Prediction of Bangladeshi Urban Children’s Mental Health for the Effect of Mobile Gaming Using Machine Learning

Tanveer Hasan a,*, Moushumi Zaman Bonny b, Rajesh Kumar Das a, Salma Sultana a, Md. Touhidul Islam a, Abdus Sattar a

a Computer Science and Engineering, Daffodil International University, DIU Rd, Ashulia, 1341, Bangladesh
b American International University – Bangladesh, 408/1, Kuratoli, 1229, Bangladesh

Corresponding author: *tanveer25-964@diu.edu.bd

Abstract—The popularity of mobile gaming is increasing globally, fueled by technological advancements, high-quality smartphones with gaming features, and widespread internet access. A large number of young people and also children are mainly involved in different types of mobile gaming, primarily online mobile gaming, which is a matter of concern for society. This addiction leads to various psychological issues, including mental health problems, loneliness, introversion, insomnia, and a lack of self-control. Bangladesh’s lack of public mental health facilities exacerbates this challenge, particularly in rural areas. To address this issue, we have developed a machine learning model to predict the level of mental health issues faced by Bangladeshi urban children due to gaming addiction. A total of 1996 data were collected from urban parents about their children’s gaming activities and categorized them as ‘Serious,’ ‘Partial,’ and ‘Normal.’ The dataset was then split into a 70:30 ratio for training and testing purposes. K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Multinomial Naïve Bayes, & Random Forest were applied as the machine learning approaches for prediction. Here, the Support Vector Machine gives good accuracy (92.75%), and other classifiers have different accuracy. The research highlights the significance of addressing gaming addiction in children and the potential of machine learning technology in predicting mental health issues, particularly in the absence of public mental health services. The study concludes by proposing the expansion of the research to include smartphone and mobile gaming addiction among Bangladeshi rural children and ensure a holistic approach to tackling the broader issue of technology-related addiction.

Keywords—Prediction; machine learning; child mental health; gaming addiction; SVM.

I. INTRODUCTION

In the fast-paced world we live in today, technology plays a crucial role in driving progress and facilitating innovation. It has enabled the transition from traditional to digital aspects of life. With the increased access to online devices during the pandemic, media reports suggest that online gaming addiction among children and young adults has intensified. One significant outcome of technological advancements is the emergence of Mobile Gaming, which has transformed gaming into a recreational activity accessible through mobile applications. While technology brings numerous benefits, it also presents a potential negative consequence: addiction [1].

With the global outbreak of the COVID-19 pandemic, the implementation of remote learning and restrictions on outdoor activities have resulted in a significant increase in the prevalence of gaming addiction among children and young adults, specifically those aged 7 to 25 and recent studies indicate a high occurrence of online gaming addiction (OGA) among this demographic, with 19 percent of boys and 7.8 percent of girls being identified as having gaming addiction [2]. Mobile game addiction refers to an excessive and irresistible compulsion to play games on portable devices, leading to negative consequences such as depression, feelings of loneliness, and academic underperformance. This addiction has become a significant societal issue, particularly among children and adolescents who have not yet fully developed their mental maturity. It is crucial to determine the underlying causes of these addictive behaviors to develop effective treatment and prevention programs [3].

There is a relationship between inattention and internet/gaming addiction. First, it is believed that compulsive online gaming contributes to sleep problems such as
insomnia. Stress, an inconsistent sleep pattern, poor sleeping habits, and mental health conditions, including anxiety and depression, are all common causes of insomnia. Second, it has been demonstrated that internet game addiction directly contributes to depression by making players anxious and stressed out [4]. According to the World Health Organization (WHO), gaming addiction is recognized as a condition similar to gaming disorder (ICD-11). In 2018, the WHO defined gaming addiction as a consistent or recurring pattern of gaming behavior, whether online or offline, characterized by a lack of control over gaming activities. This addiction is characterized by an increasing prioritization of gaming over other aspects of life and daily activities, as well as the continuation or escalation of gaming despite experiencing negative consequences[5], [6].

Our proposed model established an approach to predicting the mental health conditions of Bangladeshi urban children uniquely in this research. A machine learning approach was applied to predict mental health. SVM, KNN, Gaussian Naive Bayes, and Random Forest were some of the classifier algorithms used. Support Vector Machines have the best accuracy in this case (92.75 percent). In this article, they investigate how a month-long lockdown in the UK has changed mental health symptoms in children and adolescents in response to the pandemic. This study surveyed spiritual indication, habit problems, and delighted or absence of mind changes [7].

