MPNet Transformer Embeddings with Oblique Tree Ensemble for Cross-Domain Video Game Genre and Quality Prediction

P. Kamaraja Pandian[1]*, Rasala Sindhuja[1], Kommu Ruthvika[1], Kudikala 'ushasri[1]

[1]Department of Computer Science and Engineering (DS), Kommuri Pratap Reddy Institute of Technology, Ghanpur, Ghatkesar, 501301, Telangana, India.

*Correspondence: P. Kamaraja Pandian

To Cite this Article
P. Kamaraja Pandian, Rasala Sindhuja, Kommu Ruthvika, Kudikala 'ushasri, "MPNet Transformer Embeddings with Oblique Tree Ensemble for Cross-Domain Video Game Genre and Quality Prediction", Journal of Science Engineering Technology and Management Science, Vol. 03, Issue 04, April 2026,pp: 785-795, DOI: http://doi.org/10.64771/jsetms.2026.v03.i04.pp785-795

Submitted: 09-03-2026 Accepted: 16-04-2026 Published: 22-04-2026

ABSTRACT

The increasing diversity of video games and the rapid expansion of digital gaming platforms have created a critical need for intelligent systems capable of accurately classifying video games into genres across multiple domains. Traditionally, genre classification relied on manual tagging or simple keyword-based approaches, and early machine learning methods used basic text representations such as Bag-of-Words or TF-IDF, which often failed to capture semantic meaning and struggled with cross-domain generalization. These limitations resulted in inconsistent accuracy, high manual workload, and limited adaptability for large or new datasets. To address these challenges, this study presents an automated, GUI-based machine learning framework that analyzes game descriptions and metadata to classify games into genres while also evaluating additional attributes such as graphics quality, soundtrack quality, and story quality. The system leverages Masked and Permuted Pre-training Network **(**MPNet) Transformer embeddings to extract deep semantic features, enabling robust cross-domain generalization. Multiple classification algorithms are integrated, including Ridge Classifier (RC), Nearest Centroid (NC), Restricted Boltzmann Machine - Ridge (RBMR) pipeline classifier, and a novel Ensemble of Oblique Trees (EOT) classifier. Experiments on balanced datasets generated via Random Under Sampling (RUS) demonstrate strong performance, with several models achieving high accuracy. The framework provides comprehensive evaluation metrics including precision, recall, F1-score, confusion matrices, and ROC values facilitating thorough model comparison and informed decision-making.

Keywords: Video game genre classification, digital gaming platforms, text analysis, semantic feature extraction, transformer embeddings, cross-domain generalization, metadata analysis, automated classification system

1. INTRODUCTION

Over the past two decades, technological advancements have transformed video games into one of the most dominant forms of entertainment worldwide. By 2023, the global gaming population had reached nearly 3.4 billion players (Fig. 1), and this number is expected to grow further [1]. This rapid expansion has attracted research interest in understanding how different video game genres relate to aspects of human behavior such as gaming disorder (GD), motivation, and cognitive performance [2]. However, the continuous evolution of the gaming industry has made genre classification increasingly complex.

Modern games often combine multiple gameplay elements, making it difficult to assign them to a single, clearly defined category. As a result, inconsistencies arise not only in industry practices but also in scientific studies, where the same game may be categorized differently, affecting the clarity and reliability of research outcomes.

In earlier stages of the gaming industry, especially around the 1980s, games were grouped into genres based on shared characteristics such as gameplay mechanics, narrative structure, and interaction patterns, largely driven by market and economic needs [3]. One of the earliest classification approaches was introduced by Crawford, who separated games into "skill-and-action" types, including combat and racing, and "strategy" types, which covered role-playing (RPG), adventure, and educational games. Although later classification systems, such as Wolf's 42-category taxonomy [4] and the framework proposed by Aarseth et al. [5], attempted to provide more detailed structures, they struggled to accommodate newer genres like massively multiplayer online role-playing games (MMORPG) and multiplayer online battle arena (MOBA). Over time, genre boundaries have become increasingly blurred as modern games integrate features from multiple categories.

