Evaluation of Machine Learning Classification Algorithms for Predicting Social Media Addiction Level Among Students

Leon Marco Devela
College of Computing and Information Technologies
National University
Manila, Philippines
develala@students.national-u.edu.ph

Rae Paulos
College of Computing and Information Technologies
National University
Manila, Philippines
paulosrs@students.national-u.edu.ph

***Abstract*—With the steady rise of social media users worldwide, excessive use has become a growing concern, particularly among younger generations. Such behavior has been associated with adverse outcomes, including anxiety, depression, impaired social interactions, and declining academic performance. This study compares and evaluates several classification algorithms—K-Nearest Neighbor, Logistic Regression, Naïve Bayes, Support Vector Machine, Decision Tree, and Random Forest to predict the level of Social Media Addiction (SMA) among students. The models were trained using demographic, social media usage, and health-related variables, and assessed based on accuracy, precision, recall, and F1-score. Results show that the Random Forest algorithm outperformed all other models, achieving the highest accuracy (0.9504), precision (0.9528), recall (0.9504), and F1-score (0.9501). Its superior performance can be attributed to its ensemble learning approach, which effectively mitigates overfitting and captures complex, non-linear feature relationships. Even after hyperparameter tuning, the model's performance remained stable, indicating that the baseline configuration was already near-optimal. Cross-validation further validated its strong generalization capability. These findings establish Random Forest as a robust and reliable model for predicting social media addiction levels.**

***Index Terms*—Machine Learning, Algorithms, Social Media, K-Nearest Neighbor, Logistic Regression, Naïve Bayes, Support Vector Machine, Decision Trees, Random Forest**

I. Introduction

Social media has become an integral part of daily life, transforming how people communicate, share information, and connect with others. However, this widespread adoption has led to a concerning problem that affects millions of individuals worldwide. Recent statistics show that by April 2024, the number of social media users reached 5.07 billion, representing 62.6% of the global population, with around 88% of internet users regularly engaging with these platforms [1]. The constant connectivity and engagement with these platforms have created a new form of behavioral addiction that impacts users across all age groups, but particularly affects young people. The World Health Organization (WHO) reported that problematic social media use among adolescents increased 7% in 2018 to 11% in 2022 [2]. Furthermore, nearly half (46%) of teenagers now report being online "almost constantly," which is almost double the percentage from ten years ago. While more than half (54%) of teenagers say it is hard to quit social media [3]. This growing dependency on social media platforms raises serious concerns about mental health, productivity, and the overall well-being of users globally.

The proposed solution of this study is the use of classic machine learning classification techniques that can predict social media addiction levels based on users social media behaviors and related life outcomes. By applying classic classification algorithms such as Random Forest, Support Vector Machine, and Logistic Regression, the study aims to identify the best classification model that yields the best results. This predictive approach will analyze various features including demographic and behaviors to identify potential addiction patterns early. The advantage of using machine learning techniques lies in their ability to discover complex patterns in data that might not be obvious through traditional analysis methods [4].

This problem is significant because social media addiction has far-reaching consequences on both individual and societal levels. Studies shown that overuse of social media can rewire brains of children and teens, with approximately 42% of teens admitting that social media prevents them from connecting with friends in person [5]. The mental health implications are particularly alarming, as excessive social media use has been linked to increased rates of anxiety, depression, and even suicide among young people. Beyond personal health impacts, social media addiction affects academic performance and real-world social relationships, making it a critical issue that requires effective intervention strategies [6].

Current solutions are insufficient for several reasons. First, existing screening tools are reactive then proactive, identifying addiction only after symptoms have manifested. Next, there is a lack of personalized intervention strategies that consider individual risk factors and usage patterns. Recent research has shown that machine learning models, particularly Random Forest classifiers, can effectively predict different levels of social media addiction by analyzing behavioral and psychological characteristics [7].

The primary populations impacted by this problem include teenagers, young adults, and college students who are the most active social media users. The potential users of the proposed solution include educational institutions seeking to support student well-being, mental health clinics looking for early screening tools, parents wanting to monitor their children's digital health, and individuals who wish to self-assess their social media habits.

The proposed solution can be applied in various settings including schools and universities for student wellness programs, mental health clinics for early intervention, and mobile applications for personal digital well-being. The model could be integrated into existing social media platforms as a built-in health feature, similar to screen time monitoring tools. Furthermore, it could serve as a research tool for studying addiction patterns across different demographics and cultures, helping to develop more targeted intervention strategies.

