Journal of Theoretical and Applied Information Technology
30th April 2026. Vol.104. No.8

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

DIGITAL TECHNOLOGIES AS A TOOL FOR ADAPTIVE TEACHING OF FOREIGN LANGUAGES IN HIGHER SCHOOL BASED ON ARTIFICIAL INTELLIGENCE

OLHA TAMARKINA[1], LIUDMYLA BAIDAK[2], HANNA TSYHANOK[3], NATALIIA SHCHUR[4], ANNA ZINCHENKO[5]

[1]PhD in Pedagogical Sciences, Associate Professor, Sumy National Agrarian University, Foreign Languages Department, Sumy, Ukraine

[2]Senior Lecturer, Sumy National Agrarian University, Foreign Languages Department, Sumy, Ukraine

[3]PhD in Philology, Associate Professor, Sumy National Agrarian University, Foreign Languages Department, Sumy, Ukraine

[4]PhD in Education, Associate Professor, Ternopil Ivan Puluj National Technical University, Ukrainian and Foreign Languages Department, Ternopil, Ukraine

[5]PhD in Pedagogical Sciences, Senior Lecturer, Bohdan Khmelnytsky National University of Cherkasy, Foreign Languages Department, Cherkasy, Ukraine

E-mail: [1]olhatamarkina@gmail.com, [2]baidakliudmyla118@gmail.com, [3]hannatsyhanna@gmail.com, [4]nataliiaszczur42@gail.com, [5]annazincenkoanna@gmail.com

ABSTRACT

Current digital transformation of education necessitates the creation of intelligent solutions capable of ensuring personalized and effective learning of foreign languages in higher school. The relevance of the study is determined by the need to introduce adaptive digital technologies based on artificial intelligence (AI) into the teaching of foreign languages in higher school to personalize learning and increase its efficiency. The aim of the study is to clarify the effectiveness of the use of AI-based digital technologies as a tool for adaptive teaching of foreign languages in higher education institutions (HEIs). The object of analysis was the architecture of adaptive educational systems that integrate machine learning (ML), deep neural networks (DNNs), deep learning (DL), natural language processing (NLP), and intelligent analytics algorithms to form dynamic learning trajectories. The methodology included modelling adaptive learning scenarios in Moodle and OpenEdX environments using AI plugins, evaluation by the integral IAAL index (Acc, Stab, Latency, CogFit), as well as statistical methods (analysis of variance (ANOVA), bootstrapping, receiver operating characteristic (ROC)/ area under the curve (AUC), principal component analysis (PCA), k-means). The international perspective was provided by analysing the practices of using digital adaptive learning algorithms (ML, DL, NLP modules) in three countries — Ukraine, Germany, and Poland. The analysis identified differences in the performance of models in environments with different levels of digital maturity. The results showed that the hybrid Fusion-NLP configuration achieved the highest integral Adaptive Academic Learning Index (AALI) values (0.82–0.87), ensuring stability and cognitive relevance of learning tasks. Deep-Adapt demonstrated maximum accuracy in Germany (Acc = 0.89; AUC = 0.93), while Stat-Learn confirmed its superiority in speed in Poland and Germany (Latency = 0.92–0.93). It was found that the effectiveness of the models depends on the educational context. Fusion-NLP is the most effective among the environments with fragmented data. Deep-Adapt is the most effective among mature systems with complete data sets. Stat-Learn is optimal for rapid diagnostics in transitional infrastructures. The academic novelty is the comprehensive comparison of the three architectures using the AALI and multidimensional methods, which allowed us to determine the optimal application scenarios in different countries. Further research prospects include improving AALI through cognitive indicators, expanding the international sample, and developing AI-based monitoring dashboards for teachers, which will facilitate the practical integration of adaptive technologies into language teaching.

Keywords: *Artificial Intelligence, Digital Technologies, Adaptive Learning, NLP, Machine Learning, Higher Education, International Perspective*

1. INTRODUCTION

The rapid development of digital technologies and AI is radically transforming approaches to teaching foreign languages in higher school. Systems with elements of adaptive learning open up opportunities for building personalized educational trajectories that can take into account students' prior knowledge, motivational profile, and cognitive characteristics [1; 2]. Unlike traditional methods, AI platforms allow analysing large amounts of educational data in real time, predicting individual needs, and offering optimal educational content [3; 4].

According to international research, the use of AI in language education improves academic achievement, contributing to the development of intercultural communication skills [5; 6]. The use of adaptive models in the context of teaching English, German, and other European languages, where both communicative and technical requirements are combined, is of particular relevance [7; 8]. The practice of Ukraine [9], Germany [10] and Poland [11] shows that digital educational environments with AI modules become the basis for building sustainable educational ecosystems focused on efficiency and flexibility. At the same time, despite numerous experiments, the issues of choosing the optimal personalization algorithms and their scaling in university systems remain open. Despite the growing number of research publications on AI in education, most of them focus on either language learning specifically, or on the digital transformation of higher education in general. There is a lack of research into which type of AI learning algorithm works best for teaching foreign languages in different countries with different levels of digital maturity in their education systems. Thus, a study that performs a comparison of countries and discovers the most effective methods of AI for language learning in each country is needed. The challenges relate to both computational efficiency and the models adaptability to incomplete or heterogeneous data, which is especially relevant in transnational contexts [11; 12]. In such conditions, a combination of ML methods, DNNs, and hybrid architectures can provide a balance between prediction accuracy, response speed, and interpretability of teaching guidelines [13; 14].

The aim of this study is to determine the effectiveness of digital AI-based technologies as a tool for adaptive foreign language learning in higher school. Special emphasis is placed on comparing educational practices in Ukraine, Germany, and Poland to identify common trends and national differences in the implementation of AI approaches in language teaching.

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

1. Analyse the architectural principles of integrating AI modules into digital educational environments for foreign language teaching.

2. Test adaptive learning models (ML, DL, hybrid solutions) using university courses in English and German.

3. Develop integrated criteria for assessing the effectiveness of systems (prediction accuracy, speed, resistance to incomplete data, interpretability).

4. Compare the results in three educational contexts (Ukraine, Germany, Poland) and provide practical recommendations for higher education.

The significance of this study in relation to the current state of the art is represented by the fact that this study evaluates three different models of adaptive AI systems, comparing not just the technical elements of the systems, but also the elements related to the education system itself. Most previous research in this area has been limited to evaluating single tools, language tasks, or studies that were limited to the educational level of a single country.

So, the study aims to substantiate the appropriateness of implementing AI-driven solutions in language education and identify prospects for further integration of adaptive digital technologies in the international educational space. The countries selected for analysis – Ukraine, Germany, and Poland – reflect different stages of digital maturity of educational systems: from an environment in urgent need of modernization (Ukraine) to a mature European infrastructure (Germany) and a transitional model with integration into European standards (Poland). Such a combination allows us to assess the effectiveness of

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

AI-driven adaptive technologies in contrasting socio-economic conditions and provides grounds for cautious extrapolation of the results to other European and global contexts with similar characteristics.

