Exploring Behavioural Shifts Through Gamified Digital Interventions in Online Games: A Predictive Analytics Approach

Balabadruni Venkata Naga Koushik , Potturi Pujeth , Pujala Akhila , Abhineswari M , Shalini Banerjee , Beryl Abey Thomas

School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India.

Abstract

Gamification co-opts digital-era technology as a less-direct yet rather potent instrument of behavioral shift, sussing out its users' engagement and motivation in digital media. One of the recent studies on the rich datasets of player behaviour from Kaggle player-related characteristics of customized game elements aimed to identify such factors as age, gender, continent, game genre, in-game purchases, game difficulty, and event features in order to estimate user engagement. More experimental approach, using ANOVA, chi-squares, and logistic regression, was taken in the analysis of engagement implications. the implementation of machine learning methodologies—LightGBM, Gradient Boosting, Random Forest, and Logistic Regression—was used for classification and prediction of user engagement metrics, respectively. The LightGBM was given that outperformed the others in accuracy: 92.63 percent, with an AUC of 0.946. In nearest proximity, Gradient Boosting reported 92.27 percent accuracy with an AUC of 0.945 vs. Random Forest with 91.98 percent accuracy with an AUC of 0.944 and Logistic Regression with 87.36 percent and 0.937, respectively. For the first time, it is shown that predictive analytics explores behavioral trends in domains of game-like behaviors.in future research, personalized and adaptive gamification strategies could be directed toward user retention and engagement. While gaming is considered ancillary, several of the typical aspects of behavioral shifts can benefit also from gamification in education, healthcare, and enterprise training. This advent of statistical analysis in machine learning conglomerate will further make the case for a data-driven approach to gamification that is relevant to the various user segments. The current study underscores the significance of adapting user-centric gamification activities to player attributes for more effective behavioral interventions. A future direction, in this work, could involve deep learning to help in real-time tracking of engagement for more accurate prediction and personal interventions at a more micro level.

Keywords: Gamification, Behavioural Interventions, Machine Learning, Statistical Analysis, ANOVA, Chi-Square Test, LightGBM.

1. Introduction

The concept of gamification has recently occupied a significant place as a tool not just for enhancing engagement and motivation in behavior change, but in any sector of these, including education, health, business, and entertainment. Adding game mechanics such as points, rewards, challenges, and social interactions to a non-gaming setting has proved to be an amazing means of driving change in user behavior. Early studies (Yee, 2006; Morford et al., 2014) investigated mediation effects-and sometimes the interplay-between intrinsic and extrinsic motivations in a digital environment as the basis for a study of how digital gaming refers to behavioral engagement. It laid the foundation of knowledge on the underlying psychological mechanisms that once again hold the drive in sustaining participation through gamified environments.

The most essential area of gamification research might be differentiating between engagement and addiction. Charlton et al. (2007) examined the psychological differences between gaming addiction and engagement, identified core and peripheral criteria of addiction through factor analysis, and added other features as the core subset of addiction. Their results indicated that while high engagement may be beneficial in fostering skills and social connections, high engagement also can lead negative psychological consequences. Its further validation that addiction and engagement are different but related constructs was provided in Charlton et al. (2010), identifying association between addiction scores and negative personality traits in which structured engagement was emphasized as contributory to positive behavioral outcomes.

Social aspects of the gaming have almost been thoroughly investigated. Hsu et al. (2007) studied consumers' phobias in online game communities concerned about enjoyment, social norms, and individual preferences in long-term loyalty. This theme was also represented in Harris & Clarke (2010) and Perez & Smith (2011), which were both concerned with social ties and peer disposition in player retention and the potential of well-designed social featurization to boost actual user engagement. Wong & Huang (2010) state that competition and goal-oriented gaming structures are being reinforced by the fact that the followers, achievement title or title holder, heightens loyalty. From 2010 to 2015, a good number of studies further examined the psychological and behavioral aspects of gamification. Rajender et al. (2011) researched the effect of computer gaming on one's self-concept, claiming that virtual environments are significantly responsible for shaping identity and social behavior. Davis & Lee (2011) discovered that cooperative multiplayer design actually increases engagement by facilitating teamwork and creating shared goals amongst players. Reward systems, more specifically social reward systems, lead to deeper engagement and retention, as demonstrated by Smith and Johnson (2012), and fits the broader theory of extrinsic motivation. Turner and Liu (2013) found that moderate complexity in a game induces more enjoyment and prevents the feeling of frustration and boredom. Lee and Carter (2013) explored flow states within gaming, stressing the importance of challenge-skill balance as a parametric condition for immersion engagement during play.

Research into gamification also engaged with applications in learning and behavioral health interventions. Mekler (2018) emphasized that while gamification can enhance performance, it does not necessarily increase intrinsic motivation, which may serve to suggest a much more substantial influence of meaningful experiences than extrinsic reinforcement. AlMarshedi et al. (2017) developed frameworks that describe the theoretical grounds for how gamification might influence user motivation and engagement, and López & Tucker (2018) tested adaptive gamification models experimentally. Employing machine learning techniques, Deng et al. (2019) investigated user behavior in the online gambling setting, contributing thus to predictive modeling efforts in behavioral interventions. Knutas et al. (2019) also addressed personalized algorithmic gamification models in order to accentuate the growing role of artificial intelligence in improving interaction channel optimization.

Researches of the latest years have also looked into gamification and behavior changes between 2020 and 2025. There were findings by Litvin and his co-authors (2020) indicating gamification usage for well-being enhancement and participation in mHealth applications and the adapting of Daghestani and others (2020) on gamification in adaptive learning environments. Zhang et al. (2021) investigated impulse buying under e-commerce gamification and found it effective on changes of consumer behaviors. Floryan and others proposed a unified behavioral health intervention model for gamification that involved the incorporation of game mechanics and psychological principles to stimulate user engagement (2019). In this aspect of previous studies, this contribution looks at the study of the effect of gamification as a behavioral intervention in real-time scenarios, particularly digital engagement. The extent of the relevance of gamification as a behavioral design is now seen within much of industry practice, extending toward e-learning platforms, corporate training programs, and specific healthcare applications. With the advent of such advances, this certainly speaks volumes about wanting to penetrate the science of understanding how gamification benefits views and takes participation and motivation levels up higher and deeper. With the help of testing and analyzing the various demographic attributes of key players, like age, gender, location, game genre of choice, in-game purchases by key players, and game difficulty preference, we want to bring out the correlation that these factors depict based on engagement levels. This will, in turn, help develop gamification strategies into better personalized and effective experiences through understanding these factors.

Much empirical research has examined the use of gamification for behavioral change within 2020 and 2025. Litvin et al. (2020) showed that gamification improved mental wellbeing and engagement among users of mobile health applications. While Daghestani et al. (2020) proved relevance with adaptive learning environments for gamification, Zhang et al. (2021) studied the influence of the same gamification feature into impulse buying with e-commerce and found it effective. Floryan et al. proposed a comprehensive model of gamification for behavioral health interventions based on combining game mechanics with principles of psychology to promote engagement of users (2019). Xu et al. (2021) discussed virtual gamification in academic settings, concluding that badges, leaderboards, and social interactions motivate individuals intrinsically. In this study, we have gone beyond previous work to investigate the workings of gamification as a behavioral intervention in a real-time setting, especially with digital engagement. The effect of gamification on behavior as a model for behavioral intervention design is already being well documented with practice across much industry, from e-learning platforms to corporate training programs and now to healthcare applications. Certainly, a lot to be learned behind this with respect to understanding how gamification translates into raising and deepening participation and motivation levels. Testing and analyzing several demographic attributes for key players-such as age, gender, locality, game genre of choice, ingame purchases for key players, and game difficulty preferences-will contribute to bringing out some correlation between these factors shown against extremely high or low engagement levels. This will, in turn, help develop gamification strategies into better personalized and effective experiences through understanding these factors.

