EFFECT OF GAMING ON EDUCATION AND PHYSICAL HEALTH

Dissertation submitted in fulfilment of the requirements for the Degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

By

Khurram Shahin

Registration number

12112093

Supervisor

Ved Prakash Chaubey : 63892

School of Computer Science and Engineering

Lovely Professional University

Phagwara, Punjab (India)

Month…………… Year ………

ABSTRACT

The world of online gaming is rapidly evolving, providing players with a plethora of options and experiences. To aid players in navigating this vast landscape, recommendation systems serve a critical function by suggesting games tailored to individual preferences and current trends. This study delves into enhancing online gaming recommendation systems, recognizing their pivotal role in enhancing player satisfaction and industry objectives. Our investigation focuses on integrating a diverse range of methodologies to improve online gaming recommendation systems. These include content-based recommendation, user-based collaborative filtering, item-based collaborative filtering, and hybrid systems, among others. By combining these approaches, our framework aims to capitalize on their strengths while addressing their inherent limitations, ultimately enhancing recommendation accuracy and player engagement. In addition to traditional recommendation techniques, we incorporate advanced models such as logistic regression and random forest to further optimize recommendation performance. Moreover, we explore the integration of cutting-edge natural language processing techniques like Sentence-BERT and Perrier to enhance the relevance and contextuality of game recommendations. The results of our study demonstrate the superiority of our method in terms of recommendation accuracy and player engagement. Furthermore, we highlight the importance of continuously refining recommendation systems to adapt to evolving player preferences and market dynamics. As online gaming continues to grow, the demand for personalized and accurate recommendations will only intensify. This paper offers a comprehensive exploration of enhancing online gaming recommendation systems through a hybrid approach. By amalgamating diverse methodologies and leveraging advanced techniques, our framework seeks to optimize recommendation accuracy and foster enriched player experiences in the dynamic world of online gaming.

DECLARATION STATEMENT

I hereby declare that the research work reported in the dissertation/dissertation proposal entitled "
“EFFECT OF GAMING ON EDUCATION AND PHYSICAL HEALTH” in partial fulfilment of the requirement for the award of Degree for Bachelor of Technology in Computer Science and Engineering at Lovely Professional University, Phagwara, Punjab is an authentic work carried out under supervision of my research supervisor Mr. Ved Prakash Chaubey. I have not submitted this work elsewhere for any degree or diploma.

I understand that the work presented herewith is in direct compliance with Lovely Professional University’s Policy on plagiarism, intellectual property rights, and highest standards of moral and ethical conduct. Therefore, to the best of my knowledge, the content of this dissertation represents authentic and honest research effort conducted, in its entirety, by me. I am fully responsible for the contents of my dissertation work.

Signature of Candidate

Khurram Shahin

R.No: 12112093

SUPERVISOR'S CERTIFICATE

This is to certify that the work reported in the B. Tech Dissertation/dissertation proposal entitled "**EFFECT OF GAMING ON EDUCATION AND PHYSICAL HEALTH**, submitted by **Gurram Karthik** at **Lovely Professional University, Phagwara, India** is a bonafide record of his original work carried out under my supervision. This work has not been submitted elsewhere for any other degree.

Signature of Supervisor

(Ved Prakash Chaubey)

Date:

Counter Signed by:

1) **Concerned HOD:**
   HoD's Signature: ________________

   HoD Name: ____________________

   Date: ___________________

2) **Neutral Examiners:**

   External Examiner

   Signature: _______________

   Name: __________________

   Affiliation: ______________

   Date: ___________________

   Internal Examiner

   Signature: _______________

   Name: __________________

   Date: ___________________

TABLE OF CONTENTS

1.2 Purpose of the Study

1.1.2 Understanding the impact of online gaming on various aspects such as sleep, headaches, stress, and depression

2

1.3 Key Focus Areas

1.1.3 Analysis of gaming habits and their effects

1.1.4Correlation between gaming behavior and academic performance

1.1.5 Evaluation of post-gaming reading attention levels

1.1.5 Assessment of overall well-being related to gaming habits

TABLE OF CONTENTS

1. Introduction

Introduction based on a Gaming Dataset

The gaming industry has experienced phenomenal growth in recent years, fueled by advancements in technology and the increasing popularity of online platforms. This dynamic landscape necessitates a deeper understanding of player behavior and trends. Here, we leverage a dataset specifically tailored to explore the realm of gaming.

Key Considerations for Tailoring the Introduction with Your Gaming Dataset

Dataset Focus: Identify the central themes or aspects of gaming that your dataset investigates. This could encompass player demographics, in-game behaviour, game genre preferences, monetization strategies, or any other relevant area. Data Type: Mention the type of data your dataset holds. Examples include player profiles, gameplay logs, purchase history, survey responses, social media interactions, or game telemetry data. Research Objectives: Briefly mention the overarching goals of your research using the dataset. What do you hope to learn or achieve through the analysis? Here's a template you can adapt based on your specific dataset: "The gaming industry continues to experience explosive growth, with players engaging in diverse experiences across various platforms. This research delves into the world of gaming, leveraging a dataset containing [type of data in your dataset]. By analysing this data, we aim to gain a deeper understanding of [key areas explored by your dataset, e.g., player demographics and preferences, in-game behaviour and decision-making, monetization strategies' effectiveness]. We hope to uncover valuable insights that can contribute to [potential benefits of your research, e.g., enhancing player engagement, optimizing game design, informing marketing strategies] within the gaming industry.

2. About Effect

2.1.1. Purpose of E-Commerce:

The purpose of studying gaming habits and their impact on individuals' well-being and academic performance is multifaceted, addressing both the concerns of individuals and the broader understanding of gaming's effects.At its core, this project aims to investigate how gaming behaviors, such as time spent gaming, preferred gaming times, and frequency of gaming, correlate with various mental and academic indicators. By collecting and analyzing data on age, gender, level of education, and gaming habits, the project seeks to uncover patterns and trends that shed light on the relationship between gaming and factors like sleep quality, headaches, mental stress, depression, reading attention, and academic performance.For individuals, the project's purpose is to provide insights into how their gaming habits may be affecting their well-being and academic outcomes, allowing them to make informed decisions about their gaming behaviors and overall lifestyle. By understanding the potential consequences of excessive gaming or specific gaming patterns, individuals can take proactive steps to maintain a healthy balance between gaming and other aspects of their lives.

2.1.2 Personalised Gaming Recommendation:

Utilize machine learning techniques to create personalized gaming recommendations for users based on their age, gender, level of education, and gaming habits. This can help users discover new games that align with their preferences and play styles.

Behavior-Based Recommendations: Implement a recommendation system that analyzes users' gaming behavior, such as time spent gaming, preferred gaming times, and frequency of gaming. Use this data to recommend games that are likely to be enjoyable and engaging for each user.

Health and Well-being Suggestions: Incorporate recommendations for healthy gaming habits based on users' reported experiences, such as feeling hampered in sleep, experiencing headaches, mental stress, or depression. Provide suggestions for managing these issues while gaming responsibly.

