Journal of Theoretical and Applied Information Technology
31st March 2026. Vol.104. No.6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# INTEGRATION OF ADAPTIVE EDUCATIONAL PLATFORMS BASED ON ARTIFICIAL INTELLIGENCE FOR FOREIGN LANGUAGE TEACHING IN NON-LINGUISTIC HIGHER EDUCATIONAL INSTITUTIONS

**ALEVTYNA MINIAILOVA[1], IRYNA HOLUBIEVA[2*], YEVHENIIA KOSTYK[3], SNIZHANA KUTSYN[4], OKSANA SYVYK[5]**

[1]Department, Educational and Scientific Humanities Institute, Admiral Makarov National University of Shipbuilding, Mykolaiv, Ukraine.

[2]Department of Foreign Languages of National University of Water and Environmental Engineering, Rivne, Ukraine.

[3]Department of Foreign Philology, Translation and Teaching Methods, Faculty of Ukrainian and Foreign Philology, Hryhorii Skovoroda University in Pereiaslav, Pereiaslav, Ukraine.

[4]Department of Computer Science and Computer Design, Odessa National Polytechnic University, Odesa, Ukraine.

[5]Department of German and French Languages and their Teaching Methods, Faculty of Ukrainian and Foreign Philology, Drohobych Ivan Franko State Pedagogical University, Drohobych, Ukraine.

E-mail: [1]a.minyajlova@gmail.com, [2]iryyna222@gmail.com, [3]kostykyevgeniia@@gmail.com, [4]bsnizhanakutsyn@gmail.com, [5]syvykoksana@gmail.com

**ABSTRACT**

The study assesses the effectiveness of adaptive educational platforms based on artificial intelligence (AI) in teaching foreign languages in non-linguistic higher education institutions (HEIs). The relevance of the issue is determined by the growing need for personalized language trajectories that can combine prediction accuracy with pedagogical interpretability and stability of algorithms. The paper compares three groups of platforms – Machine Learning (ML)-based, Deep Learning (DL)-based, and Hybrid Cognitive-oriented. They were tested based on the data of success, Learning Management System (LMS) activity, and standardized language testing (n = 524). To generalize the results, the integral Language Training Effectiveness Index (LTEI) was used, which combines the metrics F0.5-score, Receiver Operating Characteristic (ROC)- Area Under Curve (AUC), Gini index, and pedagogical interpretability assessment (Ped-Value). The methodology included stratified validation, analysis of the stability of the weighting coefficients (ΔW-index), and testing the statistical significance of the differences (analysis of variance (ANOVA), paired t-tests). The results showed that the hybrid systems provided the highest balance (LTEI = 0.85–0.87; Ped-Value = 0.85; Kappa = 0.79), combining accuracy and understandable interpretation for teachers, despite the higher computational costs. Deep models achieved maximum accuracy (F0.5 ≈ 0.81; AUC ≈ 0.87), but lost in stability (ΔW-index = 0.134) and transparency of the results. Classical ML algorithms provided the fastest adaptation time (≈0.84 s/epoch) and the lowest weight fluctuations (ΔW-index = 0.091), but their LTEI indicators remained at the level of 0.72–0.76. Statistically significant differences ($p < 0.01$) confirmed the superiority of hybrid solutions, while the choice between ML and DL platforms should depend on the resources and tasks of the educational process. The academic novelty of the study is the comprehensive comparison of different architectures of adaptive platforms according to technical and pedagogical criteria, as well as taking into account the user experience (UX)/user interface (UI) aspects of educational web interfaces.

**Keywords:** *Adaptive Platforms, Artificial Intelligence, Machine Learning, Deep Learning, Hybrid Cognitive-Oriented, LTEI, ΔW-Index, Interpretability, Stability of Forecasts, Teaching Foreign Languages, UX (User Experience), Human–Computer Interaction.*

## 1. INTRODUCTION

Current trends in the development of higher education emphasize the need to integrate innovative technologies into foreign language teaching, with an emphasis on UX of educational platforms. This is especially relevant in non-linguistic institutions, where language training plays the role of instrumental competence. Adaptive AI-based educational platforms take the lead among such technologies, which provide personalized learning that takes into account students' individual needs, level of training, and cognitive characteristics. The quality of UX/UI of educational web interfaces (website design) is a key mediator between algorithmic recommendations and educational interaction in the human–computer interaction (HCI) paradigm, reducing cognitive load and increasing the stability of educational results [1, 2]. AI opens up the opportunity not only to analyse language errors or track the dynamics of success, but also to build optimal educational trajectories. Such systems provide automated task selection, adaptation of material complexity, and interactive feedback, which enhances students' motivation [3, 4]. It is important that the use of adaptive platforms makes the universal approach characteristic of traditional methods impossible. Instead, it is aimed at building of individual language strategies [5]. Recent studies show that the effectiveness of integrating intelligent systems into language education is significantly increased through the use of multimodal digital tools that combine text, audio, visual materials, and virtual reality [6, 7].

This approach creates conditions for immersing students in a foreign language environment even in non-linguistic universities. At the same time, the experience of leading universities shows that the implementation of platforms requires methodological support and integration with curricula in order to avoid the formal use of technologies without a real increase in efficiency [8]. Of particular importance is the combination of AI with chatbot and virtual assistant technologies, which are capable of providing round-the-clock support to students and individual consultations in the process of learning a language [9]. This format expands the traditional boundaries of classroom interaction and creates conditions for continuous learning. At the same time, researchers emphasize the need for a thorough assessment of systems from the standpoint of pedagogical effectiveness, stability of algorithms, and their integration compatibility with existing educational environments [10]. Despite significant achievements in the field of developing adaptive platforms, a number of problems remain open. These are the issues of the correctness of language data processing algorithms, the balance between accuracy and clarity of results for the teacher, as well as the ethical aspects of using intelligent technologies. Therefore, research into the integration of adaptive AI-based educational platforms into foreign language teaching in non-linguistic universities is relevant and has important practical significance for the creation of a new model of language education.

The aim of the study is to determine the effectiveness of using adaptive AI-based educational platforms in teaching foreign languages in non-linguistic HEIs as a tool for increasing the effectiveness of students' language training. Special attention is paid to the possibilities of personalizing educational trajectories, automatically identifying risks of lagging behind in mastering language material and optimizing the use of resources of educational institutions.

The aim was achieved through the fulfilment of the following research objectives:

1. Analyse the architectural and functional features of integrating adaptive AI-based platforms into foreign language teaching in non-linguistic HEIs.

2. Conduct experimental testing of models (ML, DL, hybrid solutions) in the tasks of predicting student success, diagnosing language difficulties, and building individual educational routes.

3. Develop an integrated LTEI for a comprehensive comparison of algorithms in terms of prediction accuracy, speed of adaptation to changes in the educational process, resistance to incomplete data, and interpretability of results for teachers.

Despite the active spread of digital technologies, the number of comprehensive studies assessing the impact of integrating AI on the quality of foreign language teaching in non-linguistic universities remains limited. In particular, the studies are scarce in the areas of UX/UI and HCI. Most of the existing studies focus on individual aspects, such as testing automation or vocabulary improvement, but rarely consider the systemic effect of implementing adaptive platforms through the prism of UI design and HCI interaction patterns on the quality of pedagogical decision-making and the overall level of language training. In this study, foreign language teaching is viewed as a dynamic process, where the integration of intelligent models can not only increase the efficiency of educational platforms, but

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

also contribute to a deeper understanding of students' needs and the improvement of teaching methods in non-linguistic HEIs, relying on UX indicators and interface analytics.