II. LITERATURE REVIEW

The rise of social media has changed the way people connect, converse, and form social relationships, but has also led to growing concerns about excessive and compulsive use. Social Media Addiction (SMA) is defined as a modern behavioral problem that mirrors the effects of traditional forms of addiction in terms of psychological dependence, withdrawal symptoms, and loss of control, similar to those experienced by people suffering from substance abuse [8]. Although the term *addiction* was once limited to the excessive use of clinical drugs, it has since evolved to include behavioral addictions, within which SMA is now recognized [9], [10].

A. The Cause of Social Media Addiction

SMA is a complex and growing issue in the digital age, with the youth population being the most affected [11]. As online platforms become more integrated into daily life, excessive use has evolved from simple habits to behavioral problems. This is influenced by various psychological, social, and behavioral factors that shape how people engage with online platforms. One major factor is the lack of friends or acquaintances, as people with fewer connections often turn to social media to find companionship and a sense of belonging [12]. With social medias instant validation through likes, comments, and shares, reinforces addictive behavior as it provides instant validation and boosts self-esteem [13]. Another significant contributor is the fear of missing out (FOMO), where users feel anxious about being excluded from trends, events, or any social gathering, prompting them to continuously check their socials. As a result of these factors, it becomes increasingly difficult for users to disconnect as social media intertwines itself in the daily routines of the users [12].

B. The Effects of Social Media Addiction on Students

Although social media offers several positive benefits, such as being used as a platform for collaboration and knowledge acquisitions, excessive use can lead to negative outcomes such as sleep problems, poor academic performance, addictive behavior, and mental health issues [14]. The effects of SMA has been linked to disrupted sleep patterns and increase screen time, particularly during the evening [15]. Studies has shown that individuals with highers levels of SMA tend to spend more time online during the evening, which in turns leads to having shorter sleeping duration and poorer sleep quality. These users are more prone to experiencing insomnia and irregular sleep cycles, which can negatively impact their day-to-day life and overall health [15].

SMA has also been closely linked to adverse mental health conditions, including but not limited to severe levels of depression, anxiety, and stress [16]. Findings indicate that those experiencing SMA had poorer mental health and that the results remain consistent regardless of demographics, such as age and gender. Moreover, social media has increased the chances in which students reported to experience a decline in their academic performance because of poor mental health conditions [17].Although, SMA did not directly influence these conditions, it is closely associated to internet addiction in general, which contributes to emotional distress and decrease in life satisfaction [16].

According to a study done in Turkey, the educational level of a student, whether they be in university or in high school, does not have any significant differences when it comes to the level of social media addiction [18]. Additionally, research revealed that students who engage in multitasking behavior, specifically the use of social media while accomplishing academic tasks often lead to worse academic performance and 20% lower grades compared to their peers who do not multitask [19]. This indicates that the excessive use of social media platforms not only distracts students regardless of level from learning and undermines their cognitive abilities and academic outcomes [14].

C. Machine Learning Approaches to Social Media Addiction

In recent years, machine learning algorithms have been utilized by researchers to analyze large-scale SMA data and identify behavioral patterns that contributes to addictive and compulsive use. These data-driven approaches provide valuable tools to understanding, predicting, and mitigating SMA. A study conducted in Bangladesh employed several ML-models, including Decision Trees, Support Vector Classifier, Random Forest, K-Nearest Neighbors, and Naïve Bayes, to analyze the social media dependence over more than a thousand participants. The study emphasized how significant addressing the problem caused by SMA, using ML techniques to estimate the prevalence of SMA and the impact it has on Bangladeshi citizens. The results shows that the accuracy or the performance of each ML model varied significantly, with RF achieving the highest accuracy of 82%, closely followed by K-Nearest Neighbors at 80% [20]. This highlights the potential of ML models, particularly Random Forest, as reliable tools for predicting SMA, which can inform early interventions and guide the selection of variables in subsequent studies, including the present research.

More recently, a study done in Indonesia developed a Streamlit based tool that predicts the level of SMA among students by utilizing machine learning algorithmes, specifically Random Forest [21]. Tegar et al. [21] reported that by by incorporating variables related to social media usage on a daily basis, mental health status, and relationship conflicts, the developed system can not only visualize represent data in real time, but also effectively identify the severity of SMA.

As this study used the same dataset as of Tegar et al., with the Streamlit application [21], it focuses on the evaluation of the performance of multiple machine learning algorithms to identify the models that provide the most accurate and generalizable predictions of SMA levels. In addition to Random Forest model employed in the Indonesian study, this research incorporates other algorithms used in similar works, including LR, DT, SVM, KNN, and NB to enable a more comprehensive comparison across different classification techniques.