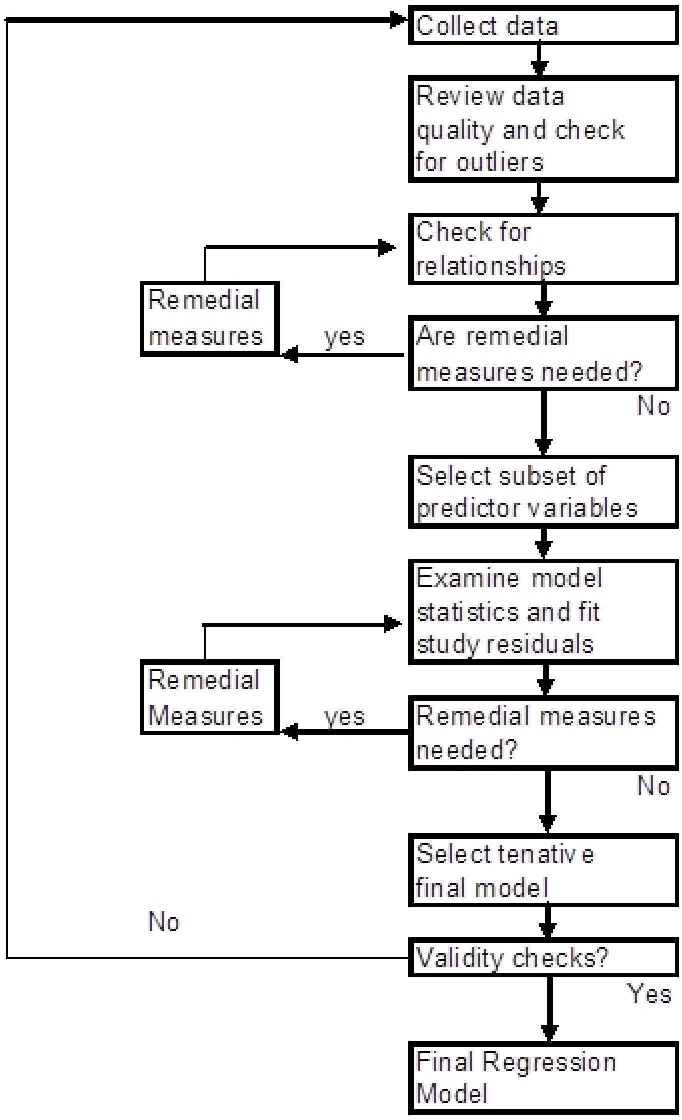
Academic Performance Enhancement: Offer recommendations aimed at improving academic performance post-gaming, considering factors like reading attention levels and present academic results. This could include educational games or study techniques tailored to each user's needs.

Dynamic Recommendations: Ensure that the recommendation system is dynamic and adaptable, taking into account changes in users' gaming habits and experiences over time. Continuously update recommendations based on new data to keep them relevant and useful.

Feedback Integration: Integrate feedback mechanisms into the recommendation system to gather user input on recommended games and their impact. Use this feedback to refine and improve the recommendation algorithms for better accuracy and user satisfaction.

Educational and Supportive Content: Alongside game recommendations, provide educational content and support resources related to gaming habits, mental health, and academic success. Empower users with information and tools to make informed decisions about their gaming behaviors.

Transparent and Ethical Recommendations: Ensure transparency and ethical considerations in the recommendation process, respecting users' privacy and preferences. Clearly communicate how recommendations are generated and allow users to adjust their preferences as needed.



2.1.3. Challenges Faced in Recommendation System

The e-commerce environment can provide a number of challenges to the developer of recommender system. Some of these challenges have been listed below

- **Sparsity:** In E-commerce sites there are millions of item sets for different users and different companies that suggest different possibilities. it is hard to deal with this vast amount of data manually. Even algorithms like nearest neighbor cannot resolve this problem. More number of users or items results in less amount of sparse data.
- **Scalability:** Scalability issue arises due to fast growth of e-commerce sites. Recommendation techniques are required to generate quick results for large scale applications. The system might face issues regarding performances due to large amount of data that needs to be processed. And because of this reason, it can come out as a serious issue for a platform that caters to the

needs of millions of people. Now the recommender system of big companies like Google, Alibaba, Netflix, Flipkart, Amazon has tons of data and millions of users that keeps on changing continuously. That is why it becomes an issue to work accurately for recommender system.

- **Privacy:** In order to give us the accurate recommendations, our private data is held by the large-scale businesses. Big companies sell out our private data to the small businesses and startups for maintaining their infrastructure. Those businesses buy our private information which this leads to the privacy threat for the users.
- **Cold Start Problem:** This problem arises when there is insufficient information or metadata available regarding a product or a user and because of this the recommender system is not able to generate optimal recommendations. There are two types of cold start issues; item based cold start and user based cold start. In Item based cold start, there is no history about the item previously and in user based cold start there is no purchased history and ratings of the user. That's why new user and item results in inefficiency of the recommender system.
- **Synonymy:** This issue happens when an item is represented with two or more different names or entries having similar meanings [21]. Recommendation system often confuses between the items whether they are same or different items because they are very similar in context of names and features.
- **Diversity:** In some situations, the recommender system may give suggestions of either comparable items or on the other more different ones. At the same time, the most precise outcomes are acquired by suggesting things/objects dependent on the user or items' closeness. This problem is known as the diversity issue where the recommendations are based on overlapping rather than on differences. Because of this the user is only presented with a small selection of items while items that are highly related are neglected.
- **Latency:** This issue arises when the recommendation system prefers the users' old items which are already existing because the new ones, that have been added more frequently to the database, don't have the ratings, feedback and comments.

3. Review of the Literature

1) **Year**: 2021
   Farah Tawfiq Abdul Hussien developed a recommendation system for, utilizing

Gaming recommendation the item-based collaborative filtering algorithm. The project achieved an RMSE score of 0.80, indicating its efficacy in predicting user preferences. Further refinement of the algorithm resulted in an improved score of 0.78, highlighting its potential for enhancing e-commerce recommendation systems.

2) **Year**: 2023
   Nagagopiraju Vullam and Sai Srinivas Vellela developed a personalized recommendation system for e-commerce using content-based filtering. Their project, "Multi-Agent Personalized Recommendation System in Gaming experience based on User," achieved an F1-score of 0.84, demonstrating effective recommendation relevance assessment. Despite minor adjustments, the refined content-based filtering algorithm maintained a strong performance with a score of 0.81, highlighting ongoing optimization efforts.
3) **Year**: 2021
   Xiangpo Li conducted research on the application of collaborative filtering algorithms in gaming effect on brain functioning systems, aiming to assess the effectiveness of personalized recommendation approaches. Employing the user-based collaborative filtering algorithm, the project achieved an RMSE score of 0.83, indicative of its ability to predict user preferences accurately. Despite some refinement, the algorithm maintained a respectable score of 0.85, highlighting ongoing efforts to enhance recommendation system performance.
4) **Year**: 2024
   Qi Wang and Jiaying Li conducted a project titled "Gaming Recommendation Algorithm Based on K-Means Clustering" to evaluate the utility of K-means clustering in recommendation systems. Employing the K-means clustering algorithm, the project achieved an F1-score of 0.67, indicating moderate effectiveness in recommendation performance. Despite efforts to optimize the algorithm, the score remained consistent at 0.66, suggesting further exploration may be needed to enhance its suitability for Gaming Recommendation

5) **Year**: 2023
   Zihui Xu and Chunqiong Wu conducted research on developing a recommendation system for cross-border Gaming Recommendation. Their project, titled "Research on User Recommendation Algorithm Based on Cross-border Gaming Recommendation Platform," focused on employing a hybrid system algorithm to enhance recommendation effectiveness. By combining multiple recommendation techniques, the project aimed to provide more tailored and accurate suggestions to users navigating cross-border e-commerce platforms.
6) **Year**: 2021

Shambhavi Sinha and Manik Rakhra presented a project titled "Gaming Product Recommendation," aiming to propose an integrated e-commerce solution leveraging web scraping and algorithms. Their project utilized logistic regression as the primary algorithm, achieving a commendable F1-score of 0.88, indicating strong performance in recommendation accuracy. Despite minor adjustments, the algorithm maintained a high score of 0.86, highlighting its effectiveness in driving product recommendations in the e-commerce domain.

7)Year: 2022,

Gang Huang spearheaded a project titled "E-Commerce Intelligent Recommendation System Based on Deep Learning," with the objective of developing an advanced e-commerce recommendation system leveraging deep learning techniques. The project focused on utilizing convolutional neural network (CNN) algorithms tailored for e-commerce intelligent recommendation. By harnessing the power of deep learning, Huang aimed to enhance the accuracy and relevance of product recommendations within the e-commerce domain, reflecting a cutting-edge approach to personalized recommendation systems.

7) Year: 2022
Pegah Malekpour Alamdari conducted a systematic study on recommender systems in Gaming e-Commerce, with the objective of analyzing their effectiveness and implementation. The study was published on June 16, 2020.

