Jurnal Masyarakat Informatika, 16(1), 2025
DOI:10.14710/jmasif.16.1.72576
ISSN: 2777-0648

Analysis of the Correlation Between Playtime, Design, and Game Mechanics to Positive Reviews on the Fighting Games Genre Using Large Language Models

Vicoriza Ramadhani, Guruh Aryotejo*, and Aris Puji Widodo
Department of Informatics, Diponegoro University, Semarang, Indonesia
* *Corresponding author: guruh2000@gmail.com*

Abstract
The video game industry has experienced rapid growth, with the fighting game genre remaining a favorite due to its tactical challenges and deep mechanics. This study explores the relationship between playtime, game design, and gameplay mechanics on positive user reviews, utilizing Large Language Models (LLM) for sentiment and emotion analysis. Data was collected from over 200,000 user reviews on Steam across 12 popular fighting games. The results indicate that the correlation between playtime and positive reviews tends to be weak, although for some titles, longer play durations are associated with more favorable sentiment. In terms of game design, players tend to prefer games with a fantasy setting (92.34%), 2D graphics (94.21%), and an anime visual style (95.12%), all of which significantly contribute to positive reviews. Regarding gameplay mechanics, features such as multiple meters (93.11%), advanced blocking (93.56%), and wall boundaries (91.72%) received higher satisfaction ratings, suggesting that greater complexity and variety in mechanics can enhance player engagement. Sentiment analysis using LLM also revealed that technical factors play a significant role in player perception. The most common complaints in negative reviews involved lag, character balancing, and the quality of additional downloadable content (DLC).

Keywords : *fighting games, playtime, game design, game mechanics, LLM, steam reviews*

1 Introduction

The video game industry is now one of the most dominant forms of entertainment in the world, attracting attention from a wide range of age groups and genders [1]. In recent years, the industry has surpassed film and music in terms of growth and revenue, making it a highly lucrative and dynamic entertainment sector [2]. One genre that remains popular with fans is fighting games, which require skills such as reaction speed, timing accuracy, and strategy in mastering movements and attacks [3].

In the digital age, distribution platforms like Steam play a crucial role in the development of this industry. Not only does Steam make it easy to access game purchases, but it also provides a space for players to interact with each other and express their opinions in the form of reviews [4]. According to [5], the review system on Steam allows players to give a "Recommended" or "Not Recommended" rating of a game. While it doesn't use a numerical rating system, these reviews are accompanied by text comments that explain the player experience in more depth, making the percentage of positive reviews a key indicator in assessing the quality and popularity of a game.

Reviews on Steam are public and can be accessed through the official API. This data is used to evaluate the gaming experience [5], assess the usefulness of reviews [6][7], and support the development of predictive models in the gaming industry. For buyers, these reviews help decision-making; For developers, it is a reflection of product quality. The duration of playtime also affects the

Received: 22 April 2025 ; Revised: 30 May 2025; Accepted: 31 May 2025; Published: 31 May 2025

depth of the review. Players with longer playing time usually give more critical and detailed reviews [4]. In addition, design elements and gameplay mechanics also make for a satisfying experience. Research shows that the visual and character design aspects increase satisfaction [8], while the combat, exploration, and item collection systems provide deeper challenges and engagement [9].

Textual reviews contain important information about the game and player perceptions. Gamers are known to have high expectations, so their opinions are valuable input for developers [10]. From the reviews, it is understandable the playability of the game [11] and the factors that support its popularity [12], as reviews often contain emotional expressions and opinions that reflect the player's judgment [13]. According to [14], common topics that appear in reviews include gaming experience, social interaction, achievements, accessories, visual value, narrative, and bugs. However, players are more likely to criticize design elements than technical bugs [15].

However, there are a number of loopholes in previous research. First, there have been no studies linking playtime, design, and game mechanics to positive reviews in the fighting games genre [5][8][9]. Second, the previous approach is generally descriptive and lacks an in-depth description of the emotions. Third, there has been no systematic use of Large Language Models (LLMs) to explore sentiments and emotions in the context of fighting games, even though this technology can provide a more accurate understanding of player perceptions [13][14]. This study aims to fill this gap through the correlation analysis between playtime, game design, and mechanics to positive reviews, as well as integrating the LLM approach in large-scale sentiment and emotion analysis.

2 Research Methods

The process carried out in the research can be illustrated in the diagram of Figure 1.

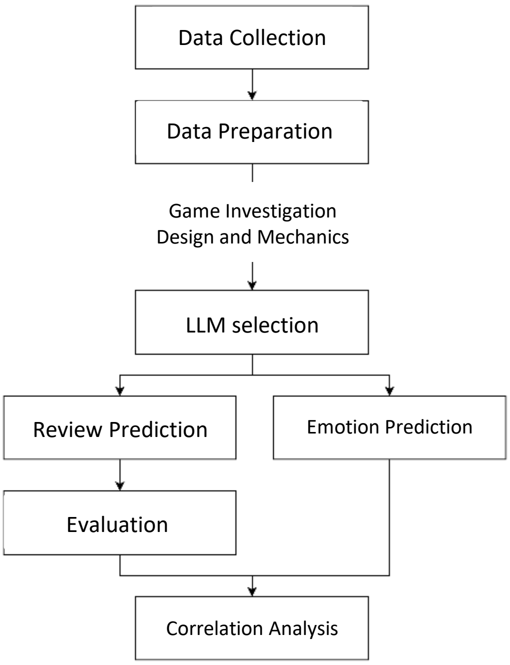

Figure 1 Research Stage Diagram

2.1 Data Collection

Games were selected based on the tags "Fighting" and "Arcade" with manual validation using the fighting game criteria according to [16] to ensure eligibility. These criteria are:

1. This game focuses on close-quarters combat
2. The characters in the game have basic attacks as well as special moves

3. Matches in fighting games are quantitatively measured on the screen
4. This game is competitive
5. Multiplayer features

The selected games are also filtered based on peak concurrent users (CCU) to focus on games with high popularity. After that, the AppID is extracted as a link to the review data. In the review data, only English-language reviews are included to facilitate sentiment analysis. The data is then focused on the selected game using AppID. Variables that are retained include: game name, review language, review text, recommendation status, total playing time, and playtime when review is created. These variables were used to analyze the relationship between playtime, game design and mechanics to positive reviews. Review text is analyzed for sentiment, while recommendation status and playtime help evaluate the correlation with players' perception of the game.

