REPUBLIC OF AZERBAIJAN

On the rights of the manuscript

ABSTRACT

of the dissertation for the degree of Doctor of Philosophy

APPLICATION OF DEEP LEARNING TECHNOLOGIES IN GAME PLATFORMS BASED ON AUGMENTED REALITY

Specialty: 1203.01 – Computer Science

Field of science: Technical sciences

Applicant: **Anar Azer Mammadli**

Baku-2026

The work was performed at "General and applied mathematics" departments of Azerbaijan State Oil and Industry University.

Scientific supervisors:	Honored scientist, doctor of mathematical sciences, prof. **Araz Rafig Aliyev**
	Doctor of Philosophy on engineering **Elviz Anvar Ismayilov**
Official opponents:	Doctor of technical sciences, prof. **Aqil Hamid Huseynov**
	Corresponding Member of ANAS, doctor of technical sciences **Ramiz Mahammad Aliguliyev**
	Doctor of Philosophy on engineering, assoc. prof. **Ramiz Balashirin Alekperov**

Dissertation council FD 2.48 of Supreme Attestation Commission under the President of the Republic of Azerbaijan operating at the Azerbaijan State Oil and Industry University

Chairman of the Dissertation council:	Corresponding Member of ANAS, doctor of technical sciences, prof. **Rafig Aziz Aliyev**
Scientific secretary of the Dissertation council:	Doctor of technical sciences, assoc. prof. **Akif Vali Alizadeh**
Chairman of the Scientific seminar:	Honored scientist, doctor of technical sciences, prof. **Latafat Abbas Gardashova**

"AZƏRBAYCAN RESPUBLİKASI ELM VƏ TƏHSİL NAZİRLİYİ AZƏRBAYCAN DÖVLƏT NEFT VƏ SƏNAYE UNİVERSİTETİ" PUBLİK HÜQUQİ ŞƏXSİ
ELMİ -KATİB

dos. Aliyev

GENERAL DESCRIPTION OF THE DISSERTATION

Relevance of the topic and degree of study. Throughout history, games have been used not only for entertainment, but also as a means of developing cognitive abilities, strategic thinking, and problem-solving skills. Game-based learning strengthens learning through experience and makes the educational process more engaging[1]. A substantial body of research indicates that gamified learning approaches can enhance motivation, support memory retention, and improve logical reasoning. As a result, games are increasingly viewed as effective learning tools in both technical disciplines such as mathematics, chemistry, and logic, and in humanities-oriented fields including language learning, history, and the social sciences. In today's context, where interactive and self-directed learning methods are rapidly expanding, games have evolved into multifunctional learning platforms suitable for a wide range of age groups.

The adoption of Artificial Intelligence (AI) and Augmented Reality (AR) in game design is widely considered one of the most significant recent shifts in the game industry[2]. Contemporary games use AI and AR to deliver more personalized, interactive, and immersive experiences, including the creation of three-dimensional environments that surround the player. AI contributes to smarter and more adaptive gameplay by enabling automated processing of game data, more sophisticated behavior models for bot characters, adaptation to player decisions, and real-time adjustments to game dynamics. These capabilities can meaningfully improve overall game

[1] Barab, S. Situationally embodied curriculum: Relating formalisms and contexts / S. Barab, S. Zuiker, S. Warren [et al.] // Science Education, –Hoboken: – 2007. 91 (5), – p. 750–782

[2] Mitsea, E., Drigas, A., Skianis, C. A systematic review of Serious Games in the era of Artificial Intelligence, Immersive Technologies, the Metaverse, and Neurotechnologies: transformation through Meta-Skills training // – Basel: Electronics, – 2025. 14 (4), – p. 649.

quality. AR strengthens immersion by connecting virtual content with the physical world. By reducing constraints in both digital and physical games, AR allows developers to provide richer media content and more interactive systems. It enables players to interact with virtual objects within real spaces, enriches storylines through real-world elements, and makes the experience more vivid and realistic[3]. When integrated with mobile devices or AR glasses, AR further supports real-time, changing environments, personalized tasks, and virtual events blended with real-world contexts, contributing to a new stage in the evolution of the industry.

This technological direction is not confined to entertainment; it is also widely applied in serious games and educational games. AI-based systems can adapt to a learner's pace, analyze strengths and weaknesses, and propose individualized development paths, creating new opportunities for education. AR supports the development of simulation-based learning environments by enabling direct interaction with real-world objects. Medical and instructional applications, driving simulators, and safety training are among the domains where such approaches have been particularly successful[4]. In practice, AI and AR have demonstrated value across a range of game-related contexts. Modern AI-driven systems can collect and analyze telemetry data such as gestures and facial expressions, movement trajectories, controller, keyboard, and touch inputs, decision times, gaze direction, behavioral sequences, and team interactions. Based on these signals, in-game responses can be adjusted dynamically, enabling experiences that are more realistic, context-sensitive, and strongly personalized. This also supports the automatic generation and classification of new content. At the same time, relationships among in-game objects and patterns

[3] Martin, J., Bohuslava, J., Igor, H. Augmented reality in education 4.0 // International Scientific and Technical Conference on Computer Sciences and Information Technologies, – Lviv: – 11-14 September, – 2018. 1, – p. 231–236.
[4] Wu, H.K. Current status, opportunities and challenges of augmented reality in education / H.K. Wu, S.W.Y. Lee, H.Y. Chang [et al.] // Computers & Education, – Oxford: – 2013. 62, – p. 41–49.

inferred from player interaction provide important sources of information for these processes. Such approaches have been especially common in simulation-oriented 2D and 3D educational game platforms and have produced promising results. Taken together, these factors highlight the growing relevance of AI and AR in game design and point to the need for further research. For instance, airlines have developed specialized simulation platforms enriched with AI for pilot training, and space agencies have adopted similar systems for astronaut training. Educational game platforms across mobile, computer, and other devices likewise employ these approaches for data analysis, content generation, and automated classification.