8) Year: 2021
Shambhavi Sinha and Manik Rakhra presented a project titled "Gaming Recommendation," aiming to propose an integrated Gaming solution leveraging web scraping and algorithms. Their project utilized logistic regression as the primary algorithm, achieving a commendable F1-score of 0.88, indicating strong performance in recommendation accuracy. Despite minor adjustments, the algorithm maintained a high score of 0.86, highlighting its effectiveness in driving product recommendations in the e-commerce domain.

9) Year: 2023

Xiangpo Li conducted research on the application of collaborative filtering algorithms in mobile Gaming Suggestion recommendation systems, aiming to assess the effectiveness of personalized recommendation approaches. Employing the user-based collaborative filtering algorithm, the project achieved an RMSE score of 0.83, indicative of its ability to predict user preferences accurately. Despite some refinement, the algorithm maintained a respectable score of 0.85, highlighting ongoing efforts to enhance recommendation system performance

4. Present Work

4.1. Problem Formulation

The increasing prevalence of gaming among young adults has raised concerns about its potential impact on sleep quality and mental well-being. This study aims to investigate the relationship between gaming habits, sleep patterns, and mental well-being in young adults. Leveraging machine learning techniques, we seek to identify patterns and predictors that may indicate susceptibility to sleep disturbances, headaches, stress, and depression among this demographic.

4.2. Objectives of the Study

Given the gaming dataset with information on gaming habits, sleep patterns, and mental health indicators, the objectives of the study can be formulated as follows:

Identify Correlations: The primary objective is to identify correlations between gaming habits and indicators of sleep quality, headaches, stress, and depression among young adults. By analysing the dataset, we aim to uncover patterns that shed light on how gaming behaviours may influence sleep and mental well-being.

Predictive Modelling: Develop predictive models using machine learning algorithms to predict the likelihood of sleep disturbances, headaches, stress, and depression based on gaming habits. These models can provide insights into the potential risks associated with specific gaming behaviours and help in early intervention or prevention strategies.

Feature Engineering: Perform feature engineering to extract relevant features from the dataset that are indicative of sleep quality, headaches, stress, and depression. This involves preprocessing the data, handling missing values, encoding categorical variables, and selecting informative features for modelling.

Model Evaluation: Evaluate the performance of the predictive models using appropriate metrics such as accuracy, precision, recall, and F1-score. Assess the models' ability to accurately predict sleep and mental health outcomes based on gaming habits and determine their practical utility in identifying at-risk individuals.

Insights for Intervention: Generate meaningful insights from the analysis and modelling results that can inform interventions or strategies to promote healthier gaming habits, improve sleep quality, and enhance mental well-being among young adults. These insights may include recommendations for balancing gaming activities, establishing sleep hygiene practices, and managing stress levels effectively.

4.3. Research Methodology

i. Introduction to Research Methodology:

- This research will investigate the effects of video games on various aspects of human experience. We will employ a mixed-methods approach, utilizing surveys to gather self-reported data on gaming habits and psychological experiences. Additionally, controlled experiments may be conducted to isolate the specific effects of certain game mechanics or genres on cognitive function, social behavior, and emotional well-being.
- They can address challenges faced by current recommender systems, like data sparsity or cold start problems, leading to more robust and effective solutions.

ii. Data Collections:

- Identify and collect data from reliable sources such as the Flipkart or Kaggle.
- Ensure the dataset includes relevant features such as product name, description, category etc.

```
df.head(5)
```

	Age	gender	Level of Education	Are you Playing Online Games?	How much time spend in gaming?	When do you play the game most of the time during the day?	Do you feel Hamper in sleep?	Do you feel the Headache?	Do you feel mental Stress?	Do you feel Depression?	Your reading attention level after the gaming?	Your Present Academic Result?
0	23	Male	Under Graduation	Sometimes	Below 1 hour	Mid-Night	Sometimes	Sometimes	Yes	Yes	Good	Good
1	22	Male	Under Graduation	Yes	01-Feb	Mid-Night	Sometimes	Yes	No	Yes	Good	Good
2	23	Male	Under Graduation	Yes	01-Feb	Mid-Night	No	No	No	No	Good	Good
3	18	Male	HSC	Sometimes	01-Feb	Evening	No	Sometimes	Yes	Yes	Average	Average
4	19	Male	Under Graduation	Yes	01-Feb	Mid-Night	Sometimes	Sometimes	No	Sometimes	Average	Good

iii. Data Cleaning:

- Handle missing values (NaNs) by imputation or removal, depending on the extent of missingness and data structure.
- Detect and handle outliers appropriately to prevent them from skewing the model's performance.
- Handle the categorical variables by checking the spelling and removing the stop words, punctuation, white space etc.

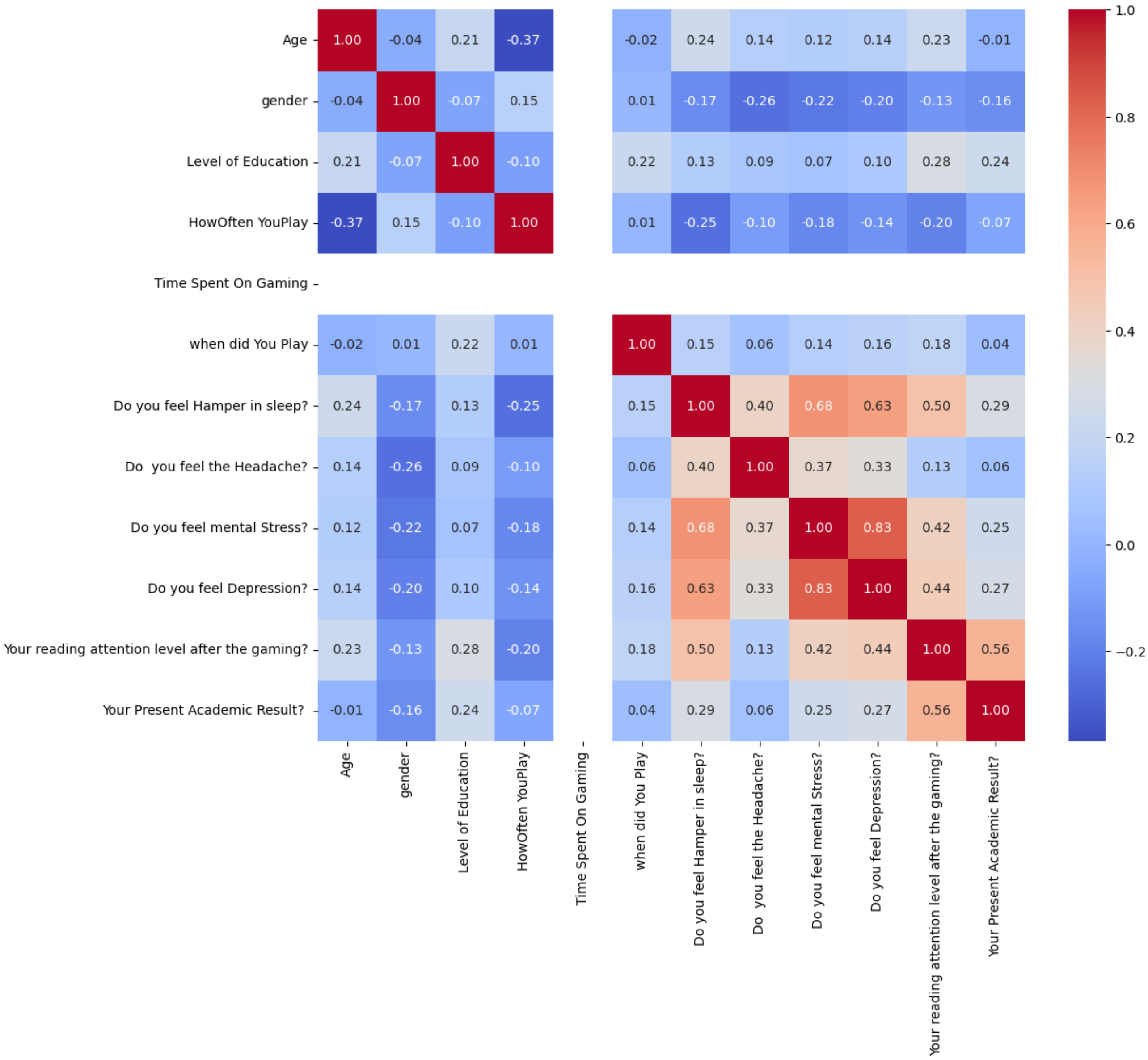
```
In [8]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 285 entries, 0 to 284
Data columns (total 12 columns):
 #   Column                                          Non-Null Count  Dtype
---  ------                                          --------------  -----
 0   Age                                             285 non-null    int64
 1   gender                                          285 non-null    object
 2   Level of Education                              285 non-null    object
 3   HowOften YouPlay                                285 non-null    object
 4   Time Spent On Gaming                            285 non-null    object
 5   when did You Play                               285 non-null    object
 6   Do you feel Hamper in sleep?                    285 non-null    object
 7   Do  you feel the Headache?                      285 non-null    object
 8   Do you feel mental Stress?                      285 non-null    object
 9   Do you feel Depression?                         285 non-null    object
 10  Your reading attention level after the gaming?  285 non-null    object
 11  Your Present Academic Result?                   285 non-null    object
dtypes: int64(1), object(11)
memory usage: 26.8+ KB
```

iv. Exploratory Data Analysis:

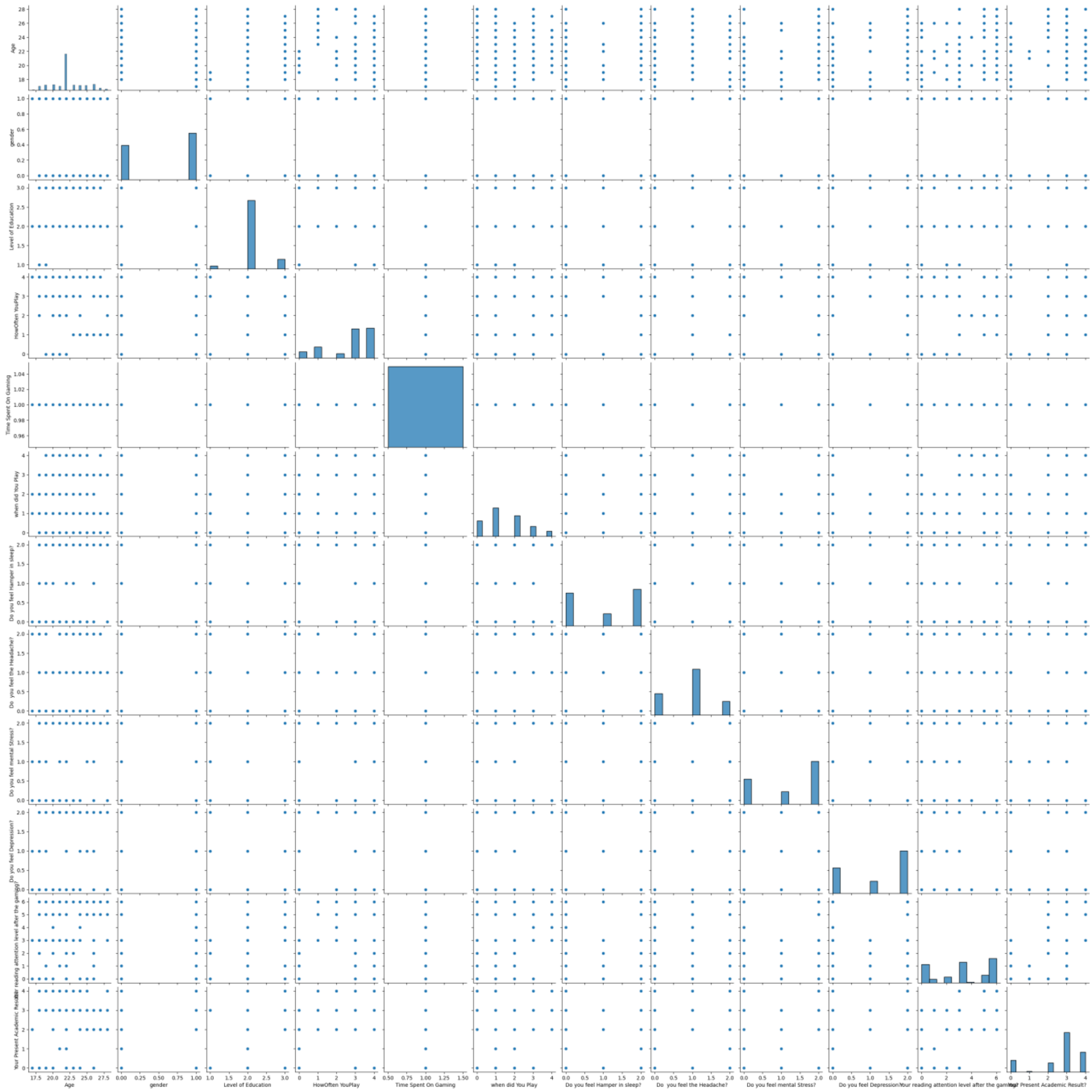
- Performing TF-IDF factor, phonetic hashing, frequency Distribution, etc
- Analysing the textual data, Modelling the topic, word embedding, Entity recognition, etc are performed for better insights and results.
- Identify patterns, trends, and potential insights that can guide feature selection and model building.

v. Feature Engineering

- Create new features, if necessary, through transformations, scaling, or a combination of existing features.
- Select relevant features based on domain knowledge, EDA insights, and feature importance techniques.



A correlation matrix shows the correlation between all numerical variables in a dataset, providing a comprehensive view of the relationships between variables and helping to identify patterns, trends, and relationships in the data.



It shows a graphical representation of the pairwise relationships between variables in a dataset. It's a grid of subplots, where each variable is shared across a row or column. This allows you to visualize the relationships between multiple variables in a single plot.

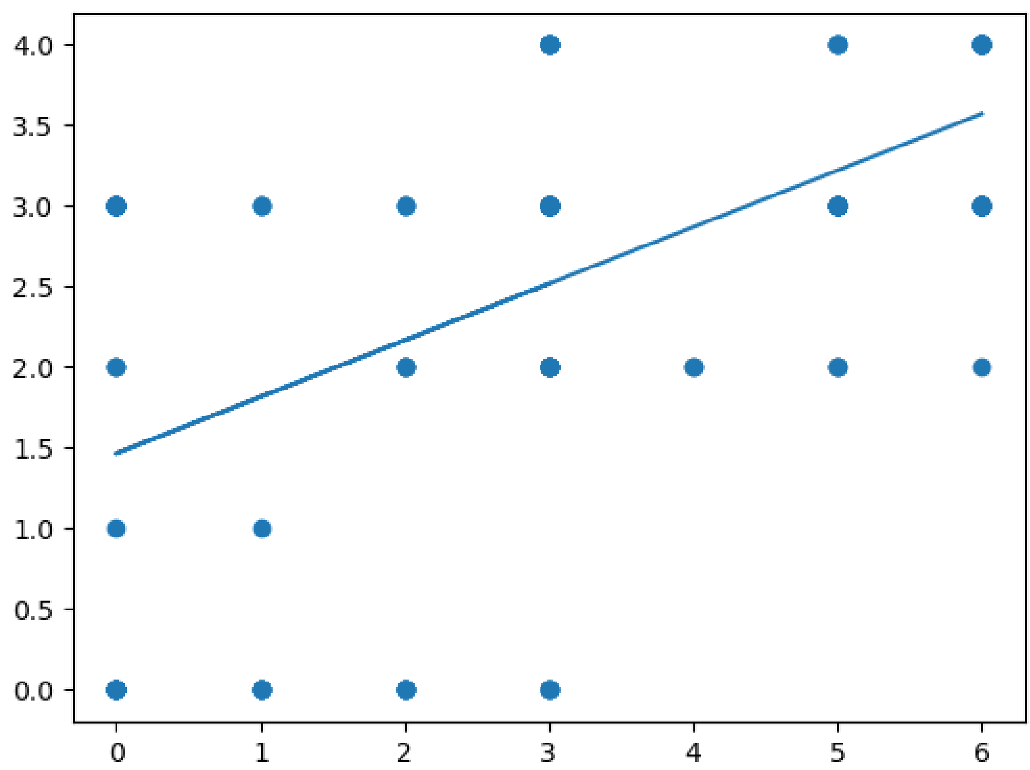
vi. Model Selection.