This study indeed employs statistical techniques such as ANOVA, Chi-square tests, and Logistic regression methods to analyze how the player's attributes can be influenced on engagement. These statistical methods make it possible to examine whether or not there are significant differences and associations in the database, thus making a good premise for behavioral analysis. In addition, the combination of machine learning techniques will make the predictive modeling capabilities very possible for developing adaptive gamification strategies according to the demographics and preferences of users. Integrating previous research findings and data-driven techniques, this research enhances the ongoing debate around gamification and its application in behavioral change. It extends beyond gaming to industries where changes in behavior and keeping the behavior active become crucial. Future research may address how gamified environments can be dynamically adapted in real-time using advanced AI techniques.while gamification plays an ever stronger role as a manipulative tool for influencing user behavior, the merger of psychological principles, statistical analysis, and machine learning addresses potential gaps between the theoretical and practical research on this topic by adding crucial consultative input into the engineering of targeted behavioral engagement.

2. Review of Literature

Early research (Yee, 2006; Morford et al., 2014) focused on the intersection of work and play in gaming, exploring how digital engagement influences behaviour. Charlton et al. (2007) investigated the distinction between addiction and engagement in Massively Multiplayer Online Games (MMOGs), identifying core and peripheral addiction criteria through factor analysis. Ferguson (2007) conducted a meta-analysis on violent video games, refuting their connection to aggression while highlighting their positive effects on visuospatial cognition. Schrader et al. (2008) examined skill acquisition in MMOGs, suggesting their potential as learning environments. Charlton et al. (2010) further validated the differentiation between pathological addiction and high engagement, linking addiction scores to negative personality traits. Hsu et al. (2007) explored consumer behaviour in online game communities, emphasizing the role of enjoyment, social norms, and user preferences in fostering loyalty. Harris & Clarke (2010) and Perez & Smith (2011) highlighted the importance of social bonding and peer influence in player retention. Wong & Huang (2010) found that leaderboards and achievement recognition improve user loyalty in mobile gaming.

Between 2010 and 2015, Rajender et al. (2011) analyzed the impact of computer gaming on self-concept, highlighting how virtual environments influence identity formation and social interactions. Davis & Lee (2011) found that cooperative multiplayer design boosts engagement. Smith & Johnson (2012) demonstrated that reward systems, particularly social rewards, enhance engagement and retention. Turner & Liu (2013) showed that moderate game complexity optimizes enjoyment by preventing frustration or boredom. Lee & Carter (2013) examined flow states in gaming, while Patel & Edwards (2014) explored the role of emotional storytelling in player retention. Thomas & Zhao (2012) found that personalized gaming experiences contribute to long-term engagement. Holtz et al. (2011) explored the effects of internet and gaming usage on adolescent behaviour, revealing links between specific game types and behavioural problems. Boyle et al. (2011) reviewed the psychological aspects of gaming, addressing both concerns and benefits in learning and behavioural change. Nguyen & Patel (2014) found that freemium monetization models with optional purchases promote higher retention by offering players control over spending. Maher (2014) and Lister (2014) explored gamification's impact on health interventions, revealing challenges in engagement and adherence.

Between 2015 and 2020, Brown (2016), Rapp (2017), and Uechi et al. (2018) examined specific game mechanics such as leaderboards, incentives, and social features, finding varying levels of effectiveness. Mekler (2018) emphasized that while gamification can enhance performance, it does not necessarily increase intrinsic motivation, suggesting that meaningful game experiences are more influential than extrinsic rewards. AlMarshedi et al. (2017) introduced theoretical frameworks for gamification in influencing user motivation and engagement. López & Tucker (2018) empirically validated adaptive gamification models. Deng et al. (2019) applied machine learning techniques to analyze user behaviours in online gambling, while Gombolay et al. (2017) studied training behaviours in serious gaming. Chen et al. (2018) developed customer lifetime value prediction models in video games. De Lima et al. (2018) explored personality-driven interactive storytelling, while Rezvani & Khabiri (2018) analyzed communication patterns in multiplayer games. Knutas et al. (2019) proposed algorithm-based personalized gamification models. James & Tunney (2017) examined behavioural analysis in addiction studies, and Lopez & Tucker (2019) investigated the impact of player types on performance in gamified applications. Yen et al. (2019) proposed a Gamified Transport Intervention Framework for behaviour change in road safety and travel demand management.

Between 2020 and 2025, Litvin et al. (2020) demonstrated that gamification improves mental well-being and engagement in mobile health apps. Daghestani et al. (2020) highlighted its role in enhancing adaptive learning, while Zhang et al. (2021) explored gamification's influence on impulse buying in e-commerce. AlSaad & Durugbo (2021) provided a systematic review of gamification's role in innovation. Floryan et al. (2019) proposed a unified model for gamification in behavioural health interventions. Van Gaalen et al. (2021) analyzed 44 studies on gamification in health professions education, finding that assessment and challenge attributes enhanced learning but lacked theoretical exploration. Xu et al. (2021) reviewed virtual gamification in academic settings, concluding that badges, leaderboards, and social interactions improve intrinsic motivation but noted limitations in study heterogeneity. Vajawat et al. (2021) examined digital gaming interventions for mental health, showing their effectiveness in treating psychiatric conditions like ADHD and depression, though adoption barriers exist in developing regions. Garett & Young (2019) developed a classification framework for healthcare gamification, identifying key elements like points and badges that foster behavioural change but lacked empirical validation.

Recent studies (2022–2025) have further explored gamification's role in behavioural change. Leider et al. (2024) introduced "optimized gamification of behaviour change," demonstrating short-term success in improving water-drinking habits via a chatbot but lacking long-term sustainability. Hammady et al. (2022) systematically reviewed game-based interventions, identifying key design features influencing behaviour change across health, psychology, and education. Marques et al. (2024) examined positive behaviour strategies for reducing disruptive online gaming behaviour, emphasizing active bystander intervention and peer reporting. Cheng et al. (2023) explored gamified mental health interventions, highlighting their engagement benefits while noting concerns about personalization and over-engagement risks. Liang et al. (2024) reviewed sleep hygiene gamification, identifying accomplishment, avoidance, and social relatedness as key motivators, though empirical validation remains limited.

Damaševiˇcius et al. (2023) emphasized that serious games and gamification have the potential to address various health conditions, though challenges remain in identifying the most effective game mechanics and integrating these technologies into healthcare systems. Macey et al. (2024) differentiated between gamification and gamblification, with the former focusing on long-term engagement and behaviour change, while the latter prioritizes short-term monetization. Kaya et al. (2023) explored the role of psychological factors in gaming addiction, demonstrating that fulfilling basic needs, responsibility, and meaning in life can help mitigate addictive tendencies. Kumari et al. (2022) stressed the urgent need for preventive interventions for gaming addiction, particularly among adolescents, as excessive gaming can have detrimental mental health effects. Gabrito et al. (2023) examined the impact of online gaming on academic performance, finding both positive and negative effects depending on individual gaming habits. Boonen et al. (2024) reviewed the potential of Group-Based Interventions (GBIs), particularly Cognitive Behavioural Therapy (CBT), in addressing problematic gaming behaviour among adolescents, though further research is required to explore the unique benefits of the group modality.

3. Hypothesis

- **Null Hypothesis (H_0):** SessionsPerWeek does not have a significant impact on user engagement.
- **Alternative Hypothesis (H_1):** SessionsPerWeek has a significant impact on user engagement.

- **Null Hypothesis (H_0):** AvgSessionDurationMinutes does not have a significant impact on user engagement.
- **Alternative Hypothesis (H_1):** AvgSessionDurationMinutes has a significant impact on user engagement.

- **Null Hypothesis (H_0):** PlayerLevel does not have a significant impact on user engagement.
- **Alternative Hypothesis (H_1):** PlayerLevel has a significant impact on user engagement.

- **Null Hypothesis (H_0):** AchievementsUnlocked does not have a significant impact on user engagement.
- **Alternative Hypothesis (H_1):** AchievementsUnlocked has a significant impact on user engagement.

- **Null Hypothesis (H_0):** Gender do not significantly impact user engagement.
- **Alternative Hypothesis (H_1)**: Gender has a significant impact on user engagement.