The application of AI and information technologies in education broadens interactive and personalized learning opportunities, allowing learners to acquire new knowledge more efficiently. AI-based systems are used across various educational functions, including:

- **Intelligent tutoring systems:** AI-driven adaptive learning systems personalize instructional materials according to students' proficiency levels and support the development of weaker competencies.
- **Educational agents and virtual classrooms:** AI-supported virtual agents and learning environments enhanced by virtual reality and AR technologies promote more active learner participation.
- **Learning analytics and prediction:** AI analyzes students' academic performance, forecasts future outcomes, and supports the construction of individualized learning plans.
- **Gamification and interactive environments:** The use of game elements within education increases motivation and makes learning processes more engaging.

These approaches demonstrate the feasibility of effectively integrating AI into teaching and learning and establish a foundation for broader adoption of information technologies in the future. In particular, AI-based educational technologies especially serious digital games incorporating AI techniques expand learners' opportunities by providing interactive and individualized learning

experiences while maintaining engagement through entertaining elements. Within the dissertation, the principles for constructing a system to classify game and gamified platform materials in Azerbaijani and English were developed; a structured database was created; and an augmented digital environment was formed using data collected from relevant platforms. In addition, the dissertation investigated the effectiveness of various Deep Learning (DL) methods for classifying game elements.

Object and subject of the research. The object of the research comprises serious digital games used for educational purposes, as well as the processing and classification of textual information within these games; the subject of the research is the application of Natural Language Processing (NLP) and explainable artificial intelligence methods in such environments.

Purpose and objectives of the research. The dissertation aims to investigate the effectiveness of applying deep learning and augmented reality in game design and to analyze how these technologies enhance the game experience. This is achieved by modelling a word game used for educational purposes and improving the quality of the game through the application of deep learning methods.

To achieve the stated aim, the dissertation identifies the following tasks::

• Investigation of game design opportunities enabled through the application of AI;

• Development of the “Yasaq” word game in the Azerbaijani language for educational and entertainment purposes;

• Application of NLP methods to enable automatic data updates in the “Yasaq” game;

• Application of deep learning methods to the classification of words in the Azerbaijani language;

• A comparative analysis of different word embedding approaches for determining word similarity in Azerbaijani;

• Use of an ontology schema to enable the explainability of “black-box” methods such as deep learning;

• Integration of new functionalities into the word game through augmented reality.

Research methods. To address the objectives of the dissertation, the study employs NLP techniques, the word game as an information source, technologies for developing serious digital games, AI and deep learning models, and augmented reality.

Key propositions submitted for defense.

- Investigation and practical implementation of opportunities to improve serious digital game design through the application of AI and NLP technologies;
- Comparative analysis of machine learning methods such as Naive Bayes, Convolutional Neural Networks (CNN), and BERT (Bidirectional Encoder Representations from Transformers) in classification-oriented decision-making, and their integration into a multilingual educational platform;
- Development and implementation of an explainable artificial intelligence approach within an educational word game;
- Construction of an interactive serious game platform integrated with AR technology and deep learning models.

Scientific novelty. Within the scope of the dissertation, the following original results with scientific novelty were obtained:

- ✓ A conceptual and applied model was developed for integrating AI into serious game platforms and supporting operational, analytics-driven decision-making;
- ✓ An efficient tokenization and vocabulary-alignment mechanism was proposed, adapted to the morphological characteristics of the Azerbaijani language;
- ✓ A Transformer-based BERT model was adapted to Azerbaijani and applied to the digital "Yasaq" game dataset, achieving higher classification accuracy than traditional methods;
- ✓ An explainable AI-based method was developed and applied to the "Yasaq" dataset to validate outcomes produced by "black-box" classification models and to provide interpretable explanations using structured data.

Theoretical and practical significance of the study and

application of results. The theoretical significance of the research lies in the development of new approaches to integrating AI, deep learning, and AR technologies into game design. Within this work, new methodological approaches were proposed for applying NLP and deep learning models in game platforms. The findings contribute to the enrichment of theoretical knowledge in the development of AI-based game platforms and the advancement of serious game mechanics. In addition, improving NLP and tokenization techniques for Azerbaijani, developing multilingual game systems, and analyzing game data through explainable AI approaches have substantial importance for both AI and game design research.

The practical significance of the research is associated with developing and implementing concrete solutions for applying AI and AR technologies in the game industry. The results obtained within the study are aimed at:

- Developing new models for serious and educational games based on the combined use of AI and AR technologies;
- Creating game platforms that support the Azerbaijani language and advancing multilingual systems;
- Integrating deep learning-based classification models into the game experience, increasing the system's level of intelligence through explainable AI technologies, and applying these models in various educational and entertainment applications;
- Developing personalized and adaptive approaches in game design, and building game mechanics that analyze player behavior and change interactively;
- Implementing effective algorithms for the automatic classification of game data and the generation of new content.

Moreover, the results of this study have broad potential for application in education, psychology, linguistics, and the gaming industry. The use of AR and AI technologies to make games more effective learning tools also provides a foundation for future research and innovative projects.