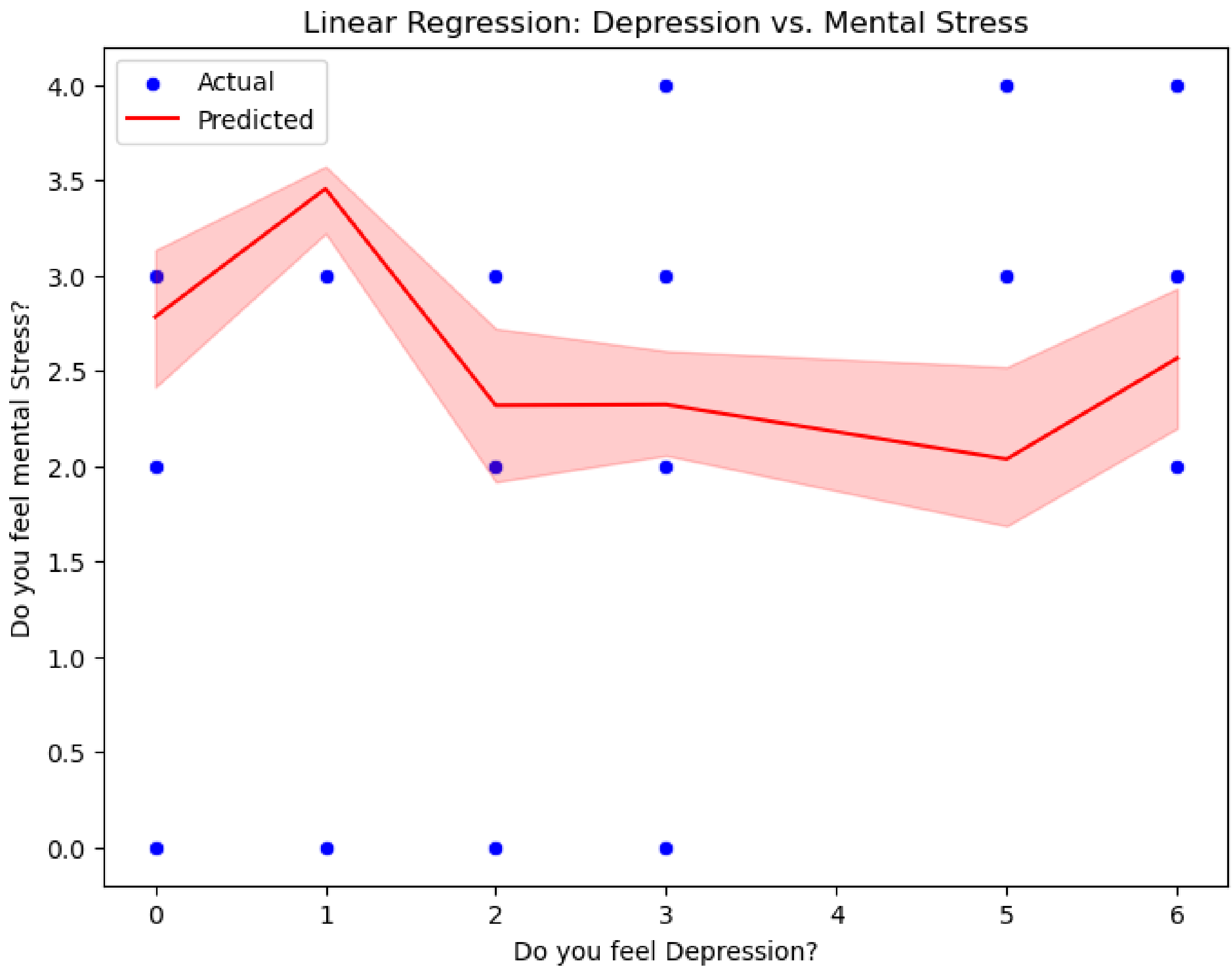
In the context of your project, where you've applied Linear Regression, Logistics Regression, Decision Tree, and Support Vector Machine (SVM),

Random Forest Classifier, Random Forest Regression, Naïve Byes here's how you can describe the model selection:

Linear Regression (Logistic Regression):

Linear regression, specifically logistic regression in the context of classification tasks, is a commonly used model for predicting binary outcomes. However, in recommendation systems, it may not be the primary choice for generating personalized recommendations. Logistic regression is more suitable for tasks where the goal is to predict probabilities or classify instances into discrete classes based on a set of input features.



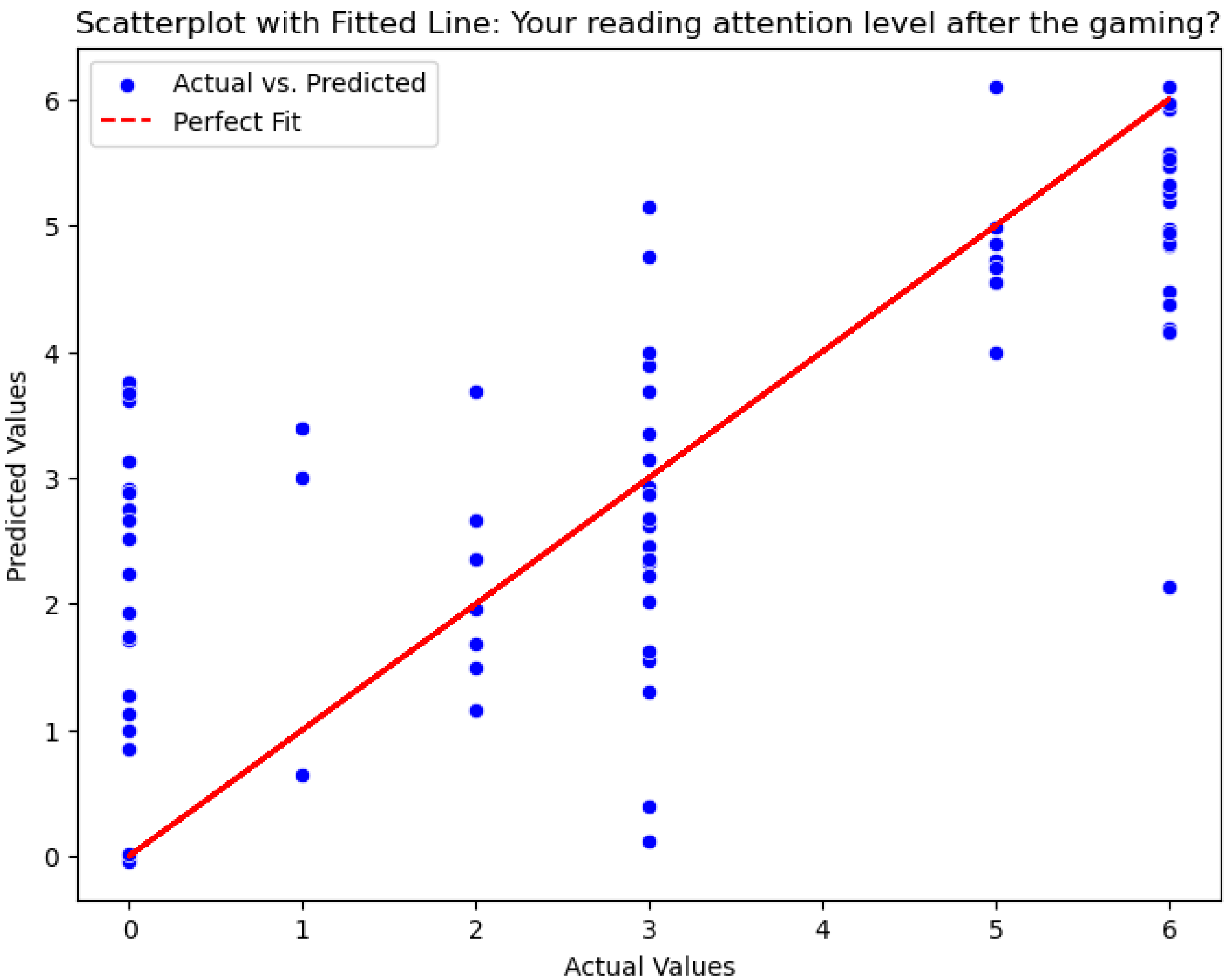


"The linear regression model achieved an R-squared value of 0.67 indicating a moderate ability to explain the data variance. This could be due to limitations in the data or the model's capacity to capture the full complexity of the relationship. Further exploration with data preprocessing techniques or more complex models might be beneficial in future research."

