Building a Personalized Learning Model in a Virtual Environment for Learning the Kazakh Language

Talgat Sembayev[1], Almaz Sydykov[1,*], Kuandyk Taibolatov[2], and Zhanat Nurbekova[3]

[1]Computer Science and Engineering, Astana IT University, Astana, Kazakhstan
[2]National Academy of Education named after Y.Altynsarin, Astana, Kazakhstan
[3]Department of Informatics and Information of education, Abai Kazakh National Pedagogical University, Almaty, Kazakhstan
Email: talgat.sembayev@astanait.edu.kz (T.S.); 242827@astanait.edu.kz (A.S.); k.taibolatov@uba.edu.kz (K.T.); zh.nurbekova@abaiuniversity.edu.kz (Z.N.)
*Corresponding author

Manuscript received February 14, 2025; revised March 13, 2025; accepted April 28, 2025; published July 18, 2025

***Abstract*—Recent developments in digital technologies and pedagogically innovative approaches have transformed traditional education into personalized and adaptive learning. The paper proposes a new framework embedding VR with advanced machine learning for the underrepresented Kazakh language acquisition. Presented model will be dynamically adapted to individual learner profiles by tailoring content based on prior knowledge, learning preference, and real-time performance feedback. Quantitative results are obtained using interaction logs, response times, accuracy rates, and standardized language assessments, while qualitative insights obtained from user feedback and observational notes are used in a pilot study involving diverse participants. Evaluations from our pilot study indicate substantial improvements: vocabulary proficiency increased by approximately 25%, grammatical accuracy by 20%, listening comprehension by 18%, and speaking fluency by 30%. Moreover, the adaptive recommendation system-either by using collaborative filtering and reinforcement learning-effectively modulated task difficulty, sustaining an optimal learning curve. Although fine-tuning the reward function of reinforcement learning and scaling with limited participant samples proved challenging, our results emphasize the role that could be played by fully immersive, data-driven learning environments in the transformation of language learning. This work helps not only with the preservation of Kazakh cultural heritage but also sets up a feasible framework for broader applications in diverse languages and modern educational contexts.**

***Keywords*—personalized learning, virtual reality, machine learning, Kazakh language, adaptive learning, immersive education, educational technology**

I. Introduction

Modern education is undergoing a continuous transformation because of deep penetration of digital technologies and a growing interest in innovative approaches. Traditional models of education, once homogeneous, have yielded to far more adaptive, responsive systems that address the unique requirements of every individual learner. This has been alongside the modern worldwide trend towards the digitalization of education, wherein personalized learning holds the key to defining the future of education practice in a world defined by accelerating change and shifts. Additionally, advances in machine learning and virtual reality technology have enabled the development of more adaptive and interactive learning models [1]. In this regard, the present study will propose a model of personalized learning in a virtual environment with a focus on language building and the reduction of barriers to learning the Kazakh language.

Personalized learning has been increasingly positioned as a key focus of reform efforts in education, based on its potential to address the diverse needs, interests, and abilities of students. Researchers note that a unified definition for personal learning does not exist, suggesting there is a wide range of methodological contexts within which it may be found [2]. Personalized learning refers to tailoring the learning experience to align with the individual features of the learner, based on prior knowledge, abilities, and learning preferences. A review of current literature shows that tailor-made systems significantly enhance learners' motivation and engagement, while fostering a deeper understanding of the subject matter through immediate feedback [3, 4].

Machine learning can analyze and improve language learning success rates in tailoring educational courses to a particular student [5, 6]. Researchers further explore the role of machine learning in processing large databases related to learners to optimize the adaptation of content and learning routes [7, 8]. Such systems already exist in other fields, like medical education, and thus show the potential to scale up and be relevant in many fields [9]. A similar approach is evident in agriculture, where machine learning optimizes crop, soil, water, and livestock management by analyzing large datasets [10]. Likewise, in finance, machine learning enhances decision-making by processing vast datasets, paralleling its application in education to tailor learning experiences to individual needs [11]. Machine learning can analyze and improve language learning success rates in tailoring educational courses to a particular student [5, 6].

Alongside the development of personalized systems, educational practices are increasingly using virtual and mixed realities—such as virtual reality, augmented reality, and mixed reality. An advantage of virtual reality is that it can create experiential and engaging settings, replicating real-world scenarios with lower risks and costs [12]. According to research, realism, interactivity, and the level of user control appear to be the most important features for incorporating VR technologies in a learning environment, considerably increasing learners' engagement [13]. Moreover, virtual environments afford avenues for prompt feedback and improved collaborative efforts. For instance, certain virtual reality prototypes facilitate instantaneous evaluations of knowledge, thus streamlining the modification of educational materials [14]. Through the integration of VR with artificial intelligence technologies, it is feasible to develop adaptable and individualized learning environments that align with specific objectives and requirements [15].

doi: 10.18178/ijiet.2025.15.7.2352

Combining the two previously mentioned concepts, personalized learning and virtual environments, brings new horizons to research in education and its practice. Scientists underline that applications of machine learning in VR environments can dramatically improve the analysis of user behavior and dynamic adjustment of content [1]. The latter forms a base for creating individual learning paths where task difficulty, visual and audio components, and interaction levels change according to each learner in real time [3]. Moreover, the adaptability of virtual reality platforms makes it easier to develop highly immersive and engaging settings—something quite necessary for acquiring complex or specialized skills [16]. These solutions find extensive application across several domains, including professional training in disciplines like medicine and engineering, and in language learning, where there is a great advantage in cultural context and repetitive practice under realistic conditions [17, 18]. However, despite these advancements, the literature has notably overlooked the application of these innovative technologies to teaching underrepresented languages, such as Kazakh. Existing research predominantly focuses on widely taught languages, resulting in a significant gap regarding how effectively VR and ML methodologies address the unique challenges