2.2 Game Design and Game Mechanics Investigation

After the data preparation was completed, the important factors of game design and mechanics were mapped based on the research by [17] with modifications of the categories to better suit the fighting game genre. Categories such as gameplay, game mechanics, and interaction elements are structured and elaborated as follows:

1. Setting (SET), the setting or thematic environment in which the game takes place.
2. Graphical Dimension (GD), a visual perspective of the game.
3. Graphical Style (GS), an art style or visual presentation of a game.
4. Character Variations (VAR), a customization of movement variance for game characters .
5. Special Gauges (GAUGE), the type of special meter or gauge system used.
6. Defensive Mechanics (DEF), a defensive option in gameplay.
7. Environmental Interactions (ENV), a type of player interaction with the game environment.
8. Assist Mechanics (ASSIST), the presence or absence of an assist system by another character.
9. Online Infrastructure (NET), a type of netcode used in online features.
10. Training Mode Quality (TRAIN), the quality of the training mode features.

Even though it's based on previous research, each game is still live tested to ensure the relevance of these elements. The factors studied are listed in Table 1:

Table 1 Game Design and Game Mechanics Studied

Categories	Possible Factors		
SET	Fantasy	Urban	Historical
GD	2D	3D	-
GS	Realistic	Anime	-
VAR	Yes	Yes	-
GAUGE	Single Meter	Multiple Meters	-
DEF	Standard Blocking	Advanced Blocking	-
ENV	Wall Boundaries	Ring-out Boundaries	-
ASSISTS	Yes	Yes	-
NET	Rollback	Delay-based	-
TRAIN	Basic Training	Advanced Training	-

2.3 LLM Selection

This study used Large Language Models (LLM) to analyze sentiment and emotion in game reviews, to examine the relationship between playtime, game design and mechanics to positive reviews. The model used is derived from the HuggingFace platform, covering two main categories: Review Prediction LLM and Emotion Prediction LLM.

1. Review Prediction LLM

   This model is used to classify reviews as positive (recommended) or negative (not recommended). The two main approaches are:

   a. Generative Model: Generates sentiment predictions based on specific prompts. The tested model, such as OPT 1.3B, is evaluated using triple retesting to ensure consistency.
   b. Predictive model: Performs a direct classification with a common model (e.g., Amazon) or a custom model (e.g., Zitroeth). The output results are encoded to binary format (1 for positive, 0 for negative) to ensure uniformity.

2. Emotion Prediction LLM

   This model is used to detect emotions in reviews based on seven categories: anger, disgust, fear, joy, sadness, surprise, and neutrality as previously conducted by [18]. Fine-tuned BERT-based models are used to analyze the text and generate numerical outputs to facilitate further analysis.

2.4 LLM Model Evaluation

The evaluation was conducted to measure the performance of the Large Language Model (LLM) used in sentiment analysis and prediction of fighting game player reviews. The metrics used include accuracy, precision, recall, and F1-score. The evaluation step begins by checking the presence of prediction columns, ground truth, and probabilities in the dataset. If the column doesn't exist, the evaluation is stopped. After that, metric calculations are carried out to assess the percentage of correct predictions (accuracy), the balance of positive predictions (precision and recall), and the harmonious average of both (F1-score).

2.5 Correlation Analysis

This analysis aims to evaluate the relationship between playtime, game design and mechanics, as well as player emotions to the percentage of positive reviews. The approach used follows methods that have been proven in previous studies [17], including distribution and non-parametric correlation tests.

Before the correlation analysis was performed, the Kolmogorov-Smirnov test was used to test the normality of the data on each variable. The results of this test are the basis for the selection of the next statistical method. For data that are not normally distributed, Spearman's Rank Correlation and Mann-Whitney U Test are used. Meanwhile, normally distributed data were analyzed using Pearson Correlation. All tests were conducted with a significance level (α) of 0.05.

1. Playtime Correlation

   The correlation between playtime and positive reviews was analyzed by calculating average playing time, total positive reviews, and percentage of positive reviews. The results are presented in a table containing the average playing time, correlation coefficient, and p-value.

2. Correlation of Game Design and Mechanics

   Design factors, such as graphics, visual style, and game mechanics, are analyzed to identify their influence on positive reviews. The data was grouped by design category, then analyzed to compare the distribution of positive reviews between categories. Results include the design categories with the highest and lowest positive reviews, as well as the statistical significance between categories.

3. Textual Review Analysis
Player emotions are analyzed to understand their impact on reviews. Emotions are grouped into six according to [19], namely positive emotions (joy, surprise) and negative emotions (anger, disgust, fear, and sadness), as well as one additional category, namely neutral. This model has been used effectively in various previous text analysis studies [20] [21]. The distribution of emotions is calculated and compared between design categories. In addition, word cloud analysis is done for each emotion to identify dominant words that reflect the player's feelings.

3 Results and Discussion

This research was conducted using a Kaggle Notebook with the support of a Tesla P100 GPU and an Intel Xeon CPU. Data processing is carried out through two main stages: dataset processing and correlation analysis.

3.1 Dataset Processing Results

The study uses two main data taken from the Steam platform: Game Information Data and User Review Data. As a result, 12 popular fighting games were selected with a total of 211,931 reviews. The variables used include: game name, review language, review text, recommendation status, total playtime, and playtime when the review is created. This final data became the basis for sentiment and correlation analysis. The games studied are:

1) TEKKEN 7
2) DRAGON BALL FighterZ
3) GUILTY GEAR -strive-
4) Mortal Kombat 11
5) Street Fighter V
6) Mortal Kombat X
7) Street Fighter™ 6
8) Mortal Kombat 1
9) SOULCALIBUR VI
10) BlazBlue Centralfiction
11) DNF DUEL
12) THE KING OF FIGHTERS XV

3.2 Game Design and Game Mechanics Investigation Results

Table 2 shows the results of game design and mechanics analyzed through live testing on selected games.