2. LITERATURE REVIEW

Recent years have been characterized by a rapid growth of interest in the implementation of AI in adaptive foreign language learning, but academic studies show both points of contact and significant differences in the interpretation of its potential. The authors [15] considered adaptability in their systematic mapping primarily through the individualization of trajectories and increased student involvement. This has something in common with the findings of the scientists [18], who emphasize personalization based on learning styles. The study [20] also demonstrated similar attention to the individualized approach, who single out personalization as the main trend of modern research in the field of AI technologies. So, these authors emphasize the benefits of AI for increasing student efficiency and motivation, but leave out the issue of institutional readiness. In contrast, the authors [16] shift the focus from the level of student interaction to the level of educational systems, emphasizing the need to transform universities through the integration of AI, ML, and Extended Reality (XR). Their position partly coincides with the view [25], which considers AI as a catalyst for strategic changes in the management of institutions. The authors [17] demonstrate a more pragmatic approach, analysing how translation practices in the context of AI become a tool for the development of communicative and intercultural competencies. This creates a certain dichotomy: some authors [15; 18; 20] focus on the micro-level of personalization, while others [16; 17; 25; 26] – on the macro-level of institutional changes.

Generative AI (GenAI) research further deepens this distinction. For example, the author [21] considers ChatGPT as a support for teaching practices. The researcher [22] emphasizes the potential of new platforms to create "smart learning," effectively moving from the instrumental role of AI to a conceptual rethinking of educational ecosystems. The scientists [23] and [24] develop this line, pointing to the integration of NLP modules, tutoring robots, and intelligent agents as the key to creating fully adaptive environments, but their approach may seem too technocratic and less focused on the pedagogical component. On the other hand, the study [19] demonstrates a more balanced perspective, combining the automation of routine tasks with the development of critical thinking in students, which makes his approach closer to educational practices than to purely technical models. Similarly, the authors [26-27] propose a vision where AI is part of intelligent learning support systems and adaptive testing, making their findings a kind of compromise between pedagogical value and technological functionality. A critical aspect in the literature is also the role of mobile solutions and regional context.

The authors [28] emphasize students' experiences in using mobile applications with integrated AI, while [29] emphasize the systematization of ML and AI approaches in digital education in general. In contrast, the authors [30] analyse practices in Latin America, demonstrating that implementation barriers (lack of resources, digital divide) can offset potential benefits, while most European and Asian studies tend to ignore these limitations. The integration of AI in foreign language teaching in higher school opens up opportunities for personalized learning, but the effectiveness of such systems depends on the quality and completeness of educational data [15; 18]. Research usually focuses on individual tasks, such as automated writing assessment or pronunciation training [17; 27], while comprehensive integration of adaptive mechanisms into the entire cycle of language education remains limited [20; 33]. At the same time, there are significant differences in the level of digital maturity between countries, which complicates the development of unified assessment criteria [3; 5]. Therefore, there is a need to create a methodology that would combine technical indicators (accuracy, speed, stability) with the pedagogical value of the results and allow comparing different algorithmic approaches in an international context.

3. PROPOSED METHODOLOGY

3.1 Research stages

The study lasted from November 2024 to April 2025 and included four consecutive stages. The first stage was a systematic review of academic literature for 2021–2025 on the application of digital technologies and AI models in adaptive foreign language learning [15–35]. The second stage (January 2025) involved creating an experimental environment that simulated the operation of the Moodle and OpenEdX platforms with integrated AI modules, in particular ML, AL, and hybrid algorithms involving NLP. The third stage (January–March 2025) involved conducting controlled experiments based on student data, which included academic records, activity in digital environments and results of standardized language tests. The fourth stage (March–April 2025) was aimed at data processing, calculation of efficiency indices, and an international perspective through statistical analysis and cross-country comparisons (Ukraine, Germany, Poland).

3.2 Research architecture and sampling

Three different architectural approaches to building adaptive foreign language teaching systems were tested in the study. The first configuration, Stat-Learn, was based on statistical ML models (Random Forest, SVM) and focused on the batch mode of processing educational data. The second, Deep-Adapt, implemented deep neural network architectures (LSTM, CNN), which support incremental learning and rapid adaptation to new educational materials. The third configuration, Fusion-NLP, combined DL with expert rules and an integrated NLP module for analysing text tasks and building dynamic learning trajectories. The main characteristics are given in Table 1. The choice of these three configurations was determined by the fact that they reflect different generations of development of adaptive systems: from classical statistical algorithms (Stat-Learn) through modern neural network solutions (Deep-Adapt) to hybrid integrated approaches involving NLP (Fusion-NLP). Unlike transformer models (BERT, GPT-like architectures), which have high computational complexity and require significant resources for training and adaptation, the selected configurations are more reproducible in the realities of university educational platforms. Thus, they allow testing the effectiveness of different adaptation paradigms while maintaining practical feasibility in higher education. The main characteristics are given in Table 1. Table 1. Configurations of adaptive foreign language teaching systems

Unlike transformer models (BERT, GPT-like architectures), which have high computational complexity and require significant resources for training and adaptation, the selected configurations are more reproducible in the realities of university educational platforms. So, they enable testing the effectiveness of different adaptation paradigms while maintaining practical feasibility in higher education. Table 1 presents the main characteristics.

Table 1: Configurations of adaptive foreign language teaching systems

Item No.	Configuration	Basic algorithms	Type of adaptation	Use of NLP	Level of integration with platforms
1	Stat-Learn	Random Forest, SVM	Package	No	High
2	Deep-Adapt	LSTM, CNN	Incremental	Partial	Medium
3	Fusion-NLP	DL + rule-based + NLP-module	Mixed	Full	High

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

The sample included three types of data:

1) students' academic records (success and attendance);
2) activity in Moodle and OpenEdX (number of activities, completed tasks, time in the system);
3) results of standardized language tests.

The standardized tests consisted of 60–80 tasks covering grammar, vocabulary, reading, and listening; about 15% of the questions were open-ended, allowing for the assessment of written and communicative skills. Testing was conducted in a mixed format: most tasks were completed in writing in the online Moodle system, while oral tasks (interviews and oral responses) were completed offline or via videoconference under the teachers' supervision. The test administration was carried out by the foreign language departments of the relevant HEIs. In total, the study involved students from five universities: two Ukrainian (Sumy National Agrarian University and Ivan Franko National University of Lviv), two German (Humboldt University in Berlin and Leipzig University), and one Polish (Jagiellonian University in Krakow). The total number of participants was 312 people (Ukraine – 124, Germany – 98, Poland – 90). The difference in number is explained by the differences in the size of the sample groups and technical accessibility to the experimental environment. The study involved several universities from Ukraine, while one pilot institution in Germany and Poland, which limited the number of participants. The sample included students of 2–3 years of studies majoring in Philology and International Relations, for whom the study of foreign languages is a mandatory component of the curricula. Students were randomly assigned to three configurations (Stat-Learn, Deep-Adapt, Fusion-NLP) to ensure comparability of results. The choice of these three types of data is explained by the willingness to capture both objective indicators of academic performance (academic records) and students' behavioural activity in digital environments (Moodle, OpenEdX), as well as an independent external assessment of language proficiency (standardized tests). This three-component sample structure allowed for comparison of algorithmic configurations in different planes – cognitive, behavioural, and academic – and ensured greater reliability of findings.

