A MACHINE LEARNING BASED APPROACH FOR PREDICTING THE IMPACT OF VIDEO GAMES ON YOUTH.

UDOY DAS

ID: 201-15-3131

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Shah Md. Tanvir Siddiquee

Assistant Professor

Department of CSE

Daffodil International University

Co-Supervised By

Md Sabab zulfiqur

Senior Lecturer

Department of CSE

Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2024

APPROVAL

This Project/internship titled **"A MACHINE LEARNING BASED APPROACH FOR PREDICTING THE IMPACT OF VIDEO GAMES ON YOUTH"**, submitted by Udoy Das, ID No: 201-15-3131 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on *24th January 2024*.

BOARD OF EXAMINERS

Dr. Md. Taimur Ahad (MTA)
Associate Professor & Associate Head
Department of Computer Science and Engineering
Daffodil International University

Chairman

Mr. Saiful Islam(SI)
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

Internal Examiner

Lamia Rukhsara(LR)
Senior Lecturer
Department of Computer Science and Engineering
Daffodil International University

Internal Examiner

Dr. Abu Sayed Md. Mostafizur Rahaman (ASMR)
Professor
Department of Computer Science and Engineering
Jahangirnagar University

External Examiner

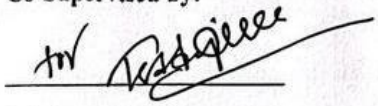
Declaration

I hereby declare that, this project has been done by me under the supervision of **Shah Md. Tanvir Siddiquee, Assistant Professor, Department of CSE** Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

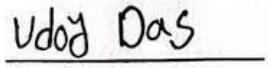
Shah Md. Tanvir Siddiquee
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised by:



Md. Sabab Zulfiqur
Senior lecturer
Department of CSE
Daffodil International University

Submitted by:



Udoy Das
ID: 201-15-3131
Department of CSE
Daffodil International University

iii

ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty God for His divine blessing makes me possible to complete the final year project/internship successfully.

I am really grateful and wish my profound my indebtedness to Supervisor **Shah Md. Tanvir Siddiquee**, Assistant Professor, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "Machine Learning" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project.

I would like to express our heartiest gratitude to **Professor Dr. Sheak Rashed Haider Noori**, Head**,** Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

I would like to thank my entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, I must acknowledge with due respect the constant support and patients of my parents.

ABSTRACT

The abstract of the topic "impact of video games prediction on youth using machine learning" would succinctly summarize the key components of the research. This research explores the intersection of video games, youth development, and machine learning, investigating how predictive models can unveil intricate patterns in gaming behaviors. By analyzing cognitive, academic, and mental health impacts, as well as social dynamics, the study aims to inform responsible gaming practices and educational strategies. Ethical considerations and long-term societal implications are integral to the examination, emphasizing the need for a balanced approach in navigating the digital landscape for the well-being of the youth. The study has the motive to predict whether a person who plays game is tensed for career or not. Thats why we collected 804 data from them 780 were used the features we used are name, gender, age, spend time in study, sleep time, wake up time, the game play most ,spend time in playing games, purpose of playing games, spend time playing games more than family , time spend for skills, tensed for career. And then preprocessed them and Prior to using certain machine learning algorithms, they were examined. The accuracy of RF is best of 97.4% which outperformed all others.

Keywords: video games, youth, machine learning, effect

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
Introduction

1.1 INTRODUCTION

In recent years, the world has witnessed an unprecedented surge in the popularity and accessibility of video games, marking a paradigm shift in entertainment consumption, particularly among the youth. As the gaming industry continues to evolve, so does the discourse around its potential impact on various aspects of society, including the cognitive, behavioral, and social development of young individuals. Amidst these discussions, the integration of machine learning in predicting the impact of video games on youth has emerged as a notable area of research.

Video games, with their immersive virtual worlds and interactive narratives, present a unique environment that can shape the experiences and behaviors of players. The youth, being particularly avid consumers of video games, are often at the forefront of these influences. Machine learning, a powerful subset of artificial intelligence, has proven instrumental in analyzing and predicting complex patterns within large datasets. When applied to the realm of video games, it opens new avenues for understanding the nuanced relationship between gaming activities and the well-being of the youth.

This study seeks to explore the intersection of video games and machine learning, focusing on predicting the impact on youth. By harnessing the capabilities of machine learning algorithms, researchers aim to decipher intricate patterns related to gaming habits, behavioral changes, and potential correlations with aspects such as academic performance, mental health, and social interactions. The predictive models developed through machine learning methodologies promise insights into the diverse ways in which video games may influence the youth population.

Prediction of The impact of video games on younger generation is a system that will help to know about a person’s mental condition like whether he/she Is tensed for career or not We are using certain data for both genders and men, such as name, age, how much time is spent studying, sleeping, waking up, how much time is spent playing games, why games are

played, how much time is spent practicing skills, and how much time is focused on a career. Here, I've attempted to calculate the percentage of male and female addicts after evaluating my data.

1.2 MOTIVATION

In an era dominated by technological advancements and digital immersion, For millions of people, especially the younger generation, video games have become an essential part of everyday life.. The growing prevalence of video game culture raises critical questions about its impact on the cognitive, social, and emotional development of individuals, with a specific focus on the younger demographic. As these questions persist, the application of machine learning emerges as a powerful tool to dissect the intricate relationships between gaming habits and their consequences on youth.

The motivation behind this study stems from the need to navigate the complex landscape of video games and their implications for the well-being of the youth. While video games offer a unique form of entertainment and engagement, concerns have been raised regarding potential adverse effects, such as addictive behaviors, academic performance fluctuations, and changes in social dynamics. Machine learning presents an opportunity to delve into the vast datasets generated by gaming platforms, providing a nuanced understanding of how different gaming experiences may influence various aspects of youth life.

According to the research, Bangladesh had 26.46 million internet users in 2016.which means a huge number of people are attached with online game .so definitely they are spending more their time playing video games. Therefore there may be impact on mental health of any person. Therefore knowing a person is tensed for career or not is much more important .so knowing the impact is more important for both male and female.

1.3 RATIONALE OF THE STUDY

Ubiquitous influence of video games on adolescents in today's changing landscape calls for a sophisticated understanding. The rise in popularity of video games among younger people requires an early intervention to understand the possible effects on social, behavioral, and cognitive domains. Conventional research methods frequently find it difficult to identify the complex patterns and forecast future trends in this dynamic environment. The amalgamation of Machine Learning (ML) offers a novel approach to scrutinizing extensive information and identifying patterns that could defy traditional techniques. This study aims to forecast and model the possible effects of video games on young people by utilizing machine learning (ML) methods. This provides a prediction framework that goes beyond the constraints of static assessments. In this situation, machine learning (ML) has the potential to uncover subtle connections, spot new patterns, and offer an outlook on the complex effects of playing video games.

1.4 RESAERCH QUESTIONS

1. which algorithm works better ?
2. Does all the algorithm work successfully?
3. what are the benefits of our model?
4. what are the requirements of this work?