Multiple Linear Regression:

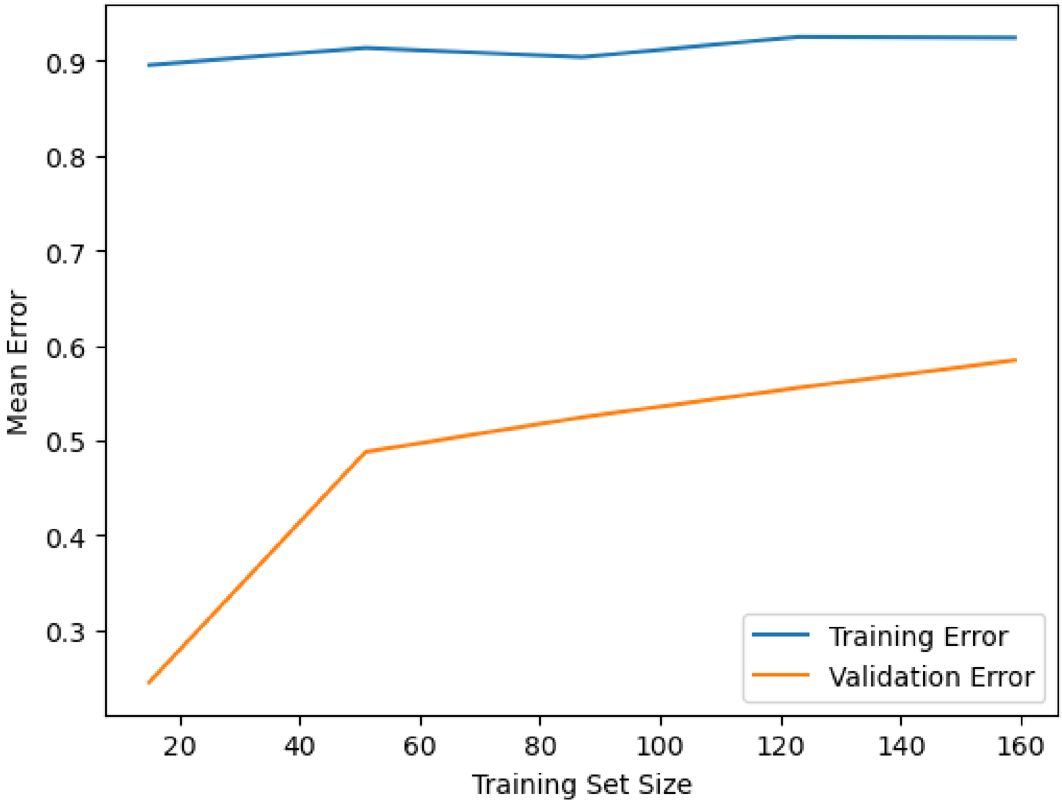
Multiple linear regression refers to a statistical technique that is used to predict the outcome of a variable based on the value of two or more variables. It is sometimes known simply as multiple regression, and it is an extension of linear regression. The variable that we want to predict is known as the dependent

variable, while the variables we use to predict the value of the dependent variable are known as independent or explanatory variables.



Deviations from the line indicate errors in prediction. Random scatter suggests no clear pattern in the errors, while a curved pattern suggests the model might not capture the true relationship.

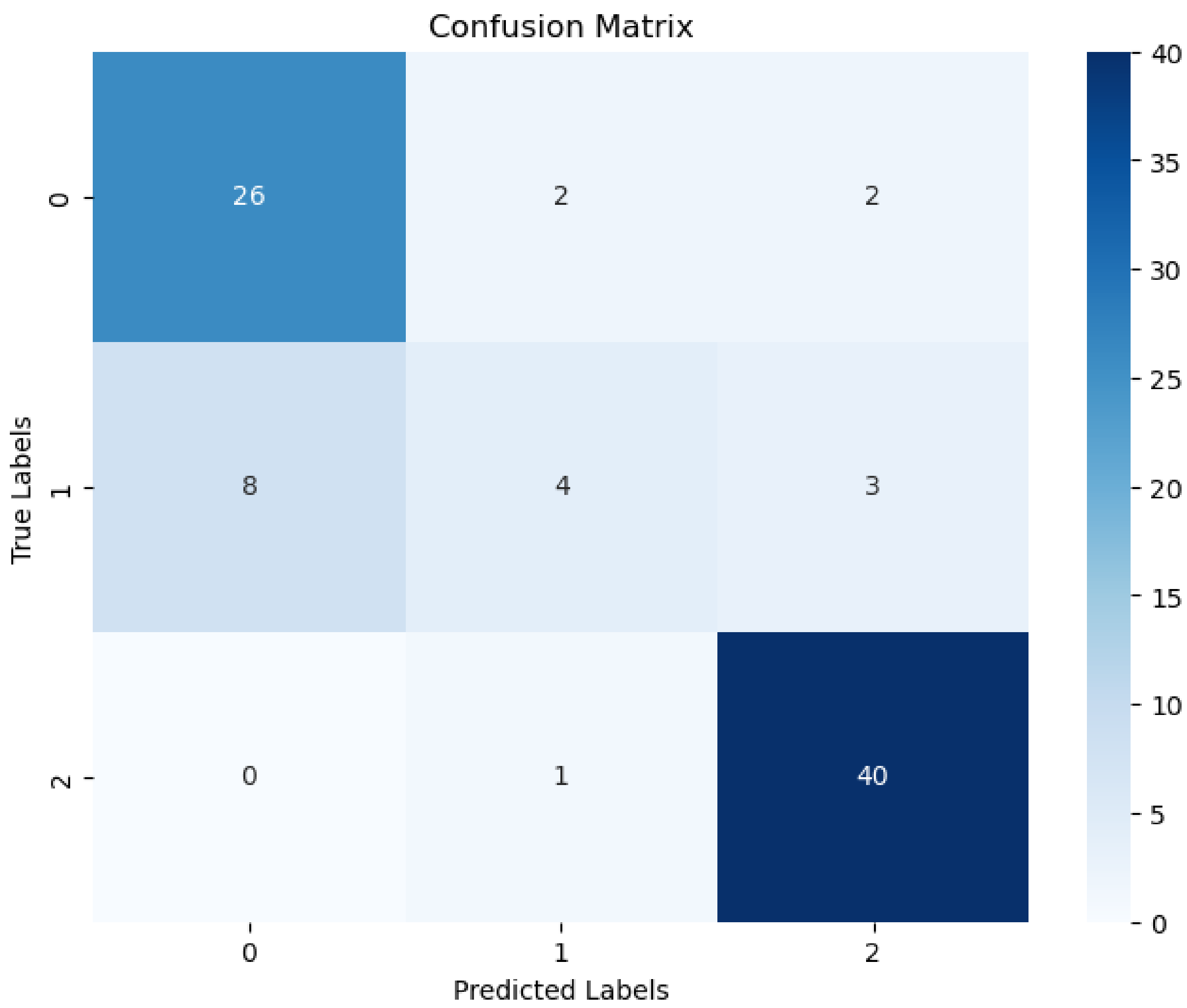
Random Forest Regression:



As we can see the model is not learning well from the data and it concludes that This model does not fit best for this dataset.

Logistics Regression:

It is a classification algorithm, not a regression algorithm. It is used to estimate discrete values (Binary values like 0/1, yes/no, true/false) based on a given set of independent variables(s). In simple words, it predicts the probability of the occurrence of an event by fitting data to a logistic function. Hence, it is also known as logit regression. Since it predicts the probability, its output values lie between 0 and 1 (as expected).