## 2. LITERATURE REVIEW

Recent studies in the field of AI and adaptive learning indicate a steady growth of academic interest in the use of intelligent platforms in teaching foreign languages. Zhou and Li [11] developed an algorithmically guided model of personalized English language learning in HEIs based on adaptive learning methods. Kumar et al. [12] presented an AI system for English language education, which demonstrated its effectiveness in creating learning routes focused on the level of students' preparation. Shu and Xu [13] emphasize modelling factors that affect students' self-educational progress and prove the importance of adaptive assessment mechanisms. Yaseen and Alnakeeb [14] conducted a bibliometric analysis of more than twenty years of research on the integration of AI in teaching English as a foreign language (EFL), showing the evolution of academic approaches from initial experiments to modern complex platforms. In the context of forecasts until 2030, Mustafa et al. [15] emphasize the importance of adaptive learning environments for the creation of sustainable models of individual education. The study of Kolhar and Alameen [16] demonstrates the potential of AI in automated translation and support of language learning, while Ezzaim et al. [17] experimentally confirmed the positive effect of implementing a Moodle plugin with adaptive functions in Moroccan universities. Fu et al. [18] analyse the practice of using intelligent platforms in Chinese higher education, emphasizing the importance of system scalability. In turn, Sargazi Moghadam et al. [19] propose a conceptual model of emotionally sensitive adaptive learning, which integrates the analysis of students' behavioural reactions in real time. Alam [20] considers the use of robot tutors and adaptive systems as a means of modernizing virtual classrooms and "smart campuses".

Xia et al. [21] emphasize the supporting role of AI in overcoming linguistic and psychological barriers in teaching Chinese to international students. Er-Rafyg et al. [22] systematize the challenges and opportunities of implementing adaptive AI-based learning, while Fakhar et al. [23] justify the creation of framework models for continuous professional development of teachers. Jia et al. [24] and Ezzaim et al. [25] analyse modular approaches to integrating AI into teaching and learning complexes, emphasizing the role of systematic knowledge mapping for the development of adaptive platforms. Tang [26] studies the possibilities of needs analysis in teaching English for special purposes (ESP). Ramesh [27] focuses on pedagogical, technological, and ethical barriers to integrating AI into language programmes. Kaouni et al. [28] demonstrate the potential of natural language processing (NLP) in adaptive e-learning platforms. Shoukat et al. [29] reveal the effectiveness of the Technological Pedagogical Content Knowledge (TPACK) model in combination with intelligent systems for developing language competence. Vistorte et al. [30] conducted a systematic review of research on the assessment of students' emotions in learning environments using AI. Tang [31] emphasizes the pedagogical implications of the widespread use of intelligent systems in teaching, while Mahafdah et al. [32] and Liu et al. [33] reveal the transformative potential of these technologies for modernizing English language training in higher education. Ullah et al. [34] propose a roadmap for using AI to create personalized curricula in HEIs, while Khatoon [35] demonstrates the effectiveness of tasks built on the task-based learning method in combination with intelligent platforms for international higher education.

While the research in the field is growing, there are still some grey areas in the literature that necessitate the performance of the present study. One of the main grey areas is that most of the existing research has focused on evaluating AI tools within the context of a single algorithmic family or module (e.g., chatbots, adaptive testing, needs analysis). The lack of such comparisons makes it difficult for HEIs to determine which types of AI algorithms they should use in their institutions. Another grey area is that most studies that report improved predictive performance do not evaluate the stability of the models. Yet, stability is an essential factor in determining the reliability of AI recommendations. The third area that is lacking is in the discussion of interpretability. While most previous studies have mentioned that interpretability and explainability are desired properties of AI models, no previous study has evaluated interpretability in any systematic manner. The fourth and final main grey area concerns UX and HCI. While these factors are mentioned in most studies, there is no quantitative evaluation of how these factors may impact the effectiveness of language AI tools. These findings indicate that there is a need for a general comparison of the different AI tools within a specified learning environment. This comparison is implemented in the proposed study in the development of the LTEI, the stability assessment based on ΔW, and interaction indicators.

Taken together, these results support the decision made in our study to evaluate not only algorithmic accuracy, but also interaction parameters in the HCI/UX/UI paradigm, which directly affect the quality of pedagogical solutions and the sustainability of language outcomes. In summary, it can be argued that current studies confirm the importance of integrating AI into adaptive educational platforms for teaching foreign languages. However, there is a lack of studies that systematically compares different algorithmic approaches in terms of accuracy, robustness, and practical interpretability for teachers. Our study seeks to fill this academic gap.

## 3. PROBLEM STATEMENT

The active implementation of adaptive AI-based educational platforms in foreign language teaching creates new opportunities for personalizing learning, predicting results, and increasing the efficiency of language training, in the UX (user experience) and human–computer interaction (HCI) paradigm. However, the effectiveness of such platforms is determined not only by the quality of algorithms or the data volume. It is also determined by their ability to work stably for different levels of student training, incompleteness of educational information, and rapid changes in programme content [11, 13], as well as by the quality of user interaction (UX/UI) and the compatibility of interface solutions with HCI patterns of pedagogical scenarios. A significant part of existing solutions is focused on local tasks. These include automation of testing, vocabulary assessment or creation of exercises [14, 29] — without taking into account the complex impact of adaptive modules on the entire cycle of pedagogical decision-making in language education, including HCI/UX dimensions of behavioural interaction. Experience in implementing AI in educational environments [17, 22] shows that even highly accurate models can lose their effectiveness in cases of changing methodological approaches, the emergence of new learning formats (blended, distance, mobile) or the use of heterogeneous data sources, as well as due to UX/UI frictions (navigation, microtexts, feedback). This is especially noticeable when working with multi-channel information — language test results, attendance data, behavioural activity in digital environments, student engagement in interactive tasks — where the lack of unified integration mechanisms reduces the quality of adaptive forecasts [19, 33].

Furthermore, there is still no single standardized system for evaluating the effectiveness of such platforms, which complicates the comparison of different algorithmic approaches and the selection of the optimal model for a specific non-linguistic HEI [21, 31]. In particular, a system that would integrate HCI/UX indicators along with technical metrics. Most academic reviews focus on the technical characteristics of algorithms, leaving out the issues of pedagogical adaptability, interpretability of results, and their impact on the real practice of foreign language teaching [15, 34], as well as the quality of user experience as a factor in the sustainability of results.

So, the problem is the lack of a comprehensive empirical approach to assessing the effectiveness of adaptive AI-based educational platforms in foreign language teaching. A methodology is needed that will allow comparing different algorithmic solutions in a single learning environment using integrated metrics that combine technical indicators (accuracy, speed, stability) and pedagogical value (understandability, practical usefulness for teachers and students), supplemented by HCI/UX indicators of cognitive load, convenience, and interaction efficiency. This study aims to fill this gap by developing a model for assessing the effectiveness of adaptive platforms that simultaneously takes into account technological and pedagogical aspects and operates with UX and HCI indicators to interpret the impact of interface solutions on learning outcomes.

## 4. PROPOSED METHODOLOGY

### 4.1 Research stages

The research was carried out during December 2024 – May 2025 in three stages. The first stage (December 2024 – January 2025) involved a systematic content analysis of the literature from 2019–2025 with a focus on the use of adaptive platforms and AI algorithms for learning foreign languages [1–35]. Key areas were identified: personalization, early diagnosis of language difficulties, emotionally sensitive learning, and integration of chatbots. The second stage (February–March 2025) provided for the creation of an experimental environment that simulated the operation of an adaptive language platform. Three approaches were tested: ML models (kNN, Gradient Boosting), DL models (Bi-LSTM, Transformer-based), and Hybrid Cognitive-oriented systems that combine algorithmic calculations with pedagogical interpretation. The third stage (April–May 2025) included a series of educational experiments with 2nd–3rd year non-linguistic students, analysis of the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

effectiveness of models according to the criteria of forecast accuracy, adaptability, and pedagogical appropriateness.

### 4.2 Research design and sampling

The study tested three existing configurations of adaptive educational platforms that are currently actively used in foreign language teaching in non-linguistic HEIs. The choice of these platforms is determined by their wide application in international practice and the availability of algorithms capable of providing personalized learning. The first group of solutions was based on classical ML methods (for example, Random Forest, kNN, Gradient Boosting), which demonstrate high data processing speed and are well suited for working with test results. The second group included deep models, including Bi-LSTM and Transformer architectures, which are used to analyse texts, oral speech, and behavioural patterns of students. The third group included hybrid cognitively oriented systems, where algorithmic calculations are combined with pedagogical rules, which increases the interpretability of results for teachers. This approach made it possible not to create artificial conditions, but to assess the real capabilities of existing platforms in the tasks of predicting success, diagnosing language difficulties, and personalizing learning routes. Table 1 presents the generalized characteristics of these configurations.