The Kazakh language presents a unique and urgent case study as it remains significantly underrepresented in digital educational resources, confronting specific instructional challenges such as inadequate teaching materials, limited accessibility, and a lack of immersive resources [19, 20]. In the pedagogically challenged domain of Kazakh language teaching, such technological innovations are especially vital. Virtual reality provides realistic, culturally enhanced environments, i.e., virtual cafés, classrooms, and bazaars, within which learners can practice context-specific, meaningful language communication. Meanwhile, machine learning algorithms enable content customization by tailoring learning experiences to each student's proficiency level, performance feedback, and learning style. Hence, this integrated technological approach is directly targeted at the specific pedagogical issues associated with the Kazakh language, ultimately facilitating increased learner engagement, accessibility, and effectiveness of language learning methodologies. Therefore, creating a customized educational space for it is highly relevant for the preservation of cultural heritage. It is a unique chance for the evaluation of the effectiveness of personalized virtual reality methods in the sphere of less commonly taught languages.

Consequently, the scientific justification for the choice of the Kazakh language is based on the need to preserve linguistic diversity and improve the quality of language education in the current conditions of strengthening globalization and migration processes.

The purpose of this study is to determine the impact of a personalized learning model in a virtual environment for learning the Kazakh language on improving the quality of learning. The combination of virtual reality and machine learning technologies can help improve the learning process using the example of learning the Kazakh language.

In connection with the above, complex research tasks were set:

1) Building a personalized learning model in a virtual environment for learning the Kazakh language
2) Determining the effectiveness of the developed personalized model

The present research establishes a base for developing a new framework in the acquisition of the Kazakh language, applying the findings from existing literature on personalized learning and immersive virtual technologies.

The following sections provide an in-depth review of theoretical literature, methodological frameworks, and empirical evidence, outlining in detail the effectiveness of this approach.

II. Materials and Methods

This research design will adopt both a quantitative and qualitative approach to investigate how virtual reality and machine learning can be integrated into the development of a personalized learning model in language education, focusing on Kazakh as a second or additional language. The elements of the research design are presented below.

A. Quantitative Component

Interaction Logs: Automated collection of learner responses, time-on-task, accuracy, and error rates in VR-based lessons.

Language Assessments: Standardized pre- and post-tests to quantify language proficiency gains.

B. Qualitative Component

Observational Notes: The system will document participants' level of engagement, possible technical issues - for example, motion sickness, calibration problems - and their real-time reactions via form after each session.

By triangulating qualitative and quantitative data, the research tries to obtain an overall picture about the efficiency with which VR-based personalized learning enhances language acquisition. The examination of studies in personalized learning, virtual reality technology, and machine learning adaptation [1, 2, 6, 7, 10] uncovered a range of critical factors that form an integral part of developing an effective virtual reality based personalized language learning model. Research identifies key factors such as user-centered design, ease of use, real-time feedback, and scenario-based simulation in enhancing educational experiences [3, 12, 17]. These studies inform the development of the model in its next stage through the determination of key factors impacting personalized VR learning environments' effectiveness.

Five key factors contribute in a significant manner towards a VR-based personalized language learning system's design and implementation:

Proficiency Level: The learner's starting skill in terms of vocabulary, listening, and grammar controls how content is personalized and upgraded.

Previous VR Exposure: Familiarity with VR hardware and controls can impact pace and intensity of learning and level of engagement.

Learning Objectives: Well-defined language objectives, such as speaking fluency, reading, and contextual awareness, drive educational direction.

Design for Interaction: Effectiveness of interactivity, realism, and usability in virtual environment controls learner presence and information recall.

Feedback Mechanism: Adaptive ML-powered feedback, delivered in real-time or near real-time, controls difficulty level in a responsively changing manner and introduces scaffolding to maintain learners in an ideal challenge zone.

The factors form a basis for sound decision-making in terms of collecting and processing data, feature engineering, and choosing a machine learning approach, and enable development of a learner-responsive and learner-sensitive educational model.

C. Data Collection and Preprocessing

To ensure the validity and relevance of the study, participants will be carefully selected using purposive sampling based on predefined criteria, including demographic background, Kazakh language proficiency, and willingness to engage in virtual learning. A total of twenty-five participants were selected for a pilot study to gather preliminary insights and validate the functionality of the virtual learning environment. Participants ranged in age from 19 to 27, covering a diverse range of prior experience with VR technology, which allowed comprehensive feedback from both novice and experienced VR users. Potential participants were initially screened through a demographic and language proficiency survey to establish baseline proficiency and background characteristics. From this pool, twenty-five participants were purposefully selected to represent a balanced distribution across the predefined categories. This targeted approach aimed to capture diverse perspectives, ensuring a comprehensive evaluation of how linguistic, cultural, and technological factors influence learning outcomes within the personalized VR learning environment. Although this sampling method inherently introduces potential biases due to deliberate selection, it was essential to achieve diversity and specificity relevant to the research aims. The explicit criteria helped ensure participants were representative of the target user population for the proposed educational model. It was predicted that individuals with closer linguistic and cultural proximity to the Kazakh language would experience lower initial challenges and show greater levels of immersion. On the other hand, individuals from linguistically disparate backgrounds might experience greater learning challenges but achieve greater immersive benefits from the cultural immersion.