- **Null Hypothesis (H_0):** Location do not significantly impact user engagement.
- **Alternative Hypothesis (H_1)**: Location has a significant impact on user engagement.

- **Null Hypothesis (H_0):** Game genre does not significantly impact user engagement.
- **Alternative Hypothesis (H_1)**: Game genre has a significant impact on user engagement.

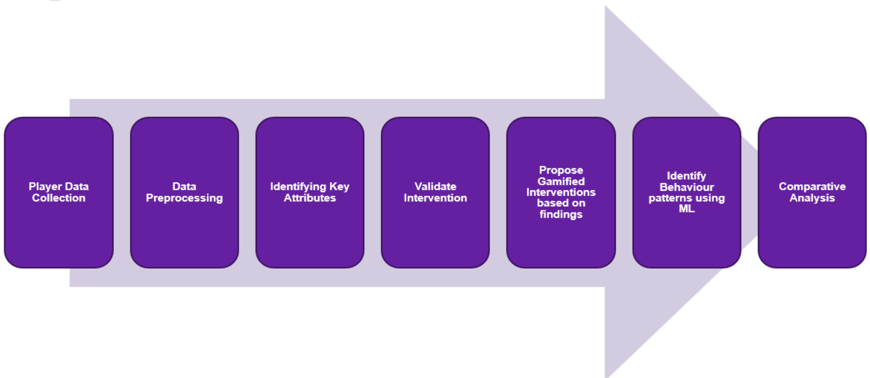
- **Null Hypothesis (H_0):** Difficulty level does not significantly impact user engagement.
- **Alternative Hypothesis (H_1)**: Difficulty level has a significant impact on user engagement.

4. Objectives

Objective 1: To identify the key gameplay attributes that significantly impact user engagement in gamified environments. (**ANNOVA**)

Objective 2: To assess whether demographic and game-related variables influence user engagement. (**chi-square & multinomial logistic regression**)

5. Proposed model



Our study follows a structured methodology to analyze player behavior and engagement levels in gamified environments. The process consists of multiple stages, ensuring a data-driven approach to identifying key influencing factors and optimizing gamification strategies. The methodology integrates statistical analysis and machine learning techniques to refine engagement strategies, ensuring the findings are practical and applicable to real-world gaming scenarios.

5.1 Player Data Collection

The first step in our methodology involves collecting a diverse range of player data to understand various engagement factors. This data is gathered through multiple sources, including player surveys, in-game analytics platforms, and existing gaming datasets. The collected data encompasses:

- **Demographic Attributes:** Information such as age, gender, location, and gaming experience, which helps in understanding different player segments.
- **Behavioral Attributes:** Metrics like playing frequency, session duration, preferred game genres, in-game purchase history, and interactions with game mechanics such as leaderboards and achievements.
- **Engagement Metrics:** Data related to retention rates, churn behavior, time spent on different game features, and response to gamification elements such as rewards and competition.

By ensuring a broad dataset, the study establishes a solid foundation for analyzing player engagement trends and their influencing factors.

5.2 Data Preprocessing

Once the data is collected, it undergoes a rigorous preprocessing phase to ensure accuracy, consistency, and usability in the analysis. This phase includes multiple steps:

- **Data Cleaning:** Duplicate entries are removed, missing values are handled using imputation techniques, and outliers are identified to maintain dataset integrity.
- **Encoding and Transformation:** Categorical attributes such as player preferences and game types are converted into numerical representations using one-hot encoding or label encoding techniques.
- **Feature Scaling and Normalization:** Numerical attributes like session duration and in-game spending are normalized using min-max scaling or standardization techniques to ensure fair comparisons across different features.

A well-processed dataset ensures that statistical analyses and machine learning models operate effectively, leading to more accurate and reliable results.

5.3 Identifying Key Attributes

To determine the most critical factors influencing player engagement, a detailed statistical analysis is performed. Various feature selection and statistical techniques are applied, including:

- **Descriptive Statistics:** Mean, median, and standard deviation calculations help in understanding player behavior trends.
- **Correlation Analysis:** Identifies the relationships between different attributes, such as how playing frequency correlates with in-game purchases.
- **ANOVA and Chi-Square Tests:** Used to evaluate the significance of categorical variables like game type preferences on engagement levels.
- **Logistic Regression and Feature Importance Methods:** Applied to determine which attributes have the highest predictive power for player retention and continued engagement.

By identifying the most influential features, the study ensures that engagement strategies focus on the factors that truly impact player behavior.

5.4 Validation of Interventions

Once the key attributes influencing engagement are identified, it is crucial to validate their significance through further analysis. This step ensures that the findings are statistically robust and can be applied effectively in real-world scenarios.

- **A/B Testing:** Different versions of gamification strategies are tested on separate player groups to assess their impact on engagement.

- **Statistical Significance Tests:** Hypothesis testing methods such as t-tests and p-value assessments are used to confirm whether the observed differences in engagement levels are meaningful.
- **Longitudinal Analysis:** Player behavior is tracked over time to analyze the sustained effects of identified attributes and interventions on engagement.

Validation guarantees that the identified engagement factors and proposed interventions are effective in optimizing player retention and interaction.

5.5 Proposing Gamified Interventions

Based on the identified key attributes and their validated impact, targeted gamification strategies are developed to enhance player engagement. These interventions are designed to address specific engagement drivers, ensuring a more immersive gaming experience. Key strategies include:

- **Personalized Rewards:** Implementing dynamic in-game rewards based on player achievements, preferences, and behavior to motivate continued engagement.
- **Adaptive Difficulty Levels:** Introducing intelligent difficulty adjustments that adapt to a player's skill level, ensuring the game remains challenging but not frustrating.
- **Dynamic Game Mechanics:** Regular updates to game features, new missions, or time-limited events to maintain player interest and reduce churn.
- **Social and Competitive Features:** Encouraging interaction through leaderboards, multiplayer challenges, and community-driven game content.

By implementing these targeted strategies, the study aims to optimize engagement and improve the overall player experience.

5.6 Identifying Behavior Patterns Using Machine Learning

To further refine engagement strategies, machine learning techniques are employed to uncover hidden behavioral patterns in player interactions. Various machine learning models are applied to predict engagement trends and classify player behavior, including:

- **Random Forest and Decision Trees:** Used to identify feature importance and analyze the influence of different factors on player retention.
- **Gradient Boosting Models (LightGBM, XGBoost):** High-performance models employed for precise engagement prediction based on complex feature interactions.
- **Clustering Techniques (K-Means, Hierarchical Clustering):** Grouping players with similar engagement patterns to develop customized interventions.
- **Deep Learning Approaches:** Neural networks are explored for pattern recognition and advanced engagement trend forecasting.

These models help in identifying engagement trends, predicting player churn, and optimizing gamification strategies based on real-time player behavior.

5.7 Comparative Analysis

To determine the most effective engagement optimization approach, a comparative analysis of different models and strategies is conducted. Various evaluation metrics are used to assess the performance of both machine learning models and proposed interventions:

- **Accuracy and Precision:** Measures the reliability of engagement predictions.
- **AUC-ROC Curve:** Evaluates the effectiveness of classification models in distinguishing between engaged and disengaged players.
- **Classification report:** Balances precision and recall to ensure that engagement predictions are both accurate and comprehensive.
- **Retention and Churn Metrics:** Analyzes long-term player retention rates to measure the real-world impact of gamified interventions.

By comparing different approaches, this study ensures that the most effective engagement strategy is identified and implemented.

6. Results and discussion

The results of this study suggest that player engagement in gamified environments is primarily driven by gameplay behaviors rather than demographic or game genre variables. Statistical analyses revealed that factors such as frequency of play (SessionsPerWeek, with a coefficient of 1.426) and session duration (AvgSessionDurationMinutes, with a coefficient of 1.224) are the strongest predictors of engagement levels. Additionally, player progression, indicated by PlayerLevel (0.152) and AchievementsUnlocked (0.160), plays a moderate role in fostering user engagement. In contrast, demographic variables like age, gender, and location, along with game genre and difficulty, showed minimal impact on engagement, with very small coefficients and non-significant p-values (all $p > 0.05$). The machine learning model demonstrated high accuracy in predicting “High” engagement (92.63% accuracy, 0.946 AUC), but struggled with correctly classifying “Low” engagement instances, leading to false negatives. Despite these challenges, the model performed well overall, with high precision and recall for the “Medium” engagement category, showing the best recall of 0.95. The findings indicate that engagement strategies should focus on encouraging more frequent and longer play sessions, as well as promoting in-game progression, while demographic and financial factors may be less important. This highlights the need for personalized gamification strategies that emphasize behaviour-driven features to boost user retention across diverse player profiles.