1.5 EXPECTED OUTPUT

It assists us in producing an expected outcome in our impact prediction system based on the data provided in accordance with our dataset. To obtain a more accurate result in this instance, I used 20 test data and 80 training data of percentage. Our result's correctness depends only on the training set. Our machine learning system will be operational once all system processes have been finished. We have employed several strategies to ensure precise

outcomes. From "RF" I got accuracy of more than 97%.

1.6 PROJECT MANAGEMENT AND FINANCE

For this work,utilized all free resources. For coding purpose I used google colab and keeping for managing resources I have used google drive .all these resources are found for free.so therefore I didn't need to face any financial issues.

1.7 REPORT LAYOUT

Chapter 1: Here, I discussed the assignment’s goal, my goals, and the outcomes i anticipate from our labor.

Chapter 2: i have discussed the background of my work and looked at previous research, relevant studies, the variety of problems, and challenges in this area.

Chapter 3 covers the study's subject, the instrument we used, data collection procedures, statistical analysis, and implementation.

The findings of my study investigation are presented in

Chapter 4, together with a summary and a descriptive analysis.

Chapter 5: Here, I provide a summary of the plan for sustainability, moral considerations, environmental effect, and societal impact.

Chapter 6: I go over an overview of our results and conclusions in this chapter.

Information regarding our extra research techniques is also provided.

CHAPTER 2

Background

2.1 PRELIMINARIES/TERMINOLOGIES

Here, the discussion is about the data pretreatment, attribute selection, algorithm classification, associated chores, and difficulties I ran into while working on the project. In addition, I will provide an overview of the complete research effort. I'll go over the steps I took to process the data in the data preprocessing section. I'll talk about the attribute I utilized in our study in the section on attribute selection. I'll discuss my methods for processing the data in the part on method categorization. I'll go over other research articles and pertinent activity in the section on related work. I shall talk about them because their approaches and correctness are pertinent to what I do. In the part on obstacles, I'll talk about the difficulties I encountered.

2.2 RELATED WORKS

They attempted to predict whether or not a particular person was addicted to video games in this paper [1]. As a result, they used a Google Form to gather 549 data creating points. They used a seven-criteria gaming addiction measurement scale when the data collection form. Following the use of several machine learning algorithms, it was discovered that adaptive boosting, or ada boosting, beat other methods with an accuracy of 93.00%. Calculations revealed that 25% of people have a game addiction.

In this paper [2] Online games can negatively impact learning achievements and concentration. Classification algorithms (Naïve Bayes, Random Forest, C4.5) used for prediction. Naïve Bayes algorithm has higher accuracy compared to Random Forest and C4.5. Naïve Bayes algorithm can predict student achievement in studying. Naïve Bayes algorithm has higher accuracy compared to Random Forest and C4.5.Naïve Bayes algorithm has an accuracy of 69.18% and AUC value of 0.771. Random Forest algorithm has an accuracy of 66.34% and AUC value of 0.738. C4.5 algorithm has an accuracy of

65.65% and AUC value of 0.686.CRISP-DM methodology (Cross-Industry Standard for Data Mining) Naïve Bayes algorithm Random Forest algorithm C4.5 algorithm

In this paper[3] The study addresses the detrimental effects of excessive video gaming, with an emphasis on Asian Player Unknown's Battlegrounds (PUBG) users. The possibility of psychiatric illnesses in Asian PUBG gamers was predicted using supervised machine learning. Plotting the ROC (Receiver Operating Characteristic) curve, calculating AUC (Area Under the Curve), increasing iterations, initializing variables, and reporting the accuracy in percentage were all processes in the classification process. In predicting Internet Gaming Disorder (IGD) and Attention Deficit Hyperactivity Disorder (ADHD), The maximum accuracy of the Logistic Regression classifier was 93.18% and 81.81%, respectively. The Decision Tree classifier, with an accuracy rate of 84.9%, was the most accurate predictor of Generalized Anxiety Disorder (GAD).The study recognizes as a major drawback the possibility of biases, such as social desirability bias, in self-report assessments, which could lead to an under- or over-reporting of psychological problems.

The paper[4] examines the potential harmful consequences of excessive gaming. - It identifies risk factors for developing problematic gaming patterns. Longitudinal study with 565 gamers assessed twice, six months apart - Used User-Avatar-Bond Scale and Gaming Disorder Test for assessment - AI classifiers (random forests) analyzed responses for GD risk prediction - Avatar immersion was the strongest predictor for GD risk - Data collected on demographics, gaming use, and social media use.- AI models accurately identified GD risk based on UABS score, age, and gaming involvement. - Random Forests outperformed other AI classifiers in predicting GD risk. - Avatar immersion was the strongest predictor for Random Forests. - Tuned versions of classifiers performed better than unturned versions. - Random Forests consistently outperformed other classifiers in both tuned and unturned versions.

this paper[5] A study looks at children and teenagers in Tamilnadu who play video games. - The survey included 406 respondents with a range of socioeconomic backgrounds. - The majority of players engaged in one weekly game session. The non-probability category

makes use of the purposeful sampling technique.- Investigated children's and teenagers' tastes and habits related to playing video games. Employed purposive sampling as a method to collect data from participants. Distribution of genders: 11.8% female and 88.2% male. The vast majority of those involved were students.

The purpose of the paper[6] is to ascertain the prevalence of pathological video gaming among Singaporean youth. We used DSM criteria to assess the construct validity of pathological video gaming. A large sample size was used to look into the prevalence of compulsive video gaming. Social competence, academic accomplishment, impulse control, and gaming habits are among the variables that are measured. It showed that 83% percent adolescents and children played video games. There is a connection between pathological video gaming and higher rates of violence, poorer social skills, poor impulse control, and poor academic achievement. Pathological gamers are more aggressive, keep video game consoles in their beds, play for longer periods of time, perform worse academically, have trouble controlling their impulses, and lack social skills. The correlational design of the study and the unclear interpretation of the answer scale.

In this paper[7]investigated the effects of computer animation on the development of children and adolescents. 687 students from three of Kerala's largest cities—Cochin, Calicut, and Trivandrum—were included in the survey. The findings indicate that children's computer animation games may impact violence and desensitize them.

A conference paper titled "Online game addiction among adolescents: motivation and prevention factors" was released by Zhengchuan Xu et al. They used an online survey to collect data, and PLS was chosen to examine the information. They worked on hypothesis testing, common method bias, and the measurement model with the largest Variance Inflation Factor (1.461). This study aimed to deter kids from playing online games by developing a pretty comprehensive model of game addiction And to provide strategies for keeping them from doing so [8][9].

Participants in this paper include Antonius J. van Rooij et al.In statistical analysis, they carried out a latent class analysis (LCA) to find significant groups of survey respondents. This survey indicates that 1.5% of people in the 13–16 age group and 3% of gamers are

Addicted to the internet. This work's primary objective was to elucidate the relationship between mental health and online video game addiction and play [10].

2.3 COMPARATIVE ANALYSIS AND SUMMARY

So Far, Plenty of investigations conducted. We had to read their papers and understand the algorithm while we looked into the same topics in order to increase accuracy. For that reason, we went through each of those articles to try to understand the experiments and conclusions. All of the previous research, though, used an API to gather its data.