Table 2 Game Design and Game Mechanics Investigation Results

Games	SET	GD	GS	VAR	GAUGE	DEF	ENV	ASSISTS	NET	TRAIN
TEKKEN 7	U	3D	R	N	BC	SB	WB	N	D	BT
DRAGON BALL FighterZ	F	2D	A	Y	BC	AB	WB	Y	R	AND
GUILTY GEAR -strive-	F	2D	A	N	MM	AB	WB	N	R	AND
Mortal Kombat 11	F	2D	R	Y	BC	SB	WB	N	R	BT
Street Fighter V	U	2D	R	N	BC	SB	WB	N	R	BT
Mortal Kombat X	F	2D	R	Y	BC	SB	WB	N	D	BT
Street Fighter™ 6	U	2D	R	N	BC	AB	WB	N	R	BT
Mortal Kombat 1	F	2D	R	N	BC	SB	WB	Y	R	BT
SOULCALIBUR VI	H	3D	R	N	BC	SB	RB	N	R	BT
BlazBlue Centralfiction	F	2D	A	N	MM	AB	WB	N	R	AND
DNF DUEL	F	2D	A	N	MM	SB	WB	N	R	BT
THE KING OF FIGHTERS XV	U	2D	R	N	BC	AB	WB	N	R	BT

3.3 Evaluation of LLM Prediction Review

The LLM Review Prediction model is used to classify player reviews as positive or negative. Model evaluation was carried out using *accuracy, precision, recall,* and *F1-score* metrics. The model achieved *an F1-score* of 94.75%, which indicates a high ability to identify positive sentiment on reviews. The results of the LLM evaluation are shown in Table 3.

Table 3 Evaluation Metric Results for LLM Prediction Review

LLM Types	LLM	Acc.	Prec.	Recall	F1 score.
Predictive	Amazon	87,23%	95,38%	89,30%	92,24%
Generative	OPT	84,81%	85,09%	99,59%	91,77%
Predictive	Zitroeth	91,12%	95,27%	94,24%	94,75%

3.4 Emotion LLM Results

The results of the LLM evaluation used to analyze specific emotions in player reviews are shown in Table 4.

Table 4 Emotional Results Based on User Reviews

Games	Emotion						
	Anger	Disgust	Fear	Joy	Sadness	Surprise	Neutral
TEKKEN 7	1930 (4.91%)	1930 (4.91%)	1930 (4.91%)	1930 (4.91%)	1930 (4.91%)	1930 (4.91%)	1930 (4.91%)
DRAGON BALL FighterZ	1274 (4.11%)	1274 (4.11%)	1274 (4.11%)	1274 (4.11%)	1274 (4.11%)	1274 (4.11%)	1274 (4.11%)
GUILTY GEAR -strive-	1600 (5.17%)	1600 (5.17%)	1600 (5.17%)	1600 (5.17%)	1600 (5.17%)	1600 (5.17%)	1600 (5.17%)
Mortal Kombat 11	1238 (4.13%)	1238 (4.13%)	1238 (4.13%)	1238 (4.13%)	1238 (4.13%)	1238 (4.13%)	1238 (4.13%)
Street Fighter V	1356 (6.90%)	1356 (6.90%)	1356 (6.90%)	1356 (6.90%)	1356 (6.90%)	1356 (6.90%)	1356 (6.90%)
MORTAL KOMBAT X	1076 (5.99%)	1076 (5.99%)	1076 (5.99%)	1076 (5.99%)	1076 (5.99%)	1076 (5.99%)	1076 (5.99%)
Street Fighter™ 6	563 (3.78%)	563 (3.78%)	563 (3.78%)	563 (3.78%)	563 (3.78%)	563 (3.78%)	563 (3.78%)
Mortal Kombat 1	428 (4.78%)	428 (4.78%)	428 (4.78%)	428 (4.78%)	428 (4.78%)	428 (4.78%)	428 (4.78%)
SOULCALIBUR VI	287 (4.00%)	287 (4.00%)	287 (4.00%)	287 (4.00%)	287 (4.00%)	287 (4.00%)	287 (4.00%)
BlazBlue Centralfiction	206 (5.60%)	206 (5.60%)	206 (5.60%)	206 (5.60%)	206 (5.60%)	206 (5.60%)	206 (5.60%)
DNF DUEL	147 (5.60%)	147 (5.60%)	147 (5.60%)	147 (5.60%)	147 (5.60%)	147 (5.60%)	147 (5.60%)
THE KING OF FIGHTERS XV	84 (3.47%)	84 (3.47%)	84 (3.47%)	84 (3.47%)	84 (3.47%)	84 (3.47%)	84 (3.47%)

3.5 Correlation Analysis

The Kolmogorov-Smirnov test shows that all variables have an abnormal distribution, Therefore, the analysis is carried out using non-parametric methods such as the Spearman Rank Correlation and the Mann-Whitney Test.

3.5.1 Playtime Correlation Analysis

The correlation between playtime and positive reviews was analyzed using Spearman Rank Correlation, as shown in Table 5 and Figure 2.

Table 5 Correlation Between Playtime Versus Positive Reviews

Games	Average playtime and review	Average percentage of positive reviews (%)	Correlation coefficient	P value
TEKKEN 7	135.09	88.34	0.0019	6.988530 e-01
DRAGON BALL FighterZ	84.07	89.24	0.0382	1.724373 e-11
GUILTY GEAR -strive-	61.40	89.67	-0.0030	5.989639 e-01
Mortal Kombat 11	47.84	85.51	0.0827	9.940448 e-47
Street Fighter V	107.45	67.15	0.1315	1.456556 e-76
Mortal Kombat X	32.74	75.48	-0.0226	2.463134 e-03
Street Fighter™ 6	43.89	88.60	0.0289	4.233861 e-04
Mortal Kombat 1	26.43	73.92	0.0157	1.378168 e-01
SOULCALIBUR VI	58.54	81.88	0.0966	2.412182 e-16
BlazBlue Centralfiction	83.64	90.60	0.0503	4.614590 e-04
DNF DUEL	23.20	70.50	0.0652	8.228255 e-04
THE KING OF FIGHTERS XV	39.78	80.32	0.0065	7.508320 e-01

Figure 2 shows a visualization of the results of the analysis showing that the correlation between playtime and positive reviews is generally weak, with a few exceptions for certain games. The results show that there are games that have high playtime with a high correlation to positive reviews, which means that the longer players play, the more likely they are to leave positive reviews, such as Street Fighter V and Mortal Kombat 11. Some games with low playtime but a high correlation to positive reviews show that even if players don't spend a long time, a satisfying brief experience is enough to give positive reviews, such as SOULCALIBUR VI and DNF DUEL.