3.3 Instruments and metrics

The experimental environment was implemented based on Python 3.11 using scikit-learn, TensorFlow, and Keras libraries to build algorithms of different levels of complexity. Pandas and NumPy were used to process large data sets, while Matplotlib and Seaborn were used to visualize the results. Deep learning models were trained on an NVIDIA RTX 3090 graphics processor, which provided faster computations compared to a conventional CPU configuration. The effectiveness of the models was assessed using the AALI, which combined four key components: accuracy of knowledge level prediction (Acc), stability of results during repeated training (Stab), average system response time (Latency), and cognitive fit of tasks to the student level (CogFit). Parameter normalization was performed relative to the best values in the sample. The CogFit component was evaluated by an expert panel of 12 English, German, and Polish language teachers (4 from each country) with at least 10 years of teaching experience in higher education. They analysed the extent to which the tasks generated by the models corresponded to the students' level and their cognitive abilities. The evaluation criteria were: (1) the task's correspondence to the language proficiency level (A1–C1), (2) the balance between complexity and learning objectives, (3) the logicality of the formulations, and the absence of redundant information. Each expert assigned scores on a 5-point scale, after which the data were normalized and integrated into the AALI. The main parameters are given in Table 1.

Table 2: Metrics for calculating the AALI

Parameter	Designation	Calculation method	Range	Weight in the Index
Prediction accuracy	Acc	Average model accuracy on test data	0–1	30%
Result stability	Stab	1 / standard deviation over 10 reruns	0–1	25%
System response time	Latency	1 / (normalized average query processing time)	0–1	25%
Cognitive relevance	CogFit	Expert rating on a 5-point scale, normalized	0–1	20%

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

So, AALI was formed as a weighted average value of the specified parameters, which made it possible to compare different algorithmic approaches in a unified coordinate system.

3.4 Data analysis methods

The differences in the performance of the three configurations were tested by using ANOVA with the subsequent post-hoc Tukey Honestly Significant Difference (HSD) test. This determined statistically significant differences between the mean values of the Acc, Latency, and CogFit indicators. The reliability of the results was confirmed by the bootstrapping method (1000 repetitions), which provided an estimate of the confidence intervals and stability of the obtained metrics. The quality of predictions in classification tasks was assessed by using ROC analysis and the calculation of the AUC.

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

This allowed us to compare the discriminatory ability of different algorithmic configurations. The dimensionality of the data was reduced and hidden patterns in the ratio of parameters were identified through the PCA, the results of which were visualized in two-dimensional space. Cluster analysis (k-means) was also used to check whether the models are grouped by the similarity of the AALI indicator profile. The practical significance of the differences between models was assessed by calculating Cohen's d, reflecting the strength of the effect in the comparison of configurations. The data were visualized in the format of diagrams showing the relationship between accuracy, response speed, and cognitive compliance.

4. RESULTS

4.1 Comparative evaluation of configurations according to the AALI

Three architectural configurations – Stat-Learn, Deep-Adapt and Fusion-NLP – were assessed using the integral AALI, which combines the accuracy of predictions (Acc), the stability of results (Stab), the average response time (Latency), and the cognitive relevance of training tasks (CogFit). The general trend showed that Fusion-NLP maintains the best balance between all components, providing high values of Acc and Stab while maintaining an acceptable processing speed. Deep-Adapt achieved maximum accuracy, but required more time for calculation, which reduced its Latency. In contrast, Stat-Learn demonstrated the fastest response and stability, but was inferior in personalization of learning trajectories. In Ukraine, the algorithms worked with heterogeneous data, which reduced the stability of deep models (Stab = 0.76 for Deep-Adapt), but Fusion-NLP remained relatively stable. In Germany, where the educational infrastructure is more mature, Deep-Adapt achieved the highest accuracy (Acc = 0.89), while Fusion-NLP showed a balance across all metrics (IAAL = 0.87). In Poland, the results were intermediate: Stat-Learn had the best Latency (0.93), while Fusion-NLP retained its lead in the overall AALI.

Table 3: Average AALI metrics by country

Configuration	Country	Acc	Stab	Latency	CogFit	AALI
Stat-Learn	Ukraine	0.76	0.81	0.90	0.63	0.77
	Germany	0.80	0.84	0.92	0.68	0.81
	Poland	0.78	0.83	0.93	0.66	0.80
Deep-Adapt	Ukraine	0.84	0.76	0.70	0.71	0.75
	Germany	0.89	0.81	0.77	0.74	0.80
	Poland	0.85	0.79	0.75	0.72	0.78
Fusion-NLP	Ukraine	0.83	0.82	0.85	0.79	0.82
	Germany	0.87	0.85	0.88	0.83	0.87
	Poland	0.85	0.83	0.87	0.80	0.84

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

The data in Table 3 demonstrate that Fusion-NLP was the most stable and versatile configuration regardless of the country. At the same time, there are differences depending on the level of development of the digital infrastructure. In Germany, deep models achieve the highest accuracy, while in Poland, the advantage of Stat-Learn is manifested in the speed of system response.

4.2 Evaluation of configurations in different educational systems

The comparative analysis made it possible to trace the specifics of the work of three configurations – Stat-Learn, Deep-Adapt and Fusion-NLP – in different educational environments. In each country, differences were observed due to the level of digital infrastructure, data consistency, and stability of access to educational platforms. In Ukraine, where the process of digital transformation is still in its early stages, Fusion-NLP turned out to be the most effective, demonstrating the ability to compensate for data fragmentation thanks to the built-in NLP module and hybrid adaptation mechanisms. In particular, the greatest gain was observed in tasks for written essays and translation of text passages, where the model formed individualized prompts and automatically selected additional materials. At the same time, Stat-Learn provided a quick response of the system, but its predictive quality was inferior to other models. In Germany, which is characterized by a highly developed educational infrastructure, the advantage was on the side of Deep-Adapt. The accuracy of predictions reached the maximum values among all configurations. This is explained by a mature approach to collecting and validating training data, which creates optimal conditions for the operation of DNNs. The listening comprehension exercise and automatic grammar test scoring showed the most noticeable results, where Deep-Adapt consistently outperformed other approaches. At the same time, Fusion-NLP remained competitive due to its high balance of indicators. In Poland, the results took an intermediate position. Stat-Learn showed the highest response speed (Latency), which is a significant factor in the context of limited resources, while Fusion-NLP provided a stable balance between accuracy and cognitive relevance of tasks. In practice, this was manifested in working with short interactive multiple-choice tests and quick diagnostic surveys, where Stat-Learn outperformed other configurations. This indicates the flexibility of the model in an environment with a transitional digital infrastructure, combining European standards with local features. The differences were visualized in Figure 1, which compares the integral values of AALI for the three configurations in each country.