Table 4.1 : comparative Analysis Table

Index	Algorithm's	Paper (1)	Paper (2)	Paper (3)	Paper (4)
1	Random Forest Classifier	90%	64.51%	-	89.7%
2	K-nearest Neighbors	80%	-	72.72%	79.3%
3	Gradient Boosting	85%	-	-	-
4	Naïve Bayes	81%	70.13%	43.2%	74.1%
5	AdaBoost	93%	-	81.81%	-
6	Extra trees	-	-	-	-
7	Support Vector Machine	65%	-	-	94.8%
8	Decision tree	84%	-	84.9%	-

2.4 SCOPE OF THE PROBLEM

Research's extent determines its territory, or the depth to which we will investigate the subject. For a number of reasons, it is impractical to expect to cover every aspect of the topic in most investigations. Additionally, the examination will take an incredibly long time possibly forever!! As such, you need to define or restrict the scope of your investigation.

For instance, you can choose to Centre your theme around a gender at a particular cause of melancholy. Individual, natural, and economical are a few instances; social, mental, and geographic are not.

In addition, you could choose to take into account some groups and ignore others because of their gender, age, or socioeconomic status.

2.5 CHALLENGES

I ran into a number of issues when conducting my investigation. Firstly one is collecting data .Bangladeshi people are very much concerned about their data. When I asked them to give me data some of them refused to give me data. Another problem I faced was selecting the attributes for the dataset. Also while coding I faced some issues and errors with programming languages and libraries.

CHAPTER 3

Research Methodology

Overview

This section will provide an overview of the methodology. and procedures part of the research paper. Lets have a quick look at it

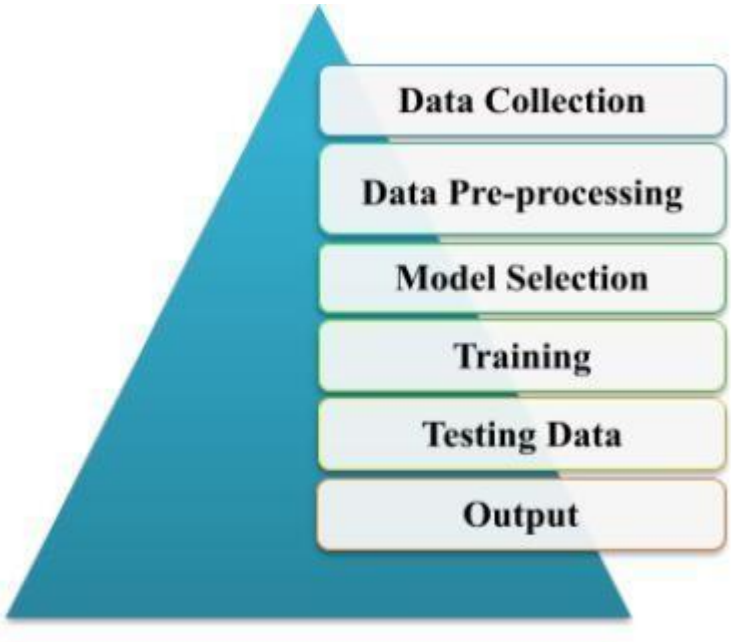


Fig 3.1: methodology on first look

3.1 RESEARCH SUBJECT AND INSTRUMENTATION

The research emphasize by predicting impact of video games on youth using Machine Learning (ML) approaches. This study aims to comprehend and forecast the social, behavioral, and cognitive implications of juvenile video game use using advanced predictive analytics. The study's instrumentation entails the deliberate use of machine learning techniques to examine substantial datasets pertaining to young people's video game consumption. The study makes use of machine learning (ML) tools, such as data mining and predictive modelling, to provide a strong framework for predicting outcomes and patterns on how video games affect young people's social dynamics, behavioral changes, and cognitive development. This equipment will make it possible to conduct a

data-driven investigation of the complex connections that exist between playing video games and different aspects of young people's development.

3.2 DATA COLLECTION PROCEDURE/DATASET UTILIZED

My work uses both offline and online methods to gather data. Most of the Data's are collected through google form and some are collected physically. Data's are mostly collected from Public and private university students .I have collected almost 800+ data but used 780 data in my research where number of males is 754(83.85%) and number of females are 126(16.15%)

3.3 STATISTICAL ANALYSIS

My dataset contains 780 records in total. fig 3.2 with 654 men and 126 women. Here, we choose 80% of the train data and 20% of the test data for our model. In fig 3.3 we have tried to see the number of male and female tensed for their career. Using ML techniques like Gradient Boosting, Rf Classifier, AdaBoost, KNN, DT, NB, and more trees, we have tried to increase the accuracy of our model.

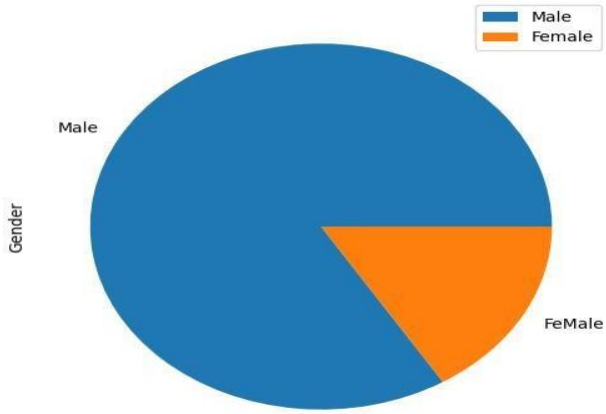


Fig 3.2:Gender based summary

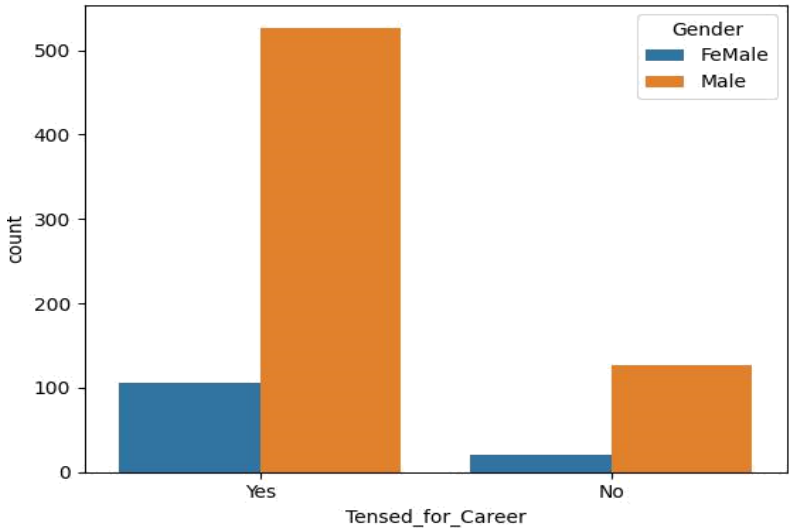


Fig 3.3: Number Of people Tensed for Career

Attributes

Dataset has 12 attributes. Those are (Name, Gender, Age, spend_time_in_study,sleeptime, Weakup_time,the_game_play_most,spend_time_in_playing_games,purpose_of_playing_ games,spend_time_playing_games_more_than_family,Time_spend_for_Skill ,Tensed_for_Career. Here I divided data's into two separate parts for training dataset to build the model and test dataset to evaluate the model.