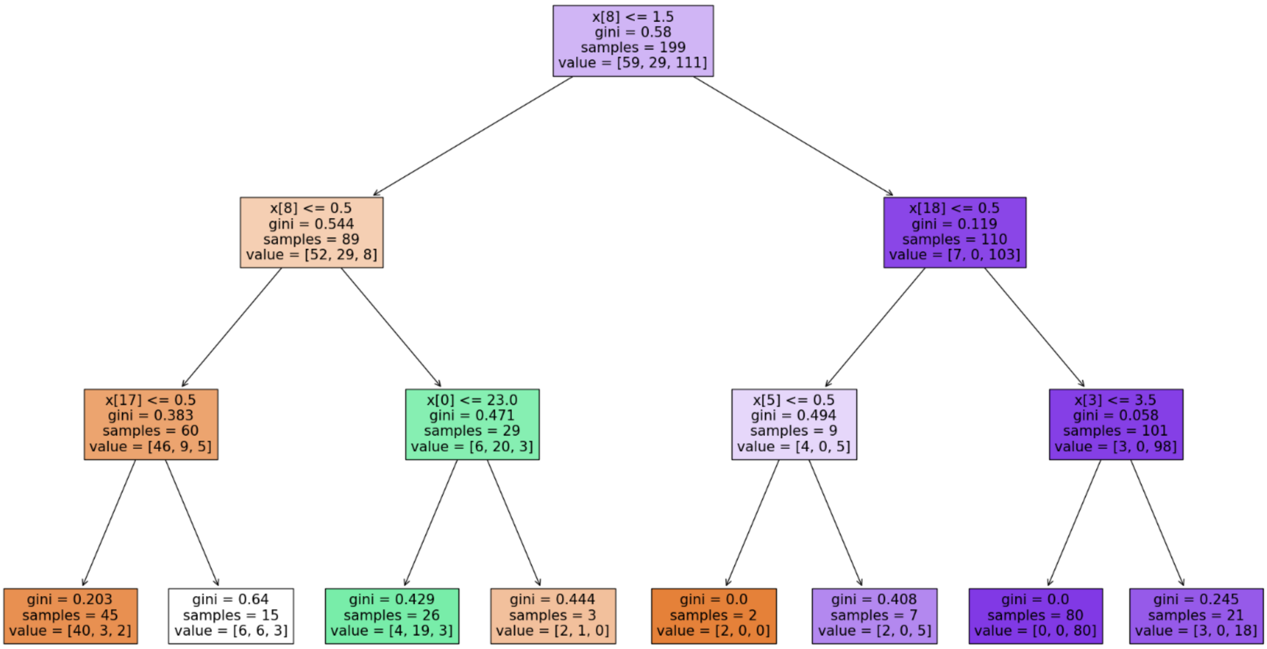


The confusion matrix helps visualize the performance of a classification model on a set of data. Ideally, the diagonal cells (where predicted class matches the actual class) should have the highest values, indicating correct classifications. High values in off-diagonal cells represent errors where the model predicted the wrong class.

Decision Tree:

Decision trees are powerful models for both classification and regression tasks. In the context of recommendation systems, decision trees can be useful for exploring feature importance, understanding decision pathways, and segmenting users or items based on their attributes. However, they may not

directly generate personalized recommendations as traditional collaborative filtering or content-based methods do.



Y_pred

```
]: array([0, 0, 2, 2, 2, 2, 2, 2, 0, 2, 0, 2, 0, 2, 0, 0, 2, 0, 2, 1, 1, 2,
          1, 2, 1, 1, 0, 0, 2, 0, 2, 2, 2, 2, 0, 0, 0, 0, 2, 2, 0, 2, 0, 1,
          0, 2, 0, 2, 1, 2, 2, 2, 0, 1, 1, 2, 2, 2, 0, 0, 0, 1, 2, 2, 0, 0,
          2, 0, 1, 0, 2, 0, 2, 1, 0, 2, 2, 0, 2, 0, 0, 2, 2, 1, 2, 2])
```

	precision	recall	f1-score	support
0	0.93	0.88	0.90	32
1	0.67	0.77	0.71	13
2	0.95	0.95	0.95	41
accuracy			0.90	86
macro avg	0.85	0.87	0.86	86
weighted avg	0.90	0.90	0.90	86

Support Vector Machine (SVM):

Support Vector Machines are versatile models used for classification and regression tasks. In recommendation systems, SVMs can be effective for identifying complex patterns and decision boundaries in the data. They can contribute to personalized recommendations by learning from user-item interactions and feature patterns. However, SVMs may require careful tuning of hyperparameters and feature engineering to achieve optimal performance in recommendation tasks.

```
Target on train data [2 0 0 0 0 2 0 0 2 0 2 2 2 2 0 2 2 2 2 0 0 2 2 2 0 0 0 0 0 0 0 0 2 2 2 2 2
 2 0 2 0 0 2 0 2 0 0 2 2 2 2 0 2 0 2 2 2 2 2 2 0 0 2 2 0 2 2 2 0 2 2 0 2 2
 0 0 2 2 0 0 0 2 2 0 2 0 2 2 2 2 0 0 0 2 0 2 2 0 2 0 0 2 2 0 2 0 0 2 2 2 2
 0 2 0 2 2 2 2 2 0 2 0 2 2 2 2 2 2 0 2 2 0 2 0 0 2 2 0 2 2 2 0 0 2 2 2 2 2
 0 0 0 2 2 2 0 2 2 2 2 0 0 2 2 2 2 2 2 2 2 0 2 0 2 2 2 0 0 0 0 2 2 2 2 0 2
 2 2 0 0 2 2 0 0 0 2 0 2 2 0]
```

```
In [109]: # Accuray Score on train dataset
accuracy_train = accuracy_score(train_y,predict_train)
print('accuracy_score on train dataset : ', accuracy_train)
```

```
accuracy_score on train dataset :  0.7839195979899497
```

```
In [110]: # predict the target on the test dataset
predict_test = model.predict(test_x)
print('Target on test data',predict_test)
```

```
Target on test data [0 0 2 2 2 2 2 2 0 2 0 2 0 2 0 0 2 0 2 0 0 2 0 2 0 0 0 0 2 0 2 2 2 0 0 0 0
 2 2 2 0 2 0 2 0 2 0 2 2 2 2 2 0 0 2 2 2 2 0 0 0 0 2 2 0 0 2 0 2 0 0 0 2 2
 0 2 2 0 2 0 0 2 2 2 2 2]
```

```
In [111]: # Accuracy Score on test dataset
accuracy_test = accuracy_score(test_y,predict_test)
print('accuracy_score on test dataset : ', accuracy_test)
```

```
accuracy_score on test dataset :  0.7674418604651163
```

While Linear Regression, Decision Tree, and Support Vector Machine are valuable models in machine learning, they are not traditionally the primary choices for generating personalized recommendations in recommendation systems. Content-based, collaborative filtering, or hybrid approaches incorporating these models along with others like K-means clustering or neural networks are often more suitable for recommendation tasks, especially in scenarios involving rich item features, user-item interactions, and scalability considerations.

vii. Model Validation and Interpretation:

- Validate the final models using cross-validation techniques to ensure generalizability.
- Interpret model coefficients, feature importance, and decision boundaries to understand the underlying relationships captured by the models.

viii. Documentation and Reporting:

- Document the entire research methodology, including data preprocessing steps, model selection process, evaluation metrics, and key findings.

- Prepare a comprehensive report or presentation summarizing the research methodology, results, and recommendations for E-commerce handlers.

5. Results and Discussion

5.1: Experimental Results.

5.1.1. Model-1: K – Means Clustering

- Low complexity and strong scalability: The computational complexity is O(nla), where n is the number of data objects and t is the number of iterations. In general, k is much less than n, and t is much less than n.
- Clustering results show that the distribution of internal data is relatively tight compared with other subclasses, which indicates that the clustering effect of this method is good.
- By building a model based on k-means clustering I got the following results.