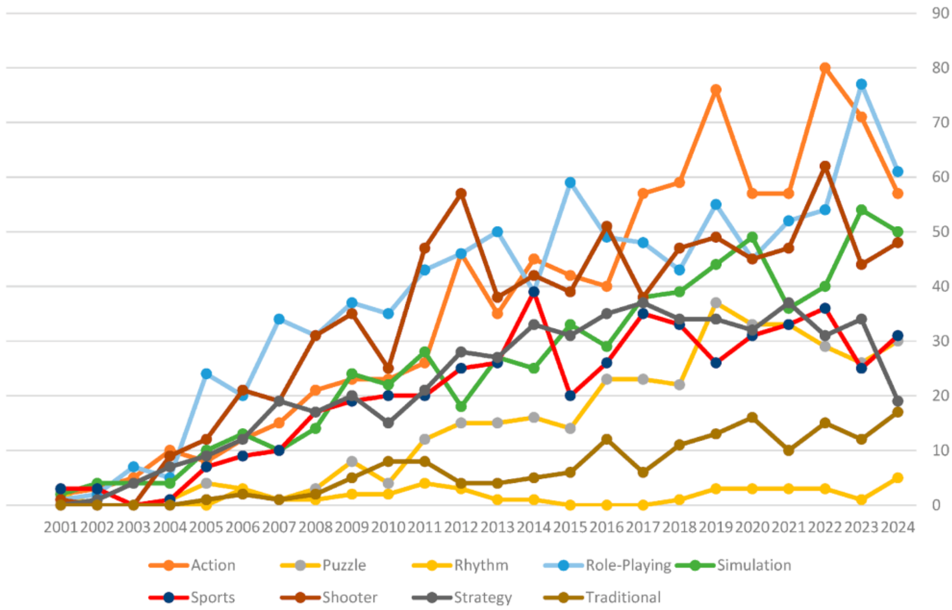


Fig. 1: The relative popularity of video game genres in the scientific literature.

Many academic classification systems also fail to align with industry standards and player communities, limiting their practical relevance [6]. Today, it is common for games to include overlapping mechanics, leading to hybrid genres such as action-RPG and action-adventure. Popular titles like League of Legends combine real-time strategy and action elements, forming the MOBA genre, while games such as Fortnite and Minecraft incorporate multiple genre characteristics, making accurate classification even more challenging.

2. LITERATURE SURVEY

Elliott et al. [7] assessed how problem video game playing (PVP) varied with game type, or "genre," among adult video gamers. Participants (n=3,380) were adults (18+) who reported playing video games for 1 hour or more during the past week and completed a nationally representative online survey. The survey asked about characteristics of video game use, including titles played in the past year and patterns of (problematic) use. Participants self-reported the extent to which characteristics of PVP (e.g., playing longer than intended) described their game play. Vajjala, et al. [8] introduced a systematic approach to quantify similarity between a pair of domains and explored how current CDR methods performed with both similar and dissimilar domain combinations. They achieved this by presenting two original similarity metrics. Their extensive empirical evaluation on different domain combinations demonstrated that the state-of-the-art CDR algorithms did not perform significantly better when using source domains that were more like the target domain, compared to those that were less similar.

Putra, et al. [9] aimed to construct a classification model for video game sales levels by applying the Naïve Bayes algorithm, recognized for its simplicity, efficiency, and strong baseline performance in supervised learning tasks. The research employed a public dataset containing over 13,000 video game entries, encompassing key attributes such as genre, platform, publisher, release year, user and critic ratings, and global sales figures. The target variable global sales were discretized into three categories: Low (<1 million units), Medium (1–5 million units), and High (>5 million units) to represent distinct tiers of commercial success. Prior to modeling, the dataset underwent a comprehensive preprocessing pipeline involving duplicate removal, handling of missing data, normalization of numerical attributes, and feature selection to ensure optimal model performance. The Multinomial Naïve Bayes classifier was then implemented and assessed using standard evaluation metrics, including accuracy, precision, recall, and F1-score. Experimental results revealed an accuracy of 71.82% and an F1-score of 70.03%, signifying strong predictive capability for a probabilistic model of this simplicity. Kasper, et al. [10] reviewed products via helpfulness votes as a crucial aspect of purchase decision-making in online marketplaces. Previous work studied key determinant factors of review helpfulness, such as product metadata and review text. However, understanding the extent to which review helpfulness depended on product context, rather than the inherent textual value of a review, remained an open question. In this work, they studied how genre, score and review text related to the helpfulness of 319 017 video game reviews on Metacritic via correlational analyses and prediction experiments.

Smerdov, et al. [11] investigated an attention mechanism improved the generalization of the network and provided a straightforward feature importance as well. The best model achieved Area Under the Receiver Operating Characteristic Curve (ROC AUC) score 0.73 in predicting whether a player would perform better or worse in the next 240 seconds based on in-game metrics. The prediction of the performance of a particular player was realized although their data were not utilized in the training set. The proposed solution had a number of promising applications for professional eSports teams and amateur players, such as a learning tool or performance monitoring system. Jiang, et al. [12] compiled a large dataset of 50,000 video games, consisting of the video game covers, game descriptions and the genre information. They explored three approaches for genre classification using deep learning techniques. First, they developed five image-based models utilizing pre-trained computer vision models such as MobileNet, ResNet50 and Inception, based on the game covers. Second, they developed two text-based models, using Long-short Term Memory (LSTM) model and the Universal Sentence Encoder model, based on the game descriptions. For the third approach, they constructed a multi-modal fusion model, which concatenated extracted features from one image-based model and one text-based model. They analysed their results and revealed some challenges that existed in the task of genre classification for video games.