Here is a figure.

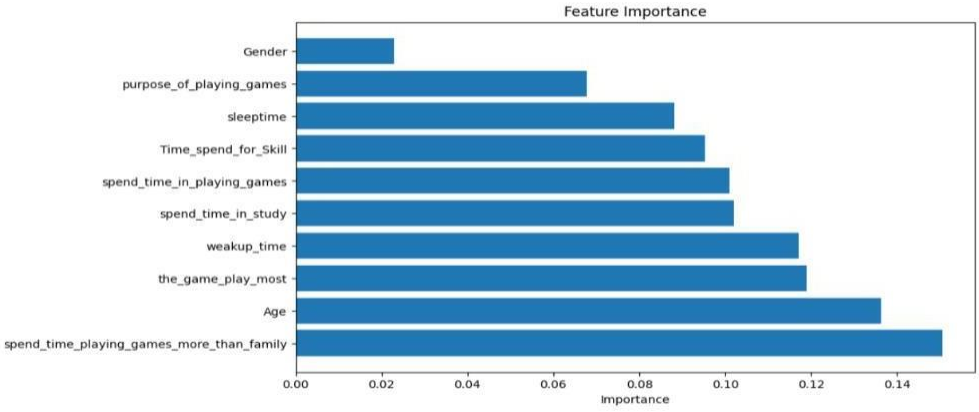


Fig 3.4: Importance of Features

Missing Data Pre-processing

The dataset have 802 values .In 3.3 where there is some missing values .hence we used mean imputation which works like when the missing value is found that gets replaced by mean value of all the available cases.

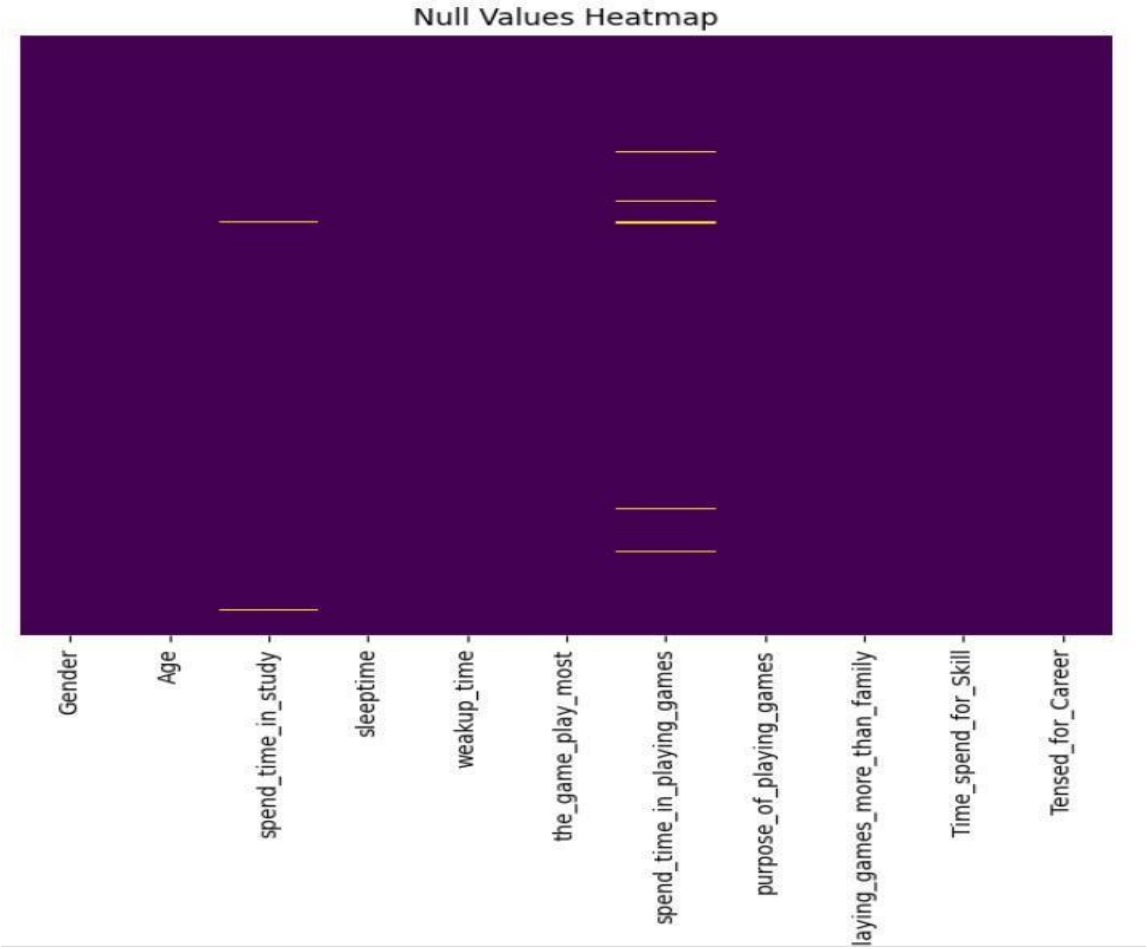


Fig 3.5: missing data showing

After handling missing values we found.