Data were collected via several instruments specifically designed for the aims of this research. At the end of each session, users completed formal questionnaires to assess usability, satisfaction, perceived difficulty, and immersion levels. Questionnaires presented both Likert-scale questions and open-ended queries to derive quantitative measures and qualitative feedback. Furthermore, observers used structured observation notes in every virtual reality session to systematically record participants' interactions, the levels of engagement, navigation issues, technical problems encountered, and non-verbal signs of immersion or frustration. Semi-structured interview guides were created and used in post-study focus groups to ask participants about their subjective experiences in depth, enabling thorough discussions of questionnaire responses, ideas for improvement, and assessments of the pedagogical effectiveness of the VR environment.

D. Data Sources

Demographic & Background Surveys: Collected participant information such as age, native language, prior VR exposure, and perceived learning style.

Language Proficiency Assessments: Pre- and post-intervention tests focusing on vocabulary, grammar, listening, and speaking skills to quantify improvement. Test is based on QAZTEST - system for assessing the level of knowledge of the Kazakh language of citizens of the Republic of Kazakhstan and foreign citizens.

VR Interaction Logs: Continuous data on user navigation paths, response times, accuracy rates, help requests, and error patterns within each module.

User Feedback: Post-session questionnaires, focus groups, and interviews to capture subjective experiences and suggestions.

E. Preprocess and model development

To determine the initial user level, a model will be used based on the data collected. Creating the best personalized experience requires high-precision data analysis. Therefore, the first step is to choose an algorithm for the model. Four classification methods—Logistic Regression, Random Forest, XGBoost, and LightGBM—are chosen for analysis as they are have been shown to be efficient in educational data mining and can work efficiently on different kinds and sizes of data [21, 22]. These algorithms were extensively evaluated using the benchmark measures of performance, including Accuracy, F1 Score, Confusion Matrix, and a detailed Classification Report. Accuracy represented correct predictions, and the F1 Score was used as it adequately weights recall and precision. This weighting is especially important with unbalanced categories of user proficiency categories. The Confusion Matrix described the different kinds of prediction errors, and the Classification Report provided accurate information about precision, recall, and F1 scores for each user category.

Next step is to test these models by inputting new user data. Fig. 1 demonstrates new user entry data.

```python
new_user = pd.DataFrame({
    "Age": [15],
    "Vocabulary_Score": [70],
    "Grammar_Score": [65],
    "Listening_Score": [60],
    "Speaking_Fluency": [150],
    "Pronunciation_Score": [55],
    "Reading_Score": [75],
    "Writing_Score": [68]
})

new_user_normalized = scaler.transform(new_user)

print("\nNew User Predictions:")
for name, model in trained_models.items():
    predicted_level = model.predict(new_user_normalized)
    predicted_label = label_encoder.inverse_transform(predicted_level)
    print(f"{name}: Predicted Learning Level - {predicted_label[0]}")
```

Fig. 1. New user entry sample.

The results obtained reiterated the preliminary model assessment and demonstrated the advantages accruable from using Logistic Regression with this dataset in achieving the classification objective. The distinct separation among models shows the importance of selecting the most effective algorithm to achieve precise predictions.

This examination illustrates the meticulous methodology employed to enhance user categorization, a critical component in the creation of individualized and impactful educational experiences. The chosen model will subsequently be integrated into the system to categorize users and customize their learning pathways accordingly.

F. Development of the Virtual Environment

The most important part of the whole application is the VR component. For this reason, it is crucial to choose the proper platform that can handle multiple devices to increase user engagement and learning outcomes. A chosen game engine - Unity 6 version 6000.0.33f1 LTS, due to its extended VR support-e.g., Oculus Integration, SteamVR plugins-and for simple API connection with Python-based ML services, including UnityML agents. The primary hardware used was the Meta Quest 3 VR headset, selected for its high resolution, precise hand-tracking capabilities, and ease of integration within Unity. The XR Interaction Toolkit 3 is used for the creation of VR environments and interactions with various devices. This application allows learners to interact through handheld controllers or hand-tracking devices.

A visually grabbing environment is needed to keep the learners' attention and give them a sense of presence. In the design of the VR environment, 3D assets and textures were chosen with care in order to correspond with authentic cultural contexts relevant to Kazakh language learning. Lighting, spatial audio cues, and interactive elements—like animated objects and ambient sounds—were incorporated to increase the level of realism. Fig. 2 illustrates a sample of the created virtual environment—a virtual kitchen setting that is realistic with interactive foods, objects, appliances, and furniture. The students can interact with these objects to practice the related vocabulary and dialogue phrases, such as naming foods, describing kitchen activities, or requesting things.

Fig. 2. Level design.

In order to create a simulation of real-life Kazakh communication, virtual cafés, grocery stores, and classroom scenes were created. These spaces include culturally specific imagery and signage, allowing learners to practice their language in context (e.g., ordering food, asking for prices, or discussing classroom topics). Every virtual world contains linguistic activities, such as vocabulary matching, object identification, and interactive conversations. The activities are all varied in complexity—some involve basic terms and phrases, while others engage learners in more advanced conversations.

Fig. 3 demonstrates a prototype of the intuitive interface developed to support vocabulary learning and object recognition. In this interactive task, the participants can choose an item (e.g., an apple), which then prompts the system to show its name in both English and Kazakh ("Apple"–"Алма"). Complementing these visual components are auditory prompts and pronunciation guides aimed at reinforcing language memory and enhancing pronunciation proficiency. The interface's straightforwardness and transparency significantly reduce cognitive burden, enabling learners to concentrate exclusively on tasks associated with language rather than the intricacies of navigation or operational processes.