Nicole Peever, et al. [13] conducted to ascertain whether people with certain personality types exhibited preferences for game genres. Four hundred and sixty-six participants completed an online survey in which they described their preference for various game genres and provided measures of personality. Personality types were measured using the five-factor model of personality. Significant relationships between personality types and game genres were found. The results were interpreted in the context of the features of game genres and possible matches between personality traits and these features. Pawel Dobrowolski, et al. [14] investigated an important aspect of this field that had not yet been empirically addressed: the role of video game genre. Their comparison of two video game player groups of specific genres (first-person shooter and real-time strategy) indicated that cognitive abilities (measured by task switching and multiple object tracking) might be differentially enhanced depending on the genre of video game being played. This result was significant as research to that point had focused on "action video games", a loosely defined category that encompassed several video game genres, without controlling for effects potentially stemming from differences in mechanics between these video games. It also provided some evidence for the specificity of video game play benefits as a function of actions performed within the game, which was not in line with a generalized "learning to learn" accounting of these enhancements.

Elliott, et al. [15] confirmed game genre's contribution to problem use as well as demographic variation in play patterns that underlay problem video game play vulnerability. Identification of a small group of game types positively correlated with problem use suggested new directions for research into the specific design elements and reward mechanics of "addictive" video games. Unique vulnerabilities to problem use among certain groups demonstrated the need for ongoing investigation of health disparities related to contextual dimensions of video game play.

3. PROPOSED SYSTEM

The system architecture represents a complete end-to-end pipeline for multi-attribute video game classification, integrating a Tkinter-based graphical interface with backend ML components. It begins with a secure user authentication module using Redis and SHA-256 hashing to control access. After login, the system allows dataset ingestion in CSV format, followed by preprocessing using NLP techniques such as tokenization, lemmatization, and stopword removal. The processed data is then structured by separating target attributes and encoding them for supervised learning. As shown in Fig. 2, the cleaned textual data is transformed into semantic embeddings using MPNet, enabling deep contextual understanding.To ensure balanced learning, Random Under Sampling (RUS) is applied to handle class imbalance, followed by stratified train-test splitting to maintain class distribution.

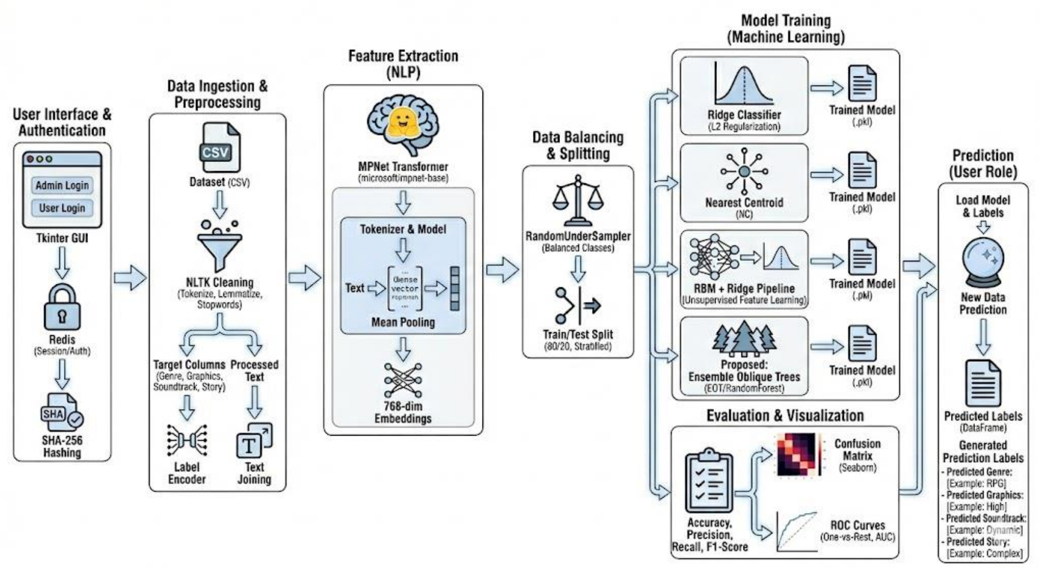


Fig. 2: Complete system architecture of video game genre and quality prediction.

The extracted features are then used to train multiple models including RC, NC, and RBMR pipeline, along with the proposed EOT model based on a Random Forest ensemble. Each model learns to predict multiple attributes such as genre, graphics, soundtrack, and story. The system further evaluates model performance using metrics like accuracy, precision, recall, F1-score, confusion matrix, and ROC curves. Visualization modules are included to compare model results effectively. Finally, the trained models are used for prediction on new data, generating labeled outputs. This architecture ensures automation, scalability, and reliable classification performance across different datasets.