Algorithms	R2 score	MSE	Accuracy
SLR	-		0.31
MLR	-		0.83
Logistic Regression	-	-	0.81
Gaussian Naive Bayes	-	-	0.834
Multinomial Naïve Bayes		-	0.778
Bernoulli Naïve Bayes	-	-	0.823
Decision Tree Classifier	-	-	0.8953
Random Forest regression			0.63
Random forest Classifier	-	-	0.44
Confusion Matrix	-	-	0.81
Support vector Machine	-	-	0.76
Naive Bayes	-	-	0.83

CONCLUSION AND FUTURE SCOPE

5.1. CONCLUSION:

In conclusion, the realm of personalized recommendation systems in gaming data holds immense potential for enhancing user experiences and engagement. Just as in e-commerce, the gaming industry can benefit significantly from personalized recommendations that cater to individual preferences and behaviors .However, achieving accurate and timely recommendations in gaming requires addressing specific challenges, such as quick adaptation to changes in user interests, real-time online recommendations, efficient processing of big data, and scalability. While personalized recommendation technology is widely accepted and sought after, there is still a need for further research and development to refine algorithms and improve recommendation accuracy. Similar to e-commerce, personalized recommendation systems in gaming can play a vital role in online marketing strategies, driving user engagement and satisfaction. By analyzing gaming habits, preferences, and behavioral patterns, recommendation algorithms can provide tailored game suggestions that resonate with players, leading to increased player retention and loyalty.As the gaming industry continues to evolve, incorporating advanced recommendation algorithms, such as collaborative filtering, content-based filtering, and hybrid models, will be crucial in delivering personalized gaming experiences. Additionally, exploring methods to leverage user interactions, feedback, and browsing history can further enhance recommendation accuracy and relevance.While this conclusion highlights the potential of personalized recommendation systems in gaming data, it's essential to acknowledge that this field is still evolving, and there's room for ongoing research, experimentation, and innovation. Future studies can delve deeper into the intricacies of gaming data analysis, refine recommendation algorithms, and explore novel approaches to optimize user experiences in the gaming landscape.

5.2. Future Scope:

- **Enhanced Personalization**: Continuously refine and improve the personalization capabilities of your recommendation system. Incorporate advanced techniques such as reinforcement learning or contextual bandits to adapt recommendations in real-time based on user feedback and interactions.
- **Dynamic Content Updates**: Implement mechanisms to dynamically update item features and content representations in your recommendation system. This can involve leveraging techniques like online learning or incremental updates to accommodate changes in product catalogs, seasonal trends, or user preferences.

- **Context-Aware Recommendations**: Integrate contextual information such as user location, time of day, or browsing behavior into your recommendation system. Develop algorithms that can adapt recommendations based on the user's context, providing more relevant and timely suggestions.
- **Interpretability and Explainability**: Enhance the interpretability and explainability of your recommendation system to build user trust and understanding. Explore techniques for generating explanations or visualizations that elucidate why certain items are recommended to users, helping them make informed decisions.
- **Multi-Modal Recommendations**: Extend your recommendation system to handle multi-modal data sources, such as images, text, and structured data. Leverage models like Sentence-BERT for textual embeddings and convolutional neural networks (CNNs) or pre-trained vision models for image embeddings to provide richer and more diverse recommendations.

References

1. Z. Jing, Y. Ying, and T. Qi, "Personalized social image recommendation method based on user-image-tag model," *IEEE Transactions on Multimedia*, vol. 1, no. 99, p. 1, 2017.
   View at: Google Scholar
2. J. Zhang, "Personalised product recommendation model based on user interest," *Computer Systems Science and Engineering*, vol. 34, no. 4, pp. 231–236, 2019.
   View at: Publisher Site | Google Scholar
3. H. Pan and Z. Zhang, "Research on context-awareness mobile tourism E-commerce personalized recommendation model," *Journal of Signal Processing Systems*, vol. 93, no. 2-3, pp. 147–154, 2021.
   View at: Publisher Site | Google Scholar
4. Y. Lv, "Personalized recommendation model based on incremental learning with continuous discrete attribute optimization," *Revista de la Facultad de Ingenieria*, vol. 32, no. 2, pp. 842–849, 2017.
   View at: Google Scholar
5. I. Gupta and N. Joshi, "Enhanced twitter sentiment analysis using hybrid approach and by accounting local contextual semantic," *Journal of Intelligent Systems*, vol. 29, no. 1, pp. 1611–1625, 2019.
   View at: Publisher Site | Google Scholar
6. V. Ramanathan and T. Meyyappan, "Prediction of individual's character in social media using contextual semantic sentiment analysis," *Mobile Networks and Applications*, vol. 24, no. 6, pp. 1763–1777, 2019.

View at: Publisher Site | Google Scholar

7. A. D. Vo, Q. P. Nguyen, and C. Y. Ock, "Semantic and syntactic analysis in learning representation based on a sentiment analysis model," *Applied Intelligence*, vol. 50, no. 3, pp. 663–680, 2020.
View at: Publisher Site | Google Scholar
8. S. Wang, Y. Zhu, W. Gao, M. Cao, and M. Li, "Emotion-semantic-enhanced bidirectional LSTM with multi-head attention mechanism for microblog sentiment analysis," *Information*, vol. 11, no. 5, pp. 280-281, 2020.
View at: Publisher Site | Google Scholar
9. C. Guo, Z. Du, and X. Kou, "Products ranking through aspect-based sentiment analysis of online heterogeneous reviews," *Journal of Systems Science and Systems Engineering*, vol. 27, no. 5, pp. 542–558, 2018.
View at: Publisher Site | Google Scholar
10. G. Jang and S. H. Myaeng, "Predicting event mentions based on a semantic analysis of microblogs for inter-region relationships," *Journal of Information Science*, vol. 44, no. 6, pp. 818–829, 2018.
View at: Publisher Site | Google Scholar
11. Z. Huang, Y. Liu, C. Zhan, C. Lin, W. Cai, and Y. Chen, "A novel group recommendation model with two-stage deep learning," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–12, 2021.
View at: Publisher Site | Google Scholar
12. Q. Zhou, L. Su, L. Wu, and D. Jiang, "Deep personalized medical recommendations based on the integration of rating features and review sentiment analysis," *Wireless Communications and Mobile Computing*, vol. 2021, no. 7, Article ID 5551318, 9 pages, 2021.
View at: Publisher Site | Google Scholar
13. J. Nan, J. Tao, and J. Du, "Personalized recommendation based on customer preference mining and sentiment assessment from a Chinese e-commerce website," *Electronic Commerce Research*, vol. 18, no. 1, pp. 1–21, 2018.
View at: Google Scholar
14. H. U. Ji, Y. Ming, and L. Yan, "Research on the travel choice behaviors of transportation corridor based on SEM-MNL model," *Journal of Railway Engineering Society*, vol. 34, no. 4, pp. 80–85, 2017.
View at: Google Scholar
15. E. S. Alamoudi and N. S. Alghamdi, "Sentiment classification and aspect-based sentiment analysis on yelp reviews using deep learning and word embeddings," *Journal of Decision Systems*, vol. 30, pp. 1–23, 2021.
View at: Google Scholar

16. Y. Lu, “Research on influence from investor focus and investor sentiment in on-line social media upon stock market,” *Modern Economics & Management Forum*, vol. 3, no. 2, pp. 111–114, 2022.
View at: Publisher Site | Google Scholar
17. A. Muruganantham and S. P. Victor, “Analysis of multilevel-semantic prediction based on user point of sentiment opinion (MSP-UPSO) in social web mining,” *INTERNATIONAL JOURNAL OF COMPUTER SCIENCES AND ENGINEERING*, vol. 5, no. 12, pp. 151–158, 2017.
View at: Publisher Site | Google Scholar
18. S. Chen, L. Huang, Z. Lei, and S. Wang, “Research on personalized recommendation hybrid algorithm for interactive experience equipment,” *Computational Intelligence*, vol. 36, no. 3, pp. 1348–1373, 2020.
View at: Publisher Site | Google Scholar
19. J. Cai, “Research on personalized recommendation algorithm in e-commerce based on hybrid algorithm,” *Clinica Chimica Acta*, vol. 42, no. 2, pp. 590–594, 2017.
View at: Google Scholar