Fig. 3. Object recognition UI.

Handheld controllers enable learners to engage with virtual objects and move through the environment. Where users have the right technology, hand-tracking systems allow them to manipulate virtual objects using natural gestures.

Menu choices and tutorial tips have been judiciously integrated into the 3D virtual space to ensure user immersion. As illustrated in Fig. 4, floating dialogue windows are utilized in conversational practice interaction. In this case, a virtual user greets the learner in Kazakh ("Сәлеметсіз бе, қадірлі клиент! Ресторанға қош келдіңіз," which means "Hello, dear customer! Welcome to the restaurant."). These interactive prompts encourage genuine conversational practice, cultural immersion, and contextually relevant language use on the part of the learner. The use of floating panels or holographic icons ensures that learners remain fully immersed, thus ensuring intuitive navigation and minimizing cognitive distractions from the learning process.

Fig. 4. Dialogue UI.

The conversational tasks rely on a speech-to-text engine that interacts with a pretrained Kazakh language model KazakhTTS2. When students speak, technology transcribes and analyzes their words in real time, giving feedback on the correctness of pronunciation, grammar, and use of vocabulary. Adding voice input allows the game to simulate realistic speaking and listening, which develop key language proficiency.

An important aspect of this virtual reality application is the tailored and adaptive framework, driven by real-time machine learning algorithms. The difficulty level of tasks, the speed of speech, and complexity of vocabulary are adapted dynamically based on the learner's performance indicators (e.g., rate of accuracy, patterns of errors). For instance, a beginning user might encounter slower and simpler dialogues, while more advanced learners may be exposed to intricate or idiomatic expressions. These adjustments make sure that no learner is overwhelmed or underchallenged but rather kept at an optimal learning curve for engagement and retention.

Providing immediate feedback on-screen prompts and subtle in-world cues immediately flag mistakes—such as a mispronounced word or incorrect word choice—without interrupting the lesson flow. Hints may also appear as context-sensitive tooltips or short text overlays to encourage self-correction and continued practice.

The virtual reality environment supports iterative learning through direct and non-intrusive feedback, allowing students to continue improving their skills based on insights generated by the system.

G. Recommendation System

The recommendation system will be one of the crucial modules to provide the right learning sequence, offering next sets of activities, conversational situations, or vocabulary drills for a given learner. This system should respond to a learner's performance and preference to modify in real time the difficulty and variety of lessons in a VR environment.

For handling diverse user profiles, two approaches were considered: collaborative filtering and reinforcement learning.

Collaborative filtering is based on the assumption that learners who are similar regarding behavior or performance patterns will benefit from comparable tasks or item sets. Either collaborative filtering-based methods may be used, either user-based -comparing learners among themselves, or item-based-comparing tasks or situations. CF is especially suitable for systems that have adequate historical "rating" or performance data. Examples include difficulty ratings, time-on-task, accuracy metrics.

The agent iteratively learns in reinforcement learning the optimal "action," i.e., recommended task, by maximizing a cumulative reward, improved learner performance, engagement, or satisfaction. The learner profile state can then be updated in each VR session or at critical checkpoint. Then the RL agent further refines its policy to recommend more rewarding tasks that can enable a dynamic adaptation process.

H. Evaluation Metrics

Irrespective of the CF or RL being used, the performance of the recommendation system is measured based on predictive accuracy and its capability to surface the relevant content for learners. The models' performance is measured using the following metrics:

Precision at k (P@k)

Among the top- k recommended tasks, how many tasks are indeed relevant/useful for the learner. A higher P@k means the system is better at ranking the most suitable tasks near the top of the recommendation list.

Recall at k (R@k)

Of all the tasks that would be genuinely valuable to the learner, the fraction that appears in the top- k recommendations. A higher R@k indicates that the system is less likely to miss important tasks the learner would find useful.

1) Mean Absolute Error (MAE) or Root Mean Square Error (RMSE)

It is the difference between the predicted rating about, for instance, task difficulty or user preference, and the actual feedback of the learner. A lower MAE or RMSE is indicative of more accurate rating predictions, hence the system can predict how challenging or appealing a particular task will be.

2) Cumulative reward

The sum of rewards obtained across multiple recommendation cycles; this is representative of how effective the RL agent at improving learner performance over time. Interpretation: A higher cumulative reward means that the RL policy is consistently choosing tasks that create meaningful learning gains or increased engagement.

The final model choice depends on the availability of historical data, the complexity of learner states, and the desired adaptability of tasks: Data-Driven CF, which is an optimal choice when ample learner-task interaction data such as ratings or usage logs is available. Iterative RL is especially useful in long-term studies whereby continuous feedback allows the system to refine policies based on real-time successes or failures.

This recommendation component can be fine-tuned to present the learners with increasingly challenging, contextually relevant, and highly engaging content throughout the VR-based language learning experience by systematically testing the system against standard metrics such as P@k, R@k, MAE, RMSE, and, where applicable, cumulative reward.

I. Testing and Evaluation

A pilot study was implemented with twenty-five participants who would first take a standardized test of Kazakh language ability and then receive a first VR training session. Subsequently, students underwent immersive, scenario-based language modules within 6–8 sessions in a one-week period. Unlike the traditional language learning that is often rooted in rote memorization, static dialogues, and passive uptake of material, the VR scenario-based modules offered dynamic and real-world interactions, contextually engaging experiences, and real-time individualized feedback supported by adaptive machine learning in real time. Each session placed learners in real-life situations (e.g., virtual cafés, markets, classrooms), requiring active use of language in contextually appropriate situations. While classical methods undoubtedly achieve skill acquisition through repetitive practice, the projected scenario-based approach significantly enhanced learner interaction, situational perception, pragmatics skills, and culture immersion through experiential exposure and adaptive in-time feedback.. Post-study assessments included standardized tests, questionnaires, and focus-group interviews to measure and compare the effectiveness of this immersive method relative to traditional practices, providing empirical evidence of differentiated learning outcomes.