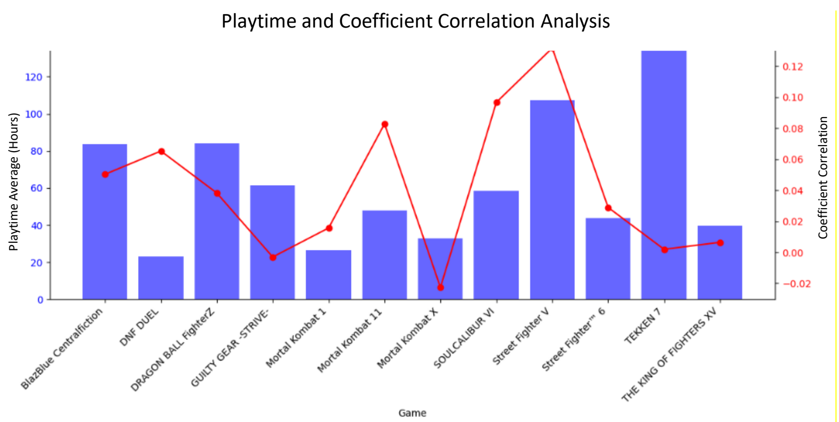

Figure 2 Playtime Visualization and Coefficient Correlation

3.5.2 Analysis of the Correlation Between Game Design and Mechanics Against Positive Reviews

The correlation between playtime and positive reviews was analyzed using Spearman Rank Correlation, as shown in Table 5 and Figure 2. Before analyzing the correlation between game design and mechanics to positive reviews, a significance test was conducted on the difference in the percentage of positive reviews between games. Of all 66 combinations analyzed in Figure 3.

The results of the analysis showed that there were three pairs of games with insignificant differences:

1. DRAGON BALL FighterZ and GUILTY GEAR -STRIVE- (p-value = 0.081573).
2. SOULCALIBUR VI and THE KING OF FIGHTERS XV (p-value = 0.088171).
3. Street Fighter™ 6 then TEKKEN 7 (p-value = 0.393432).

Furthermore, the percentage of positive reviews is analyzed with previously researched game design and mechanics factors. Figure 3 shows the percentage of positive reviews for each factor studied and Table 6 shows the value in more detail. The results are as follows:

Table 6 Correlation Between Game Design and Mechanics Against Positive Reviews

Factors	Category 1	Category 2	Mann-Whitney U Statistic	p-value	Category 1 Sample Size	Category 2 Sample Size
SET	Fantasy	Urban	4.936291 e+09	9.4544 e-47	126346	76306
SET	Fantasy	Historical	4.680486 e+08	1.6242 e-13	126346	7179
SET	Urban	Historical	2.760950 e+08	8.6821 e-02	76306	7179
GD	2D	3D	3.640407 e+09	1.9575 e-104	163307	46524
GS	Anime	Realistic	5.218157 e+09	0	69439	140392
VAR	Yes	Yes	5.117288 e+09	7.0625 e-09	130879	78952
GAUGE	Multiple Meters	Single Meter	3.471172 e+09	8.1981 e-149	38446	171385
DEF	Standard Blocking	Advanced Blocking	4.845325 e+09	0	125712	84119
ENV	Wall Boundaries	Ring-out Boundaries	7.441436 e+08	1.649970 e-07	202652	7179
ASSISTS	Yes	Yes	3.321199 e+09	3.185079 e-25	169890	39941
NET	Rollback	Delay-Based	4.357270 e+09	1.023639 e-01	152530	57301
TRAIN	Basic Training	Advanced Training	4.396394 e+09	0	143019	66812

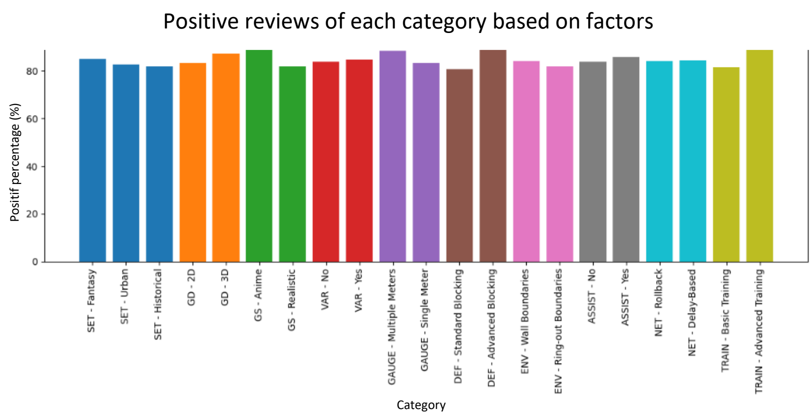

Figure 3 Percentage of Positive Reviews for Each Category

Analysis shows that some game design and mechanics have a significant influence on positive reviews. Setting Fantasy recorded the highest reviews (92.34%), with significant differences in all settings except Urban and Historical. 2D graphics (94.21%) are preferred over 3D (89.97%), while Anime Style (95.12%) is superior to Realistic. Games with feature variety (92.43%) and Multiple Meters (93.11%) get better reviews.

Advanced Blocking (93.56%) is more valued than Standard Blocking, while Wall Boundaries (91.72%) are preferred over Ring-out Boundaries. Games with assists (93.03%) and Advanced Training (92.89%) also received higher positive reviews. However, there was no significant difference between Rollback and Delay-Based networks in the review ($p = 0.102$).