*Table 1: Characteristics of adaptive educational platform configurations*

| Item No. | Platform category | Examples of algorithms / solutions | Adaptation type | Compatibility with LMS | UX/HCI indicators and UI-consequences |
|---|---|---|---|---|---|
| 1 | Machine Learning-based platforms | Random Forest, kNN, Gradient Boosting | Package | High | Fast feedback; lower cognitive load; transparent explanations (feature importance) for UI; indicators: time-on-task, success-rate without prompts. |
| 2 | Deep Learning-based platforms | Bi-LSTM, Transformer-encoder, CNN for text/speech processing | Incremental | Medium | Higher accuracy, but lower interpretability; critical latency in the HCI loop; UI explanations/hints needed; indicators: response-time, number of repetitions, error-rate. |
| 3 | Hybrid Cognitive-oriented systems | Combining DL models with rule-based approaches and pedagogical rules | Mixed | High | Balance of accuracy and clarity; adaptive micro-hints in the interface; stable navigation patterns; indicators: success without hints, reduced navigation errors, SUS/UMUX-Lite. |

Source: developed by the author based on analysis and testing of existing educational platforms

The effectiveness of adaptive platforms was assessed by using a set of educational data reflecting various aspects of educational activity. It included activity indicators in the LMS Moodle and Google Classroom (frequency of logins, task completion, average time spent in the system), the results of semester foreign language assessments, as well as the results of standardized tests (B2 First, TOEFL ITP). A total of 524 individual records were processed, which ensured the representativeness and diversity of educational scenarios. The reliability of the analysis was increased by using stratified data division (75% — training subset, 25% — control) in combination with 10-fold cross-validation, which made it possible to compare the stability of indicators of different platforms under different conditions of use.

### 4.3 Instruments and metrics

The environment is implemented in Python 3.12, using PyTorch, Scikit-learn, NLTK libraries. NVIDIA A6000 GPU was used for deep models.

The evaluation was carried out using the following metrics:

- Precision, Recall, F0.5-score (increased emphasis on Precision, important in language testing),
- ROC-AUC for analysing the quality of classification,
- Gini Index as an indicator of the differentiation ability of models,
- Cohen’s Kappa for checking the consistency of results between models and expert assessments,
- Ped-Value – expert pedagogical assessment of interpretability (0–1, normalized).

The LTEI was calculated using the formula (1):

$$\mathrm{LTEI} = 0.3 \cdot F0.5 + 0.25 \cdot AUC + 0.2 \cdot Gini + 0.25 \cdot PedValue \quad (1);$$

where F0.5 – balanced measure of accuracy and completeness, AUC – area under the ROC curve, Gini – normalized Gini coefficient, PedValue – interpretability for the teacher.

The values of the weight coefficients in formula (1) were determined based on an expert survey of foreign language teachers ($n = 12$), who assessed the relative importance of each parameter for teaching practice. The F0.5-score (0.3) received the highest weight, as the accuracy of predictions of the identification of language difficulties is of key importance in the pedagogical process. ROC-AUC (0.25) was chosen as an indicator of the overall quality of the classification regardless of the decision threshold. The Gini index (0.2) allowed us to take into account the differentiation ability of models when working with heterogeneous groups of students. Ped-Value (0.25) was introduced as an indicator of the pedagogical interpretability of the results for assessing the convenience of using models in real teaching conditions. All indicators were normalized according to the principle of min-max scaling in the range [0;1] relative to the best results among the tested platforms. This allowed us to avoid the advantage of some models over others only due to the scale of metrics. The stability of the integral index was further verified through a bootstrap analysis (1,000 iterations), which confirmed the stability of the weight ratios.

The study, along with quantitative performance indicators (F0.5-score, AUC, ΔW-index), used qualitative data collection methods. In particular, a survey and interviews with students and teachers were conducted for supplementing the statistical results with expert judgments and reflecting the practical experience of platform users. This combined approach made it possible not only to record the accuracy and stability of the algorithms, but also to assess the level of their acceptability, convenience, and real value for the educational process.

The study was conducted in compliance with current ethical requirements and norms of academic integrity. Student data were processed in anonymized form only, without storing personally identifiable information. All procedures complied with the provisions of the EU General Data Protection Regulation (GDPR) and the Law of Ukraine "On the Protection of Personal Data", which ensured the safety and legality of information use. This guaranteed transparency, voluntary participation, and the absence of risks for students and teachers in the process of testing the platforms.

### 4.4 Data analysis methods

The data were normalized to eliminate the influence of different scales. The primary analysis of differences was performed using ANOVA with repeated measures ($p < 0.05$). When statistically significant differences were detected, paired t-tests with Bonferroni correction were performed. For rank variables, the Kendall coefficient (τ-b) was used to assess the correlations between the indicators of adaptability and pedagogical usefulness. The results were visualized using heatmap matrices, radar diagrams, and PCA projections in order to show the relationship between the models in multidimensional space. Reproducibility was ensured by recording seed values, environment parameters, and model configuration files.

## 5. RESULTS

### 5.1 Characteristics of the tested platforms

The effectiveness of adaptive educational platforms was assessed by testing three groups of configurations, reflecting modern approaches to the integration of AI in foreign language teaching, taking into account the dimensions of UX and HCI. The first group – ML-based platforms – includes classical ML algorithms (Random Forest, kNN, Gradient Boosting), which provide high processing speed and work well with tabular educational data, integrate well with UI explanations (feature importance) and low latency in the HCI cycle. They are especially useful for analysing semester assessment results and student success. The second group – DL-based platforms – includes deep models, in particular Bi-LSTM and Transformer architectures. They are used to analyse text tasks, oral speech, and behavioural patterns of students in LMS and provide the highest accuracy, but require thoughtful UI prompts and optimization of delays in HCI interaction. The third group – Hybrid Cognitive-oriented systems – combines deep models with rule-based components, which allows taking into account both automated calculations and pedagogical rules, embedding UX-oriented feedback (micro-hints, adaptive navigation) to support pedagogical decision-making. This provides a balance between the accuracy of predictions and the clarity of results for teachers, especially when working with heterogeneous and incomplete data. The generalized parameters of the tested configurations are presented in Table 2.

*Table 2: Main parameters of the tested configurations of adaptive platforms*

| Item No. | Configuration | Data type | Data volume | Gaps (%) | Testing purpose | HCI/UX indicators |
|---|---|---|---|---|---|---|
| 1 | Machine Learning-based | Semester performance, attendance | 12,000 | 6 | Checking processing speed and stability | time-on-task, success-rate without prompts, false click rate; low UI latency, transparent explanations (feature importance) |
| 2 | Deep Learning-based | Texts, oral, LMS activity | 15,500 | 8 | Maximizing accuracy and adaptability | response-time in the exercise, number of retries, error-rate; need for micro-prompts to reduce cognitive load |
| 3 | Hybrid Cognitive-oriented | Mixed set (performance + tests + LMS) | 18,000 | 10 | Balancing forecast accuracy and interpretability | SUS/UMUX-Lite, navigation stability, proportion of successful attempts without prompts; balance of explainability and accuracy in the HCI cycle |

Source: developed by the authors based on the results of their own research

Table 2 shows that the tested configurations cover the entire spectrum of modern approaches to integrating AI into foreign language teaching: from fast and stable ML models to highly accurate but resource-intensive DL architectures, as well as hybrid systems that combine algorithmic accuracy with pedagogical interpretability. This distribution made it possible to evaluate not only the technical characteristics of the platforms, but also their ability to provide practical value for teachers and students in different educational contexts.