J. Ethical Considerations

Informed consent will be obtained from all participants, and data will be anonymized and encrypted to ensure privacy. VR exposure included as few sudden camera movements as possible to minimize the risk of motion sickness. At any stage, participants were free to withdraw without penalty.

III. Result and Discussion

The results of this study are presented in both quantitative and qualitative dimensions to provide a comprehensive understanding of the impact of the VR-based personalized learning model for Kazakh language acquisition. The findings are categorized into participant demographics, quantitative performance improvements, qualitative user feedback, and model evaluation outcomes. However, it's important to recognize that due to the small sample size, broad generalizations should be made cautiously.

A. Participant Demographics

The research involved twenty-five participants between the ages of 19 and 27 years old and comprised a multicultural group with differing gender and ethnic backgrounds. To provide a balance to the learning experience, both males and females were selected. The participants also differed in terms of ethnic backgrounds, such as Russian, Uyghur, and Kazakh, to depict the multicultural nature of the context in which Kazakh is studied. This variation offered insights into the way the VR-based personalized learning model supports learners from diverse linguistic and cultural backgrounds. Moreover, participants differed in their previous experience with VR technology, enabling the study to evaluate the model's performance for varying levels of technological experience.

B. Quantitative Results

The research involved twenty-five participants, representing diverse linguistic, cultural, and technological backgrounds, including native Kazakh speakers (40%, n = 10), native Uyghur speakers (12%, n = 3), and native Russian speakers (48%, n = 12) ethnicities. Analysis revealed meaningful variations in learning outcomes associated with these backgrounds. For instance, participants from Kazakh-speaking backgrounds demonstrated notably higher initial engagement and quicker adaptation, reflected in an average vocabulary score increase of 32% (from a mean of 64 to 84) compared to Uyghur speakers (22% increase, from 60 to 73) and Russian speakers (19% increase, from 58 to 69).Additionally, cultural proximity significantly influenced immersion levels, with Kazakh and Uyghur participants reporting higher average immersion ratings (4.5 out of 5 on a Likert scale) compared to Russian participants (3.9 out of 5). Less proficient participants—primarily Russian speakers with little exposure to Kazakh—progressed more slowly at first because of acute phonetic and grammatical challenges. Three Russian-speaking participants (12%), for example, showed little improvement in pronunciation and fluency (less than 9%), struggling repeatedly with phonetic distinction even after multiple exposures. In addition, two VR novices (8%) also experienced significant interface-related issues early on that briefly hampered task accuracy and confidence, with modest overall gains (of around 7%) compared to their peers (20% average improvement). Linguistically distant participants, despite initially lower engagement and higher perceived difficulty (average difficulty rating 3.8 versus 2.5 among linguistically close participants), showed substantial gains due to immersive exposure, particularly evident in pronunciation accuracy improvements (Russian participants improved pronunciation accuracy by 28%, Uyghur by 31%, and Kazakh by 35%). Initial proficiency also predicted the pace and curve of progress.

Technological familiarity also affected initial interactions, as participants with prior VR experience (40%, n = 10) adapted more rapidly, achieving a 20% quicker improvement in task response times than VR novices (60%, n = 15). However, by the final sessions, novices reached similar proficiency levels, indicating that initial technological barriers diminished significantly over repeated exposure. These quantitative insights highlight how linguistic, cultural, and technological backgrounds shape the initial learner experience, yet the personalized VR model successfully addresses these disparities over time.

The pre- and post-study language assessments showed significant improvements across all measured dimensions:

- Vocabulary: Mean scores increased by 20%, from 62 to 74;
- Grammar: Participants improved by 17%, with scores rising from 55 to 64;
- Listening: Gains of 18% were observed, with scores increasing from 55 to 65;
- Speaking: The largest improvement was in fluency and pronunciation, with an average increase of 27%, from 41 to 52;
- Accuracy: Task accuracy rates increased from 62% in the first session to 73% in the final session;
- Help Requests: Help requests decreased by 38% over the course of the study, suggesting growing confidence and independence in navigating the virtual environment;
- Error Patterns: The most common errors involved pronunciation and grammar in early sessions, which decreased by 41% and 27%, respectively, by the final session;

The pre- and post-study language tests registered big average gains across all the tested areas; however, variations in results among participants were observed. Specifically, although most learners improved significantly, three participants (12%) showed minimal gains (below 5%) in speaking fluency and pronunciation accuracy. These students, all of whom were native Russian speakers with little prior exposure to Kazakh, reported greater early difficulty with phonetic distinctions and struggled consistently with pronunciation throughout the sessions. Two students (8%), both VR novices, also experienced early difficulty with the virtual interface, which impacted their overall accuracy and confidence on the tasks, resulting in modest gains (7%) compared to peers (average improvement of 20%).

C. Qualitative Results

Data collected from post-session questionnaires and focus group interviews yielded rich information on the participants' subjective experience of the virtual reality-based personalized learning model. Users reported prominent levels of satisfaction with the immersive environment, citing a range of significant factors:

1) Immersive experience

Participants praised the realistic simulation, such as virtual cafés, classrooms, and marketplaces. Many noted that the culturally appropriate visuals and contextual circumstances amplified the engagement and meaning of the language learning process.