Table 7 Emotional Outcomes Grouped Based on Factors Studied

Factors	Categories	Emotion						
		Anger	Disgust	Fear	Joy	Sadness	Surprise	Neutral
SET	Fantasy	5969 (4.72%)	726 (0.57%)	351 (0.28%)	71285 (56.42%)	2377 (1.88%)	2875 (2.28%)	42763 (33.85%)
	Urban	3933 (5.15%)	626 (0.82%)	210 (0.28%)	43056 (56.43%)	1637 (2.15%)	2084 (2.73%)	24760 (32.45%)
	Historical	287 (4.00%)	44 (0.61%)	24 (0.33%)	4645 (64.70%)	179 (2.49%)	251 (3.50%)	1749 (24.36%)
GD	2D	7972 (4.88%)	1125 (0.69%)	486 (0.30%)	91776 (56.20%)	3347 (2.05%)	4166 (2.55%)	54435 (33.33%)
	3D	2217 (4.77%)	271 (0.58%)	99 (0.21%)	27210 (58.49%)	846 (1.82%)	1044 (2.24%)	14837 (31.89%)
GS	Realistic	6962 (4.96%)	1047 (0.75%)	432 (0.31%)	80474 (57.32%)	44540 (31.73%)	3071 (2.19%)	3866 (2.75%)
	Anime	3227 (4.65%)	349 (0.50%)	153 (0.22%)	38512 (55.46%)	1122 (1.62%)	1344 (1.94%)	24732 (35.62%)
VAR	Yes	3588 (4.54%)	441 (0.56%)	226 (0.29%)	46452 (58.84%)	1415 (1.79%)	1756 (2.22%)	25074 (31.76%)
	Yes	6601 (5.04%)	955 (0.73%)	359 (0.27%)	72534 (55.42%)	2778 (2.12%)	3454 (2.64%)	44198 (33.77%)
GAUGE	Single Meter	8236 (4.81%)	1181 (0.69%)	501 (0.29%)	98953 (57.74%)	3490 (2.04%)	4368 (2.55%)	54656 (31.89%)
	Multiple Meters	1953 (5.08%)	215 (0.56%)	84 (0.22%)	20033 (52.11%)	703 (1.83%)	842 (2.19%)	14616 (38.02%)
DEF	Standard Blocking	6462 (5.14%)	979 (0.78%)	405 (0.32%)	71722 (57.05%)	2877 (2.29%)	3534 (2.81%)	39733 (31.61%)
	Advanced Blocking	3727 (4.43%)	417 (0.50%)	180 (0.21%)	47264 (56.19%)	1316 (1.56%)	1676 (1.99%)	29539 (35.12%)
ENV	Wall Boundaries	9902 (4.89%)	1352 (0.67%)	561 (0.28%)	114341 (56.42%)	4014 (1.98%)	4959 (2.45%)	67523 (33.32%)
	Ring-out Boundaries	287 (4.00%)	44 (0.61%)	24 (0.33%)	4645 (64.70%)	179 (2.49%)	251 (3.50%)	1749 (24.36%)
ASSISTS	Yes	1702 (4.26%)	204 (0.51%)	110 (0.28%)	23279 (58.28%)	678 (1.70%)	779 (1.95%)	13189 (33.02%)
	Yes	8487 (5.00%)	1192 (0.70%)	475 (0.28%)	95707 (56.33%)	3515 (2.07%)	4431 (2.61%)	56083 (33.01%)
NET	Delay-Based	3006 (5.25%)	362 (0.63%)	162 (0.28%)	32702 (57.07%)	1166 (2.03%)	1398 (2.44%)	18505 (32.29%)
	Rollback	7183 (4.71%)	1034 (0.68%)	423 (0.28%)	86284 (56.57%)	3027 (1.98%)	3812 (2.50%)	50767 (33.28%)
TRAIN	Basic Training	7109 (4.97%)	1075 (0.75%)	443 (0.31%)	82009 (57.34%)	3147 (2.20%)	3947 (2.76%)	45289 (31.67%)
	Advanced Training	3080 (4.61%)	321 (0.48%)	142 (0.21%)	36977 (55.34%)	1046 (1.57%)	1263 (1.89%)	23983 (35.90%)

3.5.3 Textual Review Analysis (Emotions)

The results of the analysis of the distribution of emotions on game design factors show an interesting pattern related to the relationship between player emotions and positive reviews. The results of the LLM emotion prediction are shown in Table 7. Based on the results of the analysis listed in Table 7, the seven emotions can be further grouped into positive emotions, negative emotions, and neutral emotions. The results are shown in Table 8.

Table 8 Grouping of Positive, Negative, and Neutral Emotions

Factors	Categories	Emotion		
		Positive	Negative	Neutral
SET	Fantasy	58.70%	7.45%	33.8%
	Urban	59.16%	8.40%	32.45%
	Historical	68.20%	7.43%	24.36%
GD	2D	58.75%	7.92%	33.33%
	3D	60.73%	7.38%	31.89%
GS	Realistic	60.07%	8.21%	31.73%
	Anime	57.40%	6.99%	35.62%
VAR	Yes	61.06%	7.18%	31.76%
	Yes	58.06%	8.16%	33.77%
GAUGE	Single Meter	60.29%	7.83%	31.89%
	Multiple Meters	54.30%	7.69%	38.02%
DEF	Standard Blocking	59.86%	8.53%	31.61%
	Advanced Blocking	58.18%	6.70%	35.12%
ENV	Wall Boundaries	58.87%	7.82%	33.32%
	Ring-out Boundaries	68.20%	7.43%	24.36%
ASSISTS	Yes	60.23%	6.75%	33.02%
	Yes	58.94%	8.05%	33.01%
NET	Delay-Based	59.51%	8.19%	32.29%
	Rollback	59.07%	7.65%	33.28%
TRAIN	Basic Training	60.10%	8.23%	31.67%
	Advanced Training	57.23%	6.87%	35.90%

The results of the sentiment analysis showed that historical settings (68.20%), 3D graphics (60.73%), and realistic visual styles (60.07%) had the highest levels of positive emotions, indicating that players enjoyed historical settings, detailed 3D graphics, and realistic displays. Some factors are more likely to evoke neutral emotions, such as Multiple Meters Gauge (38.02%) and Advanced Blocking (35.12%), which suggests that these elements have little effect on player satisfaction. Meanwhile, the Ring-out Boundaries feature (68.20%) is preferred over Wall Boundaries (58.87%), while games that have the assist (60.23%) and Advanced Training (57.23%) features contribute to improving the gaming experience. From the network side, Netcode rollback (59.07%) has lower negative emotions (7.65%) than Delay-Based Netcode (8.19%), indicating that connection stability greatly affects player satisfaction.