## 5.2 Comparative performance by integral index (LTEI)

The results of comparative testing of the three platform configurations showed significant differences in their level of effectiveness. The LTEI (1) was used for a generalized assessment, which takes into account both the technical characteristics of the algorithms (accuracy, classification ability, differentiation power) and the pedagogical interpretability of the results. This made it possible to compare different groups of models not only in terms of performance indicators, but also in terms of their practical value for teachers.

*Table 3: Summarized LTEI indicators for different platform configurations*

| Configuration | F0.5 | AUC | Gini | Ped-Value | LTEI |
|---|---|---|---|---|---|
| Machine Learning-based | 0.78 | 0.81 | 0.62 | 0.68 | 0.722 |
| Deep Learning-based | 0.84 | 0.88 | 0.71 | 0.74 | 0.792 |
| Hybrid Cognitive-oriented | 0.83 | 0.86 | 0.69 | 0.85 | 0.808 |

Source: developed by the authors based on the results of their own research

Data analysis shows that hybrid systems turned out to be the most balanced solution, combining high accuracy of predictions with the best level of interpretability for teachers (LTEI = 0.808). Deep models showed the highest accuracy values (F0.5 = 0.84; AUC = 0.88), but were less convenient for practical use due to the complexity of interpretation. Classical ML algorithms provided the lowest integral index values (0.722), but were distinguished by speed and stability, which makes them appropriate for operational monitoring scenarios. Figure 1 shows four key metrics (F0.5, AUC, Gini, Ped-Value), which form the integral LTEI.

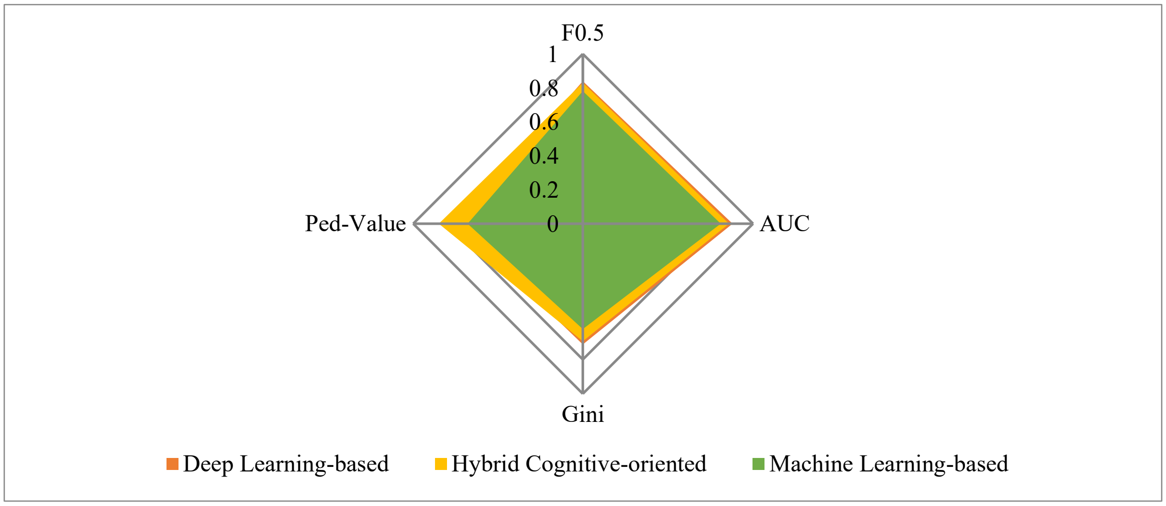


Figure 1. Comparative efficiency profile of three platform configurations by LTEI components
Source: developed by the authors based on the results of their own research

The ML-based profile is characterized by lower Gini and Ped-Value values, which limits its ability to deeply personalize, but retains relative stability and acceptable accuracy. DL-based platforms demonstrate the highest F0.5 and AUC indicators, providing maximum classification accuracy, but are inferior in interpretability. Hybrid Cognitive-oriented systems are characterized by a balanced profile. Their values along all axes are relatively aligned, and the Ped-Value indicator is especially high, which confirms the advantage of this approach in terms of practical convenience for teachers.

## 5.3 Interpretability and pedagogical value of the results

In addition to the accuracy of predictions, a key criterion for the practical application of adaptive platforms is their ability to provide understandable and consistent results that are easily interpreted by teachers. In this context, two indicators were used:

- Ped-Value – an expert pedagogical assessment of the interpretability of the results (0–1, normalized). A high Ped-Value indicates the ease of perception and the possibility of direct use of the model's conclusions in the educational process.
- Cohen's Kappa – a coefficient of consistency of the results between model predictions and teachers' expert assessments. A value of >0.75 is interpreted as high consistency, 0.4–0.75 – as moderate, <0.4 – as low.

The results show (Table 4) that Hybrid Cognitive-oriented systems that combine algorithmic calculations with expert rules demonstrated the greatest pedagogical value. They achieved a Ped-Value of 0.85 with high consistency with teachers' opinions (Kappa = 0.79). DL-based models showed high accuracy and acceptable consistency values (Kappa = 0.72), but were inferior in interpretability (Ped-Value ≈ 0.74). In contrast, ML-based algorithms, despite their stability and simplicity, received the lowest scores in terms of pedagogical usefulness (Ped-Value = 0.68; Kappa = 0.66).

*Table 4: Level of consistency and interpretability of models*

| Configuration | Ped-Value | Cohen's Kappa | Interpretation of results |
|---|---|---|---|
| Machine Learning-based | 0.68 | 0.66 | Moderate consistency, basic interpretability |
| Deep Learning-based | 0.74 | 0.72 | High precision, medium level of interpretability |
| Hybrid Cognitive-oriented | 0.85 | 0.79 | High consistency and pedagogical value |

Source: developed by the authors based on the results of their own research

The analysis showed that hybrid cognitive-oriented platforms provide the highest pedagogical value due to the combination of algorithmic power and expert rules, which increases the trust and clarity of the results. Deep models showed high accuracy, but their limited transparency complicates pedagogical interpretation. ML, although convenient for basic monitoring, demonstrated lower personalization and consistency with expert assessments. Therefore, it is hybrid solutions that provide the optimal balance of accuracy and interpretability, making them the most appropriate for adaptive language platforms in non-linguistic HEIs.

## 5.4 Stability of model training

In addition to accuracy and interpretability, an important indicator of the effectiveness of adaptive platforms is the stability of model training, which reflects the ability of algorithms to consistently update parameters during training. Two indicators were used for this purpose: ΔW-index – the average rate of change of weight coefficients between epochs, which characterizes the fluctuations and consistency of the learning process, and adaptation time (s/epoch) – the average parameter update time during one epoch. The analysis showed (Table 5) that ML-based algorithms had the lowest weight fluctuations (ΔW-index = 0.091) and short adaptation time (≈0.84 s/epoch), which makes them convenient for quick monitoring. DL-based models were characterized by greater variability (ΔW-index = 0.134) and longer adaptation time (≈1.42 s/epoch), which is due to a more complex architecture and higher resource costs. The Hybrid Cognitive-oriented systems (ΔW-index = 0.158) showed the highest variability, where the combination of deep models with rule-based approaches provides flexibility, but requires significantly more time for adaptation (≈1.97 s/epoch).

*Table 5: Average values of ΔW-index and adaptation time*

| Configu-ration | ΔW-index (average) | Adaptation time, c/epoch | Interpretation |
|---|---|---|---|
| ML-based | 0.091 | 0.84 | High stability, minimal time consumption |
| DL-based | 0.134 | 1.42 | Moderate fluctuations, acceptable adaptation time |
| Hybrid Cognitive-oriented | 0.158 | 1.97 | Highest variability, significant computational cost |

Source: developed by the authors based on the results of their own research

The chart below (Figure 2) shows the change in parameters W1 and W2 for three groups of models: ML-based, DL-based, and Hybrid Cognitive-oriented.