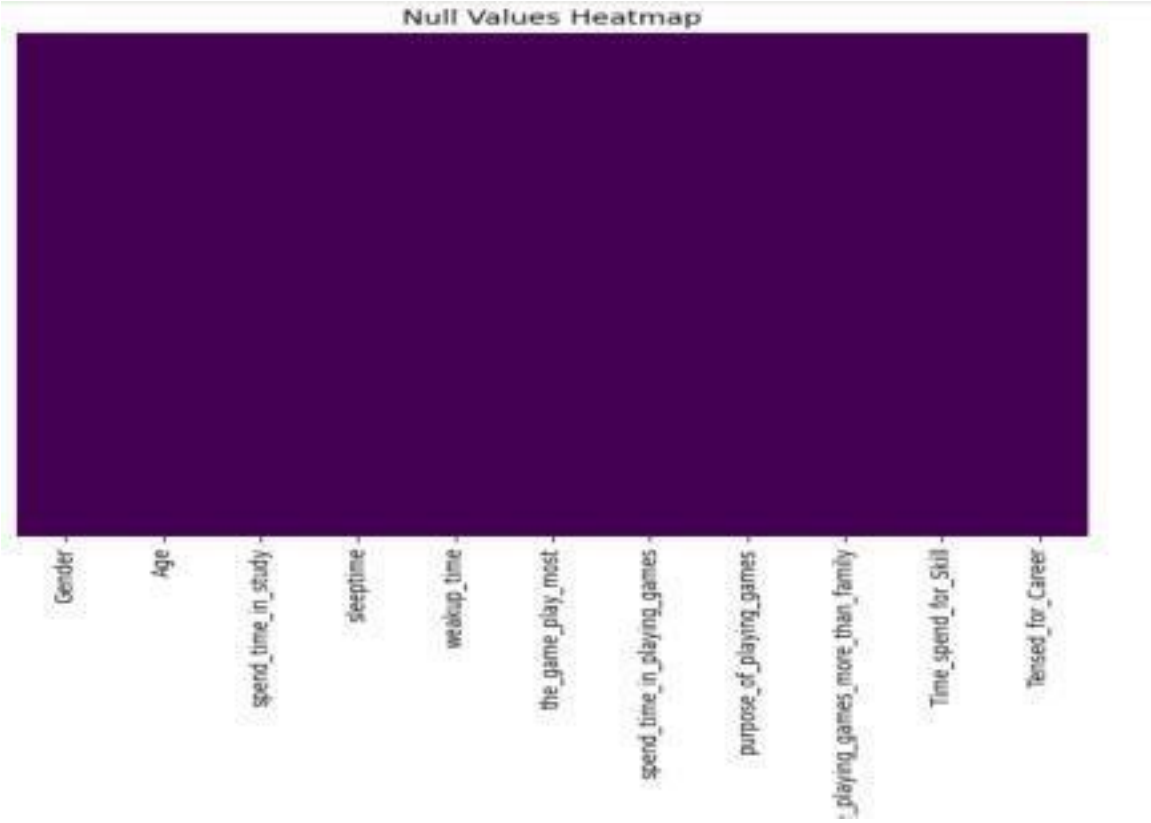


Fig 3.6: missing data handling

Data Organizing

The data has been organized into two categories after testing and training.. Additionally, to verify train data validation, we have also used the validation folder. Sub-folders have been established within the test and train folders.

Data Storing

I have all of the data saved on google drive to make my work easier. For my project, I can use the online data storage offered by Google Drive; all of the data is saved. In order to prevent them from being lost, we have since saved them on Google Drive. Afterwards, by following a few easy steps or codes, we might use those data in our project work through coding.

Machine Learning Algorithms

For their accuracy performance, we employed the Random Forest Classifier, Gradient Boosting, Naïve Bayes, AdaBoost, additional trees, Support Vector Machine, K-nearest Neighbors, and Decision Tree. In order to create a model, we applied on the training data set. All of the Python algorithms that are used, together with the packages NumPy, Pandas, Scikit-learn, and Matplotlib.

Classification Report

based on only accuracy I cannot say that that the model is performing well besides I need to calculate some parameter that can be used to make classification reports. Those are given below.

Confusion Matrix

An evaluation of a classification system's performance is done using a table known as a c onfusion matrix.It gives a summary of the actual and predicted classifications of the mod el for a given set of data.The matrix is composed of four parts:
Events that have a positive correlation but are correctly identified as negative by the mod el are called True Positives (TP).Examples that the model accurately labels as negative ev en though they are actually negative are called True Negatives (TN).Events that the mode l mistakenly labels as positive even when they are actually negative are known as False P ositives (FP) or Type I mistakes.Cases that are positive but the model incorrectly classifie s as negative (Type II error) are known as False Negatives (FN).For binary classification tasks, the confusion matrix is often represented as a 2x2 matrix, although it can be expan ded for multi-class issues.

Precision Score

Precision in classification models is a metric that indicates how well the model forecasts the desirable results. Its definition is the ratio of all of the positive forecasts made by the model to the actual positive projections. Accuracy is crucial when handling significant false positive costs.

The accuracy formula is as follows: TP / (TP+FP).

In this context, "true positives" (TP) refers to the number of cases that were correctly predicted to be positive. The number of cases that are incorrectly projected as positive is known as false positives (FP).A lower percentage of false positives is correlated with greater accuracy scores. An accuracy score has a range of 0 to 1. To properly understand a model's performance, precision must be considered in conjunction with other measures, such as recall and the F1-score.

Recall Score

Recall is a parameter used in classification models to assess how well the model can catch all positive cases. It is also known as sensitivity or true positive rate. The ratio of real positive examples to genuine optimistic expectations is how it is defined. Recall is especially important when it is costly to ignore pleasant experiences. The following formula is used to calculate recall:

From a mathematical perspective:

The formula for recall is TP divided by (TP+FN).

In this situation, TP stands for the number of true positives, or cases that were correctly predicted to be positive. Cases that were incorrectly projected to be negative, or false negatives, are quantified as FN.A lower percentage of false negatives is correlated with a higher recall score. The recall scale is 0 to 1. To provide a complete picture of a model's performance, recall should be included in addition to precision and the F1-score. This is particularly crucial in circumstances when minimizing false positives and false negatives is essential.

F1 Score

In classification models, a statistic known as the F1-score is used to combine recall and precision into a single figure. Since it provides a balance between the two measures, it is particularly useful in situations when there is an uneven distribution of classes. The F1-score is defined as the harmonic mean of recall and precision.

From a mathematical perspective:

The F1-score is 2×TP/(2×TP+ FP+ FN).

In this situation, TP stands for the number of true positives, or cases that were correctly predicted to be positive. The number of cases that are incorrectly projected as positive is known as false positives (FP).Cases that were incorrectly projected to be negative, or false negatives, are quantified as FN.An improved recall and precision balance is indicated by a higher F1-score. In cases where the dataset has an unequal distribution of classes and it is necessary to take into account both false positives and false negatives, it is especially helpful.

3.4 PROPOSED METHODOLOGY/APPLIED MECHANISM

In this section i tried to show which methods i used for the research work.