Overall, the historical setting, 3D graphics, assist feature, and ring-out boundaries mechanics contribute to enhancing a positive gaming experience, while Multiple Meters and Advanced

Blocking more often elicit a neutral reaction. These findings confirm that design elements and game mechanics play an important role in shaping players' perceptions of fighting games.

Next, an analysis was carried out from the results of the word cloud of each emotion. Word clouds for anger emotions, such as in Figure 4 highlight words such as "suck", "buy", and "time" that reflect players' dissatisfaction with the quality of the game, paid features (DLC), as well as the online gaming experience. Complaints are also related to mechanics such as combo systems and technical issues in the game.

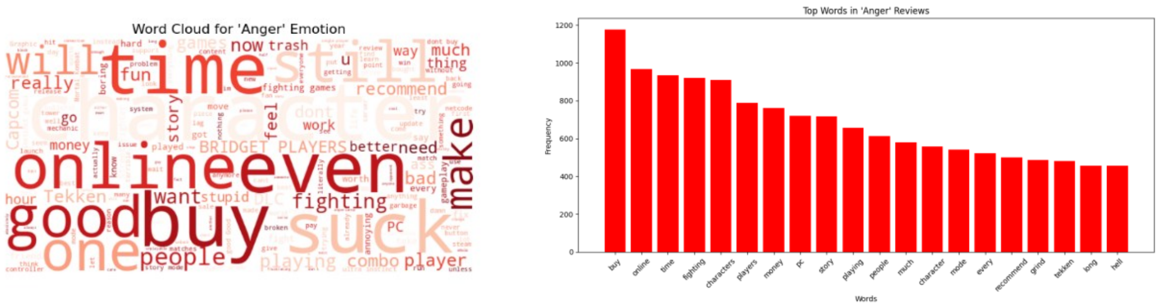

Figure 4 Word Cloud and Anger Emotion Chart Bar

The word cloud for joy emotions, shown in Figure 5, highlights words such as "good", "fun", "great", and "character", reflecting the player's positive experience of gameplay, character design, and game mechanics. Words like "combo" indicate that the combo system contributes significantly to player satisfaction.

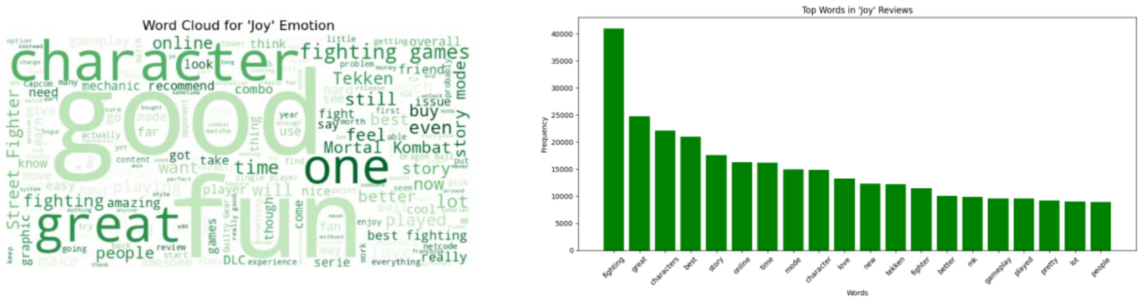

Figure 5 Word Cloud and Bar Chart Joy Emotions

Word cloud for emotion disgust, in Figure 6 highlights words such as "worst", "suck", "awful", and "garbage" that reflect the player's disappointment with the gaming experience. Dissatisfaction is often related to character design, online performance, or additional content (DLC). Words like "buy" and "played" indicate players feel the experience gained is not worth the cost incurred.

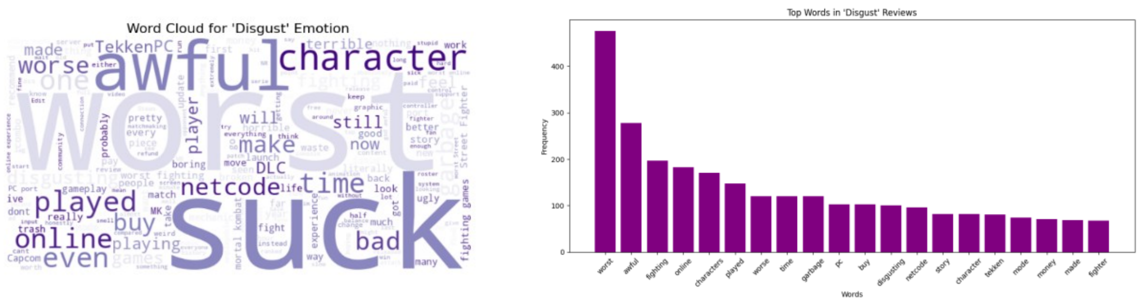

Figure 6 Word Cloud and Bar Chart of Disgust Emotions

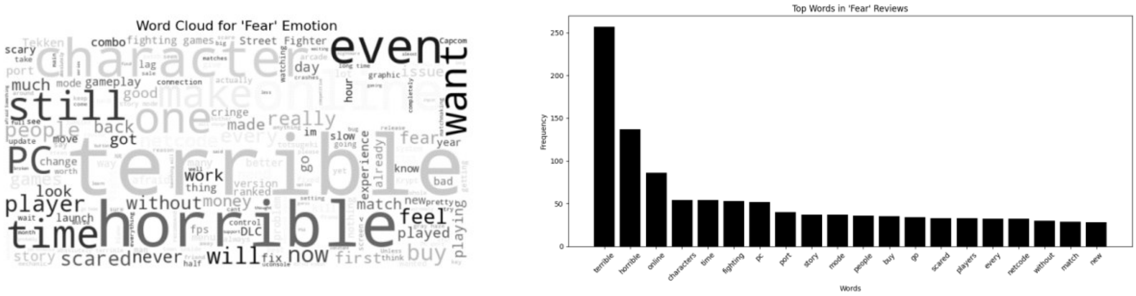

Figure 7 Word Cloud and Fear Emotion Chart Bar

Word cloud for sadness emotions, in Figure 8 highlights words such as "character" and "online" that reflect the player's disappointment with character design, or online experiences. Terms such as "feel", "still", and "time" indicate dissatisfaction with time spent in the game, while the combination of "good" and "bad" describes the contradiction between potentially good games but failing to meet expectations.