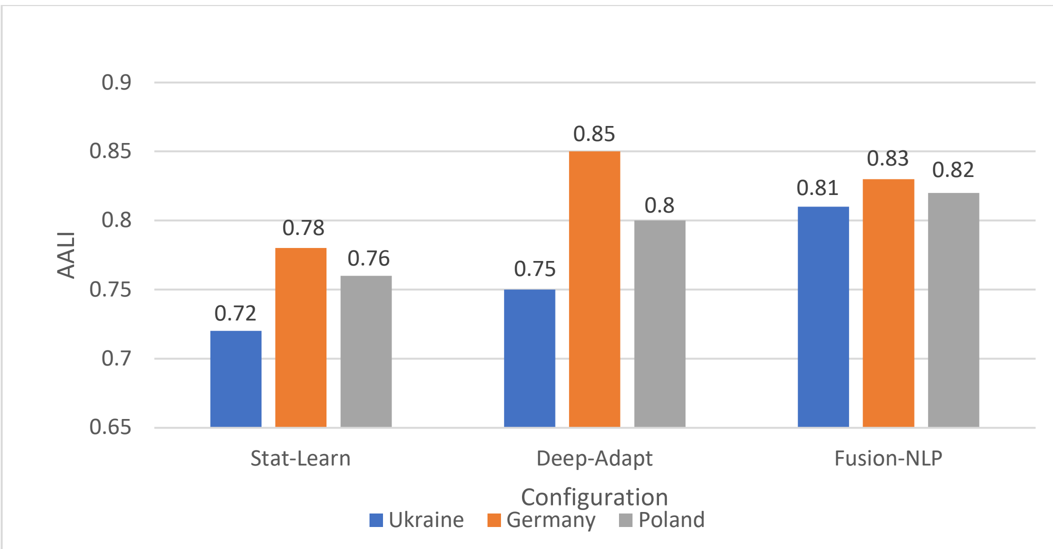

Figure 1: Comparison Of AALI Integral Values For Three Configurations In Ukraine, Germany, And Poland
Source: Developed By The Authors

Figure 1 shows a clustered bar chart that displays the average integral values of the AALI for three algorithmic configurations: Stat-Learn, Deep-Adapt, and Fusion-NLP. The hybrid Fusion-NLP model has the highest performance in Ukraine, while Deep-Adapt is in the lead in Germany due to the mature digital infrastructure. In Poland, there is a balance between all approaches with the advantage of Fusion-NLP in stability. So, the comparison confirms that the effectiveness of algorithms largely depends on the maturity of the digital infrastructure. Deep configurations work best in highly developed educational systems, while hybrid solutions are the most effective in environments with heterogeneous data.

4.3 Statistical testing of hypotheses

The differences between Stat-Learn, Deep-Adapt, and Fusion-NLP were tested by performing ANOVA separately for each country and metric (Acc, Stab, Latency, CogFit) followed by Tukey HSD (α = 0.05). Bootstrapping (1000 repetitions) was also performed to form 95% confidence interval (CI) for each metric, and the classification ability of the models was assessed via ROC/AUC. The characteristics of the statistical tests are presented in Table 4, which shows the p-value for each metric and the pairs of models between which statistically significant differences were found.

Table 4: ANOVA (p-value) and pairwise differences by Tukey HSD (α = 0.05)

Pairs are written in the format "model A > model B" (statistically significant better)

Country / Metric	Acc (p) / Tukey HSD	Stab (p) / Tukey HSD	Latency (p) / Tukey HSD	CogFit (p) / Tukey HSD
Ukraine	0,004 → Fusion-NLP > Stat-Learn; Deep-Adapt > Stat-Learn	0,031 → Fusion-NLP > Deep-Adapt	0,007 → Stat-Learn > Deep-Adapt; Fusion-NLP > Deep-Adapt	0,012 → Fusion-NLP > Stat-Learn

Germany	<0,001 → Deep-Adapt > Stat-Learn; Deep-Adapt > Fusion-NLP	0,049 → Fusion-NLP > Stat-Learn	0,015 → Stat-Learn > Deep-Adapt	0,022 → Fusion-NLP > Stat-Learn
Poland	0,018 → Fusion-NLP > Stat-Learn	0,041 → Fusion-NLP > Deep-Adapt	0,003 → Stat-Learn > Deep-Adapt; Stat-Learn > Fusion-NLP	0,027 → Fusion-NLP > Stat-Learn

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

In Germany, the Deep-Adapt deep configuration statistically outperforms the others on Acc, consistent with a more mature digital infrastructure. In Poland and Ukraine, Fusion-NLP consistently outperforms CogFit and (often) Stab, while Stat-Learn maintains its advantage on Latency. These trends are summarized in Table 5, which shows bootstrap estimates (median and 95% confidence intervals) for key AALI metrics in the three countries.

Table 5: Bootstrap estimates (median; 95% CI) for Acc, Stab, Latency, CogFit

Country	Configuration	Acc	Stab	Latency	CogFit
Ukraine	Stat-Learn	0.76 [0.73–0.79]	0.81 [0.78–0.84]	0.90 [0.88–0.92]	0.63 [0.60–0.66]
	Deep-Adapt	0.84 [0.82–0.86]	0.76 [0.73–0.79]	0.70 [0.67–0.73]	0.71 [0.68–0.74]
	Fusion-NLP	0.83 [0.81–0.85]	0.82 [0.80–0.84]	0.85 [0.83–0.87]	0.79 [0.77–0.81]
Germany	Stat-Learn	0.80 [0.78–0.82]	0.84 [0.82–0.86]	0.92 [0.90–0.94]	0.68 [0.65–0.71]
	Deep-Adapt	0.89 [0.87–0.91]	0.81 [0.79–0.83]	0.77 [0.75–0.79]	0.74 [0.71–0.77]
	Fusion-NLP	0.87 [0.85–0.89]	0.85 [0.83–0.87]	0.88 [0.86–0.90]	0.83 [0.81–0.85]
Poland	Stat-Learn	0.78 [0.76–0.80]	0.83 [0.81–0.85]	0.93 [0.91–0.95]	0.66 [0.63–0.69]
	Deep-Adapt	0.85 [0.83–0.87]	0.79 [0.76–0.82]	0.75 [0.72–0.78]	0.72 [0.69–0.75]
	Fusion-NLP	0.85 [0.83–0.87]	0.83 [0.81–0.85]	0.87 [0.85–0.89]	0.80 [0.78–0.82]

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

Table 5 presents the results of bootstrap estimates (median and 95% confidence intervals) for four key metrics – prediction accuracy (Acc), stability (Stab), latency (Latency), and cognitive fit (CogFit) – in the three countries for each configuration. The obtained data confirm the previous ANOVA/Tukey results. In Germany, Deep-Adapt achieved the highest Acc values (0.89 [0.87–0.91]), indicating the superiority of deep architectures in environments with qualitative data, while Fusion-NLP showed the highest balance between all indicators (Acc = 0.87; Stab = 0.85; CogFit = 0.83). In Poland, Stat-Learn was the fastest (Latency = 0.93 [0.91–0.95]), but was inferior in cognitive fit, where Fusion-NLP had consistently high results (CogFit = 0.80 [0.78–0.82]). In Ukraine, Fusion-NLP provided the best balance between accuracy (0.83 [0.81–0.85]) and cognitive fit (0.79 [0.77–0.81]), while Stat-Learn prevailed in speed (Latency = 0.90 [0.88–0.92]). Table 6 presents the results of the analysis of classification quality (ROC/AUC) for the three configurations of adaptive systems in language tasks.