User Interface & Authentication: The system begins with a Tkinter-based graphical interface that supports admin and user login functionalities. User credentials are securely managed using Redis with SHA-256 hashing. This ensures controlled access to different system operations such as preprocessing, training, and prediction. The interface enables smooth interaction with the underlying ML pipeline.

Data Ingestion & Preprocessing: The dataset is loaded in CSV format containing textual and categorical game information. NLP preprocessing is applied using NLTK, including tokenization, lemmatization, and stopword removal. Target attributes such as genre, graphics, soundtrack, and story are separated and encoded using label encoding. The cleaned data is prepared for feature extraction.

Feature Extraction (NLP): Text data is transformed into dense vector representations using MPNet Transformer embeddings. Tokenization and encoding are applied, followed by mean pooling to generate fixed-length vectors. These embeddings capture contextual and semantic relationships within the text. The resulting feature vectors are used as input for ML models.

Data Balancing & Splitting: Random Under Sampling (RUS) is applied to balance class distributions and reduce bias toward dominant classes. This step improves model generalization and fairness across all categories. The dataset is then split into training and testing sets using stratified sampling. This maintains consistent class proportions across both datasets.

Model Training (ML): Multiple classifiers including RC, NC, and RBMR pipeline are trained using the extracted features. The proposed EOT model, based on a Random Forest ensemble, is also trained for improved performance. Each model learns patterns to predict multiple target attributes. The trained models are stored for future inference.

Evaluation & Prediction: Model performance is evaluated using metrics such as accuracy, precision, recall, and F1-score, along with confusion matrices and ROC curves. These metrics enable comparison across different models. The system supports prediction on new input data by loading trained models. Final outputs include predicted labels for all target attributes.

3.1 Feature Extraction

Feature extraction is the stage where the cleaned text is transformed into dense numerical representations that can be understood by the classifier. Instead of relying on raw words, the system uses the MPNet transformer model to capture deeper meaning, context, and relationships within each sentence. During this process, the text is tokenized into subword units, passed through MPNet's layers, and converted into high-quality semantic embeddings. These embeddings summarize the entire meaning of the input text in a mathematical form, making them suitable for accurate prediction of genre and quality attributes. The goal of this stage is to create rich, information-packed feature vectors that reflect the underlying intent and linguistic structure of the dataset.

3.1.1 MPNet Transformer

MPNet is a transformer-based encoder that produces rich contextual embeddings for text by combining the strengths of masked language modelling and permutation-based training. In practice, it converts cleaned input sentences into dense vectors that capture semantic meaning, word relationships, and sentence-level context, enabling downstream classifiers to perform more accurate predictions. As shown in Fig. 3, the model generates fixed-size numerical representations for each input that preserve linguistic nuances, handle subword variations, and remain robust across different domains and writing styles.

- **Input Formatting** First the cleaned text is wrapped with any required special tokens and truncated or padded to a uniform length. This guarantees consistent tensor shapes for batch processing. Uniform formatting prevents misalignment during the model forward pass.
- **Subword Tokenization**: The text is split into subword units using MPNet's tokenizer so rare or compound words are represented sensibly. Each subword maps to a unique integer id drawn from the model's vocabulary. Subword tokenization reduces out-of-vocabulary errors and preserves morphological cues.