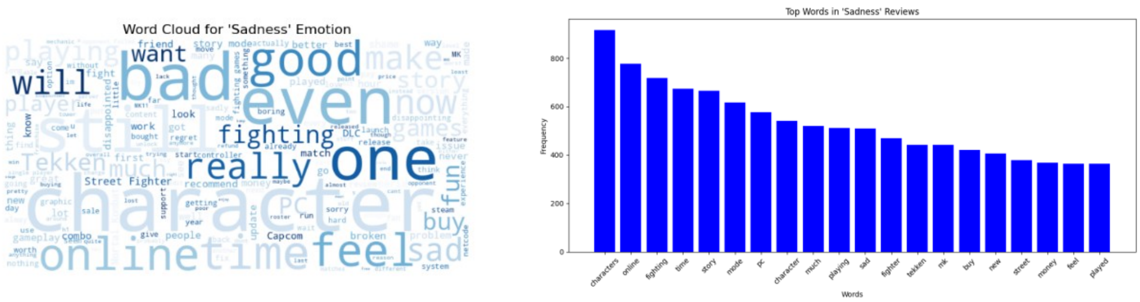

Figure 8 Word Cloud and Sadness Emotion Chart Bar

Word cloud for emotion surprise, in Figure 9 features words such as "one", "time", "online", and "character" reflecting elements that surprise players in game mechanics, online experience, or character design. Terms like "good" and "make" indicate the positive surprises of innovative gameplay, while "will" and "really" reflect a response to unexpected elements, both innovation and problem.

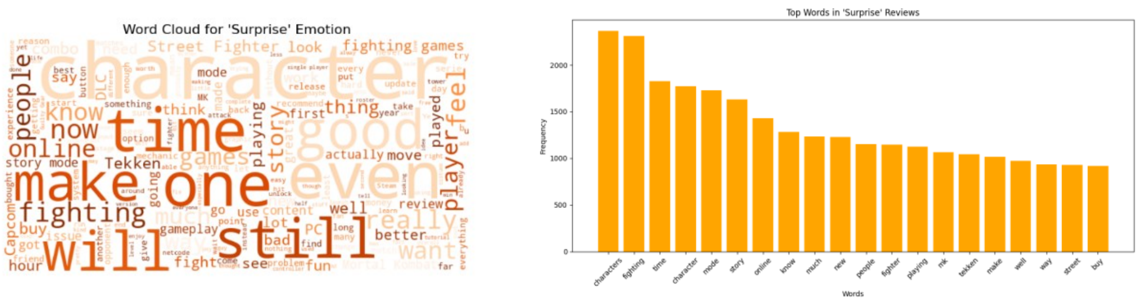

Figure 9 Word Cloud and Bar Chart Emotions Surprise

Finally, the word cloud for neutral emotions, shown in Figure 10, features words such as "character", "fighting", and "time" that reflect the gaming experience that is considered normal or balanced. Elements such as term words such as "TOTSUGEKI" and character names (e.g., "Bridget", "Chun-Li") reflect elements of the community's culture that are accepted without significant emotional reactions.

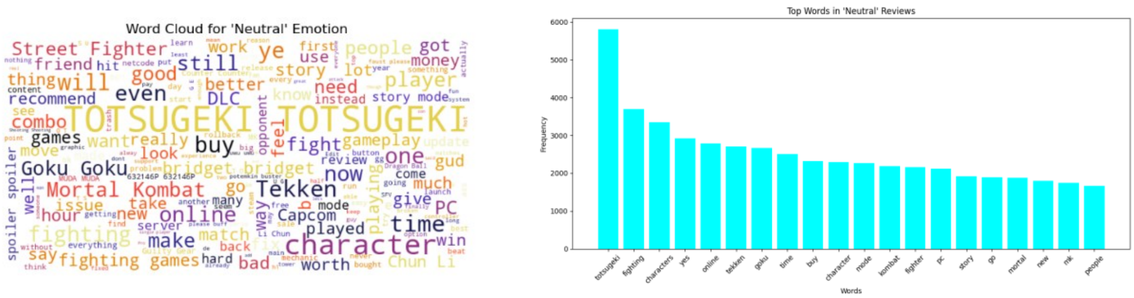

Figure 10 Word Cloud and Neutral Emotion Chart Bar

4 Conclusion

The results of the analysis show that the correlation between playtime and positive reviews is generally weak, with the exception of Street Fighter V and Mortal Kombat 11, where players with longer playtime tend to give positive reviews. In contrast, in Guilty Gear -STRIVE- and Mortal Kombat X, playtime is not significantly related to player satisfaction, indicating that factors such as gameplay mechanics, character balancing, and online features are more influential than playtime.

Several aspects of game design and mechanics also affect positive reviews. Fantasy settings (92.34%), 2D graphics (94.21%), and Anime visual styles (95.12%) are preferred over other categories. In terms of mechanics, Multiple Meters (93.11%), Advanced Blocking (93.56%), and Wall Boundaries (91.72%) received higher reviews, indicating players' preference for more complex and varied features. Historical settings (68.20%), 3D graphics (60.73%), and realistic visual styles (60.07%) had the highest proportion of positive emotions, indicating that players enjoy more detailed visual appearances and attractive settings. Ring-out Boundaries (68.20%) are preferred over Wall Boundaries (58.87%), indicating a preference for opponent elimination mechanisms over fixed arena boundaries.