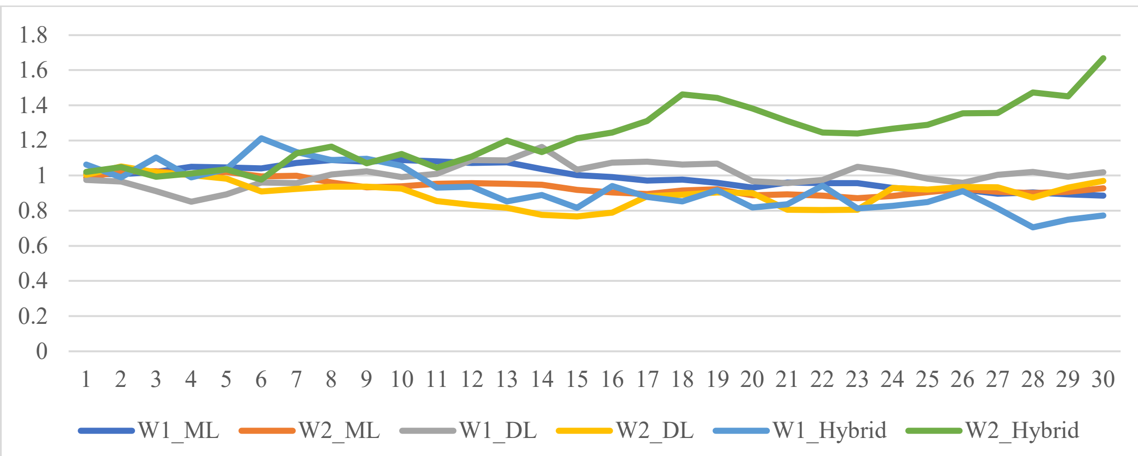


*Figure 2: Phase diagram of weight trajectories over 30 training epochs*
Source: developed by the authors based on the results of their own research.

Figure 2 shows that ML algorithms (W1_ML, W2_ML) exhibit the smallest fluctuations, quickly reaching stability. Deep models (W1_DL, W2_DL) are characterized by a significantly higher amplitude of changes, especially in the second half of training, which indicates greater sensitivity to data and a longer adaptation period. Hybrid systems (W1_Hybrid, W2_Hybrid) revealed the most pronounced variability: their trajectories fluctuate more intensively, reflecting the balance between flexibility of adaptation and increased computational costs. In general, the stability of models decreases in the direction from ML to Hybrid, while hybrid solutions provide the greatest potential for flexible response to changes in training data, but require more computational power.

### 5.5 Statistical confirmation of differences

The hypothesis of statistically significant differences in the performance of the tested platforms was tested by using ANOVA with repeated measures. The obtained results indicate a significant difference between the groups ($p < 0.01$), which gives grounds to reject the null hypothesis of the same distribution of LTEI indicators. The sources of differences were identified using paired t-tests with Bonferroni correction. The analysis showed that Hybrid Cognitive-oriented systems significantly outperform both other groups, while

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the difference between ML-based and DL-based models was less pronounced and in some cases statistically insignificant (Table 6).

*Table 6: Results of statistical testing of differences between groups*

| Comparison | Criterion | Statistics (F/t) | p-value | Bonferroni correction | Interpretation |
|---|---|---|---|---|---|
| ML vs DL vs Hybrid | ANOVA | F = 9.87 | 0.0004 | – | There are statistically significant differences |
| Hybrid vs DL | t-test | t = 3.14 | 0.002 | 0.006 | Hybrid > DL |
| Hybrid vs ML | t- test | t = 3.68 | 0.001 | 0.003 | Hybrid > ML |
| DL vs ML | t- test | t = 1.42 | 0.078 | 0.078 | The difference is not statistically significant |

Source: developed by the authors based on the results of their own research

Therefore, statistical tests confirmed the superiority of hybrid platforms, which provide the best balance between accuracy, stability, and pedagogical interpretability. The difference between classical ML algorithms and DL models turned out to be insignificant, which indicates the feasibility of their application depending on resource conditions and teaching tasks.

### 5.6 General analysis and practical conclusions

The obtained results showed that the effectiveness of adaptive educational platforms depends on their architecture and application conditions. ML-based solutions provided the lowest fluctuations in weight coefficients (ΔW-index = 0.091) and the shortest adaptation time (≈0.84 s/epoch). Their values of the integral LTEI ranged from 0.72 to 0.76, which indicates stability and speed, but at the same time limitations in deep personalization. Such platforms are advisable to use for regular monitoring of semester performance and prompt detection of deviations in learning. DL-based systems demonstrated the highest accuracy of predictions (F0.5-score ≈ 0.81; AUC ≈ 0.87), but their stability was lower (ΔW-index = 0.134; adaptation time ≈1.42 s/epoch). The LTEI values for this group were 0.78–0.81, which confirms their superiority in complex data processing tasks – texts, oral speech and LMS activity. At the same time, the level of interpretability remained average (Ped-Value ≈ 0.74), which may complicate the pedagogical use of the results. So, DL platforms are most appropriate for large universities with powerful technical resources, where the accuracy of analysis is critically important. The Hybrid Cognitive-oriented systems demonstrated the best results, as they combine deep algorithms with rule-based approaches. Their integral indicator (LTEI) reached 0.85–0.87, Ped-Value was 0.85, and Cohen’s Kappa was 0.79, which corresponds to high consistency with the expert assessments of teachers. Although these systems had the largest weight fluctuations (ΔW-index = 0.158) and the longest adaptation time (≈1.97 s/epoch), they provided the optimal balance between accuracy, flexibility, and pedagogical clarity. Their effectiveness is especially pronounced when working with incomplete and mixed data, making them the best choice for non-linguistic HEIs.

In general, it can be noted that ML platforms (IEMP 0.72–0.76) are most suitable for rapid monitoring and basic diagnostics of success. DL platforms (LTEI 0.78–0.81) are advisable to use for high-precision tasks and analysis of complex language data. Hybrid platforms (LTEI 0.85–0.87) provide systemic personalization and pedagogical interpretation, which makes them the best option for integration into language training of non-linguistic majors. Using a combined strategy will allow HEIs to achieve maximum efficiency: ML – for daily monitoring, DL – for in-depth analysis, Hybrid – for complex adaptive language education. The analysis demonstrated a significant advantage of hybrid systems over other approaches. At the same time, the difference between ML and DL platforms in a number of cases turned out to be statistically insignificant ($p\_adj > 0.05$). This means that both groups of models can be effective in different contexts. ML algorithms are more suitable for fast monitoring and basic diagnostic tasks, while DL approaches are justified in cases where maximum accuracy in processing complex language data is critical. This balance emphasizes that the choice of model should be made not only on the basis of accuracy, but also taking into account the resources and practical needs of the educational institution.

## 6. DISCUSSION

The obtained results confirm that the effectiveness of adaptive educational platforms is determined not only by the choice of algorithm, but also by the context of their application – the type of educational data, the level of their completeness and the dynamics of changes in the educational process, the quality of UX and the compliance of UI patterns with the HCI principles. The differences identified in the stability of predictions, the speed of adaptation, pedagogical interpretability, and convenience of user interaction (UX/UI) indicate that there is no universal solution: each group of models has its own strengths and weaknesses. In particular, ML configurations benefit from the latency of the HCI loop, DL models require thoughtful UI prompts to reduce cognitive load, and hybrid systems provide a balance of explainability and UX stability. These findings are consistent with earlier studies on the role of adaptive mechanisms in AI systems for education and emphasize the need to evaluate platforms taking into account HCI/UX metrics along with technical indicators [19, 22]. The analysis showed that Machine Learning-based platforms provide the smallest fluctuations in weight coefficients (ΔW-index = 0.091) and the shortest adaptation time (≈0.84 s/epoch). This makes them optimal for fast monitoring and detection of basic trends, but limits them in more complex personalized learning scenarios.