III. Methodology

This section presents the methodology employed in conducting the study. It describes the data collection, tools, and the procedures used to achieve the objectives of the research. This section also explains in detail the process involved in data preparation, model training, and evaluation of results.

A. Data Collection

The data used in this study was from an online survey conducted during the first quarter of 2025. The dataset named *Students' Social Media Addiction* was uploaded by Adil Shamim on Kaggle and it captures usage intensity, preferred social media platform, and relationship dynamics, making it suitable for analyzing the severity or level of SMA among students. It represents students from the ages of 16 to 25 who are in high school, undergrad, or graduate programs across different countries, including but not limited to Bangladesh, India, the United States, the United Kingdom, Canada, Australia, Germany, Brazil, Japan, South Korea, and a hundred others (Total of 110 countries). Those who participated in the survey were contacted through official university mailing lists and other social media networks in order to ensure diversity in academic levels and country representation. Quality control was done by including validation checks, range restrictions (usage hours between 0-24), de-duplication using the unique student ID verification, and anonymization of those who participated in the survey.

Key variables include Age, Gender, Academic Level, Country, Average Daily Usage, Most Used Platform, Sleep Hours Per Night, Mental Health Score, and Addicted Score. These features enable exploration of correlations between social media use and factors such as academic performance, mental health, and sleep behavior. However, since the dataset is self-reported and cross-sectional, results may be influenced by reporting bias and cannot establish causality. Table I provides a detailed description of each variable, including data type and description to clarify how each feature is represented in the dataset.

TABLE I
VARIABLE DESCRIPTION

Attribute Name	Data Type	Description
Student_ID	Integer	Unique numerical identifier assigned to each participant.
Age	Integer	Age of respondent.
Gender	String	Gender of the respondent (Male/Female).
Academic_Level	String	Educational level.
Country	String	Country of residence.
Avg_Daily_Usage_Hours	Float	Number of hours spent daily on social media platforms.
Most_Used_Platform	String	Social media platform most frequently accessed by the respondent.
Affects_Academic_Performance	Boolean	Social media use impacts academic performance (Yes/No).
Sleep_Hours_Per_Night	Float	Average hours of sleep obtained per night.
Mental_Health_Score	Integer	Self-assessed mental well-being score.
Relationship_Status	String	Current relationship category (e.g., Single, In Relationship).
Conflicts_Over_Social_Media	Integer	Reported number of disagreements arising from social media usage.
Addicted_Score	Integer	Composite score representing the level of social media addiction.

B. Data Pre-Processing

To predict the level of SMA among students based on self-reported data, the dataset was first cleaned and preprocessed. The following steps were performed to prepare the data for analysis and model training:

- Understand the dataset.
- Check for duplicate rows and null values.
- Convert categorical values to numeric values with label encoding.

Understanding the dataset began with examining its features, checking the count of non-null values, and the identification of the data types for each column. This helped with determining the necessary actions needed to take for data cleaning and preprocessing. The dataset was also inspected for duplicated rows as this can skew or lead to biases when training the different models if left unchecked. Finally, categorical variables such as Academic Level, Gender, Most Used Platform, Affects Academic Performance, Country, and Relationship Status were converted into numerical values using label encoding to make it compatible with machine learning algorithms.

C. Experimental Setup

Visual Studio Code was used as the main integrated development environment (IDE) for data cleaning, visualization, preprocessing, and model training. All experimental models were done using Python 3.10, with libraries such as NumPy (2.3.3), pandas (2.3.3), matplotlib (3.10.7), scikit-learn (1.7.2) and seaborn (0.13.2) for data manipulation, visualization, and model implementation.

The dataset contained 705 rows and was split into training and testing sets, where 80% of the data was allocated for training and the remaining 20% was used for testing and measuring the accuracy of the models. Each model was trained using data that used the same preprocessing techniques to maintain fairness across comparisons.

In order to identify the most optimal configuration for each algorithm, hyperparameter tuning was performed to establish the baseline models used for comparison. For the K-Nearest Neighbors (KNN) model, the value of k was set to $k = 5$. The Logistic Regression model was trained with a maximum iteration count of 1000 and a random state of 5 to ensure reproducibility. The Support Vector Machine model employed a radial basis function (RBF) kernel and a random state of 5. The Naïve Bayes model used the GausianNB implementation form scikit-learn library with the default parameters. The Decision Tree classifier was configured with a random state of 5, while the Random Forest model utilized 100 estimators and the same random seed for consistency. The hyperparameters used for each model are summarized in Table II.

