond simple language conversion and provides focus on understanding culture and on means for interaction across cultures. Texts that use narrative forms are not considered simply as objects with features relating to aesthetic value. These texts are considered as dense sources containing cultural content that allow students to examine symbols and values that are not made explicit and also to examine patterns in memory relating to history. Such examination is important for providing services that involve language at a high level. The importance given to features relating to the work of translation and to features that involve practical application through experience is also substantial.

Using these findings, the study provides a framework for implementing an approach that uses narrative texts for teaching services relating to language, as shown in Table 2. The framework does not present separate lists that show courses or conditions relating to resources. The framework is presented as steps that show connections between them. These steps include design of courses in a form that uses modules, an approach to teaching that uses scenarios, work on translation that involves practice and examination of cases, work that involves collaboration between institutions that provide teaching and organizations that involve industry, and an approach to support that uses both forms involving technology and resources. The approach that uses scenarios and work that involves translation in settings that provide practical experience provide means to develop skills that involve understanding across cultures and skills relating to work that involves acting as an intermediary between cultures. Work that involves collaboration with organizations in the area designated as the Hainan Free Trade Port provides connections between these skills and situations that occur in professional settings that are real.

The analysis also shows limitations and areas for further work. The pattern of differences between predictions and data suggests

Table 2: Application of Literary Narrative and Language Service Education Model.

Steps	Specific Measures	Objectives
Modular Course Design	Offer modules such as "Literature and Language Education" and "Cultural Symbols and Translation" to organically integrate literary analysis with language skills teaching.	Enable students to master language skills
Scenario-based Teaching	Design international language service cases, such as simulating translation tasks for international literary festivals and cultural symbol translation in cross-cultural business negotiations.	Enhance students' cross-cultural adaptability and practical abilities
Translation Practice and Case Analysis	Let students complete translation tasks for literary works and analyze cultural adaptation choices and strategies during the translation process.	Strengthen students' cultural sensitivity and train their language service abilities in complex cultural scenarios.
University-Industry Collaboration Practice	1.Collaborate with multinational enterprises and cultural institutions in the Hainan Free Trade Port, allowing students to participate in actual translation and cultural promotion activities, such as tour guide translation and translation of promotional materials. 2. Establish a cultural education practice platform, leveraging the international advantages of the Hainan Free Trade Port, organizing students to participate in international literary festivals, cross-cultural reading clubs, international conferences, and cultural exhibitions.	Develop students' language service abilities in real-world contexts while providing talent support for the international development of the Hainan Free Trade Port.
Technology Support and Resource Development	1.Develop AI-based literary narrative cultural analysis tools and scenario simulation translation platforms. 2. Use AI technology to develop literary narrative analysis tools and translation case libraries to assist students in efficient learning. 3. Develop regionally specialized programs, leveraging the geographical advantages of the Hainan Free Trade Port.	Provide personalized learning platforms to assist students in efficient learning and practice, adapting to diverse needs.

that larger differences occur at higher levels of support. This indicates that other factors may affect responses. Such factors include the structure of the organization or previous work in other countries. These factors may separate individuals showing high support from others. The measure of variance that the model accounts for is moderate. This suggests that the six features in the analysis provide the main patterns but do not include all influences. Future research should prioritize longitudinal designs to assess the long-term retention of cultural competence, moving beyond cross-sectional surveys to track graduates' career trajectories. Furthermore, to enhance the model's transferability, scholars must investigate implementation barriers-such as curriculum rigidity or resource deficits-and test adaptive strategies, like tailoring literary corpora to local linguistic demographics, ensuring the model's viability across diverse educational settings. This approach allows measurement of the learning that occurs from particular approaches using narratives.

The study shows that combining narratives from literature with analysis using the method that builds decision structures provides a framework for language service programs. Narratives from literature provide material on culture and situations involving communication. The method using decision structures provides a way to measure factors and to focus on dimensions in programs. For areas such as the Hainan Free Trade Port, this approach that combines the two elements offers a means to develop individuals working in language services. These individuals show skill in language, understanding of culture, knowledge of technology, and ability to work between languages and cultures.

5 CONCLUSION

This study used a survey and a method building decision structures to examine the use of narratives from literature in programs for language services. The focus was on the case of the Hainan Free Trade Port. The analysis shows that factors indicating learning about culture and skills in communication between cultures are the main influences on support from individuals involved in the programs. Factors relating to skills in moving between languages, design that combines different elements, and factors affecting motivation provide further influence. The study proposes a framework using narratives and separate units that link analysis of literature, work on situations, work between the organization and industry, and learning supported by technology. This framework provides a model that uses both theory and data for developing individuals in language services. The application of the method building decision structures shows the value of such methods in finding important dimensions and in using data to develop programs in language services. The study also has limitations from the sample size, use of data that individuals report on themselves, and use of a single method. Future work should increase the sample, include studies following individuals over time or using assignment to groups to measure learning, and compare other methods and other ways of representing features. This work would further establish and improve the framework in different settings involving multiple languages and cultures.

Acknowledgments

Grand Number: 2025H20984, Innovative Pathways for Integrating English Literature Teaching under the Perspective of Digital-Intelligent Transformation toward "Understanding Contemporary China" (2025) sponsored by Collaborative Research Project of the Shaanxi Provincial Special Program for Foreign Language Disciplines.

References

[1] Wang, Shumin, Xin Jiang, and Muhammad Bilawal Khaskheli. "The role of technology in the digital economy's sustainable development of Hainan Free Trade Port and genetic testing: Cloud computing and digital law." Sustainability 16, no. 14 (2024): 6025.

[2] Zhang, Ping, and Bob Adamson. "Multilingual education in minority-dominated regions in Xinjiang, People's Republic of China." Journal of Multilingual and Multicultural Development 44, no. 10 (2023): 968-980. https://doi.org/10.1080/01434632.2020.1850744

[3] Krulatz, Anna, MaryAnn Christison, Eliane Lorenz, and Yeşim Sevinç. "The impact of teacher professional development on teacher cognition and multilingual teaching practices." International Journal of Multilingualism 21, no. 2 (2024): 711-727.

[4] Zhu, Yaoshun, Zhitao Zhu, and Wanxian Xu. "Cross-border higher education cooperation under the dual context of artificial intelligence and geopolitics: opportunities, challenges, and pathways." Frontiers in Education, vol. 10, p. 1656518. Frontiers, 2025.

[5] Dong, Jianwei, Yumin Chen, Bingyu Yao, Xiao Zhang, and Nianfeng Zeng. "A neural network boosting regression model based on XGBoost." Applied Soft Computing 125 (2022): 109067.

[6] Baddane, Khadija, and Abdelghanie Ennam. "Measuring pedagogical transformation: A quantitative analysis of critical thinking integration in literary criticism for heightened student engagement and learning outcomes." International Journal of Linguistics, Literature and Translation 7, no. 1 (2024): 39-50.

[7] Wu, Zihe, and Ye Zhang. "The data visualization and intelligent text analysis for effective evaluation of English language teaching." Scientific Reports 15, no. 1 (2025): 22737.

[8] Guan, Zhengzhi, Fangfang Hou, Boying Li, Chee Wei Phang, and Alain Yee-Loong Chong. "What influences the purchase of virtual gifts in live streaming in China? A cultural context-sensitive model." Information Systems Journal 32, no. 3 (2022): 653-689.