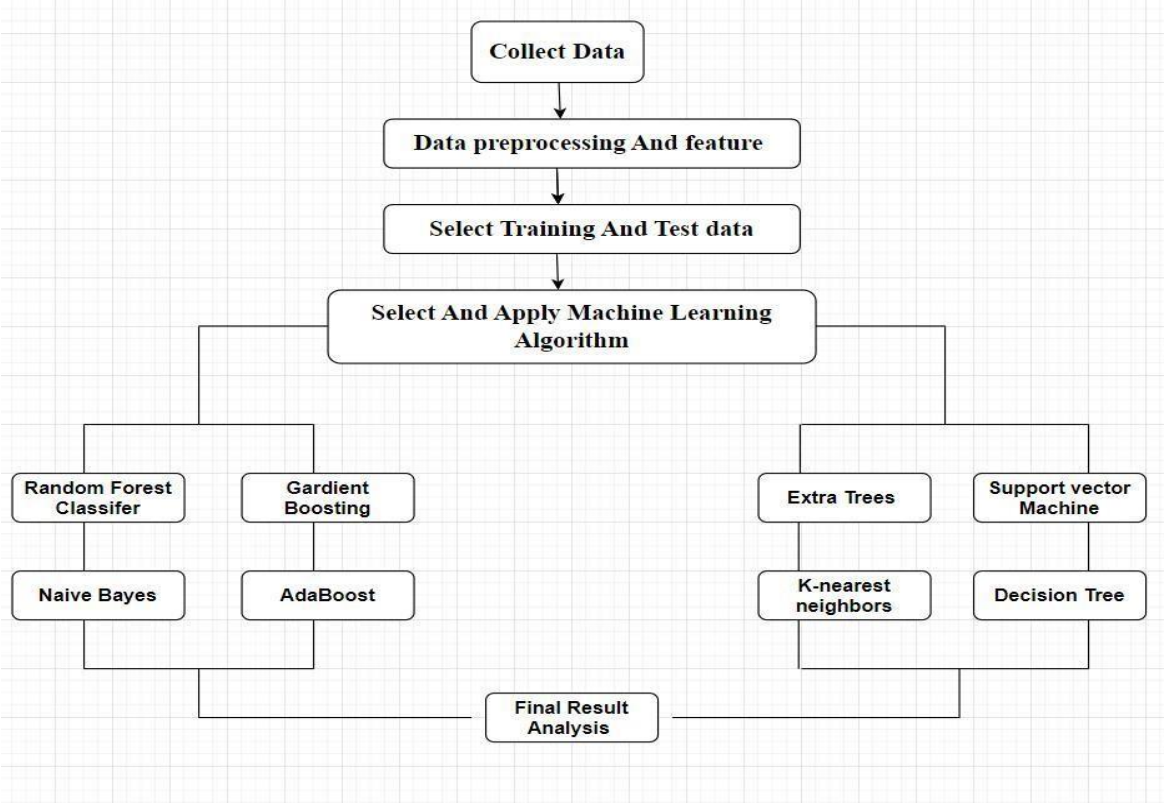


Figure 3.7: suggested structure

3.5 IMPLEMENTATION REQUIREMENTS

● Python 3.10

A version of Python is 3.10. This programming language is advanced. In their research, the majority of academics utilize it. Because of its simplicity in learning and comprehension, this programming language is fairly popular among younger programmers and comes highly recommended for AI-based work.

● Google Collab

The Python programming language can be downloaded for free and open-source via Google Collab. Both the Jupyter notebook and our web browser can be used to work here. However, The fact that Google Collab offers us free virtual GPU access online is by far its greatest benefit.

● Conditions for Hardware and Software

- RAM (more than 4 GB)
- Hard Disc (minimum 4 GB)
- Web Browser (Chrome preferred)
- Operating System(Os) (Windows 7 or above)

CHAPTER 4

Experimental Results and Discussion

Overview

I'll attempt to examine my experiment's setup in this chapter. Next, I'll attempt to examine my work's outcomes, including some of the model evaluation terms F1 score, precision etc. Before moving on to the Confusion Matrix. At last I will try to discuss the topics on which I will work on.

4.1 EXPERIMENTAL SETUP

- Initially, we have the data that we need to use our model and run our code. The following is a list of the system:
- We have collected data from a variety of public and private university students as we have worked with the forecast of the impact of video games on younger generations.
- The majority of our allotted time for the exploration project was devoted to collecting data on gaming students from various types of workplaces through review.
- We have also used Google Forms to collect information online. The data are perfect for further use once we have marked them.
- At that time, the information has been completed and standardized, allowing us to begin the preparation.

By then, our data has been preprocessed.

4.2 EXPERIMENTAL RESULTS AND ANALYSIS

After collecting data preprocessing them and applying different machine learning algorithm it's time to know about the experimental result. Here I can see that random forest outperformed every other algorithms.

Table 4.2: Performance Table

Algorithm Name	Accuracy	Precision	Recall	F1-Score
Random Forest Classifier	97.4	97.51	97.43	97.37
K-nearest Neighbors	82.69	81.18	82.69	79.81
Gradient Boosting	87.82	88.54	87.82	86.10
Naïve Bayes	76.92	71.54	76.92	72.77
AdaBoost	83.33	82.43	83.33	80.33
extra trees	96.15	96.33	96.15	96.00
Support Vector Machine	78.84	83.32	78.84	69.52
Decision tree	94.23	94.13	94.23	94.12

Confusion Matrix

A popular method for demonstrating the efficacy of a machine learning model is a performance metric table. Data from test output is gathered to create it. To validate the performance, four terms are computed: False Negative (FN), False Positive (FP), True Negative (TN), and True Positive (TP). I'll go into more detail about this subject in the section that covers the experiment and findings. Numerous methods, such as random forests, K-nearest-neighbors, gradient boosting, naïve bayes, ads boost, additional trees, support vector machines, and decision tree repeating confusion matrix, are shown in Figures 4.1 to 4.8.