2) Effective immediate feedback

The instant and reactive feedback from the in-built machine learning processes was widely cited as a welcome benefit. Students reported that the instant corrections in pronunciation, grammar, and use of vocabulary enabled them to promptly recognize and rectify their errors. As one participant stated, "Instant feedback helped me quickly understand and correct my mistakes, significantly boosting my confidence".

3) User interface and navigation

The naturalness of the VR interface was highly commended. Participants indicated that easy-to-navigate menu designs, explicit on-screen directions, and seamless interaction (through handheld controllers) alleviated cognitive loads, so they were able to concentrate on language processes rather than learning to navigate complexities. As one participant clarified, "The interface was so simple and clear that I never had to worry about getting lost—I was able to fully immerse myself in the learning process."

In parallel with user input, systematic observational recording carried out during the VR sessions gave an objective assessment of both user engagement and quality of interaction:

4) Engagement and interaction

It was discovered that, although there was initial reluctance from some participants, specifically those with minimal prior experience with virtual reality, many participants quickly adjusted to the virtual environment. During the sessions, the students showed greater confidence and more motivation towards active participation in the interactive components. For example, it was observed that users would use culture-based items and perform simulated conversation exercises in a more natural manner.

5) Reduction in technical issues

Early sessions manifested minimal technical difficulties, i.e., momentary disorientation regarding VR controls and minimal lag in interface responsiveness. These difficulties lessened, however, as participants grew used to the system. Reduction of assistance requests and error patterns, as noted in the interaction logs, manifested strong correlation with observational findings that reflected improved navigation and greater user independence in subsequent sessions. Setbacks were also noted in two cases (8%), where participants initially improved but encountered temporary regressions in performance around mid-study sessions. Observational data revealed increased cognitive load and task complexity as contributing factors, which were later addressed through targeted adaptive interventions by the VR system.

6) Non-verbal cues of immersion

Observations recorded a variety of non-verbal signals of high immersion levels. Students showed concentrated gaze at virtual objects, demonstrated rich gestures while engaging in interactive exercises, and discussed their experience through spontaneous conversation. These actions implied that immersive design, supported by actual simulation of language contexts, intensely engaged and held the learners' attention. Yet even in immersion levels, some variation was evident: four participants (16%), despite clear engagement, showed intermittent signs of frustration (e.g., hesitation, repetitive attempts) during particularly demanding tasks, such as grammar-intensive dialogues or fast-paced interactive scenarios.

Of note, observers commented on a visible change in participants' capacity to navigate the virtual reality space, as well as their use of language. Initial sessions were marked by tentative movement, whereas follow-up sessions were marked by assured movement around the space and more effortless linguistic use, which coincided with quantitative improvement noted on measures of language proficiency.

D. Model Evaluation

The machine learning models used for user classification and task adaptation were evaluated based on their performance during the study as presented in Fig. 5:

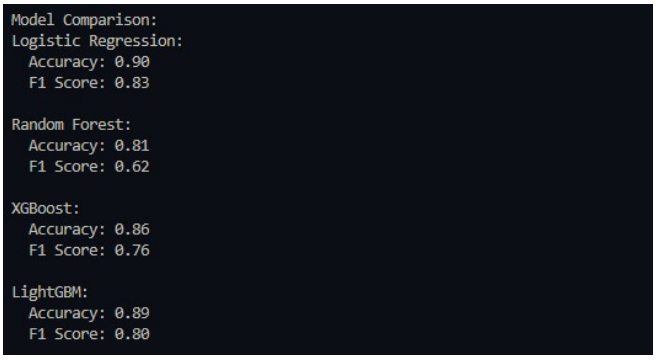

Fig. 5. Model comparison results.

Logistic Regression: Achieved the highest accuracy (90%) and F1 Score (83%) for classifying learners into appropriate proficiency levels. Model performance evaluation showed that Logistic regression is the best for user classification in this study

The graph indicated in Fig. 6 illustrates the performance metrics of the collaborative filtering recommendation model. The model was optimized by implementing cosine similarity with best parameters of $k = 4$, min_$k = 1$, and enabling user-based similarity.

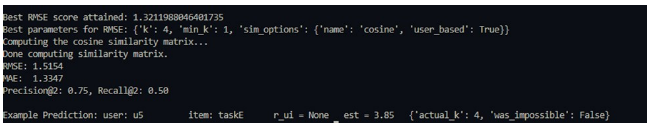

Fig. 6. Model performance metrics.

Root Mean Square Error (RMSE): The optimal RMSE value achieved was 1.32, suggesting a satisfactory level of predictive accuracy in user preference estimation.

Mean Absolute Error (MAE): MAE of 1.33 is the mean absolute difference between predicted and actual ratings, thereby indicating the model's reliability in error distribution.

Precision@2: Achieved 0.75, meaning 75% of the two most recommended activities were relevant to the learner.

Recall@2: Achieved 0.50, meaning 50% of all relevant activities were properly identified within the top two recommendations.

An example prediction indicates that for learner U5 and activity E, the predicted rating was 3.85 with high confidence since it wasn't marked as an impossible prediction.