The analysis showed that ML-based platforms provide the smallest fluctuations in weight coefficients (ΔW-index = 0.091) and the shortest adaptation time (≈0.84 s/epoch). This makes them optimal for fast monitoring and detection of basic trends, but limits them in more complex personalized learning scenarios. Similar results are consistent with the conclusions about the suitability of classical algorithms for operational educational analytics [21, 28]. DL-based models demonstrated the highest accuracy (F0.5 ≈ 0.81; AUC ≈ 0.87), but required more resources and time for convergence (ΔW-index = 0.134; ≈1.42 s/epoch). At the same time, their interpretability remained at an medium level (Ped-Value ≈ 0.74), which partially reduces the practical value in the pedagogical environment. This is consistent with the findings of other researchers, which emphasize the high performance of DL architectures in working with complex language data, but limited transparency of the results [26, 30]. The most balanced were Hybrid Cognitive-oriented systems, which combine algorithmic calculations with pedagogical rules. They achieved the highest LTEI values (0.85–0.87) and pedagogical interpretability assessment (Ped-Value = 0.85; Kappa = 0.79), which confirms their ability to provide both accuracy and understandable interpretation for teachers. However, this balance was accompanied by higher adaptation time costs (≈1.97 s/epoch), which is consistent with international research on hybrid solutions as an optimal compromise between performance and explainability [27, 34]. Correlation analysis revealed a statistically significant inverse relationship between the frequency of changes in learning scenarios and the stability of models (ρ = –0.58…–0.25), which confirms the thesis about the need to control the dynamism of the educational environment to maintain the accuracy of predictions. Statistical tests (ANOVA, $p < 0.01$; paired t-tests) confirmed these results, demonstrating a significant advantage of hybrid platforms over the other two groups. So, the study confirms that the choice of platform type should be based not only on technical characteristics, but also on the educational context. ML approaches are most suitable for rapid monitoring; DL models – for the analysis of complex language data in large samples; hybrid systems – for complex personalization and pedagogical interpretation. This creates the basis for the flexible implementation of adaptive platforms in non-linguistic HEIs, where not only accuracy is important, but also teachers' trust in the system.

The findings of the presented study can be interpreted critically based on the obtained indicators and the existing literature on AI-based language learning. The results show that the cognitive-oriented approach of the hybrid model achieved the best results in terms of accuracy (LTEI 0.85–0.87; Ped-Value 0.85; Kappa 0.79). However, this result was achieved due to the rule-based components included into the model, which emphasizes the interpretability of the model rather than its predictive performance. The stability results require some caution. The outcomes of the stability indicators suggest that the hybrid model has a higher ΔW-index and a longer adaptation time for the users. This finding indicates that the increased interpretability of the model comes with increased volatility and computational costs. Furthermore, the quantitative difference between the ML and DL models was not statistically significant. This result suggests that the choice of model is more related to available resources and latency constraints in the HCI context than to the model's predictive performance.

In comparison with existing studies, one key difference in this research is the focus on the relationship between the model's predictive, stability, and interpretability properties and the resulting UX and HCI implications. Most of the existing studies focus on one of the following topics:

the performance of specific AI tools in language learning, the conceptual discussion of adaptive learning environments, or the observation of UX aspects of such environments. However, the current study is unique in that it proposes a measurement model that allows for the comparison of the three types of models within a single educational setting. The inclusion of the LTEI indicator with a Ped-Value component, the stability of the weighting of the indicator terms as validated by the bootstrap, and the inclusion of UX and HCI indicators in the evaluation model all contribute to the extension of previous studies and findings.

When compared to the most recent 2024 literature, the present findings help to confirm the general direction of the results and point to the specific mechanism that explains the increased implementation readiness of certain AI architectures in non-linguistic higher education institutions (HEIs). For instance, the studies by Er-Rafyg et al. [22] and Fakhar et al. [23] confirm that AI-enabled adaptive solutions can indeed be effective in the context of higher education. However, the studies also reveal the various constraints that may impede the implementation of such technologies. Similar results are drawn by Xia et al. [21]. However, as will be described below, the experiments that were performed revealed that while the three platform types achieved acceptable levels of predictive performance, the usefulness of certain implementations over others became evident. The results showed that the DL-based models achieved the highest accuracy (higher F0.5/AUC). This is consistent with the findings of other recent studies, such as Mustafa et al. [15] and Zhou and Li [11], which also indicated the potential of AI-enabled adaptive solutions in higher education. However, the lower stability of these models (higher ΔW-index) indicates that while they may be accurate, they are less likely to be immediately helpful to teachers. In contrast, the results also indicated that the Hybrid cognitive-oriented models achieved the best overall balance. The integration of the cognitive model that enabled the system to provide interpretable, teacher-facing explanations led to higher interpretability and higher agreement with expert judgments (Ped-Value and Cohen's Kappa), even if the higher costs associated with using this approach. This is consistent with the findings of Ezzaim et al. [17], who experimentally confirmed that implementing an AI-based adaptive solution in such a way had a positive impact on the evaluation of higher education learning environments. However, they did not compare different architectures (ML, DL, and hybrid) or investigate the extent to which the proposed model included considerations of HCI and interpretability. Finally, Vistorte et al. [30] also provide information that is crucial to understanding the findings of this study. Their research reveals that aspects such as human interaction and user experience are crucial to the effective use of AI in educational environments. As such, the inclusion of indicators related to UX and HCI (including the Ped-Value component) confirms the importance of such aspects and reveals that interpretability has a direct and measurable impact on the ranking of those platform configurations.

In order to explicitly address each of the proposed objectives, a critical analysis of the achieved results is provided below. Regarding the first objective, which sought to analyse the architectural features of the integration of AI-based platforms into the foreign language teaching in non-linguistic HEIs, it became evident from the analysis that the different architectural features of the AI systems led to different features and possibilities in the language teaching process. For instance, the models based on ML ensured stability and fast monitoring of the AI-based language teaching platform, while the models based on DL maximized the prediction of student success. Finally, the results also indicated that the architectural feature of a system is essential to the language teaching process, as no specific system architecture is universally best for all language teaching environments. Concerning the second objective, which aimed to experimentally compare the three types of AI models in terms of their ability to predict student success and difficulties in foreign language learning, the results indicated that the configurations of the Hybrid model were the most significant in terms of the LTEI index score. The comparison of ML and DL models, in contrast, did not show statistically significant differences. This suggests that the differences in performance between the model types are not easily generalizable to other language learning environments. With respect to the third objective, which concerned the development of the LTEI index itself, the study was successful in developing and validating the index. The use of the bootstrap method indicated that the coefficients of the weights of the index components were stable. However, a critical reflection of the index also indicates that the weight of the Ped-Value component is subject to a certain degree of subjectivity. Future research in this area can focus on finding other ways to weight the components of this index. Beyond the objectives that had been proposed, there are also some issues that emerged as unresolved and require further investigation. For example, it was not possible to assess the long-term impact of such an approach on foreign language proficiency. Additionally, ethical issues need to be addressed. Finally, a balance between the sophistication of the AI recommendations and the time required for interaction between the student and the AI indicates that future research in this area should consider these factors.

## 7. LIMITATIONS

The results of this study should be interpreted with certain limitations in mind.

1) The analysis covered only three groups of adaptive platforms (ML-based, DL-based, and Hybrid Cognitive-oriented), which narrows the range of models to a limited set of architectures (Random Forest, kNN, Gradient Boosting, Bi-LSTM, Transformer-encoder, and rule-based approaches). The study did not test other promising solutions, including next-generation transformers, graph neural networks, reinforcement learning methods, or modern optimizers (e.g., AdamW), which limits the generalizability of the findings.

2) The sample included only those educational platforms that have integrated tools for personalizing language training and are actively used in international practice. So, the results cannot be fully transferred to all types of educational systems, especially to specialized or narrowly focused solutions that do not support adaptive algorithms.

3) The integrated LTEI was created based on a uniform combination of technical and pedagogical metrics. Although this allows for a comprehensive assessment of the platforms, some subjectivity remains in the Ped-Value component, which is based on teachers' expert assessment. Alternative weighting schemes for indicators could change the final results of the platform ranking.