Based on ongoing research of elementary school pupils (4–8 years old) identified by teachers as being "at-risk" for mental health issues, this analysis was conducted. Before the COVID-19 pandemic, low-income households that had previously been classified as being in poverty were more likely to report experiencing financial hardship during the barrier [8]. This study was ruled out to determine the ubiquity of mental illness between the ages of 3 and 4 and to find possible correlations at baseline. Their study supports different findings on mental pollution in preschool children and ultimately helps in the application of appropriate screening strategies and the subsequent management of intellectual pollution [9].

The study's goal is to link the indications and symptoms of inattention and hyperactivity in children with attention-deficit/hyperactivity disorder to language development, and cognitive, environmental, socioeconomic, and subtle lifestyle changes [10]. Researchers found that 6% of children under five and 8% of children under nine had above-threshold scores when utilizing the strengths and problems questionnaire as their primary indicator. They identified the children and households most at risk and could provide central early intervention services [11].

This study uses online game playing and YouTube video viewing to identify any common behavioral patterns among moderate and heavy Internet users. Children aged eight to seventeen are the target audience for this study. It is a qualitative study, and the case study and questionnaire survey methods were employed. A disparity between participants' actual and assumed Internet usage time was found in the case study. Parents and many academics will gain from this study because it will alter their perspectives on this significant topic [12].

This study determines which educational circumstances impact children's digital gaming addiction, both in-person and online. This paper defines parenting as encouraging and supporting a child's physical, emotional, social, and intellectual development from infancy to adulthood. According to the World Health Organization, gaming addiction is similar to gaming disorder (ICD-11) [13]. This study aims to determine the prevalence and driving forces behind mobile gaming addiction among Bangladeshi college students. The results show that loneliness, smartphone usage time and mobile gameplay time, and sources of entertainment are the leading causes of cell phone addiction [14]. Asked 1803 Bangladeshi children from 21 districts about their usable electronic gadgets and how much they spent on outdoor activities. They found that 67.11% of children use mobile phones among all the gadgets, and because of the COVID-19 pandemic, 24.48% of students use mobile phones or computers to join online classes. So that’s why they face problems like headaches, sleeping disturbance, and anxiety [15].

In this study on Bangladeshi high school students' involvement with the use of technology devices and the pervasiveness of health-related problems, they said that if students used technology devices, health conditions would decrease. 1803 student samples were taken in this study. They collected data by questionnaire. 21 districts of Bangladesh students are discussed in this study [16]. Look at Facebook addiction and related difficulties among university students in Bangladesh. The information collection included information on sociodemographic traits, actions, and online pursuits. They said that students’ addiction to Facebook is because of failure of love, family problems, depression, sleep disturbance, and so on. Their study noted that 36.9% of students are marked as they are at risk for Facebook addiction [17].

This study guarantees significant advantages in those psychological status efforts' observance, diagnosis, and intervention style. Psychological status or mental health status is related to depression, suicidality, anxiety, Sleeping disturbance, etc. [18]. This article discusses several risk factors related to adolescent disorder. Potential risk factors for teenage gaming disorders include attention issues, emotional dysfunction, family relationship issues, and meeting play needs that are prioritized over meeting daily life needs. They added that these elements are connected to gaming disorders and a few other mental health conditions that gaming disorders have been linked to [19]. During the COVID-19 pandemic Situation Lockdown, the objective is to research children's and teenagers' video game behaviors, and the goal is to learn about children's lives, monitor video game habits, use of parental video games, and other parent-related issues, such as when children play video games and GD (Gaming Disorder) [20].

The study was dependent on compensation theory. It collected data based on public school students in Istanbul who continue to study. In this study, 461 samples of boys were collected. The findings indicate a strong correlation between game addiction and narcissism as well as the child's level of well-being [21].

Collected data using the PubMed database so there are no duplicate values. incorporated information from 15 studies involving 1034 ASD sufferers. Males with ASD made up 85.72 percent of the data gathered [22]. A study examining
differences in problematic video games between urban (Toronto) youth and non-urban youth in northern Ontario (Northern Ontario). They collected 2175 samples of her and showed that 76.6% of his youth lived in urban areas and 80.3% of young people lived in non-urban areas [23].