Approbation of the work. The results of the research were presented and discussed at international and national-level

conferences and symposia:

- Ümummilli lider Heydər Əliyevin anadan olmasının 100 illik yubileyinə həsr olunmuş gənc tədqiqatçı və doktorantların respublika elmi konfransı (4-5 May 2023, Bakı, Azərbaycan);
- Kibernetika və İnformatika Problemləri üzrə 5-ci Beynəlxalq Konfrans (28-30 Avqust 2023, Bakı, Azərbaycan);
- 16th International Conference on Applications of Fuzzy Systems, Soft Computing and Artificial Intelligence Tools - ICAFS-2023 (14-15 Sentyabr 2023, Antalya, Türkiyə);
- İnformasiya və kommunikasiya texnologiyalarının tətbiqi üzrə 17-ci IEEE beynəlxalq konfransı (18-20 oktyabr, 2023, Bakı,Azərbaycan);
- Ümummilli lider Heydər Əliyevin anadan olmasının 101 illiyinə həsr olunmuş gənc tədqiqatçı və doktorantların respublika elmi konfransı (4-5 May, 2024, Bakı, Azərbaycan);
- Biliyə əsaslanan və intellektual informasiya və mühəndislik sistemləri üzrə 28-ci beynəlxalq konfrans (11-13 Sentyabr 2024, Sevilya, İspaniya);
- Süni İntellekt:nəzəriyyədən praktikaya beynəlxalq konfrans (19-20 Sentyabr 2024, Naxçıvan, Azərbaycan);
- 3rd International conference on data science and management-2024 (13-14 dekabr 2024, Odişa, Hindistan);
- 6-cı beynəlxalq Boğaziçi elmi araşdırma seminarı (18-19 yanvar 2025, Boğaziçi, Türkiyə);
- Mühəndis problemlərinin həllində idarəetmə və nəzarət məsələsində beynəlxalq konfrans, (13-14 Mart 2025, Bakı, Azərbaycan).

Published scientific works. Within the scope of the dissertation topic, nine scholarly works have been published, including five journal articles and four conference proceedings. Of these publications, four appeared in journals indexed in the Scopus databases.

The name of the organization where the dissertation was carried out. The dissertation work was carried out at the "General and Applied Mathematics" department of Azerbaijan State Oil and Industry University.

Composition of the dissertation work. The dissertation consists of an introduction, three chapters, a conclusion, a list of references, and appendices. The main body of the work comprises 159 pages and includes 18 figures and 18 tables. The reference list contains 111 sources. The overall character count of the dissertation and its structural sections is distributed as follows: Total–229998 characters, Introduction–18695 characters, Chapter one–61771 characters, Chapter two–69321 characters, Chapter three–78981 characters, Conclusion–1207 characters.

CONTENT OF THE DISSERTATION

The introduction substantiates the relevance of the topic and the conducted research, defines the aims and directions of the study, and presents the research object and subject, research methods, the main propositions submitted for defense, the scientific novelties obtained, the practical significance of the results, the approbation of the work, as well as the dissertation's length and structure.

The first chapter The first chapter analyzes issues related to the application of intelligent methods in game design.

In section 1.1, the classes of problems associated with applying intelligent systems to game design are examined, and the current state of applying intelligent systems to in-game elements is analyzed based on the literature and practical experience. This is followed by an analysis of two widely used problem areas in the application of intelligent methods in games. The discussion addresses challenges involved in creating dynamic in-game elements and developing interactive game environments.

The generation of game elements is widely used in the classification of categories of text-based games, the creation of new dialogues, obstacles, and similar elements based on a player's gameplay style, as well as in intent detection of players.

In section 1.2, a comparative analysis of the application of several existing text-processing methods in games is provided. Considerable work has been carried out internationally in the area of text-processing and classification. Among these, approaches applied to games with "Taboo"-style mechanics, as well as to other game types, are highlighted.

Practice shows that using traditional methods in the mathematical formulation of NLP tasks is accompanied by certain difficulties[5]. The mathematical formulation of NLP problems involves several challenges. Because language is informal and ambiguous, it is difficult to represent it through precise mathematical models[6]. Since texts are context-dependent, models must correctly capture and interpret this context. In addition, language data are high-dimensional and discrete in nature, which complicates optimization and analytical procedures. Furthermore, for low-resource languages, limited availability of corpora, dictionaries, and labeled data makes training and deploying such models even more difficult. Addressing these issues requires the development of new approaches; in particular, the introduction of modified text-processing methods remains highly relevant.

In section 1.3, the use of AR technology in game design is analyzed with a focus on overcoming existing limitations and increasing immersion. The combined use of AI and AR methods enriches game platforms with new capabilities, creating immersive experiences and interactive environments. In this case, certain operational constraints of the platform are removed, high-quality control is achieved more rapidly, and overall platform efficiency increases. These constraints may stem from the characteristics of both

---

[5] Mammadli, A.A., Zeynalabdinov, M.E. Integrating word embeddings into textual serious games: techniques and outcomes // Ümummilli lider Heydər Əliyevin anadan olmasının 101 illiyinə həsr olunmuş gənc tədqiqatçı və doktorantların respublika elmi konfransı – 4-5 May, –2024 – Baku,Azerbaijan, –p.52-55

[6] Zhang, Y., Jin, R., Zhou, Z.H. Understanding bag-of-words model: A statistical framework // – Berlin: International Journal of Machine Learning and Cybernetics,– 2010. 1 (1), – p. 43–52

physical and digital games. By mitigating certain size-related limitations on physical or digital platforms, AR enhances interactivity and collaboration by allowing players to use their surrounding environment in combination with augmented elements[7].

The second chapter analyzes methodologies used for applications of explainable artificial intelligence and text processing in games.

In section 2.1, different word embedding methods are examined and the rationale for selecting BERT embeddings within the project is explained. During the study, the following embedding methods were tested: Word2Vec (CBOW and Skip-gram), GloVe, FastText, BERT. When these methods are compared, it is observed that:

- Word2Vec and GloVe produce static embeddings, meaning they provide a single vector per word and cannot distinguish different meanings of the same word depending on sentence context.
- FastText is robust to rare words and unseen word forms, but still does not reflect contextual differences sufficiently well.
- BERT, by contrast, produces contextual embeddings: the vector representation changes according to the word's role within the sentence.