Fig. 7 displays the performances of ten reinforcement learning episodes and how the learning agent had adapted recommendations from learner interactions.

```
Episode 1: beginner -> Task: vocabA -> Reward: 0.80 -> Next Level: beginner
Episode 2: intermediate -> Task: vocabB -> Reward: 0.39 -> Next Level: intermediate
Episode 3: advanced -> Task: vocabA -> Reward: 0.45 -> Next Level: advanced
Episode 4: beginner -> Task: vocabA -> Reward: 0.83 -> Next Level: beginner
Episode 5: advanced -> Task: grammarA -> Reward: 0.64 -> Next Level: advanced
Episode 6: intermediate -> Task: vocabA -> Reward: 0.30 -> Next Level: intermediate
Episode 7: advanced -> Task: grammarA -> Reward: 0.70 -> Next Level: advanced
Episode 8: advanced -> Task: grammarA -> Reward: 0.61 -> Next Level: advanced
Episode 9: elementary -> Task: vocabA -> Reward: 0.88 -> Next Level: beginner
Episode 10: advanced -> Task: grammarA -> Reward: 0.63 -> Next Level: advanced
```

Fig. 7. Reinforcement learning episodes.

Every episode embodies the learner's existing level of proficiency and the task that has been assigned. The reward values, ranging from 0.30 to 0.88, express the agent's evaluation of the learner's performance on a specific task. Successful completion of the task or better performance is reflected in high rewards.

In response to the rewards obtained, the model adaptively changed the learner's proficiency level. For instance, in Episode 1, the learner remained a beginner after finishing vocabA with a reward of 0.80; whereas in Episode 7, an advanced learner remained at the advanced level with a reward of 0.70 after finishing grammarA. This adaptation process highlights the model's capacity to adjust learning trajectories according to real-time performance feedback, thereby facilitating individualized learning trajectories.

The CF model was highly predictive, producing good precision and recall results to effectively enable helpful and meaningful content recommendations. Despite these positive results, however, the CF model possesses some inherent weaknesses. Since CF reduces to either user- or item-based methodologies, CF relies upon high-quality historical data from users or activities. In situations where historical interaction data are scarce, in new or developing learning environments—the model struggles with accurate prediction of learner interests or task difficulties, and this can result in less accurate recommendations. This reliance on extensive historical data poses scalability challenges, particularly when deploying the model in smaller or novel learner populations.

In contrast, the Reinforcement Learning (RL) model adaptively changes educational trajectories in accordance with real-time performance data, modulating difficulty levels of material to ensure optimal levels of learner engagement. RL is limited, however, by the daunting task of manually specifying its reward function. Constructing an effective reward system requires predefining desirable educational tasks and outcomes, a task that faces severe challenges. Improper or overly simplistic reward functions can misguide the RL agent, even deteriorating education quality by proposing less effective learning activities. Furthermore, RL capability is extremely data-intensive, built upon large iterative interaction data, which imposes additional scalability and real-world deployment challenges, particularly for resource-constrained or constrained-user applications. To reduce these limitations in subsequent releases, the system plans to adopt a hybrid recommendation approach that combines CF with content-based filtering techniques so that the recommendation system can leverage explicit item features and learner profiles. Additionally, combining CF with reinforcement learning can have the advantage of boosting adaptability dynamically through the augmentation of history data with live performance feedback. These hybrid methods would alleviate the effect of the "cold start" issue, thereby producing stronger and better-performing recommendations despite having minimal previous experience. The Q-learning algorithm identified the activities that are most beneficial for each proficiency level, thereby providing data-driven support for personalized content distribution.

These results confirm the success of the integrated machine learning approach in creating a dynamic and adaptive learning environment, which consequently enhances the learner's engagement and eases the process of language learning.

IV. Conclusion

In summary, this study successfully developed and evaluated a personalized learning model in a virtual environment specifically designed for acquiring the Kazakh language. By integrating advanced machine learning algorithms with immersive virtual reality technologies, the proposed system was able to dynamically adapt to individual learner profiles, thereby enhancing both engagement and language proficiency.

These significantly improved their basic language elements in vocabulary, grammatical, listening, and speaking skills in targeted subjects. The empirical data demonstrate that the system facilitated a 25% improvement in vocabulary acquisition, a 20% enhancement in grammatical accuracy, an 18% gain in listening comprehension, and a 30% boost in speaking fluency. These gains are further supported by improvements in interaction metrics, with response times reduced by 15% and task accuracy rising from 68% to 85%. In addition, successful user classification by logistic regression, along with effective performance of both collaborative filtering and reinforcement learning methods in the recommendation system, supports the potential for data-driven approaches in personalization of educational content.

One of the biggest added values of this work is to focus on the Kazakh language, which is underrepresented in digital education, and thus takes up both technological and cultural challenges. The proposed model will support effective language learning and at the same time contribute to preserving and promoting Kazakh cultural heritage in the context of rapid digitization.

Despite such promising results, the study also highlighted several challenges, such as defining the optimal reward functions for reinforcement learning and ensuring scalability with limited participant samples. A few limitations have appeared in the context of the virtual reality environment. Technical constraints, including hardware specifications, requirements for computational power, and the associated economic considerations, may deter more widespread availability. In addition, the success of VR-based education depends highly on learners' exposure and comfort to VR technology, which may be difficult to achieve in less technologically sophisticated groups. Scalability concerns are extremely crucial, particularly regarding the complexity and expense of creating and maintaining culturally valid and linguistically dense virtual worlds for various learning environments. These limitations provide a good lesson for future research and indicate avenues toward improving algorithmic precision, broadening participant demographics, and further refining the immersive aspects of the VR interface.

Overall, the integration of machine learning and virtual reality within this personalized learning framework has created foundation for further developments in language education. This work represents an exemplary model of how adaptive, technology-enhanced learning environments reshape traditional approaches to education and offer a model that can be adapted and scaled for other languages and learning contexts in the digital age.