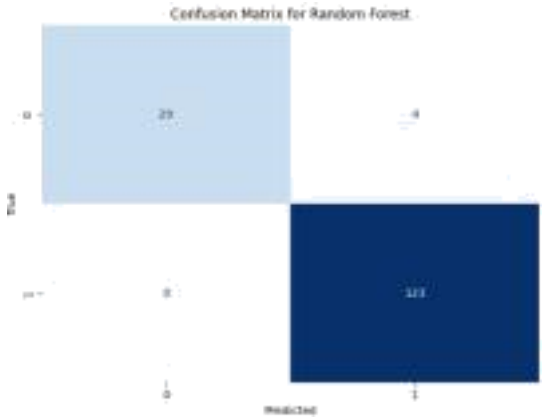


Figure 4.1: Random Forest Confusion Matrix

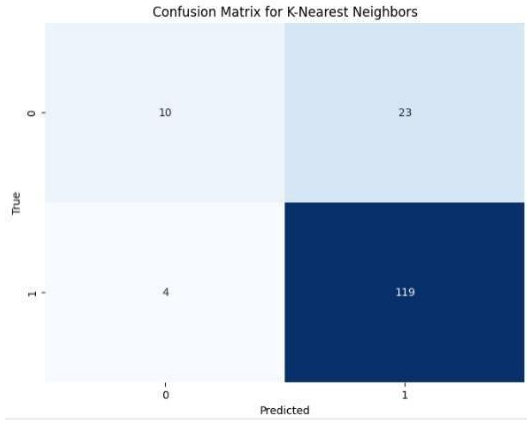


Figure 4.2: K-Nearest Neighbors Confusion Matrix

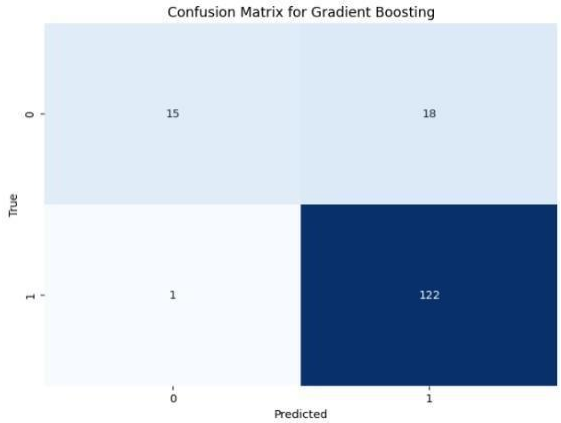


Figure 4.3: Gradient Boosting Confusion Matrix

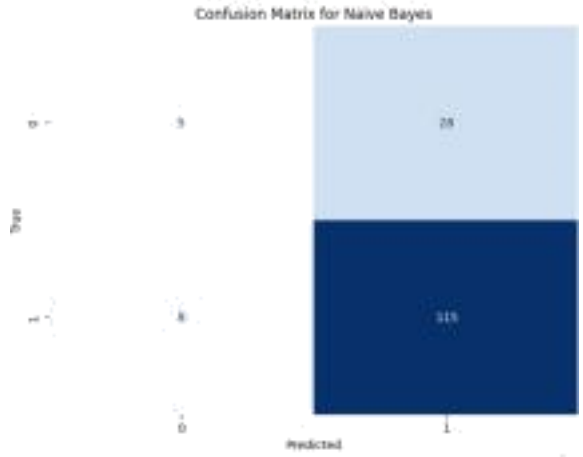


Figure 4.4: Naïve Bayes Confusion Matrix

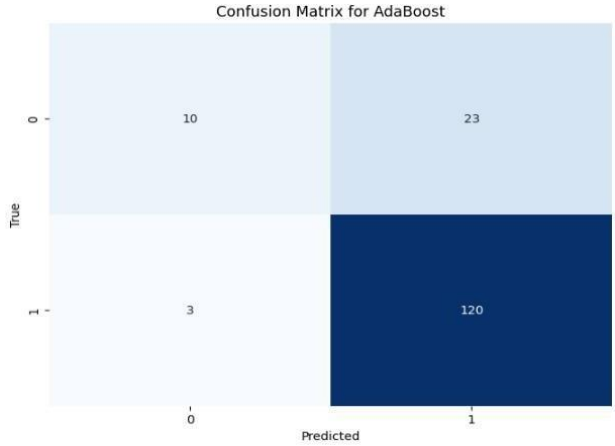


Figure 4.5: Adaboost Confusion Matrix

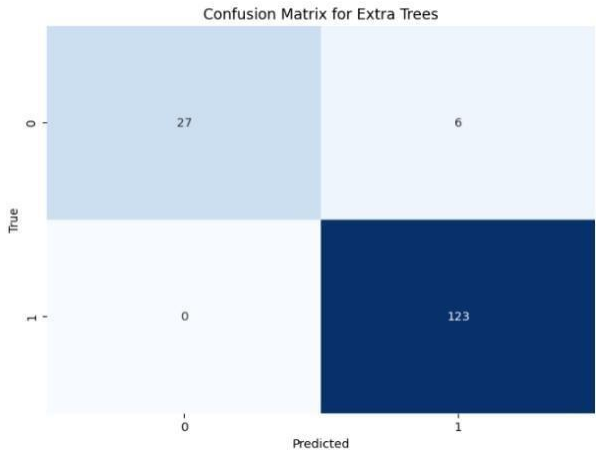


Figure 4.6: Extra Trees Confusion Matrix

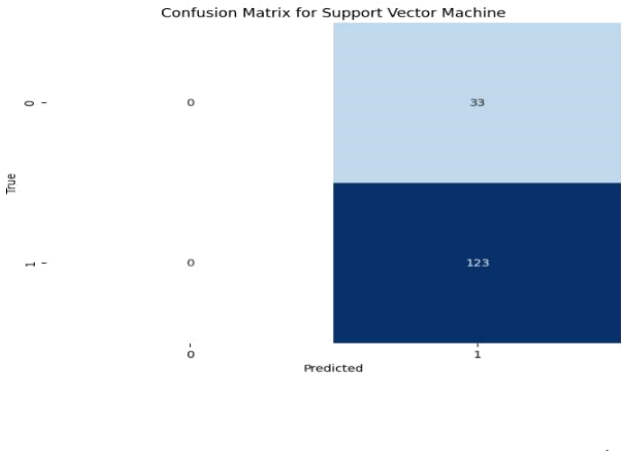


Figure 4.7: Support Vector Machine Confusion Matrix

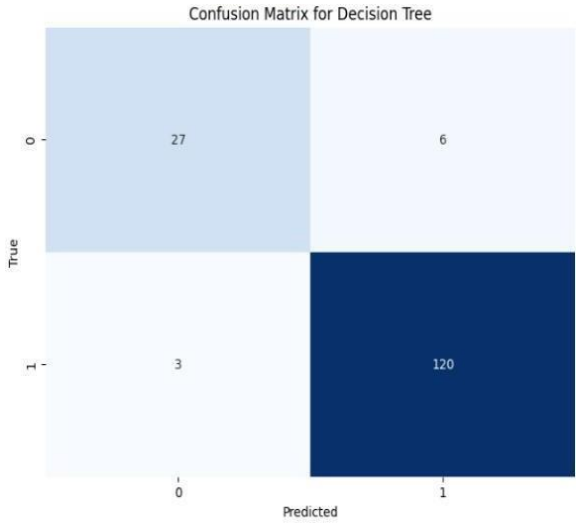


Figure 4.8: Decision Tree Confusion Matrix

4.3 DISCUSSION

The integration of machine learning in predicting the impact of video games on youth opens avenues for nuanced discussions regarding the complex dynamics between gaming habits and various facets of adolescent development. The kind of model i want to build is prediction of impact on mental health of young people. For that first i needed to gather data for applying machine learning algorithms. After applying multiple algorithms we have got a fruitful result of 97.8%.i think it would help people to know the outcomes of the research paper and also they can apply in their lives.

CHAPTER 5

Impact on Society, Environment and Sustainability

5.1 IMPACT ON SOCIETY

Especially in the field of machine learning prediction, has the potential to have a big impact on a lot of different societal issues, like ethical business practices, public awareness campaigns, evidence-based policy development, educational interventions, responsible parenting, research advancements, youth health and well-being concerns, and potential laws to protect young players. Taken together, the study could aid in the growth of a more objective and knowledgeable knowledge of the ways that video games influence young people's life, influencing both individual choices and larger society viewpoints. The impact of predicting video game effects on youth through machine learning reverberates through societal, educational, and psychological realms. On one hand, machine learning algorithms provide a lens into the intricate gaming behaviors of the youth, allowing for a comprehensive understanding of their engagement patterns. This insight aids in the identification of potential cognitive and academic impacts, unraveling correlations between gaming habits and mental health outcomes. In society, the knowledge derived from predictive models offers a foundation for informed discussions and policy decisions. By acknowledging the positive and negative dimensions of video game engagement, society can foster responsible gaming practices and create a balanced approach that harnesses the educational and recreational benefits of video games while mitigating potential risks. From an educational perspective, machine learning predictions provide an opportunity to tailor interventions and support systems for students. Understanding how gaming influences cognitive functions and academic performance enables educators to design educational strategies that leverage the positive aspects of gaming for learning enhancement. Psychologically, the impact is profound as predictive models delve into the emotional and social dynamics of youth gaming. The identification of potential risk factors for conditions like gaming addiction and the exploration of multiplayer interactions contribute to a nuanced understanding of the intricate web of digital and social influences on young people's mental health.

5.2 IMPACT ON ENVIRONMENT

The effect of predicting video game effects on youth through machine learning extends beyond the psychological realm, influencing the environment in diverse ways. As machine learning algorithms analyze gaming patterns, behavioral data, and their consequences, they contribute to the growing digital landscape's ecological footprint. The proliferation of data centers, essential for processing and storing the vast datasets required for such predictions, demands substantial energy consumption, leading to increased carbon emissions. Additionally, the continuous evolution and upgrading of machine learning models contribute to electronic waste, as older hardware becomes obsolete. This creates challenges for responsible e-waste disposal and recycling, impacting ecosystems and communities. Moreover, the development and deployment of machine learning models in this context involve the consumption of computational resources, contributing to the overall carbon footprint associated with the technology sector. Efforts to address the environmental impact of video game prediction models should encompass strategies for optimizing algorithms, implementing energy-efficient computing practices, and promoting transparency in data usage to foster a more sustainable and responsible integration of machine learning in understanding the influence of video games on youth.