We surveyed 300 students at Dhaka University, and this survey asked some questions related to socio-demographics, health, and behavior. After regression analysis, they found that those people who are at risk of Facebook addiction have very little physical activity and sleeping problems (less than 7 hours), and those who use Facebook for more than 5 hours daily and have symptoms of depression but there were no gender differences [24].

They find out that there is a relationship between sleep quality and internet addiction. They experimented on 390 medical college students in Dhaka according to the index of Pittsburg Sleep Quality. To identify their sleep quality and to determine internet addiction, they used Orman’s Internet Addiction Test. This experiment said that the sleep quality of around 69.5% of students is either moderate or severely addicted to the Internet [25].

The study introduces a method for predicting the presence of IGD, ADHD, and Generalized Anxiety Disorder (GAD) in players of the renowned MOBA game, PUBG. This approach incorporates player and game information, as well as self-esteem assessments. Focusing on PUBG players from Asian countries, the study extracts player and game statistics and utilizes advanced supervised machine-learning models for accurate prediction. The initial findings demonstrate impressive prediction accuracies of 93.18% for IGD, 81.81% for ADHD, and 84.9% for GAD, showcasing promising results [26].

The study examines a variety of smartphone addiction-related topics, such as the physical, mental, and behavioral changes that can occur as well as the influence on performance and related health problems like cancer, oxidative stress, and neurodegenerative diseases. The results underline the necessity for additional study in this field and offer objectives for more important scientific improvements [27].

In this study, 405 adolescents aged 9 to 14 were tested for obesity, levels of physical activity, and digital game addiction. It discovered a negative association between physical activity and BMI, as well as a negative relationship between game addiction and physical activity. Girls, particularly those who were obese, exhibited higher levels of game addiction. Game addiction was predicted by age, gender, BMI, and physical activity [28].

This comprehensive evaluation of the research focuses on how to prevent and cure pediatric obesity using computerized decision support (CDS) tools and ML algorithms. The study offers insights into the features and results of CDS and ML therapies. Electronic Health Records (EHRs) and alerts for monitoring Body Mass Index (BMI) were frequently used in CDS treatments. Decision trees and artificial neural networks showed potential in correctly forecasting childhood obesity [29].

Every aspect of life has been impacted by the pandemic, and mental health is no exception. Bangladesh has been experiencing a resource crisis, which has caused and continues to cause a governance priority conundrum. The endless breaks from school are hard on our pupils’ mental health, and teenagers are particularly susceptible. The issues with mental health that rural children experienced during. It has been determined that COVID-19 significantly impacts a number of variables that affect children's healthcare in ways that are not just medical but also social, psychological, economic, and educational. All of these factors may have affected the mental development of kids, particularly those who live in rural areas [30], [31], [32].

II. MATERIALS AND METHOD

Supervised learning techniques were used to create a prediction model that can predict child mental health at three levels. A machine learning classifier that differentiates between objects based on particular characteristics was also applied.

![Fig. 1 Overall Methodological Flow Chart](Image)

A classifier predicts an object from a set of feature objects. A model will predict a "mental hamper" in the dataset. The dataset was categorized into three categories: "Serious," "Partial," and "Normal," which indicates the level of mental health hampered condition, and the dataset was split before the training model. 70:30 is the split ratio, meaning 70% of the data was used to train the model, and the remaining 30% was used to test the model. The overall methodological flow chart is shown in Fig. 1. It used a variety of machine-learning algorithms to solve the problem.

A. K-Nearest Neighbors

K-NN is the most basic machine learning algorithm. This method encourages us to choose a set of training samples that are the closest in Distance to a new location and predicted label. The user can set the number of samples or change it depending on the point density in the area. The standard Euclidean distance is the most common method for determining the distance between two points. K-Nearest Neighbors has performed admirably in various classification and regression applications, such as analyzing handwritten digits or satellite data. Fig. 2 shows KNN.
B. Gaussian Naïve Bayes

It is a Bayes theorem-based technique. The Naïve Bayes classifier assumes that a property's inclusion in a class does not affect the property's existence. This model's ease of construction and practicality shine through when working with enormous datasets. It is well known that naive Bayes performs better than complex classification algorithms since it is more straightforward.