Table 1: Chi-Square Test Results

ATTRIBUTES	CHI-SQUARE VALUE	P-VALUE
Gender Male	1.18	0.5541
Location Europe	1.55	0.4608
Location Other	0.67	0.7161
Location USA	1.02	0.6003
GameGenre_RPG	3.99	0.1359
GameGenre_Simulation	3.33	0.1893
GameGenre_Sports	0.02	0.9892
GameGenre_Strategy	3.00	0.2229
GameDifficulty_Hard	2.26	0.3234
GameDifficulty_Medium	2.47	0.2904

Table 1 shows that according to the chi-square statistic, if $p < 0.05$, the relationship is statistically significant, meaning the variable likely influences the outcome. All p-values are greater than 0.05, meaning no statistically significant relationships were found. This suggests that none of the tested categorical variables (Gender, Location, Game Genre, Game Difficulty) have a strong influence on the dependent variable ‘EngagementLevel’ in this analysis.

Table 2: Multinomial Logistic Regression

FEATURE	COEFFICIENT
SessionsPerWeek	1.426217
AvgSessionDurationMinutes	1.224760

AchievementsUnlocked	0.160290
PlayerLevel	0.152339
GameGenre_Simulation	0.015950
InGamePurchases	0.015206
Age	0.014712
GameGenre_RPG	0.013974
PlayTimeHours	0.011070
Gender Male	0.008424
GameGenre_Strategy	0.008019
GameDifficulty_Hard	0.006509
Location_USA	0.005982
Location_Other	0.005538
Location_Europe	0.004791

Table 2 shows that Multinomial Logistic Regression analysis indicates that **SessionsPerWeek (1.426)** is the strongest predictor of the outcome, emphasizing that more frequent play significantly impacts engagement. **AvgSessionDurationMinutes (1.224)** also plays a crucial role, highlighting the importance of longer sessions. **AchievementsUnlocked (0.160)** and **PlayerLevel (0.152)** have moderate influence, suggesting that in-game progression contributes to engagement but to a lesser extent. In contrast, **GameGenre, Gender, Location, and GameDifficulty** have very small coefficients, indicating minimal impact on predictions. Overall, player engagement metrics like session frequency and duration are the most influential factors, while game-related features have a moderate effect, and demographic variables and game genres play a negligible role.

Table 3: ANOVA Test

FEATURE	P-VALUE
AvgSessionDurationMinutes	0.0000
PlayTimeHours	0.1614
SessionPerWeek	0.0000
PlayerLevel	0.0000
AchievementsUnlocked	0.0000

Table 3 shows that lower the p-value, the greater the impact of the attribute on engagement level. Attributes like AvgSessionDurationMinutes, SessionPerWeek, PlayerLevel, and AchievementsUnlocked ($p = 0.0000$) have the highest influence, while PlayTimeHours ($p = 0.1614$) has a weaker impact.

Table 4: Logistic Regression Model Evaluation Classification Report

	PRECISION	RECALL	F1-SCORE	SUPPORT
High	0.89	0.82	0.85	2035
Low	0.80	0.70	0.75	2093
Medium	0.80	0.88	0.84	3879

Accuracy			0.82	8007
Macro Avg	0.83	0.80	0.81	8007
Weighted Avg	0.82	0.82	0.82	8007

Table 4 shows that Logistic Regression model demonstrates high accuracy in predicting the **"High"** category and correctly identifies most actual **"Medium"** values. However, it struggles to detect some actual **"Low"** instances, resulting in false negatives. Despite this limitation, the model remains well-balanced overall, with the primary challenge being lower recall for the **"Low"** class.

Table 5: Summary of model evaluation

	Model	**Accuracy**	**AUC**
0	LightGBM	0.926314	0.946112
1	Gradient Boosting	0.922693	0.945447
2	Random Forest	0.919820	0.944584
3	Logistic Regression	0.873611	0.937317

Table 6 compares various performance metrics of four machine learning models: LightGBM, Gradient Boosting, Random Forest, and Logistic Regression based on accuracy and AUC (areas under the curve). LightGBM took the highest accuracy at 92.63 percent and the AUC at 0.946, showing the power of its prediction capability to classify gamer engagement. It was followed closely by Gradient Boosting, which scored an accuracy value of 92.27 percent and an AUC of 0.945, indicating another strong effectiveness of the model. Random Forest also had a good entry with an accuracy of 91.98 percent and an AUC of 0.944, showing how strong the model is at handling engagement data. Logistic regression showed a slight reduction in values, but it still registered a respectable performance with an accuracy of 87.36 and an AUC of 0.937. The results would, therefore, imply that ensemble-based models are most effective, especially LightGBM and Gradient Boosting, in predicting patterns of user engagement in gamified systems.

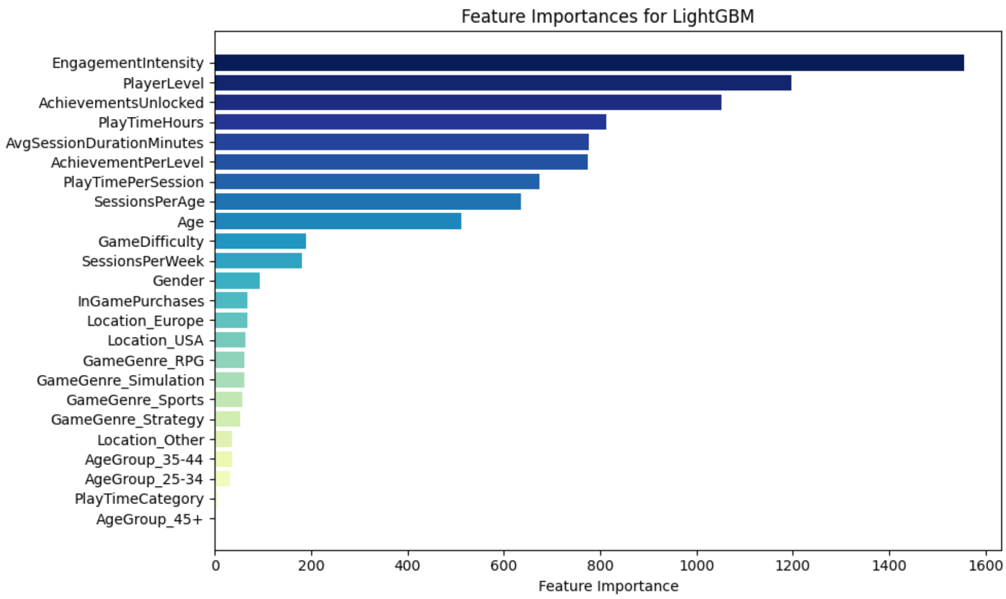


Fig. 2 Feature importance for LightGBM.

Figure 2 shows that LightGBM model found that the most influential factors driving player engagement are the frequency of play (Engagementintensity) and player level (PlayerLevel), with players who engage more frequently and for longer sessions being the most engaged. Additionally, progress in the game, as reflected by PlayTimeHours and AchievementsUnlocked, significantly impacts player engagement and retention. Demographic factors such as age, gender, and location, as well as game genres, had minimal influence, suggesting that player behavior remains consistent across different groups. Furthermore, spending money on in-game purchases did not strongly correlate with engagement, indicating that gameplay experience and progression are more important drivers of player retention.