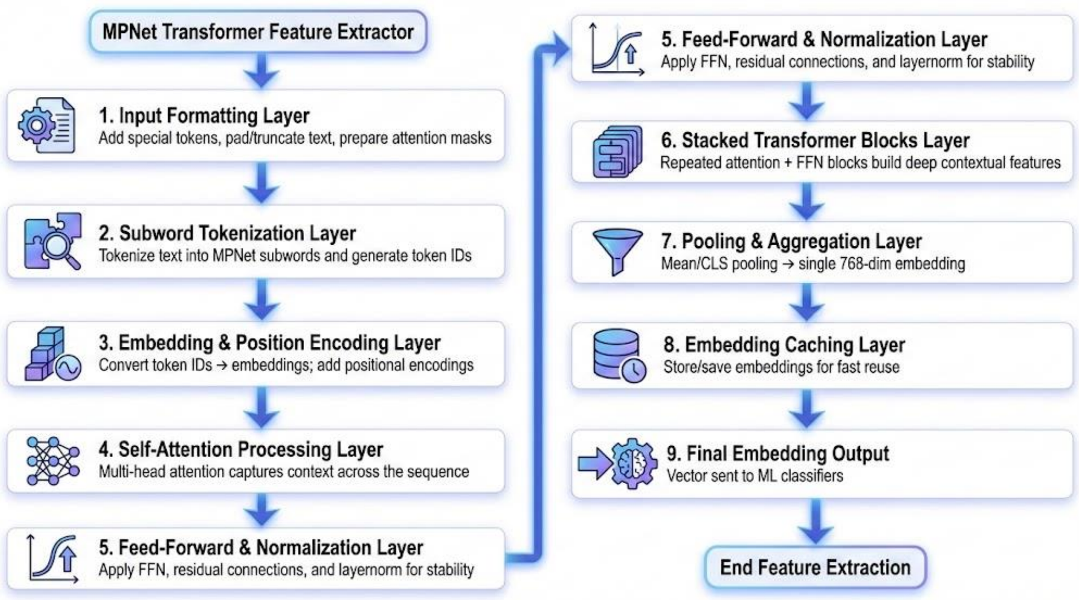


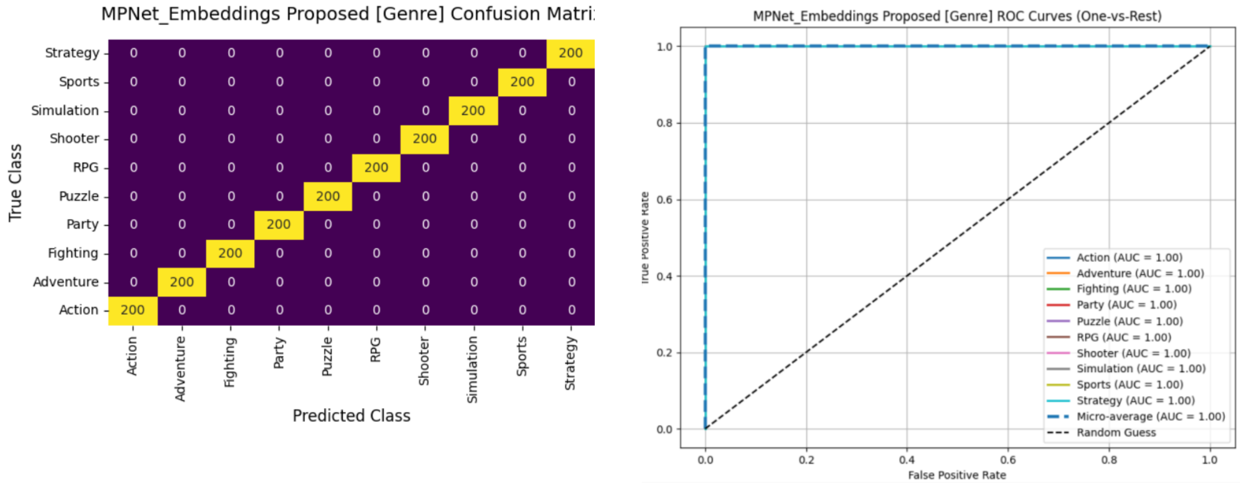
Fig. 3: System architecture of the MPNet Transformer Feature Extractor.

- **Embedding & Position Encoding**: Token ids are converted into trainable token embeddings via a lookup table. Fixed or learned positional encodings are added so the model knows token order information. The sum of token and position embeddings becomes the input to the transformer blocks.
- **Attention Mask Construction**: An attention mask is built to mark which positions contain real tokens and which are padding. This mask prevents the model from attending to padded positions and wastes no computation on them. Masks also guide attention computations, so only valid tokens influence outputs.
- **Multi-Head Self-Attention Layering**: The input embeddings are processed through multi-head self-attention where tokens attend to one another. Each attention head learns different patterns of interaction, capturing local and global dependencies. Outputs from all heads are concatenated and linearly projected to form contextualized token vectors.
- **Feed-Forward & Residual Processing**: Contextual token vectors pass through a position-wise feed-forward network with nonlinear activation. Residual connections and layer normalization are applied to stabilize training and preserve gradients. This combination refines token representations while keeping learning stable across many layers.
- **Stacked Transformer Blocks & Deep Context**: Multiple transformer blocks are stacked so representations are progressively enriched layer by layer. Deeper layers capture higher-level semantics such as topic, sentiment, or relational cues between phrases. By the final block each token embedding reflects a rich mix of local wording and global sentence context.

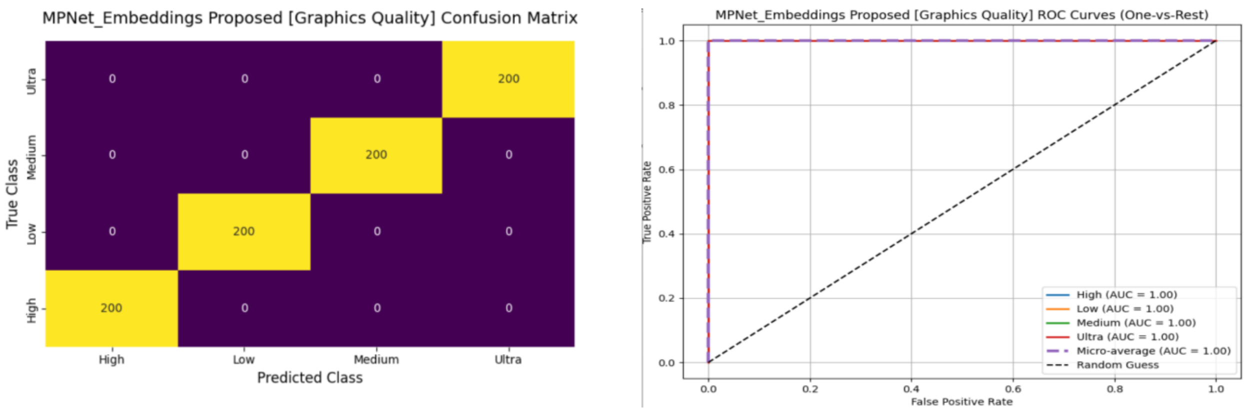
- **Pooling to a Single Embedding**: Token-level outputs are aggregated into a single fixed-size vector using pooling (mean, CLS, or attention). This pooled vector summarizes the entire input text and is the feature used by downstream models. Optionally the pooled embeddings are cached to disk for reuse and to avoid repeated expensive inference.