Table 6: ROC/AUC of classification tasks (mean; 95% CI)

Country	Stat-Learn	Deep-Adapt	Fusion-NLP
Ukraine	0.86 [0.84–0.88]	0.90 [0.88–0.92]	0.89 [0.87–0.91]
Germany	0.89 [0.87–0.91]	0.93 [0.91–0.95]	0.92 [0.90–0.94]
Poland	0.88 [0.86–0.90]	0.91 [0.89–0.93]	0.90 [0.88–0.92]

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

Deep-Adapt demonstrates the best discrimination ability in classification subtasks (AUC) in all three countries, however, the advantage of Fusion-NLP over CogFit and Stab makes it more practical for adaptive language trajectories in tasks with a text component.

Summarizing the conducted statistical test, it can be noted that the effectiveness of the three configurations varies significantly depending on the country and digital infrastructure conditions. Deep-Adapt confirmed its superiority in classification tasks and achieved the highest AUC values, especially in Germany, where the quality of data and the stability of educational platforms create favourable conditions for the operation of deep networks. Fusion-NLP consistently demonstrated better stability and cognitive fit in all countries, making it the most versatile for mixed and text-centric learning scenarios. Stat-Learn maintained its advantage in speed (Latency), in particular in Poland and Germany, which indicates the feasibility of its use where the key factor is the processing speed. Therefore, a combination of models taking into account the educational context allows achieving the optimal balance between accuracy, speed, and adaptability of systems.

4.4 Multivariate analysis of models

Multivariate analysis explored the relationships between the configurations in more depth and confirmed the conclusions of previous statistical tests. The PCA made it possible to visualize the location of the models in two-dimensional space according to the integral indicators of AALI. In the resulting diagram, Stat-Learn was clearly separated due to high Latency values, while Deep-Adapt and Fusion-NLP formed close clusters due to similar levels of accuracy and stability. Further cluster analysis using the k-means method (k=3) confirmed these observations: Stat-Learn formed an independent cluster, while Fusion-NLP and Deep-Adapt grouped together, reflecting their functional proximity, but with an internal division according to the criterion of cognitive correspondence. The visualization of the results of PCA and k-means clustering is presented in Figure 2.

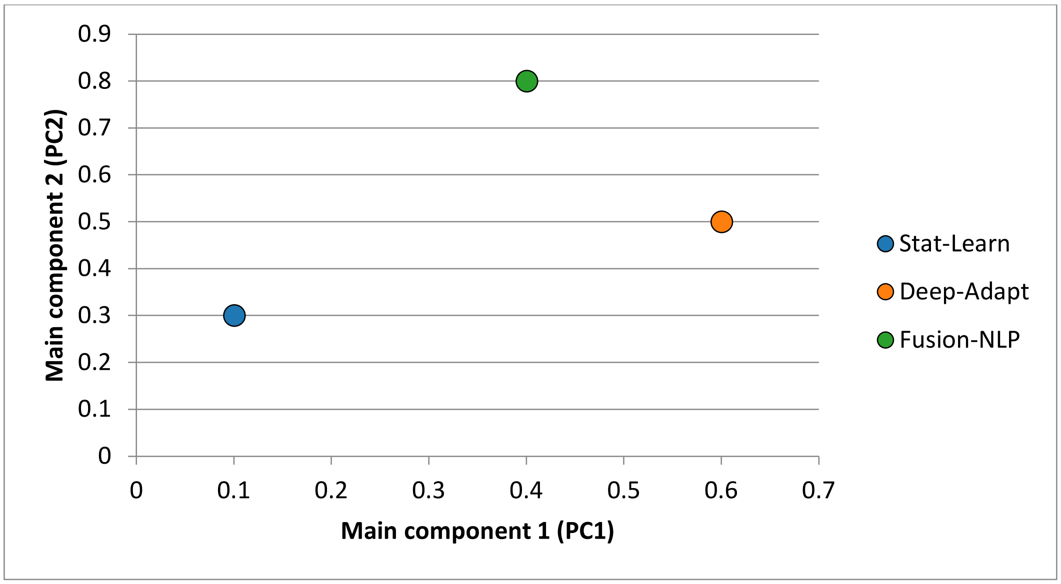

Figure 2: Visualization of PCA and k-means clustering results for three adaptive learning configurations (Stat-Learn, Deep-Adapt, Fusion-NLP) in principal component space

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

The obtained results indicate that the multivariate methods are not only consistent with the results of ANOVA and Tukey HSD, but also clarify the picture of the relationships between the models. Visualization in the space of principal components (Figure 2) showed a clear grouping of three configurations: Stat-Learn forms a separate cluster dominated by the speed indicator (Latency), while Deep-Adapt and Fusion-NLP are located closer to each other, demonstrating similar profiles of accuracy (Acc) and cognitive fit (CogFit). This indicates that hybrid and deep models share common structural features of adaptability, while statistical algorithms remain more isolated in the efficiency space. Clustering using the k-means confirmed the presence of three stable groups corresponding to different types of learning environments: fast (Stat-Learn), high-precision (Deep-Adapt), and balanced (Fusion-NLP).

4.5 Practical significance of the differences

The calculation of Cohen's d coefficient identified the strength of the effect revealed during statistical tests and assessed its pedagogical significance. The largest effect was observed in the comparison of Fusion-NLP with other configurations. It demonstrated high d values in all three countries (>0.8), which indicates significant advantages of this architecture for working with language tasks of increased cognitive complexity, in particular in Germany, where the level of digital infrastructure supports the full integration of NLP modules. At the same time, Stat-Learn showed medium effect values ($d \approx 0.5–0.6$) in Poland and Ukraine, which indicates its appropriates for quick decisions in environments with incomplete or uneven data. Deep-Adapt had the largest advantage in Germany, where it demonstrated $d \approx 0.7$ when compared to ML-oriented Stat-Learn. This emphasizes its value in stable conditions for accurate prediction and gradual adaptation. So, the difference between the configurations turned out to be not only statistically significant, but also pedagogically relevant: Fusion-NLP is more effective in countries with developed digital ecosystems, Stat-Learn is optimal in transitional systems, while Deep-Adapt occupies an intermediate position, providing a balance between stability and accuracy.