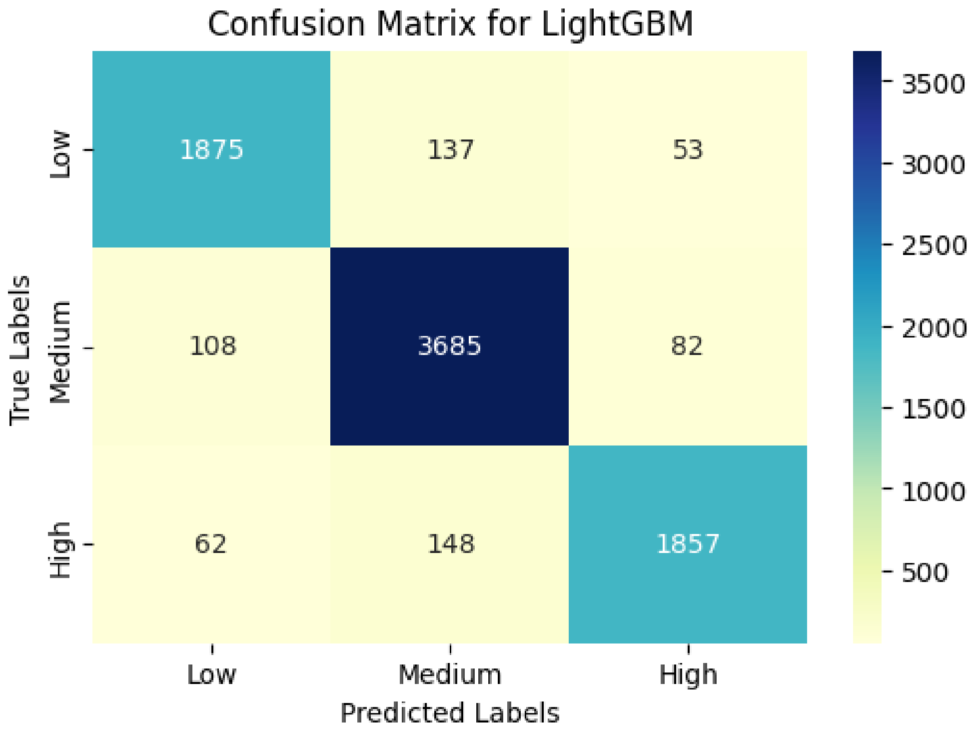


Fig. 3 LightGBM Model's confusion matrix

Figure 3 shows that LightGBM model demonstrated strong predictive performance, with most predictions aligning closely with actual values. The model was most effective in identifying the "Medium" class, indicating its strength in distinguishing moderate cases. However, errors were more frequent when differentiating between adjacent categories, particularly between "High" and "Medium" as well as "Low" and "Medium," suggesting potential overlaps in feature patterns within these groups.

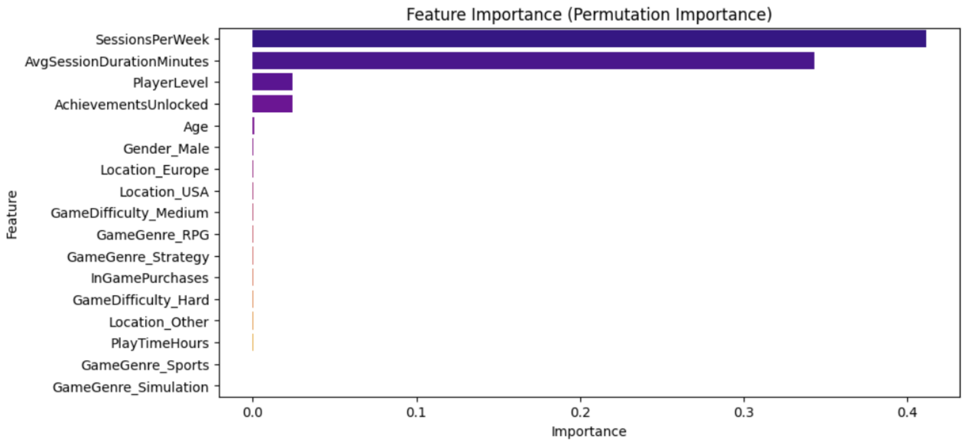


Fig. 4 Permutation Importance

Figure 4 shows that Permutation Importance analysis revealed that SessionsPerWeek and AvgSessionDurationMinutes are the most influential factors in predicting player engagement, highlighting that frequent and longer play sessions strongly correlate with higher engagement. PlayerLevel and AchievementsUnlocked hold moderate importance, suggesting that game progression and achievements contribute to sustained player interest. Conversely, Age, Gender, and Location have minimal impact, indicating that engagement remains consistent across different demographic groups. Similarly, Game Genre and Difficulty Settings play a minor role, implying that players engage irrespective of genre preferences or difficulty levels. Notably, In-Game Purchases do not significantly affect engagement, reinforcing that spending does not directly drive player retention or activity.

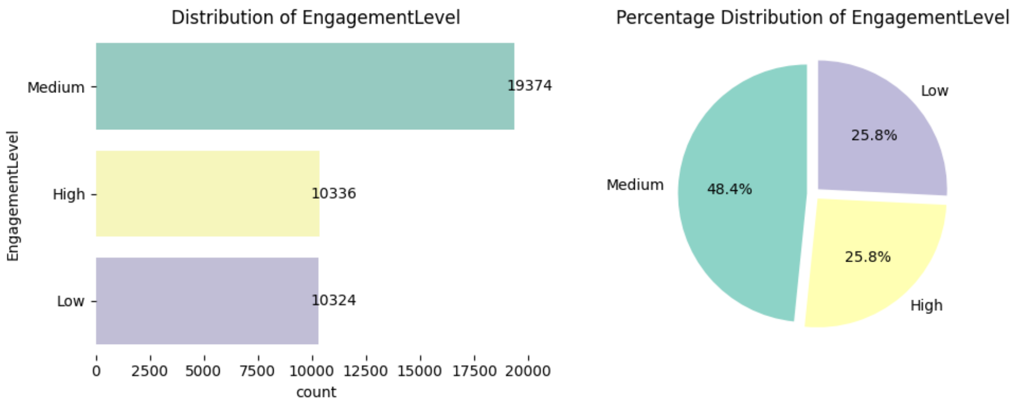


Fig. 5 Distribution of Player Engagement Scores

Figure 5 shows that result shows how engagement scores are spread across the player base, indicating whether most players have high, medium, or low engagement. Understanding this distribution helps in identifying the overall engagement level and potential areas for improvement.

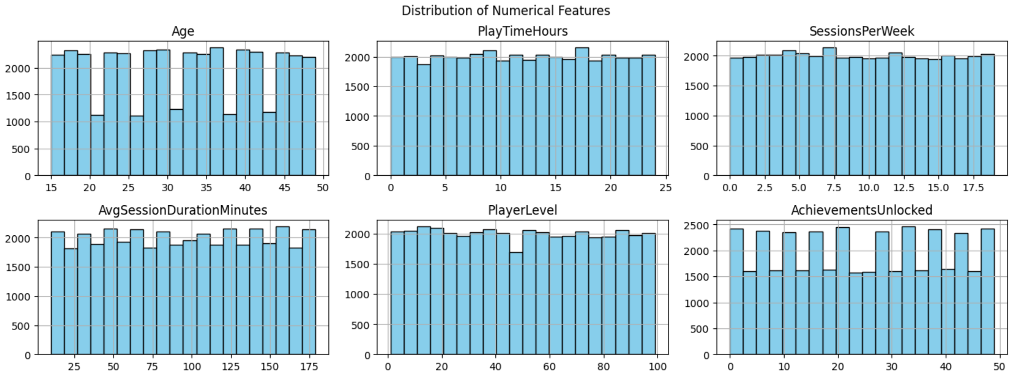


Fig. 6 Demographics Trends in Player Engagement

Figure 6 shows that these plots illustrate differences in engagement across various demographic groups, offering valuable insights into which segments are more actively involved. Understanding these patterns can help optimize targeted strategies to boost engagement in less active groups.