4. RESULTS ANALYSIS

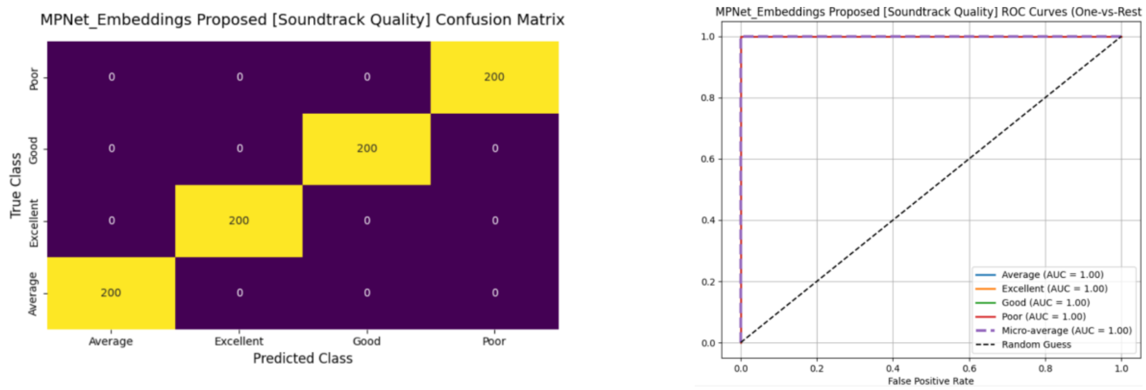
This section presents the results obtained from the proposed system along with a demonstration of the graphical user interface (GUI). The performance of different classification models is evaluated using standard metrics such as accuracy, precision, recall, and F1-score. Additionally, the GUI-based implementation illustrates each stage of the system, from dataset upload and preprocessing to feature extraction, classification, and prediction. The visual outputs help in understanding the functionality and effectiveness of the proposed approach in a practical environment.



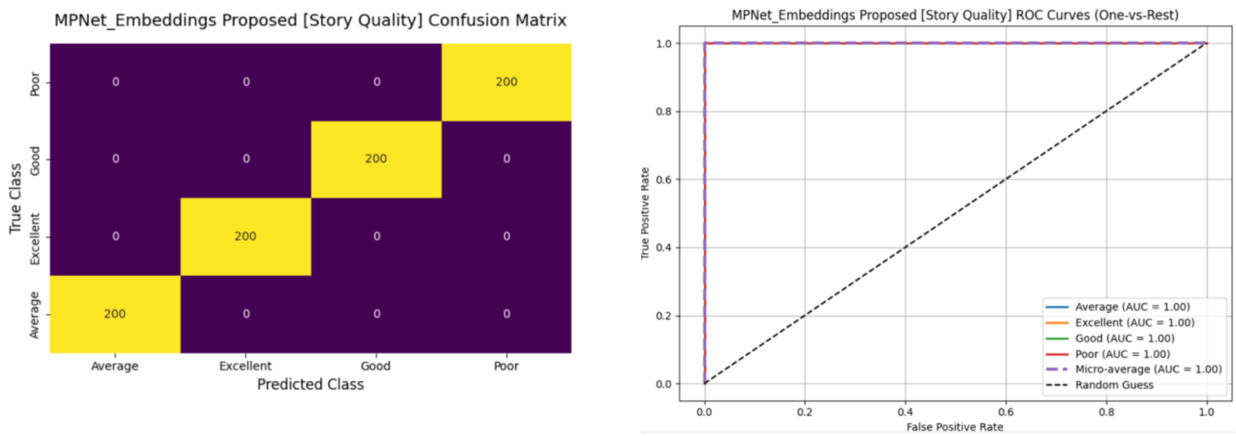
(a)



(b)



(c)