C. Random Forest

The Random Forest technique for classification and regression learning is well-known. During the training period, it develops several decision trees. It will be sent to each tree and classified as a new case. Each tree categorizes and provides outputs in a specific category. Because of the Random Forest, A majority vote determines the class corresponding to the most comparable classes produced by different trees.

D. Support Vector Machine

Classification and regression issues can be resolved using the Support Vector Machine (SVM) algorithm. It is an algorithm for supervised learning. It may also be used to separate linear and nonlinear data. Using kernel methods can improve the efficiency of the outcome. Both high-dimensional and low-dimensional data spaces can be handled by it with success. A support vector machine uses hyperplanes to divide a dataset into groups. Hyperplanes are maximal marginal dividing lines generated from clusters to classify datasets, or an (n-1) dimension hyper-plane drawn with a maximum margin when splitting an n-dimensional hypercube. Fig. 4 shows a two-dimensional hypercube split by a one-dimensional hyper-plane (2-1=1).

E. Data Collection

To collect data, we created a Google form and went door to door collecting information and looking for parents who have children and are studying in class from one to ten, ranging in age from five to seventeen. They mainly focused on non-adult students and asked their guardians various questions based on worldwide mental health and gaming addiction-related questions to better understand their situation. Simply provided the questions in both English and Bengali, Children's guardians were able to understand them quickly. Some of the questions asked to the parents were related to the impact of playing online games, and they answered by selecting three options: No, Partial, and Yes. ‘Yes’, denotes they agree with the question, ‘Partial’ means they partially agree, and ‘No’ means they are completely mismatched.

Collected over 1200 data from the children. Most of these kids, 65 percent, were males, and the remaining 35% were girls. 645 children attended primary school, and 555 attended secondary school. This data was collected from five district municipalities in Bangladesh.

F. Data Processing Analyzing

There are 1196 occurrences and 23 attributes in the collection. Some characteristics didn't apply to our research. So those columns were removed. There were only 19 residual attributes after removing those columns, such as ‘gender’, ‘class’, ‘has own smartphone’, ‘is an introvert’, ‘how much hamper sleep’, ‘want to stay alone’, ‘make abnormal behavior when defending in-game’, ‘sudden weight loss’, ‘want to play games as special wishes’, ‘like a fighting game’, ‘spend a vacation with gaming’, etc. relevant with child mental health and our work.

Data was collected via a Google form, and all questions and answers were marked as required. That’s why missing values handling is not needed for this purpose. Label encoding is used in this database features a column by importing the label encoding library. After that, use Standard Scalar to scale down the dataset and prepare for training and testing. Calculate the impact column in our dataset mathematically and manually evaluate the impact on the children. Serious, Partial, and Normal consequences are what can be anticipated. In this case, "severe" denotes the profound effect this child has on mental health. In the case of "Partial," this child also has a moderate effect on mental health. And for "Normal," this child's impact on mental health is little or nonexistent.
III. RESULTS AND DISCUSSION

The main objective of our research is to predict the effect of mobile gaming addiction on Bangladeshi urban children’s mental health. Classifiers based on several machine learning models were used to determine this. Predicted output has three categories: “Serious”, “Partial”, and “Normal”. There were 1196 data for training each of the models. Different accuracy is found according to the various classifier models. Of the four models, SVM and Random Forest performed the best accuracy.

The dataset had 23 feature columns, and each column had three response options: yes, no, and partial. The dataset modified these responses to be encoded as 2, 0, and 1 using a label encoding function. Each respondent’s score was calculated by separating them into primary and secondary school levels.