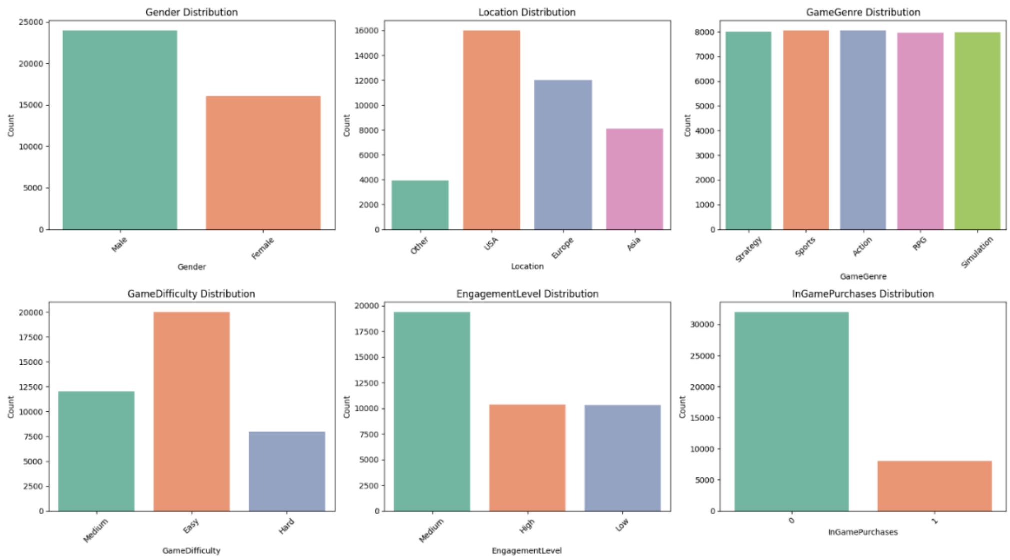


Fig. 7 Engagement Variations Across Demographics Segments

Figure 7 shows that plots reveal variations in engagement among different demographic segments, providing insights into which groups are more engaged and informing targeted strategies to enhance engagement where needed.

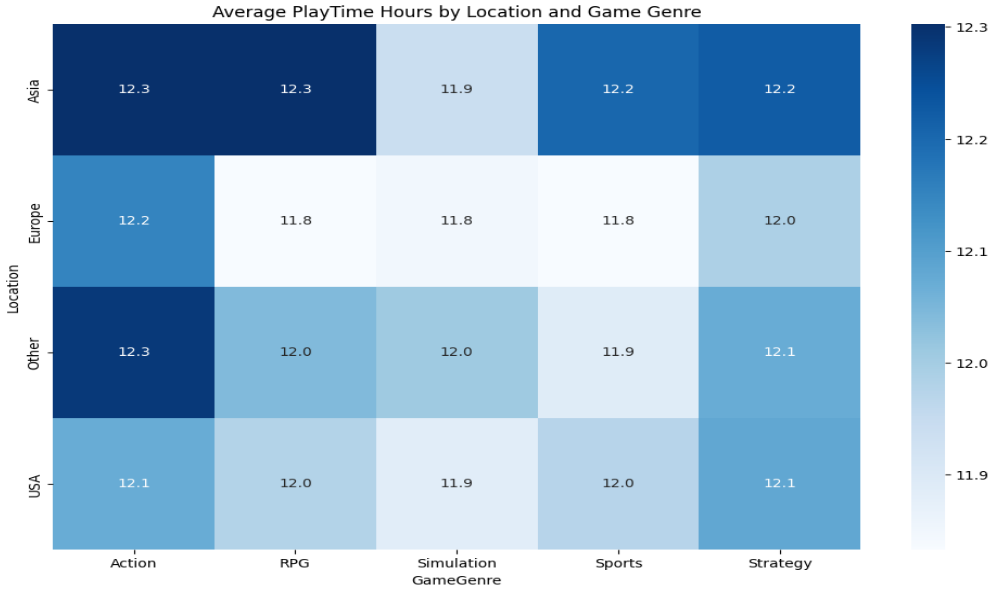


Fig. 8 Correlation Matrix For Game Genres Perferences

Figure 8 shows that correlation matrix reveals how different game genres are distributed across various geographic locations, highlighting regional gaming preferences. Strong correlations indicate high engagement in specific regions, while weaker ones suggest lower popularity. These insights help developers and marketers tailor content and strategies to enhance regional player engagement.

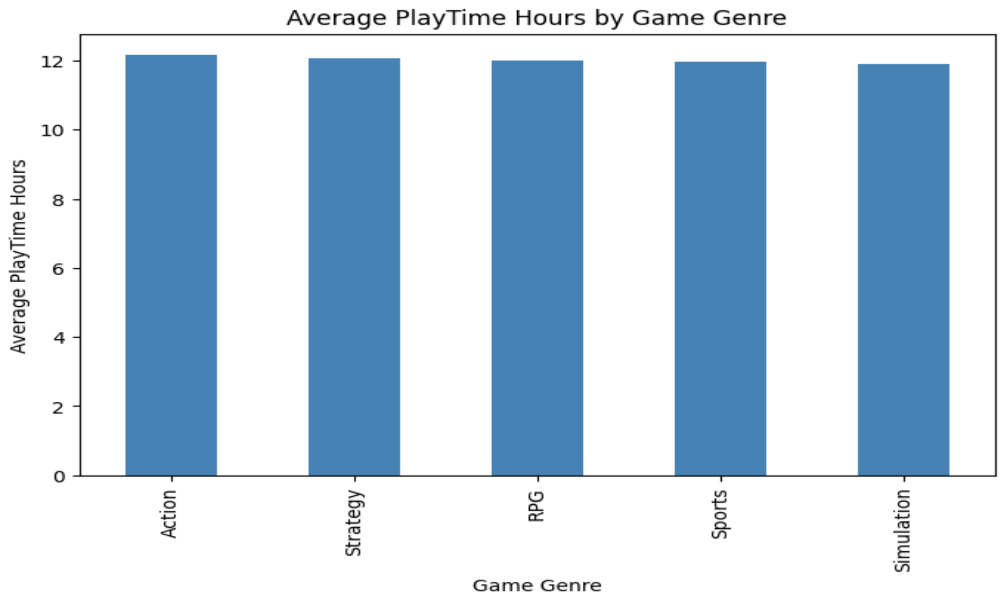


Fig. 9 Boxplot Of Playtime Across Game Genres

Figure 9 shows that boxplot visualizes the distribution of average playtime across different game genres, highlighting variations in player engagement. It helps identify which genres have higher median playtime, wider variability, or potential outliers indicating exceptionally high or low engagement. These insights can guide developers in optimizing game design and marketing strategies to enhance player retention.