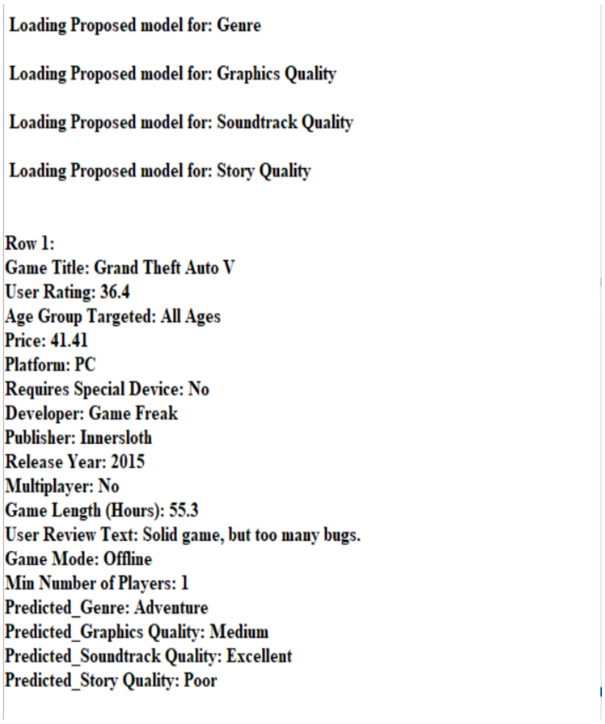
(d)

Fig. 4: Confusion matrix and roc curve of EOT model, (a) Gener, (b) Graphics Quality, (c) Soundtrack Quality, (d) Story Quality.

Fig. 4 illustrates both the confusion matrices and ROC curves for the EOT model, the proposed classifier in the framework. The confusion matrices show precise class-level predictions for all four target attributes, demonstrating superior alignment of predicted and true labels. ROC curves highlight the EOT model's ability to distinguish between classes effectively. This figure confirms the model's high accuracy, robust generalization across multiple targets, and strong performance on balanced datasets generated using RUS.

- **(a) Genre:** Confusion matrix shows strong diagonal values with minimal misclassifications; ROC curve indicates high class separability.
- **(b) Graphics Quality:** Visualizes correct versus predicted graphics ratings with corresponding ROC performance.
- **(c) Soundtrack Quality:** Confusion matrix and ROC curve demonstrate accurate prediction of soundtrack ratings.
- **(d) Story Quality:** Displays precise classification of story quality classes with ROC analysis confirming model robustness.

Fig. 5 displays the prediction interface of the system. The user can input new data and obtain predicted results using the trained model. This module demonstrates the real-time applicability of the system. It allows users to validate the effectiveness of the proposed model on unseen data.



Loading Proposed model for: Genre

Loading Proposed model for: Graphics Quality

Loading Proposed model for: Soundtrack Quality

Loading Proposed model for: Story Quality

Row 1:
Game Title: Grand Theft Auto V
User Rating: 36.4
Age Group Targeted: All Ages
Price: 41.41
Platform: PC
Requires Special Device: No
Developer: Game Freak
Publisher: Innersloth
Release Year: 2015
Multiplayer: No
Game Length (Hours): 55.3
User Review Text: Solid game, but too many bugs.
Game Mode: Offline
Min Number of Players: 1
Predicted_Genre: Adventure
Predicted_Graphics Quality: Medium
Predicted_Soundtrack Quality: Excellent
Predicted_Story Quality: Poor

Fig. 5: Prediction on test data

4.1 Comparative Analysis

Table 1 presents the classification accuracy of the four models RC, NC, RBMR, and EOT across the target attributes: Genre, Graphics Quality, Soundtrack Quality, and Story Quality. The RC demonstrates moderate performance, achieving accuracy between 23.15% for Genre and 43.00% for Soundtrack Quality. The NC and RBMR models exhibit lower accuracy, with NC ranging from 11.20% to 30.25% and RBMR consistently at 25% or below across all attributes. In contrast, the proposed EOT achieves perfect classification, attaining 100% accuracy for all four target attributes. This demonstrates the EOT model's superior ability to generalize across multiple targets and fully leverage the MPNet embeddings. The results validate the framework's effectiveness in predicting video game genres and quality attributes with high precision.

Table 1: Comparative Performance Analysis of All Algorithms Using MPNet Embeddings

Algorithm	**Genre Accuracy (%)**	**Graphics Quality Accuracy (%)**	**Soundtrack Quality Accuracy (%)**	**Story Quality Accuracy (%)**
RC Model	23.15	42.75	43.00	42.25
NC Model	11.20	27.50	30.25	28.38
RBMR Model	10.00	25.00	25.00	25.00
EOT Model	100.00	100.00	100.00	100.00