4.6 General conclusions for adaptive language learning

The final rating showed that Fusion-NLP configuration demonstrated the highest efficiency according to the AALI. This configuration combines deep algorithms, NLP modules, and expert rules, ensuring a balance of forecast accuracy and stability of recommendations. It was best manifested in Germany, where a developed digital infrastructure facilitates the integration of such solutions. Deep-Adapt ranked second, which is effective in mature educational systems with complete data, which is confirmed by the results of Poland and partly Ukraine. Stat-Learn turned out to be the most suitable for quick knowledge testing and work with incomplete data, which corresponds to the conditions of Ukraine with an uneven infrastructure. International results confirmed the dependence of productivity on the level of digital maturity. In Germany, Fusion-NLP is promising, in Poland — Deep-Adapt with the transition to hybrid systems, in Ukraine — Stat-Learn with the gradual introduction of NLP modules. Therefore, the further development of adaptive learning should take into account the technical potential of the platforms and the context: Stat-Learn is focused on diagnostics, Deep-Adapt – on stable forecasting, and Fusion-NLP – on individualization of learning and complex text tasks.

5. DISCUSSION

The results showed that the effectiveness of adaptive AI systems depends largely on the level of the country's digital infrastructure and the nature of the educational data. The hybrid Fusion-NLP configuration demonstrated stability and balance between all AALI metrics in an environment with heterogeneous and fragmented data, such as in Ukraine. At the same time, Stat-Learn provided the fastest system response but was inferior in cognitive relevance. In Germany, where educational platforms are characterized by high data quality and stability, Deep-Adapt became the leader. It achieved maximum prediction accuracy, but Fusion-NLP remained competitive due to balanced results. Poland demonstrated an intermediate scenario: Stat-Learn was optimal for rapid diagnostics in mass groups, while Fusion-NLP turned out to be most suitable for individualized learning. Such findings are consistent with international reviews that emphasize that the combination of NLP modules and deep algorithms expands the adaptability of educational systems in different contexts [27]; [28]. Additional statistical analyses confirmed this pattern: Deep-Adapt provided the highest AUC values on the classification subtasks, especially in Germany, while Fusion-NLP consistently led in terms of CogFit and Stab in all three countries. This suggests that models with hybrid adaptive mechanisms perform better in scenarios with increased data variability, which is also supported by other studies on the use of AI in language education [4]; [22]. The findings also suggest another implication related to the implementation of AI in adaptive learning. The effectiveness of such technology is not just related to the complexity of the technology itself, but also in how the technology relates to the existing educational environment. Thus, it is not necessarily in the introduction of more complex models of AI that institutions should focus their efforts, but instead in ensuring that the models that are introduced are compatible with the educational environment in which they are to be implemented.

In the context of Ukraine, an additional indirect support for such an interpretation is provided by the finding that the modernization of Moodle-based learning in the country during the crisis was associated with improved academic performances of the students [36]. This finding further supports the conclusions of the present paper as it indicates that even under the conditions of the crisis in education, the effectiveness of the AI-based system also depends on the environment in which it is applied.

The results have practical significance from a pedagogical perspective. It is appropriate to use Stat-Learn in systems where speed and mass are key factors (as in Ukraine). In mature digital ecosystems (Germany), it is worth relying on Deep-Adapt for highly accurate personalization. Fusion-NLP has the highest versatility, as it combines accuracy, stability. and cognitive flexibility. Therefore, it appears to be the most promising for the development of adaptive learning in long-term scenarios [10]; [11]. At the same time, the conceptual article [31] questions the thesis of AI as the "next big breakthrough", but emphasizes its inevitable role in the future of educational practices, which is consistent with our conclusion that the effectiveness of models depends on the maturity of the educational infrastructure. The authors [32] demonstrated another approach, who proposed a synthetic index to measure the impact of AI on students. This has much in common with our AALI and confirms the appropriateness of quantifying pedagogical effects. The tested configurations can be integrated into the university LMS environments (Moodle, OpenEdx) according to their functionalities. Stat-Learn can be used for the entry diagnostic or for assessing the students in large groups quickly. Deep-Adapt can be used in an environment where highly accurate assessment of the students' knowledge is required. Finally, Fusion-NLP is better suited for foreign language courses where text processing is essential and where individual feedback can be provided to the students.

At the same time, the researchers [33] draw attention to the ethical challenges of using AI in language classrooms, emphasizing equity of access and fairness, which is an important caveat for countries with less developed infrastructure. In contrast to these cautious assessments, the study [34] presents a vision of an "optimistic future" for AI in higher education. In turn, the authors [35] empirically demonstrate that automatic speech assessment technologies can be successfully integrated into students' speaking practice, confirming the pedagogical value of systems with a high level of cognitive relevance. Therefore, the results of our study not only confirm previous empirical observations, but also clarify academic debate: from sceptical positions on the role of AI to optimistic predictions, from quantitative assessment models to ethical challenges. In this context, Fusion-NLP can be considered the most versatile and practically feasible configuration, while Deep-Adapt and Stat-Learn have niche advantages in their respective educational scenarios.

6. LIMITATIONS

The results of the study should be interpreted with methodological and contextual limitations in mind. The study has several limitations that should be taken into account when interpreting the findings. These include the restricted set of tested architectures, the fixed weighting structure of AALI, the one-cycle sampling design, cross-country contextual differences, and the absence of direct long-term pedagogical and ethical measurements.

The study covered only three configurations (Stat-Learn, Deep-Adapt, Fusion-NLP) and basic ML/DL algorithms, without modern transformer architectures, which limits the possibility of extrapolating the findings. The AALI was created using fixed weights and contained a subjective component (CogFit). So, other weighting schemes could change the ranking of the models. Latency and Stab metrics were evaluated in a standardized environment with fixed parameters, which does not reflect the diversity of real-world scenarios. Statistical procedures (ANOVA, bootstrapping, ROC/AUC) were applied to a limited number of sub-scenarios, which could reduce sensitivity to small effects. The data sample covered only one academic cycle and three types of sources (academic records, LMS activity, test results), without long-term time intervals or inter-university comparisons. The international sample (Ukraine, Germany, Poland) was based on normalized scales, but residual differences in data collection practices and curricular features may have affected the Acc and Stab scores. The expert assessment of CogFit may also have contained cultural and institutional biases. In addition, the study focused mainly on technical metrics, without directly assessing pedagogical effects or ethical aspects. Further study should include a broader range of models, longitudinal observations, alternative IAAL schemes, and the assessment of real educational impact.