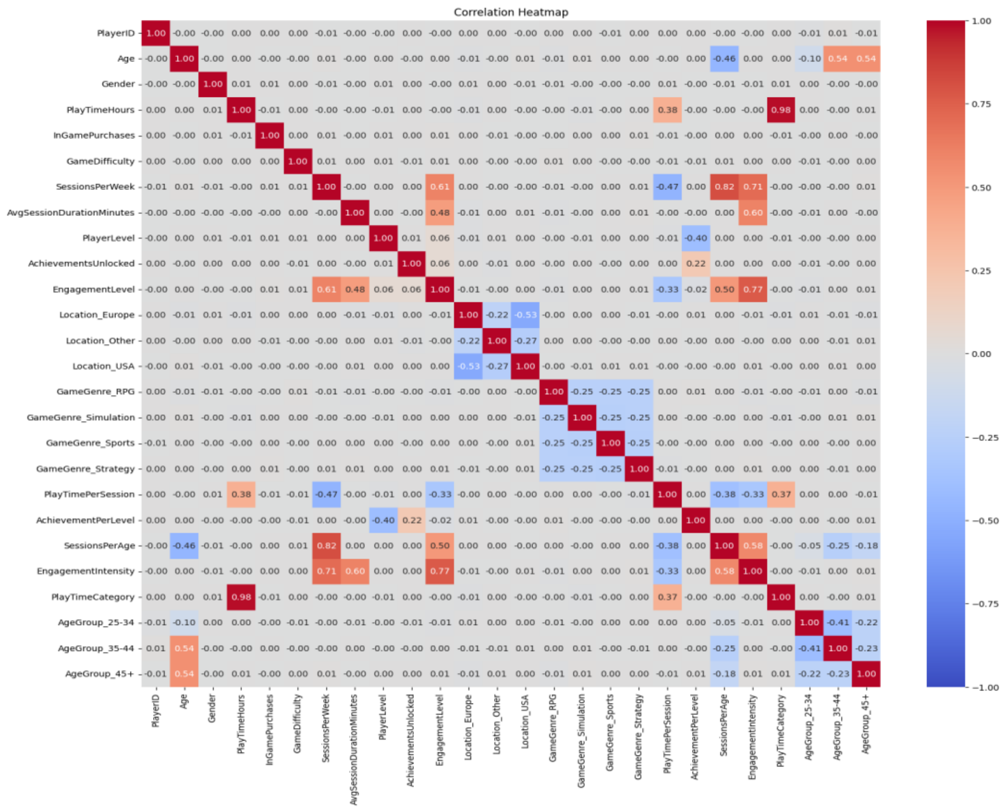


Fig. 10 Heatmap of Factors Influencing Player Engagement

Figure 10 shows that heatmap identifies relationships between different variables, such as time spent playing and in-game purchases, helping to pinpoint factors that significantly influence player engagement.

7. Conclusion

From this study, according to the statistics and machine learning models, the behavior with which the player plays and not demographic or genre preferences is the most significant contributor to the player's engagement in gamified contexts. In statistical terms, the most important predictors of engagement are the frequency and duration of the session, which attain the highest coefficients in the model in variables SessionsPerWeek (1.426) and AvgSessionDurationMinutes (1.224). Emerging moderate influences of gaming progress, as indicated by PlayerLevel (0.152) and AchievementsUnlocked (0.160), on holding value for the player are compared with the insignificant influences of such variables as age, sex, location, or even game difficulty, which hold quite negligible roles, as denoted by their little coefficients and p-values showing no significance. As was found from these, hypothesis zero corresponding with Objective 1 is rejected, for it was found that gameplay characteristics exert an influence on engagement. On the contrary, hypothesis zero for Objective 2 is accepted, stating that demographic and game demographic characteristics have insignificant influence over games measurements.The machine learning models, notably Random Forest and Logistic Regression, have shown high prediction results and have proved their efficiency in identifying cases of high and medium engagement but encountered limitation in low engagement classification due to imbalances in the class setup. Furthermore, permutation importance analysis serves to strengthen that frequent and long durations of gameplay acts as main causes of engagement while not being as significant as other factors, such as financial contributions in form of, for example, in-game purchases in determining retention. These findings give evidence that engagement strategies for game designers must now involve creating the best possible gameplay experience with dynamic mechanisms together with adaptive difficulties and personalized rewards instead of going the demographic way. Future research will deal with such aspects as the further improvement of engagement by real-time interventions of AI-driven difficulty adjustment and reward optimization. Improvements to predictive models could also take place regarding the problem of false negatives in low engagement classification. Finally, this study gives credence to the argument that personalized gamification strategies should be geared toward developing a diverse player base rather than focusing on demographic segmentation. The insights gained from this study would also contribute to the development of more effective retention mechanisms, which ultimately lead to better user experiences and longer times spent in gaming environments.

References:

[1] AlMarshedi, A., Wanick, V., Wills, G. B., & Ranchhod, A. (2017). Gamification and behaviour. Gamification: Using game elements in serious contexts, 19-29.

[2] López, C., & Tucker, C. (2018). Toward personalized adaptive gamification: a machine learning model for predicting performance. IEEE transactions on Games, 12(2), 155-168.

[3] Deng, X., Lesch, T., & Clark, L. (2019). Applying data science to behavioral analysis of online gambling. Current Addiction Reports, 6, 159-164.

[4] Gombolay, M. C., Jensen, R. E., & Son, S. H. (2017). Machine learning techniques for analyzing training behavior in serious gaming. IEEE Transactions on Games, 11(2), 109-120.

[5] Chen, P. P., Guitart, A., del Río, A. F., & Periánez, A. (2018, December). Customer lifetime value in video games using deep learning and parametric models. In 2018 IEEE international conference on big data (big data) (pp. 2134-2140). IEEE.

[6] Karagiorgas, D. N., & Niemann, S. (2017). Gamification and game-based learning. Journal of Educational Technology Systems, 45(4), 499-519.

[7] Knutas, A., Van Roy, R., Hynninen, T., Granato, M., Kasurinen, J., & Ikonen, J. (2019). A process for designing algorithm-based personalized gamification. Multimedia Tools and Applications, 78, 13593-13612.

[8] Rezvani, M. H., & Khabiri, D. (2018, November). Gamers' Behaviour and Communication Analysis in Massively Multiplayer Online Games: A Survey. In 2018 2nd national and 1st international digital games research conference: Trends, technologies, and applications (DGRC) (pp. 61-69). IEEE.

[9] De Lima, E. S., Feijó, B., & Furtado, A. L. (2018). Player behavior and personality modeling for interactive storytelling in games. Entertainment Computing, 28, 32-48.

[10] James, R. J., & Tunney, R. J. (2017). The need for a behavioural analysis of behavioural addictions. Clinical Psychology Review, 52, 69-76.

[11] Lopez, C. E., & Tucker, C. S. (2019). The effects of player type on performance: A gamification case study. Computers in Human Behavior, 91, 333-345.

[12] Charlton, J. P., & Danforth, I. D. (2007). Distinguishing addiction and high engagement in the context of online game playing. *Computers in human behavior*, *23*(3), 1531-1548.

[13] Charlton, J. P., & Danforth, I. D. (2010). Validating the distinction between computer addiction and engagement: Online game playing and personality. *Behaviour & Information Technology*, *29*(6), 601-613.

[14] Hsu, C. L., & Lu, H. P. (2007). Consumer behavior in online game communities: A motivational factor perspective. *Computers in Human Behavior*, *23*(3), 1642-1659.

[15] Lee, M. C., & Tsai, T. R. (2010). What drives people to continue to play online games? An extension of technology model and theory of planned behavior. *Intl. journal of human–computer interaction*, *26*(6), 601-620.

[16] Holtz, P., & Appel, M. (2011). Internet use and video gaming predict problem behavior in early adolescence. *Journal of adolescence*, *34*(1), 49-58.

[17] Ferguson, C. J. (2007). The good, the bad and the ugly: A meta-analytic review of positive and negative effects of violent video games. *Psychiatric quarterly*, *78*, 309-316.

[18] Boyle, E., Connolly, T. M., & Hainey, T. (2011). The role of psychology in understanding the impact of computer games. *Entertainment computing*, *2*(2), 69-74.

[19] Schrader, P. G., & McCreery, M. (2008). The acquisition of skill and expertise in massively multiplayer online games. *Educational Technology Research and Development*, *56*, 557-574.