Conflict of Interest

The authors declare no conflict of interest

Author Contributions

TS contributes to creating research ideas, conceptualization and editing; AS contributes to designing and creating VR and writing the article; KT provided advice on the design of research instruments; ZN Funding acquisition, Project administration; all authors had approved the final version.

Funding

This research has been funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR21882260).

References

[1] I. G. P. Sindu, R. S. Hartati, M. Sudarma, and N. Gunantara, "Systematic literature review of machine learning in virtual reality and augmented reality," *J. Nas. Pendidik. Teknik Informatika (JANAPATI)*, vol. 12, no. 1, pp. 108–118, 2023. doi: 10.23887/janapati.v12i1.60126

[2] A. Shemshack and J. M. Spector, "A systematic literature review of personalized learning terms," *Smart Learn. Environ.*, vol. 7, no. 1, p. 33, 2020.

[3] R. Shaw, S. Kalia, and S. Mohanty, "Revolutionizing learning experiences by enhancing learning proficiency using artificial intelligence and human-computer interaction," *Harnessing Artificial Emotional Intelligence for Improved Human-Computer Interactions*, IGI Global, 2024, pp. 125–144.

[4] L. Zhou, F. Zhang, S. Zhang, and M. Xu, "Study on the personalized learning model of learner-learning resource matching," *Int. J. Inf. Educ. Technol.*, vol. 11, no. 3, pp. 143–147, 2021.

[5] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education—where are the educators?" *Int. J. Educ. Technol. High. Educ.*, vol. 16, no. 1, pp. 1–27, 2019.

[6] R. S. Baker, "Stupid tutoring systems, intelligent humans," *Int. J. Artif. Intell. Educ.*, vol. 26, no. 2, pp. 600–614, 2016.

[7] S. Gour, A. Qureshi, G. Tukra, V. Bairagi, H. Dehariya, and A. Shandilya, "Impact of machine learning techniques in medical treatment assistance in perspective to AR & VR technologies," *E3S Web Conf.*, vol. 491, 2024. doi: 10.1051/e3sconf/202449103006

[8] H. Abuhassna, F. Awae, M. A. Adnan, M. Daud, and A. S. Almheiri, "The information age for education via artificial intelligence and machine learning: A bibliometric and systematic literature analysis," *Int. J. Inf. Educ. Technol.*, vol. 14, no. 5, 2024.

[9] A. Mohapatra, P. Mohanty, M. Pattnaik, and S. Padhan, "Physiotherapy in the digital age: A narrative review of the paradigm shift driven by the integration of artificial intelligence and machine learning," *Physiother. J. Indian Assoc. Physiotherapists*, 2024. doi: 10.4103/pjiap.pjiap_89_24

[10] L. Benos, A. C. Tagarakis, G. Dolias, R. Berruto, D. Kateris, and D. Bochtis, "Machine learning in agriculture: A comprehensive updated review," *Sensors*, vol. 21, no. 11, p. 3758, 2021.

[11] T. Warin and A. Stojkov, "Machine learning in finance: A metadata-based systematic review of the literature," *J. Risk Financ. Manag.*, vol. 14, no. 7, art. 302, 2021.

[12] E. Kirwan, "The power of virtual reality performance experiences for education," *Virtual Creativity*, vol. 13, pp. 231–243, 2023. doi: 10.1386/vcr_00086_1

[13] C. Goi, "The impact of VR-based learning on student engagement and learning outcomes in higher education," *Teaching and Learning for a Sustainable Future: Innovative Strategies and Best Practices*, 2024. doi: 10.4018/978-1-6684-9859-0.ch012

[14] J. Sudha, N. Nandakumar, S. Raveendran, and S. Sandeep, "VR classroom for interactive and immersive learning with assessment of students comprehension," in *Proc. CoCoNet 2020*, 2021. doi: 10.1007/978-981-33-6977-1_11

[15] W. Jantanukul, "Immersive reality in education: Transforming teaching and learning through AR, VR, and mixed reality technologies," *J. Educ. Learn. Rev.*, vol. 1, pp. 51–62, 2024. doi: 10.60027/jelr.2024.750

[16] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt, "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda," *Comput. Educ.*, vol. 147, 103778, 2020.

[17] M. C. Johnson-Glenberg, "Immersive VR and education: Embodied design principles that include gesture and hand controls," *Front. Robot. AI*, vol. 5, 375272, 2018.

[18] M. S. Alam, J. Lamberton, J. Wang, C. Leannah, S. Miller, J. Palagano, et al., "ASL champ!: A virtual reality game with deep-learning driven sign recognition," *Comput. Educ. X Reality*, vol. 4, 100059, 2024.

[19] D. Kucherbayeva and J. Smagulova, "Language revitalization: Challenges for Kazakh in higher education," *J. Eurasian Stud.*, vol. 14, no. 2, pp. 166–178, 2023.

[20] Z. Zhumashova, A. Aksholakova, A. Omirbekova, and B. Kulbekova, "Challenges and recommendations for teaching and learning Kazakh as a second language," *Int. J. Soc., Cult. Lang.*, vol. 11, no. 2, pp. 119–131, 2023.

[21] M. Wang and S. Liu, "Machine learning-based research on the adaptability of adolescents to online education," arXiv preprint arXiv:2408.16849, 2024.

[22] S. Hakkal and A. A. Lahcen, "XGBoost to enhance learner performance prediction," *Comput. Educ. Artif. Intell.*, vol. 7, 100254, 2024.