4) The experiments were conducted in a unified hardware environment with fixed training parameters (number of epochs, batch-size), which does not cover all possible variants of the platforms' operation in real conditions. In practice, differences in hardware or settings can affect the stability and speed of algorithm adaptation.

5) Statistical tests (ANOVA, t-tests with Bonferroni correction) were performed for a limited number of platforms and scenarios, which reduces the sensitivity to detecting weak effects. Although the results confirmed significant differences between groups, the scale of the platform sample does not allow conclusions to be drawn about the entire ecosystem of adaptive educational solutions.

6) The study focused on analytical metrics (accuracy, stability, interpretability) and did not include direct measurement of pedagogical impact – changes in student performance, reduced teacher workload, or timeliness of instructional interventions. It also failed to address the ethical and legal implications of using personal data in such systems. Further research should expand the sample of platforms, include new algorithm architectures, test other weighting schemes in the LTEI formula, and integrate analysis of direct pedagogical impact and ethical responsibilities in the use of adaptive solutions.

## 8. CONCLUSIONS

The study addressed its main aim, to assess the effectiveness of AI-based adaptive educational platforms for foreign language teaching in non-linguistic HEIs, by jointly considering the three main dimensions: predictive quality, stability, and interpretability. The achievement of the first objective led to the confirmation of the impact of platform architecture on its effectiveness. Specifically, the results showed that systems based on machine learning (ML) are advantageous in scenarios requiring low latency, deep learning (DL)–based systems are appropriate when dealing with complex linguistic artefacts, and hybrid cognitive-oriented systems are best suited when teacher interpretability is an essential feature. The achievement of the second objective provided insights into the relative effectiveness of ML, DL, and hybrid solutions. The results of the experiments made it possible to conclude that all three configurations achieved acceptable levels of success in the prediction task, but the usability of these platforms for teaching purposes differed. Finally, the achievement of the third objective led to the development and validation of the LTEI index.

The results of the experimentation revealed that the most effective architecture was that of a hybrid cognitive-oriented system. The results showed that such systems achieved the highest values for interpretability and agreement with expert assessments while also achieving high levels of integrated effectiveness. A trade-off that emerged from this analysis, however, was that the architecture also had the highest training variability and adaptation time. Thus, it is most suitable for environments in which interpretability and decision support from teachers are required. The results also showed that the DL-based models achieved the highest accuracy indicators. However, they exhibited relatively low levels of stability and interpretability. The ML-based systems were the most stable and fast. However, their lower level of LTEI and interpretability indicates that these systems have limitations in terms of deep personalization of educational experiences for learners. The results of the statistical testing confirmed that the three groups differed significantly (ANOVA F = 9.87, p = 0.0004). The configuration of the groups with a hybrid architecture outperformed the configurations with deep learning (p_adj = 0.006) and machine learning (p_adj = 0.003). However, the difference in effectiveness between the DL and ML configurations was statistically insignificant (p_adj = 0.078). This finding suggests that the decision between using models based on machine learning and deep learning is primarily driven by the available resources, while the advantage of the hybrid architecture is decisive when such interpretability and decision support from teachers are required.

Although the study achieved a lot, it is also important to discuss the limitations of the work. One limitation of the LTEI index is that the component based on Ped-Value is partially dependent on the expert judgments of the individuals calculating it. The finding that the systems of the hybrid cognitive-oriented architecture are the most effective should not be generalized to the entire class of such platforms. The outcome is dependent on the specific conditions under which they were measured, and it can be generalized with caution to other contexts. Another limitation of this study is that while the user experience and human computer interaction indicators were used, metrics such as learning gains over multiple semesters, teacher workload, and intervention timeliness were outside the scope of this study.

The implications of the findings are that ML-based AI platforms are most appropriate for situations involving continuous monitoring and when low latency is required. For other situations, such as when the accuracy required of an adaptive educational platform is of the highest order, a DL-based configuration may be appropriate. However, for HEIs that need to provide stable, interpretable, and actionable decision support to foreign language teachers, the choice of a hybrid cognitive-oriented platform is the most appropriate. Thus, this study presents a framework that can be used to compare the effectiveness of different architectures in terms of the LTEI index, the stability measure indicated by the ΔW-index, and the UX/HCI indicators. Furthermore, the study also discusses the trade-offs and open issues that must be addressed in future studies to improve the transferability and effectiveness of these approaches in foreign language teaching in non-linguistic HEIs.

From the authors' perspective, the main value that this work provides to higher education institutions is in providing a decision logic. The authors consider that using a hybrid model that is cognitive-oriented is the best approach to implementing these methods in non-linguistic higher education institutions. The authors also consider that using machine learning models is the most rational solution for monitoring, and that using deep learning models is the best approach for institutions that can afford the computational costs and require high accuracy in recognizing complex linguistic elements. In the authors' opinion, the best way to implement these methods is to consider the nature of the scenario in which the models will be used, as well as the requirements that it places upon the user experience and interaction design.

## REFERENCES:

[1] I. I. Kharchenko, "Application of adaptive educational platforms in the process of teaching professional language", *Pedagogical Academy: scientific notes*, Vol. 18, 2025. https://doi.org/10.5281/zenodo.15476679

[2] N. Scherba, "Digital technology systematization in the formation of information and digital competence of future teachers of English", *Science and Technology Today*, Vol. 44, 2025, pp. 811-821. https://doi.org/10.52058/2786-6025-2025-3(44)-811-821

[3] C. Yang, J. Chen and D. Zou, "Artificial Intelligence in Interpreting Education Curriculum: A Delphi Study for Interpreter Competencies", *International Journal of Education and Humanities*, Vol. 5, No. 3, 2025, pp. 387-403.

[4] V. Amaro and X. Zhang, "Intercultural Interfaces: Artificial intelligence and the challenges of cultural sensitivity", *E-Revista de Estudos Interculturais*, Vol. 13, 2025. https://doi.org/10.34630/e-rei.vi13.6325

[5] L. Liao, "Artificial intelligence-based English vocabulary test research using log analysis with virtual reality assistance", *Computer-Aided Design and Applications*, Vol. 20, No. S9, 2023, pp. 23-39. https://doi.org/10.14733/cadaps.2023.S9.23-39

[6] T. N. Nguyen Q. and T. T. Nguyen, "AI literacy in language learning through digital multimodal composing: a systematic review", *VNU Journal of Foreign Studies*, Vol. 41,

No. 3, 2025, pp. 142-156. https://doi.org/10.63023/2525-2445/jfs.ulis.5517

[7] M. J. Shodiq and R. A. N. Ziihah, “Development of Arabic Language Learning Media with The Content-Based Instruction for Students with Disabilities in Inclusive Schools”, *Arabi: Journal of Arabic Studies*, Vol. 9, No. 2, 2024, pp. 153-164.