7. CONCLUSIONS

The study confirmed that the use of AI models in adaptive foreign language learning provides a significant improvement in the quality of the educational process due to the personalization of learning trajectories, the accuracy of predictions, and the stability of results. Fusion-NLP was the most versatile configuration, which achieved the highest integral values of the AALI in all three countries (0.82 in Ukraine, 0.87 in Germany and 0.84 in Poland). This confirms the ability to combine accuracy, stability, and cognitive relevance of learning tasks. Deep-Adapt demonstrated the highest accuracy in Germany (Acc = 0.89; AUC = 0.93), but was inferior in speed (Latency = 0.77), which makes it more suitable for stable educational systems with complete data sets. Stat-Learn proved to be optimal for working with incomplete or heterogeneous data, providing the highest performance in Poland (Latency = 0.93) and Germany (Latency = 0.92), as well as stability in Ukraine (Stab = 0.81). International comparison confirmed the dependence of the effectiveness of models on the level of digital maturity of education systems. Fusion-NLP became the most stable in environments with fragmented data, while Deep-Adapt worked best in mature systems (Germany). In Poland, Stat-Learn and Fusion-NLP provided a balance between accuracy and speed. This indicates the appropriateness of a context-oriented approach to choosing algorithmic solutions. From a practical perspective, universities seeking to integrate AI-driven adaptive systems can take the following steps: phased testing of models in local training courses with subsequent scaling, training, and retraining of teachers to work with AI panel interfaces, integration of selected configurations with existing LMSs (Moodle, OpenEdX, etc.), as well as the development of internal data collection and protection protocols to improve the quality of educational analytics.

In general, it can be stated that Fusion-NLP is the most promising for individualized language learning, Deep-Adapt is appropriate for highly accurate personalization in stable systems, while Stat-Learn is optimal for rapid diagnostics and work in transitional educational environments. Further research should be directed towards improving the AALI by expanding cognitive indicators, scaling the international sample, and creating cloud-based monitoring dashboards for teachers, which will strengthen the practical integration of adaptive technologies into language training in a global context. The novelty of the study is the comparative application of the proposed framework to three different cases of adaptive AI within higher education and foreign language teaching in education systems of different digital maturity. The practical impact of the study is the possibility of basing the decision of which AI solution to employ in a specific educational context on the results of the study.Future research can investigate the influence of transformer-based architectures, alternative weighting schemes for the AALI, and utilize longitudinal study designs to assess the impact of the AI teacher over time. Furthermore, it would also be useful to incorporate evaluations of the communicative and retention skills of the learners, as well as to develop explainable AI dashboards for the instructors of these classrooms.

REFERENCES:

[1] I. Semenets-Orlova, R. Shevchuk, B. Plish, A. Moshnin, Y. Chmyr, & R. Poliuliakh, "Human-Centered Approach in New Development Tendencies of Value-Oriented Public Administration: Potential of Education", *Economic Affairs (New Delhi),* Vol. 67, No. 5, 2022, pp. 899–906. https://doi.org/10.46852/0424-2513.5.2022.25

[2] O. Smolinska, I. Koval, M. Pavliuk, & D. Shulha, "Research of Assertiveness in Organization of the Time Perspective of Degree-Seeking Students in Higher Education", *Insight: The Psychological Dimensions of Society*, Vol. 11, 2024, pp. 222-237. https://doi.org/10.32999/2663-970X/2024-11-12

[3] E. Smyrnova-Trybulska, N. Morze, & L. Varchenko-Trotsenko, „Adaptive Learning in University Students' Opinions: Cross-Border Research", *Education and Information Technologies,* Vol. 27, No. 5, 2022, pp. 6787-6818. https://doi.org/10.1007/s10639-021-10830-7

[4] R. Sajja, Y. Sermet, M. Cikmaz, D. Cwiertny, & I. Demir, "Artificial Intelligence-Enabled Intelligent Assistant for Personalized and Adaptive Learning in Higher Education", *Information,* Vol. 15, No. 10, 2024, Art. No. 596. https://doi.org/10.3390/info15100596

[5] T. Schmidt, & T. Strasser, "Artificial Intelligence in Foreign Language Learning and Teaching: A CALL for Intelligent Practice", *Anglistik: International Journal of English Studies*, Vol. 33, No. 1, 2022, pp. 165-184. https://doi.org/10.33675/ANGL/2022/1/14

[6] I. Gligorea, M. Cioca, R. Oancea, A. T. Gorski, H. Gorski, & P. Tudorache, „Adaptive Learning Using Artificial Intelligence in E-learning: A Literature Review", *Education Sciences*, Vol. 13, No. 12, 2023, Art. No. 1216. https://doi.org/10.3390/educsci13121216

[7] M. Kruk, & A. Kałużna, „Investigating the Role of AI Tools in Enhancing Translation Skills, Emotional Experiences, and Motivation in L2 Learning", *European Journal of Education*, Vol. 60, No. 1, 2025, Art. No. e12859. https://doi.org/10.1111/ejed.12859

[8] J. B. Son, N. K. Ružić, & A. Philpott, "Artificial Intelligence Technologies and Applications for Language Learning and Teaching", *Journal of China Computer-Assisted Language Learning*, Vol. 5, No. 1, 2025, pp. 94-112. https://doi.org/10.1515/jccall-2023-0015

[9] I. Semenets-Orlova, A. Klochko, O. Tereshchuk, L. Denisova, V. Nestor, & S. Sadovyi, "Special Aspects of Educational Managers' Administrative Activity under Conditions of Distance Learning", *Journal of Curriculum and Teaching*, Vol. 11, No. 1, 2022, pp. 286–297. https://doi.org/10.5430/jct.v11n1p286

[10] E. Hanstein, "Adaptation Strategies of Humboldt University Berlin and University of Leipzig to the Challenges of Artificial Intelligence Technology in the Context of Language Teaching", In *2nd Annual Conference of Islamic Education 2023 (ACIE 2023)*, Jember (Indonesia), 6-7 September 2023, pp. 127-129. https://doi.org/10.2991/978-2-38476-182-1_22

[11] A. Jarkiewicz, A. Krawczyk, & E. Szulewicz, „The Development of the Participatory Approach to Digital Learning: Research Findings from Poland", *Technology, Knowledge and Learning*, 2025, pp. 1-24. https://doi.org/10.1007/s10758-025-09848-4

[12] K. Kuddus, "Artificial Intelligence in Language Learning: Practices and Prospects", *Advanced Analytics and Deep Learning Models*, 2022, pp. 1-17. https://doi.org/10.1002/9781119792437.ch1

[13] T. N. Fitria, "The Use Technology Based on Artificial Intelligence in English Teaching and Learning", *ELT Echo: The Journal of English Language Teaching in Foreign Language Context*, Vol. 6, No. 2, 2021, pp. 213-223. https://doi.org/10.24235/eltecho.v%vi%i.9299

[14] H. Yang, & S. Kyun, "The Current Research Trend of Artificial Intelligence in Language Learning: A Systematic Empirical Literature Review from an Activity Theory Perspective", *Australasian Journal of Educational Technology*, Vol. 38, No. 5, 2022, pp. 180-210. https://doi.org/10.14742/ajet.7492

[15] T. Kabudi, I. Pappas, & D. H. Olsen, "AI-Enabled Adaptive Learning Systems: A Systematic Mapping of the Literature", *Computers and Education: Artificial Intelligence*, Vol. 2, 2021, Art. No. 100017. https://doi.org/10.1016/j.caeai.2021.100017