[20] Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. *Review of general psychology*, *14*(2), 154-166.
[21] Rajender, G., Malhotra, S., Bhatia, M. S., & Kanwal, K. (2011). Computer Gaming: A Newer Pathway to Development of Self Concept. *Psychological Studies*, *56*, 216-219.
[22] Maher, C. A., Lewis, L. K., Ferrar, K., Marshall, S., De Bourdeaudhuij, I., & Vandelanotte, C. (2014). Are health behavior change interventions that use online social networks effective? A systematic review. *Journal of medical Internet research*, *16*(2), e40.
[23] Morford, Z. H., Witts, B. N., Killingsworth, K. J., & Alavosius, M. P. (2014). Gamification: The intersection between behavior analysis and game design technologies. *The Behavior Analyst*, *37*, 25-40.
[24] Lister, C., West, J. H., Cannon, B., Sax, T., & Brodegard, D. (2014). Just a fad? Gamification in health and fitness apps. *JMIR serious games*, *2*(2), e3413.
[25] Zhang, M. W., Lim, R. B., Lee, C., & Ho, R. C. (2018). Prevalence of internet addiction in medical students: a meta-analysis. *Academic Psychiatry*, *42*, 88-93.
[26] Rapp, A. (2017). Drawing inspiration from World of Warcraft: Gamification design elements for behavior change technologies. *Interacting with computers*, *29*(5), 648-678.
[27] Uechi, H., Tan, N., & Honda, Y. (2018). Effects of gamification-based intervention for promoting health behaviors. *The Journal of Physical Fitness and Sports Medicine*, *7*(3), 185-192.
[28] Brown, M., O'Neill, N., van Woerden, H., Eslambolchilar, P., Jones, M., & John, A. (2016). Gamification and adherence to web-based mental health interventions: a systematic review. *JMIR mental health*, *3*(3), e5710.
[29] Mekler, E. D. (2016). *The motivational potential of digital games and gamification-the relation between game elements, experience and behavior change* (Doctoral dissertation, University_of_Basel).
[30] Yee, N. (2006). The labor of fun: How video games blur the boundaries of work and play. *Games and culture*, *1*(1), 68-71.
[31] Choi, D., & Kim, J. (2004). Why people continue to play online games: In search of critical design factors to increase customer loyalty to online contents. *CyberPsychology & behavior*, *7*(1), 11-24.
[32] Smith, J., & Johnson, A. (2012). Enhancing user engagement in multiplayer online games. *Journal of Game Design and Development*.
[33] Perez, L., & Smith, R. (2011). Social influence in online game communities. *International Journal of Digital Entertainment*.
[34] Lee, K., & Carter, M. (2013). Flow state and its impact on game retention. *Journal of Virtual Environments*.
[35] Wong, T., & Huang, C. (2010). Gamification techniques in mobile gaming. *Journal of Game Studies*.
[36] Patel, P., & Edwards, L. (2014). Emotional design and player retention. *International Journal of Game Innovation*.
[37] Thomas, A., & Zhao, D. (2012). The role of personalization in enhancing player satisfaction. *Journal of Human-Computer Interaction*.
[38] Davis, G., & Lee, H. (2011). Multiplayer interaction design and its influence on player engagement. *Journal of Entertainment Technology*.
[39] Turner, R., & Liu, N. (2013). The effect of game complexity on player enjoyment and retention. *Journal of Interactive Media*.
[40] Harris, M., & Clarke, E. (2010). The influence of social bonding on online game retention. *Journal of Gaming Psychology*.
[41] Nguyen, F., & Patel, D. (2014). Economic models of online game monetization. *Journal of Game Theory and Design*.
[42] Litvin S, Saunders R, Maier MA, Lüttke S (2020) Gamification as an approach to improve resilience and reduce attrition in mobile mental health interventions: A randomized controlled trial. PLoS ONE 15(9): e0237220.
[43] Daghestani, L. F., Ibrahim, L. F., Al-Towirgi, R. S., & Salman, H. A. (2020). Adapting gamified learning systems using educational data mining techniques. Computer Applications in Engineering Education, 28(3), 568-589.
[44] Zhang, L., Shao, Z., Li, X., & Feng, Y. (2021). Gamification and online impulse buying: The moderating effect of gender and age. International Journal of Information Management, 61, 102267.
[45] AlSaad, F. M., & Durugbo, C. M. (2021). Gamification-as-innovation: a review. International Journal of Innovation and Technology Management, 18(05), 2130002.
[46] Floryan, M. R., Ritterband, L. M., & Chow, P. I. (2019). Principles of gamification for Internet interventions. Translational behavioral medicine, 9(6), 1131-1138.
[47] van Gaalen, A. E., Brouwer, J., Schönrock-Adema, J., Bouwkamp-Timmer, T., Jaarsma, A. D. C., & Georgiadis, J. R. (2021). Gamification of health professions education: a systematic review. Advances in Health Sciences Education, 26(2), 683-711.

[48] Xu, J., Lio, A., Dhaliwal, H., Andrei, S., Balakrishnan, S., Nagani, U., & Samadder, S. (2021). Psychological interventions of virtual gamification within academic intrinsic motivation: A systematic review. Journal of Affective Disorders, 293, 444-465.
[49] Vajawat, B., Varshney, P., & Banerjee, D. (2021). Digital gaming interventions in psychiatry: evidence, applications and challenges. Psychiatry research, 295, 113585.
[50] Yen, B. T., Mulley, C., & Burke, M. (2019). Gamification in transport interventions: Another way to improve travel behavioural change. Cities, 85, 140-149.
[51] Garett, R., & Young, S. D. (2019). Health care gamification: A study of game mechanics and elements. Technology, Knowledge and Learning, 24, 341-353.
[52] Lieder, F., Chen, P. Z., Prentice, M., Amo, V., & Tošić, M. (2024). Gamification of Behavior Change: Mathematical Principle and Proof-of-Concept Study. *JMIR serious games*, *12*, e43078. https://doi.org/10.2196/43078
[53] Hammady, R., & Arnab, S. (2022). Serious Gaming for Behaviour Change: A Systematic Review. Information, 13(3), 142. https://doi.org/10.3390/info13030142
[54] Marques, T.G., Schumann, S. & Mariconti, E. Positive behaviour interventions in online gaming: a systematic review of strategies applied in other environments. Crime Sci 13, 14 (2024). https://doi.org/10.1186/s40163-024-00208-8
[55] Cheng, C., & Ebrahimi, O. V. (2023). Gamification: a Novel Approach to Mental Health Promotion. Current psychiatry reports, 25(11), 577–586. https://doi.org/10.1007/s11920-023-01453-5
[56] Liang, Zilu & Melcer, Edward & Hoang, Nhung & Khotchasing, Kingkarn. (2024). Sleep Hygiene Games and Gamification: Where Are We and Where Are We Heading?. 10.13140/RG.2.2.32935.56481.
[57] Damaševičius, R., Maskeliūnas, R., & Blažauskas, T. (2023). Serious Games and Gamification in Healthcare: A Meta-Review. Information, 14(2), 105. https://doi.org/10.3390/info14020105
[58] Macey, Joseph & Adam, Martin & Hamari, Juho & Benlian, Alexander. (2024). Examining the Commonalities and Differences Between Gamblification and Gamification: A Theoretical Perspective. International Journal of Human-Computer Interaction. 1-14. 10.1080/10447318.2024.2346690.
[59] Kaya, A., Türk, N., Batmaz, H., & Griffiths, M. D. (2023). Online Gaming Addiction and Basic Psychological Needs Among Adolescents: The Mediating Roles of Meaning in Life and Responsibility. International journal of mental health and addiction, 1–25. Advance online publication. https://doi.org/10.1007/s11469-022-00994-9
[60] Kumari, S., & Dhiksha, J. (2022). Effect of intervention for gaming addiction among adolescents: A systematic review. International Journal of Health Sciences, 6(S3), 10707 10717. https://doi.org/10.53730/ijhs.v6nS3.8718
[61] Ibañez, Roger & Velza, Jacob & Grabrito, Ricky. (2023). Impact of Online Gaming on the Academic Performance of DEBESMSCAT-Cawayan Campus Students. Scientific Journal of Informatics. 10. 423 - 434. 10.15294/sji.v10i4.45007.
[62] Boonen, H., Vanderplasschen, W., Sinclair, D. L., Flayelle, M., & Billieux, J. (2024). Group-based interventions for adolescents with Gaming Disorder or problematic gaming behavior : a systematic review. CURRENT ADDICTION REPORTS, 11(3), 551–564. https://doi.org/10.1007/s40429-024-00570-2
[63] Kaya, A., Türk, N., Batmaz, H., & Griffiths, M. D. (2023). Online Gaming Addiction and Basic Psychological Needs Among Adolescents: The Mediating Roles of Meaning in Life and Responsibility. International journal of mental health and addiction, 1–25. Advance online publication. https://doi.org/10.1007/s11469-022-00994-9
[64]] Sun, X., Yu, V., & Chen, V. H. H. (2024). Toxic behavior in multiplayer online games: the role of witnessed verbal aggression, game engagement intensity, and social self-efficacy. Chinese Journal of Communication, 1–19. https://doi.org/10.1080/17544750.2024.2425662