The primary objective of BERT is to encode words by considering context bidirectionally. Whereas traditional Word2Vec and GloVe provide a fixed vector for a given word, BERT outputs dynamic vectors conditioned on context. For example, the word "bank" is represented by different vectors in "river bank" and "financial bank". In serious games and text-processing systems, correctly capturing such

[7] Mammadli, A.A., Sadikhova, G.N., Akbarov, N.Kh. Using image detection technologies in AR based gamification platforms // Ümummilli lider Heydər Əliyevin anadan olmasının 100 illik yubileyinə həsr olunmuş gənc tədqiqatçı və doktorantların respublika elmi konfransının materialları, –Baku: ASOIU,– 4-5 May,– 2023, – p. 351–356

meaning shifts is critically important. The BERT embedding model consists of the following stages:

Stage 1. The text tokenization and a vocabulary is constructed. The goal is to split the text into sub-tokens in an optimal manner so that all words, including rare and new ones, can be handled. The tokens [CLS] (sentence start), followed by the words, and [SEP] (sentence end) are added. Here, [CLS] and [SEP] are special tokens. Each token is assigned an index;

Stage 2. Each word is represented through three different embedding vectors:

- *Token embedding*: Converts each token ID into a numeric vector. Each token ID is replaced by its corresponding vector from the token-embedding matrix. It is learned during training via Masked Language Modeling and other tasks.
- *Position embedding*: Encodes a word's position in the sentence. The self-attention mechanism in the Transformer architecture is used. Self-attention does not inherently encode word order, so additional positional vectors are used to inform the model where each token occurs. In other words, whether "play" occurs at the beginning, middle, or end of a sentence is represented differently, which helps the model capture order and contextual sequence.
- *Segment embedding*: Used to differentiate two sentences. If the input contains two sentences:
    - Sentence 1 tokens receive Segment A (ID 0).
    - Sentence 2 tokens receive Segment B (ID 1).

  If there is only one sentence, all tokens are marked as Segment A.

These three embedding vectors are summed and passed to the Transformer Encoder layer.

Stage 3. To stabilize the final embedding vectors and facilitate learning, Layer Normalization is applied. Standard normalization for each element is computed as:

$$\hat{x}_i = \frac{x_i - \mu}{\sqrt{\sigma^2 + \varepsilon}}$$

- ε – a very small constant
- σ^2 – variance
- μ – mean

As a result, a normalized vector is obtained. To provide the model with flexibility, two learnable parameters are applied:

$$y_i = \gamma \hat{x}_i + \beta$$

- γ – to scale each element up or down;
- β – to shift each element left or right;

These γ və β parameters are learned during training and add additional model capacity. Consequently, each token embedding is normalized internally, layer outputs become more stable, gradients do not explode, and faster convergence is achieved.

Because BERT embeddings are contextual, the same word is represented by different vectors in different sentences. This distinction enables a more accurate understanding of deep meaning and contextual relationships in text.

In section 2.2, the automatic classification of in-game elements using AI is examined. In-game elements may include visual objects, words, or any other objects usable within the game. Using AI and NLP, text is analyzed in terms of semantic meaning, structure, and context, and then assigned to specific categories. This plays an important role, especially in the development of educational games, platforms, and serious games. In addition, the issue of classifying in-game visual objects (images) is also addressed [8]. This process considers feature extraction, processing, and assignment of images to appropriate categories, which is essential for analyzing multimodal

[8] Mammadli, A.A., Ismayilov, E.A. Combining fuzzy deep learning with HPC to classify images // Lecture Notes in Networks and Systems, Springer, – 2024, 1141, – p. 250–257

data alongside text[9]. For text classification, several approaches are analyzed: traditional machine learning algorithms, deep learning models, and Transformer-based methods. To automate text classification, text is first converted into numeric vectors and then passed to classification algorithms. Text classification is used not only for categorization but also for deeper understanding of content and personalization.

Initially, probabilistic and graph-based models such as Naïve Bayes and Graph Convolutional Networks were used. The Naïve Bayes model assumes that there are no dependencies among words and that each word independently contributes to the classification outcome. For this reason, it is fast and computationally lightweight. It performs well on small- and medium-sized datasets because it requires fewer computations and produces results quickly[10]. However, because it does not account for semantic and contextual relationships among words, it performs weakly on long and complex sentences. In tasks where word order and context are important, Naïve Bayes faces significant limitations. Graph Convolutional Networks (GCN), on the other hand, model data over a graph structure, enabling deeper learning of relationships among words and concepts. Here, nodes (words or concepts) and their relations (semantic, syntactic, or conceptual links) are considered. This approach is advantageous for preserving context and discovering complex semantic relationships. However, GCNs are computationally heavier and require more

[9] Mammadli, A.A., Ismayilov, E.A. Application of fuzzy deep learning to classify ımages // IEEE Publisher, Application of Information and Communication Technologies, Proceedings,– 2023. – p. 1–5

[10] Mammadli, A.A. Unlocking educational insights: integrating Word2vec embeddings and naive bayes classifier for serious game data analysis and enhancement // Baku: Azerbaijan Journal of High Performance Computing, – 2023. 6 (2), – p. 191–198

resources. Nevertheless, for long texts and tasks requiring semantic richness, they can yield stronger results than Naïve Bayes[11].

CNN models are mainly used to capture local structures and patterns in text. Convolutional filters detect short patterns in text, which is particularly effective in recognizing local dependencies such as phrases and expressions. CNNs support parallel computation and allow faster training. Their key limitation is that they do not fully capture global context and long-range dependencies between distant words. Thus, although CNNs are strong in learning local features, they may not fully represent the overall semantic meaning of a text. For this reason, running CNNs together with suitable word embedding methods is more appropriate.