[16] M. P. Ilić, D. Păun, N. Popović Šević, A. Hadžić, & A. Jianu, "Needs and Performance Analysis for Changes in Higher Education and Implementation of Artificial Intelligence, Machine Learning, and Extended Reality", *Education Sciences*, Vol. 11, No. 10, 2021, Art. No. 568. https://doi.org/10.3390/educsci11100568

[17] J. Muñoz-Basols, C. Neville, B. A. Lafford, & C. Godev, "Potentialities of Applied Translation for Language Learning in the Era of Artificial Intelligence", *Hispania*, Vol. 106, No. 2, 2023, pp. 171-194. https://doi.org/10.1353/hpn.2023.a899427

[18] S. G. Essa, T. Celik, & N. E. Human-Hendricks, "Personalized Adaptive Learning Technologies Based on Machine Learning Techniques to Identify Learning Styles: A Systematic Literature Review", *IEEE Access*, Vol. 11, 2023, pp. 48392-48409. https://doi.org/10.1109/ACCESS.2023.3276439

[19] Y. Wang, "Artificial Intelligence Technologies in College English Translation Teaching", *Journal of Psycholinguistic Research*, Vol. 52, No. 5, 2023, pp. 1525-1544. https://doi.org/10.1007/s10936-023-09960-5

[20] S. Hashim, M. K. Omar, H. Ab Jalil, & N. M. Sharef, "Trends on Technologies and Artificial Intelligence in Education for Personalized Learning: Systematic Literature", *Journal of Academic Research in Progressive Education and Development*, Vol. 12, No. 1, 2022, pp. 884-903. http://dx.doi.org/10.6007/IJARPED/v11-i1/12230

[21] K. Nikolopoulou, "Generative Artificial Intelligence in Higher Education: Exploring Ways of Harnessing Pedagogical Practices with the Assistance of ChatGPT", *International Journal of Changes in Education*, Vol. 1, No. 2, 2024, pp. 103-111. https://doi.org/10.47852/bonviewIJCE42022489

[22] D. Kem, "Personalised and Adaptive Learning: Emerging Learning Platforms in the Era of Digital and Smart Learning", *International Journal of Social Science and Human Research*, Vol. 5, No. 2, 2022, pp. 385-391. https://doi.org/10.47191/ijsshr/v5-i2-02

[23] T. Alqahtani, H. A. Badreldin, M. Alrashed, A. I. Alshaya, S. S. Alghamdi, K. Bin Saleh, ... & A. M. Albekairy, "The Emergent Role of Artificial Intelligence, Natural Learning Processing, and Large Language Models in Higher Education and Research", *Research in Social and Administrative Pharmacy*, Vol. 19, No. 8, 2023, pp. 1236-1242. https://doi.org/10.1016/j.sapharm.2023.05.016

[24] 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,* Singapore: Springer Nature Singapore, 2022, pp. 395-406. https://doi.org/10.1007/978-981-19-2980-9_32

[25] B. George, & O. Wooden, "Managing the Strategic Transformation of Higher Education through Artificial Intelligence", *Administrative Sciences*, Vol. 13, No. 9, 2023, Art. No. 196. https://doi.org/10.3390/admsci13090196

[26] T. Wang, B. D. Lund, A. Marengo, A. Pagano, N. R. Mannuru, Z. A. Teel, & J. Pange, "Exploring the Potential Impact of Artificial Intelligence (AI) on International Students in Higher Education: Generative AI, Chatbots, Analytics, and International Student Success", *Applied Sciences,* Vol. 13, No. 11, 2023, Art. No. 6716. https://doi.org/10.3390/app13116716

[27] X. Huang, D. Zou, G. Cheng, X. Chen, & H. Xie, "Trends, Research Issues and Applications of Artificial Intelligence in Language Education", *Educational Technology & Society*, Vol. 26, No. 1, 2023, pp. 112-131. https://doi.org/10.30191/ETS.202301.261.0009

[28] C. L. Yuen, & N. Schlote, "Learner Experiences of Mobile Apps and Artificial Intelligence to Support Additional Language Learning in Education", *Journal of Educational Technology Systems*, Vol. 52, No. 4, 2024, pp. 507-525. https://doi.org/10.1177/00472395241238693

[29] H. Munir, B. Vogel, & A. Jacobsson, "Artificial Intelligence and Machine Learning Approaches in Digital Education: A Systematic Revision", *Information*, Vol. 13, No. 4, 2022, Art. No. 203. https://doi.org/10.3390/info13040203

[30] S. Z. Salas-Pilco, & Y. Yang, "Artificial Intelligence Applications in Latin American Higher Education: A Systematic Review", *International Journal of Educational Technology in Higher Education*, Vol. 19, No. 1, 2022, Art. No. 21. https://doi.org/10.1186/s41239-022-00326-w

[31] X. O'Dea, & M. O'Dea, „Is Artificial Intelligence Really the Next Big Thing in Learning and Teaching in Higher Education?: A Conceptual Paper", *Journal of University Teaching and Learning Practice*, Vol. 20, No. 5, 2023, pp. 1–17. https://search.informit.org/doi/10.3316/informit.T2024112700021091343163252

[32] A. Grájeda, J. Burgos, P. Córdova, & A. Sanjinés, "Assessing Student-Perceived Impact of Using Artificial Intelligence Tools: Construction of a Synthetic Index of Application in Higher Education", *Cogent Education*, Vol. 11, No. 1, 2024, Art. No. 2287917. https://doi.org/10.1080/2331186X.2023.2287917

[33] D. Dakakni, & N. Safa, „Artificial Intelligence in the L2 Classroom: Implications and Challenges on Ethics and Equity in Higher Education: A 21st Century Pandora's Box", *Computers and Education: Artificial Intelligence*, Vol. 5, 2023, Art. No. 100179. https://doi.org/10.1016/j.caeai.2023.100179

[34] M. Á. Escotet, "The Optimistic Future of Artificial Intelligence in Higher Education", *Prospects,* Vol. 54, No. 3, 2024, pp. 531-540. https://doi.org/10.1007/s11125-023-09642-z

[35] B. Zou, Q. Lyu, Y. Han, Z. Li, & W. Zhang, "Exploring Students' Acceptance of an Artificial Intelligence Speech Evaluation Program for EFL Speaking Practice: An Application of the Integrated Model of Technology Acceptance", *Computer Assisted Language Learning*, Vol. 38, No. 5-6, 2025, pp. 1366-1391. https://doi.org/10.1080/09588221.2023.2278608

[36] M. Ostrovska, O. Tretiak, H. Smolnykova, O. Studenets, & M. Ivanchuk, "Moodle-Based Development of Primary Education under Martial Law", *International Electronic Journal of Elementary Education*, Vol. 15, No. 5, 2023, pp. 399–407. https://doi.org/10.26822/iejee.2023.308