For students, the impact on mental health was assessed as "Serious" (18-12), "Partial" (11-6), and "Normal" (5-0). Similarly, a score of (19-13) indicates "Serious," (12-5) indicates "Partial," and (5-0) indicates "Normal" for high school students. After analyzing the data in Figure 6, the rate of mental retardation is higher in students who read in classes six to nine. Students from classes one to four are comparatively less affected than them. On the other hand, gender-based analysis has shown in Figure 7 that it affects boys more than girls.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN</td>
<td>83.28%</td>
</tr>
<tr>
<td>Gaussian Naive Bayes</td>
<td>82.45%</td>
</tr>
<tr>
<td>SVM</td>
<td>92.75%</td>
</tr>
<tr>
<td>Random Forest</td>
<td>85.23%</td>
</tr>
</tbody>
</table>

TABLE I
ACCURACY OF CLASSIFIERS
Performance was used to determine how well the classifiers could make decisions. Accuracy, precision, recall, and F1-score were used to evaluate classifier performance. The overall accuracy was seen as a sufficient requirement for classifiers. The concept of the samples that were correctly identified must be present in the test set. Table 1 displays the built-in classifiers’ accuracy values. In our model, SVM gives the highest accuracy, 92.75%, and Random Forest gives the second highest accuracy, approximately 85.23%. Because of this, Additional performance metrics were needed to classify the dataset correctly.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Normal</th>
<th>Partial</th>
<th>Serious</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN</td>
<td>0.91</td>
<td>0.77</td>
<td>0.85</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>0.91</td>
<td>0.77</td>
<td>0.85</td>
</tr>
<tr>
<td>SVM</td>
<td>1.00</td>
<td>0.85</td>
<td>0.97</td>
</tr>
<tr>
<td>Random Forest</td>
<td>0.91</td>
<td>0.77</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Precision measures how well the classifier's positive labels match the data label in terms of classification. Since the precision scores for each of the three class marks are directly related to them, we must compute them. The data for each classifier and the three labels applied in this analysis are shown in Table 2. For children who significantly impact their mental health, the classifier Gaussian Naive Bayes provides a score of 0.85. SVM also offers an outstanding score of 0.97. A score close to 1 was considered to determine the data sample that affected mental health.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Normal</th>
<th>Partial</th>
<th>Serious</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN</td>
<td>0.70</td>
<td>0.69</td>
<td>0.94</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>0.70</td>
<td>0.69</td>
<td>0.94</td>
</tr>
<tr>
<td>SVM</td>
<td>0.72</td>
<td>0.95</td>
<td>0.96</td>
</tr>
<tr>
<td>Random Forest</td>
<td>0.70</td>
<td>0.69</td>
<td>0.94</td>
</tr>
</tbody>
</table>

An identifier for the class label Measurement sensitivity, often known as Recall, is a metric that evaluates categorization efficiency and also concentrates on achieving a Yes Class Label score as close to 1 as feasible. Table 3 displays the classifier and label recall scores for three classes. The memory scores for KNN and SVM for the mentally impaired (Serious) were 0.94 and 0.96, respectively.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Normal</th>
<th>Partial</th>
<th>Serious</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN</td>
<td>0.79</td>
<td>0.73</td>
<td>0.90</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>0.79</td>
<td>0.73</td>
<td>0.90</td>
</tr>
<tr>
<td>SVM</td>
<td>0.84</td>
<td>0.89</td>
<td>0.96</td>
</tr>
<tr>
<td>Random Forest</td>
<td>0.79</td>
<td>0.73</td>
<td>0.90</td>
</tr>
</tbody>
</table>

The purpose of this study is to develop a support vector machine-based prediction model that will more accurately identify mentally challenged people (SVM). Random Forest, SVM, and KNN perform admirably among the classifiers. However, accuracy is not a perfect measure, and precision, F1 score, and recall become proper measures to identify the best model.

Figure 9 compares accuracy and precision for every classifier, and the higher accuracy, 92.75%, and precision, 97%, comes from SVM. KNN, Naive Bayes, and Random Forest work nicely, and the precision is the same (85%) for those three classifiers. Figure 10 compares precision, recall, and F1 scores; the highest performance comes from SVM.

IV. CONCLUSION

The internet can change from productive to disruptive efforts. It allows people to play online games, which disrupts daily life, personal relationships, and work. In this study, we faced different types of problems with children's mental health when they played games on mobile or various types of devices, then developed a methodology for mental health in a particular way and applied a machine learning approach to prediction—used different types of classifier algorithms like SVM (Support Vector Machine), KNN, Gaussian Naive Bayes, and Random Forest. Here, the Support Vector Machine has good accuracy (92.75%). In high-dimensional spaces, SVM is effective. Gaussian Naive Bayes and KNN have performed with almost the same accuracy. However, even Random Forest also has good accuracy. Random Forest can handle large, high-dimensional datasets.

REFERENCES