BERT is a Transformer-based approach that processes text bidirectionally (left-to-right and right-to-left), allowing each word to be interpreted using both preceding and following context simultaneously[12]. Because BERT embeddings depend on context, the same word is represented by different vectors across different sentences. This provides BERT with exceptional strength in capturing deep semantic relationships and processing overall text structure[13]. BERT is highly effective in modeling long-range dependencies. The BERT model consists of the following stages:

Stage 1.Word embeddings are obtained as in Section 2.2;

---

[11] Mammadli, A.A., Ismayilov, E.A. Application of deep learning technologies in serious games // International Conference on Problems of Cybernetics and Informatics, – Baku: – 28-30 August, – 2023, – p. 1–4

[12] Devlin, J. BERT: Pre-training of deep bidirectional transformers for language understanding / J. Devlin, M.W. Chang, K. Lee [et al.] // . NAACL HLT: Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, – Minneapolis: –2019. 1, – p. 4171–4186

[13] Mammadli, A.A. Application of deep learning for procedural content integration for learning serious games // – Baku: Proceedings of Azerbaijan High Technical Educational Institutions, – 2025. 48 (6), – p. 455–469

Stage 2. The embeddings are transformed into three different learnable projections:

$$Q = XW^Q, K = XW^K, V = XW^V$$

Here:

- $W^Q, W^K, W^V \in \mathbb{R}^{dxd_k}$
- d_k= d/h (h – number of heads).

Stage 3. The Attention Score is computed, measuring each token's compatibility with all other tokens.

$$Attention\ Score = \frac{QK^T}{\sqrt{d_k}}$$

- Q — Query vectors (size: $n \times d_k$),
- K^T— Transpose of Key vectors (size: $d_k \times n$).

As a result, compatibility scores are obtained for each token against all others, producing an $n \times n$ matrix. The $\sqrt{d_k}$ term is used to prevent scores from becoming too large and to improve numerical stability when applying softmax.

Stage 4. The softmax function is applied to the compatibility scores:

$$Attention\ Weights = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)$$

This allows the model to determine probabilistically how much attention to allocate to each word.

Stage 5. A new vector is computed using the obtained weights. This vector carries both self-information and contextual information from other words:

$$Attention\ Output = Attention\ Weights \times V$$

- V — Value vectors (size: $n \times d_k$),

That is, each token receives the weighted sum of the Value vectors of surrounding tokens.

Stage 6. These processes are performed in parallel for each head, and the outputs are concatenated:

$$MultiHead(Q, K, V) = Concat(head_1, \dots, head_h)W^O$$

- $W^O \in \mathbb{R}^{dxd}$ — an additional learnable matrix,

This structure enables the model to:

- Capture the full meaning of a token within context.
- Attend to different semantic aspects (e.g., one head captures grammatical relations while another captures semantic proximity).
- Process both short and long range dependencies simultaneously.

Having multiple heads supports parallel learning of different perspectives.

In section 2.3, the use of ontologies to apply the explainable artificial intelligence (xAI) approach employed in the study is analyzed. An ontology, a structured representation of concepts and their relations is used. Ontologies are employed to structure knowledge domains and to define relationships among concepts. The section explains ontology types, core components (classes, individual instances, attributes, and relations), as well as the tools and methodologies used. Ontologies are also widely applied in medical

diagnostics, education, semantic search systems, and the game industry[14].

The third chapter investigates how AI and AR technologies were applied in developing the text-based "Yasaq" game. The aim is to make the game experience more interactive and personalized by enabling players to classify texts, generate adaptive content, and integrate AI through explainable approaches. At the same time, AR is used to merge the physical and virtual environments and introduce new dynamic elements into the game. This chapter analyzes text classification, explainable AI, and augmented reality aspects in a staged manner.

In section 3.1, the database of the "Yasaq" game is constructed and various AI methods are applied to text classification. The data used for the "Yasaq" game were prepared to improve gameplay through a rich collection of words and associated taboo-term sets. The dataset was collected from various educational resources and user-generated content. First, a dedicated database is created for the game, and the categories to which the words belong are recorded. The full dataset is initially broad and diverse, containing words and taboo expressions from multiple topics. Experiments show that using the full dataset significantly increases computational demands and introduces resource constraints. For this reason, smaller subsets were prepared and experiments were continued mainly on the most frequently used and semantically clearer classes. A second reason is that creating multiple variants makes it possible to measure how models perform under different conditions and levels of complexity. For example, when a three-class dataset is used, the difficulty of multiclass classification is lower, whereas in versions expanded to five or six classes, the model's ability to capture semantic relationships is tested more rigorously. Finally, creating different variants allows more

[14] Raies, K., Rebhi, K., Khemeja, M. Towards ontology of gameplay: Application to game based learning systems // Proceedings of the 10th International Conference on Semantic Systems, – Leipzig: – 4-5 September, – 2014. 1288, – p. 1–4

objective comparisons across models in the scientific literature. On small datasets, classical methods yield stronger results, but on the full dataset and on ontology-enriched variants, the advantage shifts to more advanced “black-box” models. Variants of the Yasaq datasets allow:

- Models to be tested across different sizes and difficulty levels;
- Lighter experiments to be conducted under resource constraints;
- Explainable AI approaches to be developed by adding ontological relations;
- Fair and objective model comparisons to be conducted.

In its final version, the “Yasaq” database contains 3,000 cards, each consisting of a main word, five related taboo words, and a class label. In the final version, the cards are divided into six classes: human, animal, food, object, place, and profession. The database was translated via the Google Application Programming Interface and obtained in two languages. It is currently available in English and Azerbaijani.