[8] M. Yu, “Application of an Artificial Intelligence-based Adaptive Learning System to Chinese Language Education in Universities”, In *Proceedings of the 2024 International Symposium on Artificial Intelligence for Education* (pp. 416-421), 2024. https://doi.org/10.1145/3700297.3700368

[9] S. P. J. Chiriboga, A. L. T. Burgos, M. M. R. Avila, J. L. C. Chida, K. J. A. Macias, Y. E. C. Morante and A. P. B. Párraga, “Artificial Intelligence and Personalized Learning in Foreign Languages: An Analysis of Chatbots and Virtual Assistants in Education”, *Revista Científica de Salud y Desarrollo Humano*, Vol. 6, No. 1, 2025, pp. 882-905. https://doi.org/10.61368/r.s.d.h.v6i1.515

[10] O. O. Ayotunde, D. I. Jamil and N. Cavus, “The impact of artificial intelligence in foreign language learning using learning management systems: a systematic literature review”, *Information Technologies and Learning Tools*, Vol. 95, No. 3, 2023, p. 215. https://doi.org/10.33407/itlt.v95i3.5233

[11] Z. Zhou and W. Li, “Personalized college English learning based on artificial intelligence: Algorithm-driven adaptive learning method”, *International Journal of High Speed Electronics and Systems,* 2024, art. no. 2540102. https://doi.org/10.1142/S0129156425401020

[12] A. P. Kumar, G. Fatma, S. Sarwar and K. S. Punithaasree, “Adaptive Learning Systems for English Language Education based on AI-Driven System”, In *2025 International Conference on Intelligent Systems and Computational Networks (ICISCN)* (pp. 1-5). IEEE, 2025. https://doi.org/10.1109/ICISCN64258.2025.10934601

[13] X. Shu and C. Xu, “Artificial Intelligence-Based English Self-Learning Effect Evaluation and Adaptive Influencing Factors Analysis”, *Mathematical Problems in Engineering*, Vol. 2022, No. 1, 2022, art. no. 2776823. https://doi.org/10.1155/2022/2776823

[14] M. G. Yaseen and S. S. Alnakeeb, “Exploring the evolution of ai integration in English as a foreign language education: A Scopus-based bibliometric analysis (1997-2023)”, *Mesopotamian Journal of Computer Science*, Vol. 2023, 2023, pp. 143-158. https://doi.org/10.58496/MJCSC/2023/019

[15] G. Mustafa, T. Urooj and M. Aslam, “Role of Artificial Intelligence for adaptive learning environments in higher education by 2030”, *Journal of Social Research Development*, Vol. 5, No. 3, 2024. https://doi.org/10.53664/JSRD/05-03-2024-02-12-22

[16] M. Kolhar and A. Alameen, “Artificial Intelligence Based Language Translation Platform”, *Intelligent Automation & Soft Computing*, Vol. 28, No. 1, 2021. http://doi.org/10.32604/iasc.2021.014995

[17] A. Ezzaim, A. Dahbi, A. Haidine and A. Aqqal, “The Impact of Implementing a Moodle Plug-in as an AI-based Adaptive Learning Solution on Learning Effectiveness: Case of Morocco”, *International Journal of Interactive Mobile Technologies*, Vol. 18, No. 1, 2024. https://doi.org/10.3991/ijim.v18i01.46309

[18] X. Fu, K. Lokesh Krishna and R. Sabitha, “Artificial intelligence applications with e-learning system for China’s higher education platform”, *Journal of Interconnection Networks*, Vol. 22, No. (Supp02), 2022, art. no. 2143016. https://doi.org/10.1142/S0219265921430167

[19] T. Sargazi Moghadam, A. Darejeh, M. Delaramifar and S. Mashayekh, “Toward an artificial intelligence-based decision framework for developing adaptive e-learning systems to impact learners’ emotions”, *Interactive Learning Environments*, Vol. 32, No. 7, 2024, pp. 3665-3685. https://doi.org/10.1080/10494820.2023.2188398

[20] A. Alam, “Employing adaptive learning and intelligent tutoring robots for virtual classrooms and smart campuses: reforming education in the age of artificial intelligence”, In *Advanced computing and intelligent technologies: Proceedings of ICACIT 2022* (pp. 395-406). Singapore: Springer Nature Singapore, 2022. https://doi.org/10.1007/978-981-19-2980-9_32

[21] J. Xia, Y. Ge, Z. Shen and Rahman M. Najar, “The auxiliary role of artificial intelligence

applications in mitigating the linguistic, psychological, and educational challenges of teaching and learning chinese language by non-chinese students", *International Review of Research in Open and Distributed Learning*, Vol. 25, No. 3, 2024, pp. 116-133. https://doi.org/10.19173/irrodl.v25i3.7680

[22] Er-A. Rafyg, H. Zankadi and A. Idrissi, "AI in adaptive learning: Challenges and opportunities", *Modern artificial intelligence and data science 2024: Tools, techniques and systems*, Vol. 1116, 2024, pp. 329-342. https://doi.org/10.1007/978-3-031-65038-3_26

[23] H. Fakhar, M. Lamrabet, N. Echantoufi and L. Ajana, "Towards a New Artificial Intelligence-based Framework for Teachers' Online Continuous Professional Development Programs: Systematic Review", *International Journal of Advanced Computer Science & Applications*, Vol. 15, No. 4, 2024. https://doi.org/10.14569/ijacsa.2024.0150450

[24] L. Jia, B. Santhosh Kumar and R. Parthasarathy, "Research and application of artificial intelligence based integrated teaching-learning modular approach in colleges and universities", *Journal of Interconnection Networks*, Vol. 22, No. (Supp02), 2022, art. no. 2143006. https://doi.org/10.1142/S0219265921430064

[25] A. Ezzaim, A. Dahbi, A. Haidine and A. Aqqal, "AI-based adaptive learning: A systematic mapping of the literature", *Journal of Universal Computer Science*, Vol. 29, No. 10, 2023, p. 1161. https://doi.org/10.3897/jucs.90528

[26] J. Tang, "Artificial intelligence-based needs analysis for english specific purposes in digital environment", *Learning and Motivation*, Vol. 83, 2023, art. no. 101914. https://doi.org/10.1016/j.lmot.2023.101914

[27] M. R. Ramesh, "Challenges of Integrating Artificial Intelligence With Language Curriculum: Addressing Pedagogical, Technological, and Ethical Barriers", In *Modern Methods for AI-Integrated Language Curriculum* (pp. 205-230). IGI Global Scientific Publishing, 2025. https://doi.org/10.4018/979-8-3693-9606-3.ch008

[28] M. Kaouni, F. Lakrami and O. Labouidya, "The use of Artificial Intelligence and Natural language processing In E-learning Platforms: An Exploration of Potentials and Limitations", *International Journal of Information Science and Technology*, Vol. 8, No. 4, 2024, pp. 1-9. http://doi.org/10.57675/IMIST.PRSM/ijist-v8i4.258

[29] S. Shoukat, R. Mamoon and M. F. Arif, "Enhancing language proficiency through TPACK model and AI applications a study on effective integration strategies in english language instruction", *Pakistan Languages and Humanities Review,* Vol. 8, No. 2, 2024, pp. 540-554. https://doi.org/10.47205/plhr.2024(8-II)47

[30] A. O. R. Vistorte, A. Deroncele-Acosta, J. L. M. Ayala, A. Barrasa, C. López-Granero and M. Martí-González, "Integrating artificial intelligence to assess emotions in learning environments: a systematic literature review", *Frontiers in psychology*, Vol. 15, 2024, art. no. 1387089. https://doi.org/10.3389/fpsyg.2024.1387089

[31] K. H. Tang D., "Implications of artificial intelligence for teaching and learning. *Acta Pedagogia Asiana,* Vol. 3, No. 2, 2024, pp. 65-79. https://doi.org/10.53623/apga.v3i2.404

[32] R. F. Mahafdah, S. Bouallegue and R. Bouallegue, "Applying Artificial Intelligence in the E-Learning Field", In *International Conference on Advanced Information Networking and Applications* (pp. 392-403). Cham: Springer Nature Switzerland, 2024. https://doi.org/10.1007/978-3-031-57931-8_38

[33] J. Liu, X. Bao and L. Chen, "Artificial intelligence in educational technology and transformative approaches to English language using fuzzy framework with CRITIC-TOPSIS method", *Scientific Reports*, Vol. 15, No. 1, 2025, p. 25542. https://doi.org/10.1038/s41598-025-09844-9

[34] R. S. Ullah, F. Hashim, M. M. Bandeali and A. Akbar, "Artificial Intelligence in Curriculum Design a Roadmap for Adaptive and Personalized Learning in Higher Education", *The Critical Review of Social Sciences Studies*, Vol. 3, No. 3, 2025, pp. 304-322. https://doi.org/10.59075/8j71b437

[35] R. Khatoon, "AI-Powered Task-Based Learning for Cross-Border Higher Education", In *Bridging Global Divides for Transnational Higher Education in the AI Era* (pp. 141-164). IGI Global Scientific Publishing, 2025. https://doi.org/10.4018/979-8-3693-7016-2.ch007