Utilizing Large Language Models (LLMs), the study identified patterns of emotions in player reviews. Positive sentiment is dominant in more optimal game design and mechanics , while negative reviews are more due to technical issues, such as lag, character balancing, and inadequate DLC quality. This research confirms that in fighting games, design and mechanics play a greater role in shaping player perception compared to playtime. Features such as visual design, more interesting combat mechanics, and stable connectivity in online mode contribute to increasing positive reviews. On the other hand, technical factors such as feature limitations, character imbalances, or a poor online experience can lead to negative reviews, regardless of the amount of time spent in the game.

Acknowledgments

The author would like to express his deepest gratitude to Guruh Aryotejo, S.Kom., M.Sc. and Dr. Aris Puji Widodo, S.Si., M.T. for their guidance, support, and invaluable input during this research. Gratitude was also expressed to Diponegoro University, which has provided facilities and resources that greatly support the implementation of this research. Finally, the author would like to thank the family for always providing moral support and motivation, so that this research can be completed properly.

Bibliography

[1] L. R. Cunha, A. A. B. Pessa and R. S. Mendes, "Shape patterns in popularity series of video games", 2024 http://arxiv.org/abs/2406.10241.

[2] E. Goh, O. Al-Tabbaa and Z. Khan, "Unravelling the complexity of the Video Game Industry: An integrative framework and future research directions", Tele. and Info. Repo., 12, 2023. https://doi.org/10.1016/j.teler.2023.100100.
[3] A. Jhon,. "The influence of playing 3D fighting games in practicing martial arts", Jour. of Phys.: Confer. Series., 1175(1), 2019. https://doi.org/10.1088/1742-6596/1175/1/012253.
[4] L. Eberhard, Kasper, P, P. Koncar and C. Gutl, “Investigating Helpfulness of Video Game Reviews on the Steam Platform”, 2018 5th Inter. Conf. on Soci. Netw. Anal, man. and Secu., SNAMS 2018, 43–50, 2018 https://doi.org/10.1109/SNAMS.2018.8554542.
[5] H.-N. Kang, H.-R. Yong and H.-S. Hwang. "A Study of Analyzing on Online Game Reviews Using a Data Mining Approach: STEAM Community Data". Internal. Jour. of Innovation, Manag. and Techn., 90–94 2017. https://doi.org/10.18178/IJIMT.2017.8.2.709.
[6] Baowaly, M. K., Tu, Y. P., & Chen, K. T. (2019). Predicting the helpfulness of game reviews: A case study on the steam store. Journal of Intelligent and Fuzzy Systems, 36, 4731–4742. https://doi.org/10.3233/JIFS-179022.
[7] Wang, Z., Chang, V., & Horvath, G. (2021). Explaining and predicting helpfulness and funniness of online reviews on the steam platform. Journal of Global Information Management, 29, 1–23. https://doi.org/10.4018/JGIM.20211101.oa16.
[8] P. Skalski and R. Whitbred, "Image versus Sound: A Comparison of Formal Feature Effects on Presence and Video Game Enjoyment", Psych. Day., 2010.
[9] B. Bostan and U. Kaplancali, "Explorations in Player Motivations: Game Mechanics". Euro. GAME-ON Conf., 2009.
[10] Chambers, C., Feng, W. C., Sahu, S., & Saha, D. (2005). Measurement-based characterization of a collection of on-line games. In Proceedings of the 5th ACM SIGLO conference on internet measurement (IMC) (pp. 1–14).
[11] Busurkina, I., Karpenko, V., Tulubenskaya, E., & Bulygin, D. (2020). Game experience evaluation. a study of game reviews on the steam platform. In International conference on digital transformation and global society (pp. 117–127). Cham: Springer. https://doi.org/10.1007/978-3-030-65218-0_9.
[12] Lin, D., Bezemer, C. P., Zou, Y., & Hassan, A. E. (2019). An empirical study of game reviews on the Steam platform. Empirical Software Engineering, 24, 170–207. https://doi.org/10.1007/s10664-018-9627-4.
[13] Li, X., Zhang, Z., & Stefanidis, K. (2021). A data-driven approach for video game playability analysis based on players' reviews. Information, 12. https://doi.org/ 10.3390/info12030129. Article 129.
[14] Ahn, S., Kang, J., & Park, S. (2017). What makes the difference between popular games and unpopular games? Analysis of online game reviews from steam platform using word2vec and bass model. ICIC Express Letters, 11, 1729–1737.
[15] Wase, K., Ramteke, P., Bandewar, R., Badole, N., Kumar, B., & Bogawar, M. P. S. (2018). Sentiment Analysis of product review. International Journal of Innovations in Engineering and Science, 3, 8–13.
[16] Harper and Todd, "The culture of digital fighting games: performance and practice", 172, 2014
[17] T. Guzsvinecz, "The correlation between positive reviews, playtime, design and game mechanics in souls-like role-playing video games". Multi. Tool. and Appli., 82(3), 4641–4670, 2023 https://doi.org/10.1007/S11042-022-12308-1/FIGURES/8.

[18] S. Pan, G. J. W. Xu, K. Guo, S. H. Park and H. Ding. "Cultural Insights in Souls-like Games: Analyzing Player Behaviors, Perspectives, and Emotions Across a Multicultural Context". IEEE Transac. On Gam., 2024. https://doi.org/10.1109/TG.2024.3366239.

[19] P. Ekman *et al.*, "An argument for basic emotions," *Cognition and emotion*, vol. 6, no. 3-4, pp. 169–200, 1992.

[20] S. Butt, S. Sharma, R. Sharma, G. Sidorov, and A. Gelbukh, "What goes on inside rumour and non-rumour tweets and their reactions: A psycholinguistic analyses," *Computers in Human Behavior*, vol. 135, p. 107345, 2022.

[21] D. Rozado, R. Hughes, and J. Halberstadt, "Longitudinal analysis of sentiment and emotion in news media headlines using automated labelling with transformer language models," *Plos one*, vol. 17, no. 10, p. e0276367, 2022.