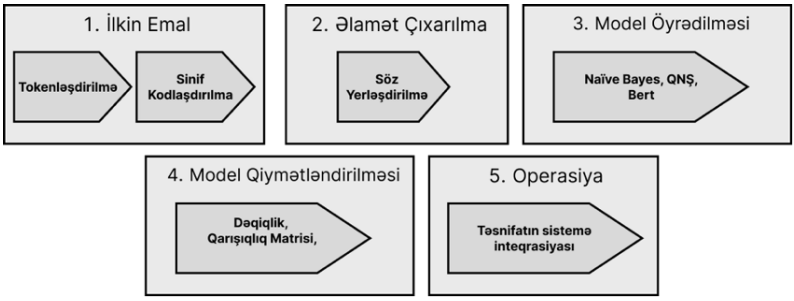

Figure 1. Machine learning stages for text classification

Text classification consists of several stages. Figure 1 illustrates the machine learning pipeline used for text classification:

- Preprocessing
- Feature Extraction
- Model Training

- Model Evaluation
- Operation

At the initial stage, words and classes are processed: data are selected, and the main word together with its five taboo words is taken as the input. The input is tokenized and formatted for machine learning, after which class encoding is performed. After preprocessing, the process moves to the word embedding stage. This stage converts textual units (a word, sentence, or document) into high-dimensional numeric vectors. These vectors represent meanings and relationships in digital form and enable machine learning models to better capture linguistic structure. For Azerbaijani, applying word embeddings has produced important outcomes in terms of semantic relationships and deeper text analysis. Moreover, these approaches have been successfully integrated into serious games to improve the quality of text-based tasks and to create richer learning environments. Next, different text-classification models are applied and their effectiveness is compared. Models were first trained on an Azerbaijani subset of the "Yasaq" dataset and then tested in English to examine cross-lingual performance differences. As a result, a generalized multilingual model capable of operating in both Azerbaijani and English was built. This approach allowed the model to learn structures of different languages in parallel, and BERT was observed to achieve strong results in both languages. In study, applying the three-category version of "Yasaq_1.5K" produced 81% accuracy, while 74% accuracy was obtained on the Movie Review dataset. CNN-based models were also applied in studies, and the performance of this approach was observed. Test accuracies for text classification are shown in Table 1. In a subsequent study, an expanded six-category dataset was used to apply ANN and BERT models; according to the results, the ANN achieved an accuracy of 78%, whereas BERT reached 94% . In another approach, Multilayer Perceptron, Naïve Bayes, and BERT were applied to DBpedia-enriched texts, and it was observed that semantic relationships were represented more effectively. Table 2 presents the text classification test accuracy

obtained on the “Yasaq_3K” datasets. To prevent overfitting, an early stopping strategy with a patience of 3 epochs was applied, and validation accuracy was used as the primary monitoring criterion.

Table 1

Test accuracy of text classification on the “Yasaq_1.5K” dataset

N	*Classification model*	*Word embedding*	*Test accuracy*
1	Multilayer Perceptron	CountVec	0.97
2	Multilayer Perceptron	Word2Vec	0.95
3	Multilayer Perceptron	TFIDF	0.98
4	Naïve Bayes	CountVec	0.94
5	Naïve Bayes	Word2Vec	0.87
6	Naïve Bayes	TFIDF	0.93

By achieving high accuracy and F1 scores, the study confirms the effectiveness of transformer-based models for multilingual text classification, even in low-resource languages. These results demonstrate the model’s ability to handle diverse linguistic structures and its suitability for multilingual use.

Table 2

Test accuracy of text classification on the “Yasaq_3K” dataset

N	*Classification model*	*Word embedding*	*Test accuracy*
1	BERT	BERT	0.94
2	CNN	GloVe	0.81
3	Naïve Bayes	TFIDF	0.90
4	Multilayer Perceptron	TFIDF	0.93

In section 3.2, the improvement of the “Yasaq” game through explainable artificial intelligence is discussed. Systems that can explain AI decisions and justify the words presented to players make the user experience more transparent and comprehensible. Text

classification and analysis within the game constitute one of the main research areas of this section. The study aims not only to achieve correct classification but also to provide explanations for decisions made by "black-box" models. For this purpose, ontologies and structured knowledge bases such as Wikipedia and DBpedia are used. Through ontologies, relationships among words and concepts are modeled, enabling the system to validate its decisions logically using structured knowledge and provide detailed explanations. In this way, both players and researchers can understand why the model produced a particular result. DBpedia is a database that stores Wikipedia data in the form of a structured ontology[15]. Its purpose is to enable humans and machines to query and use information in a semantic form. It stores data in Resource Description Framework (RDF) format and supports querying via the SPARQL query language. RDF is a semantic web standard that represents data in a subject–predicate–object (triple) structure, for example: "Paris – isCapitalOf – France." In the study, experiments were conducted for three classes and a structured inference tree was constructed for each class. This tree is ontology-based and supports deriving logical answers. The ontology-based inference tree stages for the Person class are shown in Figure 2. Based on DBpedia data, if a term belongs to the Person category, the concept should have properties such as birthYear, deathYear, and Specialty. As sufficient conditions, if the term has properties in the ontology such as family members, achievements, and similar attributes, it may belong to the Person class. In this work, results were validated against necessary conditions, and it was experimentally shown that the method can be extended using sufficient conditions. The explainable method correctly classifies results produced by "black-box" machine learning methods with 80% accuracy and can indicate why a term belongs to a

[15] Lehmann, J. DBpedia - A large-scale, multilingual knowledge base extracted from Wikipedia / J. Lehmann, R. Isele, M. Jakob [et al.] // Semantic Web, – Amsterdam: – 2015. 6 (2), – p. 167–195

given class[16]. This result pertains only to sufficient conditions and can be improved further by extending with additional properties. Inference trees were created for the person, animal, food, object, place, and profession categories.