5.3 ETHICAL ASPECTS

The moral implications of video games for youth raise questions about their effects on social interactions, potential desensitization to violence, and addictive tendencies. Video game critics contend that playing too many games might impair life skills development and result in poor academic achievement. Video games also present a moral conundrum since they can promote damaging stereotypes and in-game purchases, which can make young people more gullible. It is imperative to strike a balance between the advantages of improving cognitive and motor abilities and these moral issues in order to ensure the welfare and proper growth of youth gamers. This can be accomplished with the aid of ethical game design, parental supervision, and cultural knowledge.

5.4 SUSTAINABILITY PLAN

When parents, educators, and game producers work together, they can successfully reduce the negative effects that video games have on young people. It is crucial to incorporate educational activities into games in order to encourage responsible gaming, critical thinking, and positive values. Gamers should use ethical design principles to steer clear of exploitative tactics and provide information that is acceptable for their age. It is recommended that educational institutions undertake media literacy programs to equip young people with the necessary abilities to handle gaming content appropriately. It is essential that parents take an active role in setting guidelines and encouraging candid communication. With industry self-regulation and ongoing study, video games can be progressively developed as a medium for young people's amusement.

CHAPTER 6

Summary

6.1 SUMMARY OF THE STUDY

In a nutshell, this project describes the impact of gaming among the young generation of our country and the result shows whether they are tensed for career or not. For completing the process first I had to gather data's from various places where there were numerous number of gamers available who played games constantly. Then after having dataset I preprocessed them and trained them to apply multiple machine learning algorithms also I must make a system that can give me output.

6.2 CONCLUSIONS

My main goal from this investigation were the effects of video games on mental health. Ascertain whether this is the case as well as the most effective solution. I also gained knowledge on how to obtain data for research-based initiatives, extract datasets, and use data science practically. How to classify data for testing, choose pertinent datasets, use machine learning algorithms, test the datasets, and ascertain the findings from the testing procedure. In addition, this thesis work gave me a solid understanding of inequality and taught me how to use data science to tackle real-world problems.

6.3 IMPLICATION FOR FURTHER STUDY

I would like to work with a larger dataset for future lessons. Here I worked with a small dataset. Also I would love to integrate the effect of it in the social skills. Therefore the work would look much more interesting.

REFERENCES

[1] S. Islam, A. N. Tusher, M. Sabuj Mia And M. S. Rahman, "A Machine Learning Based Approach To Predict Online Gaming Addiction In The Context Of Bangladesh," 2022 13th International Conference On Computing Communication And Networking Technologies (Icccnt), Kharagpur, India, 2022, Pp. 1-7, Doi: 10.1109/Icccnt54827.2022.9984508.

[2] Gata, Windu, Hasan Basri, Rais Hidayat, Yuyun Elizabeth Patras, Baharuddin Baharuddin, Rhini Fatmasari, Siswanto Tohari And Nia Kusuma Wardhani. "Algorithm Implementations Naïve Bayes, Random Forest. C4.5 On Online Gaming For Learning Achievement Predictions." *Proceedings Of The 2nd International Conference On Research Of Educational Administration And Management (Icream 2018)* (2019): N. Pag.

[3] Swati Aggarwal, Shivin Saluja, Varshika Gambhir, Shubhi Gupta, Simrat Pal Singh Satia,Predicting Likelihood Of Psychological Disorders In Playerunknown's Battlegrounds (Pubg) Players From Asian Countries Using Supervised Machine Learning,Addictive Behaviors,Volume 101,2020

[4] Stavropoulos, V., Zarate, D., Prokofieva, M., Van De Berg, N., Karimi, L., Gorman Alesi, A., Richards, M., Bennet, S., & Griffiths, M. D. (2023). Deep Learning(S) In Gaming Disorder Through The User-Avatar Bond: A Longitudinal Study Using Machine Learning. *Journal Of Behavioral Addictions*, *12*(4), 878–894.

[5] [M. Saravanan. (2017); A Study On Video Game Playing Among Children And Adolescent In Tamilnadu With Reference To Chennai. Int. J. Of Adv. Res. 5 (Jun). 1563-1567] (Issn 2320-5407).

[6] Choo, Hyekyung & Gentile, Douglas & Sim, Timothy & Li, Dongdong & Khoo, Angeline & Liau, Albert. (2010). Pathological Video-Gaming Among Singaporean Youth. Annals Of The Academy Of Medicine, Singapore. 39. 822-9. 10.47102/Annals-Acadmedsg.V39n11p822.

[7] Sebastian, Wilson T. "A Study On The Impact Of Computer Animation On Children In Three Cities Of Kerala (Calicut, Cochin & Trivandrum)." (2010).

[8] Zhengchuan Xu, Ofir Turel & Yufei Yuan (2012) Online Game Addiction Among Adolescents: Motivation And Prevention Factors, European Journal Of Information Systems, 21:3, 321-340,

[9] André, F., Munck, I., Håkansson, A., & Claesdotter-Knutsson, E. (2022). Game Addiction Scale For Adolescents-Psychometric Analyses Of Gaming Behavior, Gender Differences And Adhd. *Frontiers In Psychiatry*, *13*, 791254.

[10] Van Rooij, A. J., Schoenmakers, T. M., Vermulst, A. A., Van Den Eijnden, R. J., & Van De Mheen, D. (2011). Online: Identification Of Addicted Adolescent Gamers. *Addiction (Abingdon, England)*, *106*(1), Video Game Addiction 205–212.

APPENDIX RESEARCH REFLECTION

During the research work i faced three major problems first one is featute selection for the research paper. Second one is collecting the datas.third one is choosing which algorithm do i choose for the research work.therefore we tried to apply various algorithms to see which one suits for our reseach work .Another huge setback was creating survey questions.while collecting data i faced a lot of problems because people are concerned about their data and afraid to give them.so we had to suffer a lot for gathering information.At last after a lot of hard work i finally succed of complete my work.

Plagiarism Report:

ORIGINALITY REPORT

21%	19%	9%	14%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	7%
2	Submitted to Intercollege Student Paper	2%
3	www.ncbi.nlm.nih.gov Internet Source	1%
4	download.atlantis-press.com Internet Source	1%
5	Abdur Nur Tusher, Anika Anjum, Md. Tariqul Islam, Shudipta Das, Sakira Rezowana Sammy, Gahangir Hossain. "Automated Phrasal Verb and Key-Phrase Checking with LSTM-Based Attention Mechanism", 2023 3rd International Conference on Computing and Information Technology (ICCIT), 2023 Publication	1%
6	worldwidescience.org Internet Source	1%
7	Submitted to Huntington Beach Union High School District	1%