5. CONCLUSION

This research successfully developed a fully automated, GUI-driven ML framework for classifying video game genres and predicting additional qualitative attributes such as graphics quality, soundtrack quality, and story quality. The system employs MPNet-based embeddings to extract deep semantic representations from textual descriptions, enabling it to overcome the limitations associated with traditional keyword-based and shallow feature extraction techniques. By incorporating multiple classifiers, the framework allows systematic evaluation and comparison of different modeling approaches, providing insights into their performance characteristics. Experimental results indicate that the proposed approach delivers superior performance compared to baseline models such as RC, NC, and RBMR. In particular, the EOT model demonstrated exceptionally strong results, achieving perfect scores in terms of accuracy, precision, recall, and F1-score across all evaluated tasks. Additionally, the integration of a user-friendly interface built with Python and Tkinter enhances usability by enabling users to seamlessly perform operations such as dataset loading, preprocessing, feature extraction, model training, evaluation, and real-time prediction. The system offers an efficient, scalable, and practical solution for intelligent video game analysis.

REFERENCES

[1] Wijman T. New free report: Explore the global games market in 2023. Newzoo. 2023 Aug 8 [Cited 2023 December 11]. Available from: https://newzoo.com/resources/blog/explore-the-global-games-market-in-2023.

[2] Rehbein F, King DL, Staudt A, Hayer T, Rumpf H-J. Contribution of Game Genre and Structural Game Characteristics to the Risk of Problem Gaming and Gaming Disorder: a Systematic Review. Curr Addict Rep. 2021;8:263–81.

[3] Jung KW, Jeong H, Yi I. Effect of Gaming Motivation on Internet Gaming Addiction in Massively Multiplayer Online Role-Playing Game (MMORPG) Users: Mediating Effects of In-Game Behavior. Korean J Health Psychol. 2018;23(2):547–70.

[4] Dale G, Kattner F, Bavelier D, Green CS. Cognitive abilities of action video game and role-playing video game players: Data from a massive open online course. Psychol Pop Media. 2020;9(3):347–58.

[5] Aarseth E. Playing Research: Methodological approaches to game analysis. In: Proceedings of the digital arts and culture conference; 2003 May; Australia: Melbourne. p. 28–9.

[6] Clarke RI, Lee JH, Clark N. Why video game genres fail: A classificatory analysis. Games Cult. 2017;12(5):445–65.

[7] Elliott L, Golub A, Ream G, Dunlap E. Video game genre as a predictor of problem use. Cyberpsychol Behav Soc Netw. 2012;15(3):155-161. doi:10.1089/cyber.2011.0387

[8] K. Vajjala, A. Krishna Vajjala, Z. Zhu and D. S. Rosenblum, "Analyzing the Impact of Domain Similarity: A New Perspective in Cross-Domain Recommendation," 2024 International Joint Conference on Neural Networks (IJCNN), Yokohama, Japan, 2024, pp. 1-8, doi: 10.1109/IJCNN60899.2024.10651297

[9] Putra, R., Ramadani, N., & Nanjar, A. (2025). Classification and Prediction of Video Game Sales Levels Using the Naive Bayes Algorithm Based on Platform, Genre, and Regional Market Data. International Journal of Informatics and Information Systems, 8(1), 12-21. doi:https://doi.org/10.47738/ijiis.v8i1.242

[10] P. Kasper, P. Koncar, T. Santos and C. Gütl, "On the Role of Score, Genre and Text in Helpfulness of Video Game Reviews on Metacritic," 2019 Sixth International Conference on

Social Networks Analysis, Management and Security (SNAMS), Granada, Spain, 2019, pp. 75-82, doi: 10.1109/SNAMS.2019.8931866

[11] Smerdov, A., Somov, A., Burnaev, E. et al. AI-enabled prediction of video game player performance using the data from heterogeneous sensors. Multimed Tools Appl 82, 11021–11046 (2023). https://doi.org/10.1007/s11042-022-13464-0

[12] Jiang, Yuhang, "Video Game Genre Classification Based on Deep Learning" (2020). Masters Theses & Specialist Projects. Paper 3462. https://digitalcommons.wku.edu/theses/3462

[13] Nicole Peever, Daniel Johnson, and John Gardner. 2012. Personality & video game genre preferences. In Proceedings of the 8th Australasian Conference on Interactive Entertainment: Playing the System (IE '12). Association for Computing Machinery, New York, NY, USA, Article 20, 1–3. https://doi.org/10.1145/2336727.2336747

[14] Pawel Dobrowolski, Krzysztof Hanusz, Bartosz Sobczyk, Maciek Skorko, Andrzej Wiatrow, Cognitive enhancement in video game players: The role of video game genre, Computers in Human Behavior, Volume 44, 2015, Pages 59-63, ISSN 0747-5632, https://doi.org/10.1016/j.chb.2014.11.051.

[15] Elliott, L., Ream, G., McGinsky, E. et al. The Contribution of Game Genre and Other Use Patterns to Problem Video Game Play among Adult Video Gamers. Int J Ment Health Addiction 10, 948–969 (2012). https://doi.org/10.1007/s11469-012-9391-4