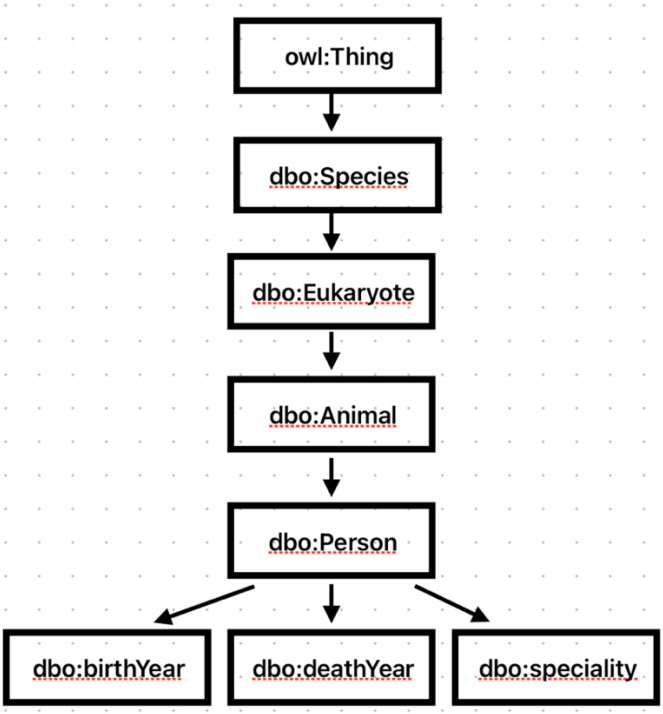

Figure 2. DBpedia ontology-based inference tree stages for the Person class

In section 3.3, the use of AR in the "Yasaq" game is analyzed. Considering the current development of technology, it can be stated that almost all mobile devices support AR, although differences still exist among devices. As a result, the integration of augmented reality

[16] Mammadli, A.A., Ismayilov, E.A, Zanni-Merk, C. Explainability of text classification through ontology-driven analysis in serious games // – Amsterdam: Procedia Computer Science, – 2024. 246 (1), – p. 2128–2137

and the enrichment of traditional games with deep learning help overcome the limitations of existing games. Making such applications accessible to players is expected to further increase their demand. The application of these games in education and their integration with various subjects can enhance students' learning habits and help them retain knowledge in a more engaging and enjoyable way. This paragraph provides a detailed description of the technical implementation of an augmented reality–based prototype developed for a "Yasaq"-type game. The classification model first determines the classes of the cards. Subsequently, these cards are visualized in both physical and virtual environments using AR technology. With AR support, players gain access to virtual information overlaid on cards, objects, and spaces in the real world. AR technology not only makes the gaming experience more immersive but also enables the integration of textual content with the physical environment, allowing players to interact in a more natural and realistic manner[17]. In addition, the system can present images, videos, and other types of media. By using this system, the narrative of the physical cards is enriched in real time through AR markers and mobile devices. To improve the system's accuracy, several preprocessing techniques were applied. At this stage, the following steps were carried out:

- Increasing the contrast and sharpness level of the image,
- Application of Gaussian and Median filters to reduce random pixel variations and unnecessary noise originating from the background environment,
- Improving text separation from the background through binarization (Otsu's thresholding).

As a result of these optimization steps, the accuracy of recognized texts was increased from 72% to 91%, ensuring suitability of the

[17] Mammadli, A.A. Enhancing educational board games with augmented reality and deep learning // International Conference on Artificial Intelligence: from theory to practice, Nakchivan, – 4 November –2024, –p. 341-346

system for real-world use. As shown in Figure 3, the system pipeline architecture is presented. The system begins with player input. Unity uses the Vuforia computer vision package to identify features of the scanned card in real time.

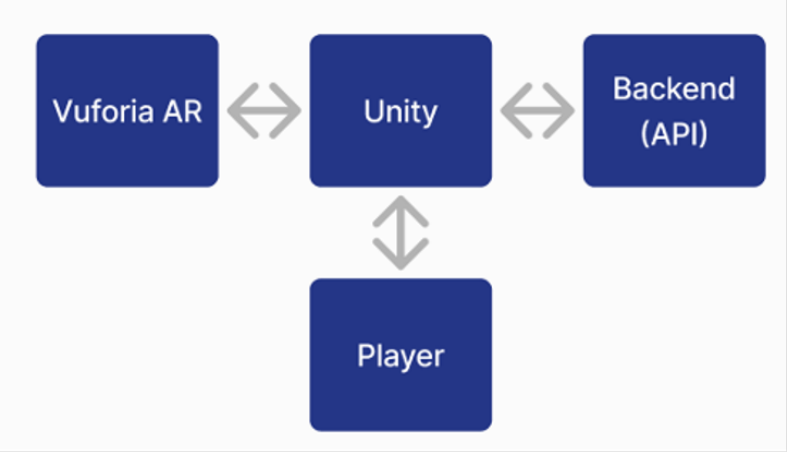

Figure 3. System architecture for the AR application

The solution uses Vuforia's method to ensure continuous tracking and proper alignment of digital content with the physical card. To obtain full information, the system sends explanation and category queries via an API. Once the requested data are retrieved, the API returns a response to Unity, and the information is displayed dynamically within the augmented layer[18]. This structure enables players to access essential information in real time without interrupting gameplay. One of the main reasons for selecting Unity is that its AR Foundation module is directly compatible with both ARKit (iOS) and ARCore (Android). Through this module, it is possible to develop a platform-independent application that runs on different

[18] Mammadli A.A., Ismayilov E.A. Enhancing user interaction through augmented reality interfaces // Mühəndis problemlərinin həllində idarəetmə və nəzarət məsələsində beynəlxalq konfrans – Baku, Azerbaijan – 13-14 March 2025, 2 (1),– p.1-4

mobile devices using the same logic. Unity also supports broad integration with additional AR SDKs such as Vuforia, Wikitude, and EasyAR, which facilitates adapting the game environment to different devices. Although Unreal Engine is known for high-quality 3D graphics and physics engines, its AR usage is more oriented toward large-scale, resource-intensive projects. Moreover, the complexity of Unreal's user interface and its steep learning curve were not suitable for the educational, mobile-focused application considered in this dissertation. Unity is also distinguished by the following advantages:

- Simplified prototyping: rapid development of interactive elements based on Prefabs, UI Toolkit, and a component-based design model;
- Multiplatform support: the ability to run the application on Android, iOS, and even WebGL using the same source code;
- Extensive documentation and community: thousands of sample projects and a broad developer community for rapid problem resolution;
- Mobile performance: an optimized rendering engine and low latency to support lightweight applications.

Considering these characteristics, Unity proved to be the most optimal and functional choice for developing the AR-based "Yasaq" game. The system built with this engine demonstrates technical stability while also supporting user-centered design principles aligned with educational and gameplay contexts. The mobile application can leverage the semantic power of the BERT model in real time with minimal computational resources. This architectural approach not only improves the system's performance and stability but also facilitates the flexible updating and expansion of the BERT model. Figure 4 illustrates the AR visualization of the system's output.

Figure 4. Player's view of the card: AR and real-world example.

Tests on the Android platform were conducted using Samsung Galaxy A52 and Xiaomi Redmi Note 11 devices. The application displayed AR visuals smoothly, and the OCR recognition accuracy averaged 91%. Performance testing showed that the application's response time varied between 300–500 ms. Tests on iOS were carried out using the Xcode simulator as well as iPhone 11 and iPhone 15 devices. AR and OCR functionality were successfully integrated through an Objective-C bridge. Version 10.12 of the Vuforia SDK was used for the iOS version. As a result, output files for both platforms (APK and IPA) were installed on real devices, and the user experience was observed.

RESULTS OF THE DISSERTATION

1. A deep learning–based classification model for the Azerbaijani language was developed. Multiple word embedding

methods were tested to assign the “Yasaq” game texts to semantic classes, and the superiority of BERT was confirmed, achieving test accuracy of up to 94%.

2. To explain the decisions produced by “black-box” methods, an ontology-based explainable AI approach integrated with DBpedia was developed. Using this approach, decisions were justified through semantic relations over knowledge bases such as DBpedia and presented in a form that can be interpreted by the user.

3. A multilingual corpus for the digital serious game “Yasaq” was created, covering Azerbaijani and English and incorporating translation and optimization.

4. Experiments were conducted on text classification using Naive Bayes, CNN, Multilayer Perceptron, and BERT models, together with various word embedding methods, and their performance was compared. The results showed that while classical models yield satisfactory outcomes on small datasets, BERT achieves the highest performance in tasks requiring multi-class classification and contextual analysis.

5. A preprocessing strategy adapted to the morphological characteristics of Azerbaijani mitigated low-resource language constraints.

6. For a game prototype built in Unity, augmented reality and deep learning models were integrated, resulting in an interactive version of the “Yasaq” game.

The main content of the dissertation is published in the following works:

1. Mammadli, A.A., Sadikhova, G.N., Akbarov, N.Kh. Using image detection technologies in AR based gamification platforms // Ümummilli lider Heydər Əliyevin anadan olmasının 100 illik yubileyinə həsr olunmuş gənc tədqiqatçı və doktorantların respublika elmi konfransının materialları, –Baku: ASOIU,– 4-5 May,– 2023, – p. 351–356.

2. Mammadli, A.A., Ismayilov, E.A. Application of deep learning technologies in serious games // International Conference on Problems of Cybernetics and Informatics, – Baku: – 28-30 August, – 2023, – p. 1–4 **(Scopus)**.

3. Mammadli, A.A., Ismayilov, E.A. Application of fuzzy deep learning to classify images // IEEE Publisher, Application of Information and Communication Technologies, Proceedings,– 2023. – p. 1–5 **(Scopus)**.

4. Mammadli, A.A. Unlocking educational insights: integrating Word2vec embeddings and naive bayes classifier for serious game data analysis and enhancement // Baku: Azerbaijan Journal of High Performance Computing, – 2023. 6 (2), – p. 191–198.

5. Mammadli, A.A., Ismayilov, E.A, Zanni-Merk, C. Explainability of text classification through ontology-driven analysis in serious games // – Amsterdam: Procedia Computer Science, – 2024. 246 (1), – p. 2128–2137 **(Scopus)**.

6. Mammadli, A.A. Enhancing educational board games with augmented reality and deep learning // International Conference on Artificial Intelligence: from theory to practice, Nakchivan, – 4 November –2024, –p. 341-346.

7. Mammadli, A.A., Ismayilov, E.A. Combining fuzzy deep learning with HPC to classify images // Lecture Notes in Networks and Systems, Springer, – 2024, 1141, – p. 250–257 **(Scopus)**.

8. Mammadli, A.A. Advancing serious games with multilingual deep learning // 6th International Boğaziçi Scientific Research Congress, – Boğaziçi: – 18-19 January, – 2025, – p. 1100–1106.

9. Mammadli, A.A. Application of deep learning for procedural content integration for learning serious games // – Baku: Proceedings of Azerbaijan High Technical Educational Institutions, – 2025. 48 (6), – p. 455–469.

Author's personal role in co-published works:

[1] - Author of the main idea, problem statement, and analysis of the results;

[2] - Author of the main idea, solution method, conducting the experiments, and analysis of the results;
[3] - Solution method and computer simulation, data processing, and analysis of the results;
[5] - Author of the main idea, solution method, data processing, conducting the experiments, and analysis of the results;
[7] - Author of the main idea, solution method, computer simulation, conducting the experiments, and analysis of the results.

The defense will be held on 28 April 2026 at 14.00 at the meeting of the Dissertation Council FD 2.48 of Supreme Attestation Commission under the President of the Republic of Azerbaijan operating at Azerbaijan State Oil and Industry University.

Address: Baku, Azadlıq avenue 16/21, AZ1010
e-mail: info@asoiu.edu.az

Dissertation is available at the library of the Azerbaijan State Oil and Industry University.

Electronic versions of dissertation and its abstract are available on the official website of the Azerbaijan State Oil and Industry University.

The abstract was sent to the necessary addresses on 17.03.2026.

Signed for print: 11.03.2026
Paper format: A5
Volume: 38203
Number of hard copies: 20