TABLE II
HYPERPARAMETERS USED TO ESTABLISH BASELINE MODELS

Algorithm	Hyperparameter	Value
K-Nearest Neighbors	Number of Neighbors (k)	5
Logistic Regression	Max Iterations	1000
	Random State	5
Support Vector Machine	Kernel	RBF
	Random State	5
Naïve Bayes	N/A	Default
Decision Tree	Random State	5
Random Forest	Number of Estimators	100
	Random State	5

Hyperparameter tuning using grid search combined with 5-fold cross-validation was performed to improve the performance of each model. Each algorithm was optimized using a specific range of hyperparameters to identify the most suitable configuration. These hyperparameters are summarized in Table III.

TABLE III
HYPERPARAMETERS USED FOR MODEL OPTIMIZATION

Algorithm	Hyperparameter	Value
K-Nearest Neighbors	Number of Neighbors (k)	3, 5, 7, 9, 11
	Weights	uniform, distance
	Metric	euclidean, manhattan
Logistic Regression	Regularization Strength (C)	0.001, 0.01, 0.1, 1, 10, 100
	Solver	lbfgs, saga
	Penalty	l2
	Max Iterations	1000, 2000
Support Vector Machine	Regularization Parameter (C)	0.1, 1, 10, 100
	Gamma	scale, auto, 0.001, 0.01, 0.1
	Kernel	rbf, poly, sigmoid
Naïve Bayes	Var Smoothing	1e-9, 1e-8, 1e-7, 1e-6, 1e-5, 1e-4, 1e-3
Decision Tree	Max Depth	3, 5, 7, 10, 15, 20, None
	Min Samples Split	2, 5, 10, 20
	Min Samples Leaf	1, 2, 4, 8
	Criterion	gini, entropy
Random Forest	Number of Estimators	50, 100, 200, 300
	Max Depth	5, 10, 15, 20, None
	Min Samples Split	2, 5, 10
	Min Samples Leaf	1, 2, 4
	Max Features	sqrt, log2

D. Algorithm

The ML algorithms employed in this study to classify the level of SMA among students include K-Nearest Neighbor, Logistic Regression, Naïve Bayes, Support Vector Machine, Decision Tree, and Random Forest.

K-Nearest Neighbor (KNN)

This is a non-parametric classification algorithm that assigns a label to a data point based on the majority class among its k nearest neighbors or within its proximity [22]. Although KNN can be applied for both regression and classification tasks, it is primarily used for classification and operates on the idea that similar data points exists close to one another [23].

KNN uses the distance of a data point to other data points in order to find the nearest neighbor. To identify these neighbors, the algorithm computes the distance between the test instance x and each training instance X_i using a distance metric, commonly the Euclidean distance formula, which can be mathematically representedd as shown in (1) [22].

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{1}$$

Where:
- $d(x, y)$ is the distance between points x and y.
- x_i and y_i are the ith feature of the points.
- n is the total number of features.

Logistic Regression (LR)

Despite its name, Logistic Regression is a supervised learning algorithm primarily used for classification tasks rather than regression. It predicts the probability that a given input belongs to a particular class rather than estimating a continuous value [22]. Logistic Regression models the relationship between the dependent and independent binary variable by applying the logistic (sigmoid) function as shown in (2) to limit the output between 0 and 1 [24], [25].

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

Where:
- $f(x)$ is the output of the sigmoid function.
- e represents the Euler's number at a constant of 2.71828.
- x is the input.

Naïve Bayes (NB)

This is a supervised ML algorithm that uses probability of a data point to predict its category. It operates on the assumption that all features in the dataset are independent from one another, hence the term ”naive”. The algorithm further assumes that each feature in the dataset contributes equally to the prediction outcome [26], [27]. This is represented mathematically in (3).

$$P(C_k \mid X) = \frac{P(X \mid C_k)\, P(C_k)}{P(X)} \tag{3}$$

Where:

- $P(C_k|X)$ represents the probability of class C_k given the features X.
- $P(X|C_k)$ represents the probability of the features given class C_k.
- $P(C_k)$ is the prior probability of class C_k
- $P(X)$ is the probability of the features.

Support Vector Machine (SVM)

This can be applied to both regression and classification tasks but primarily for classification problems in which it attempts to identify the hyperplane or the boundary that separates the classes present within the dataset. The algorithm maximizes the margin between classes to improve generalization and classification accuracy [28], [29].

When the data points are not linearly separable, SVM employs kernel functions that transforms the input data into a higher dimensional feature space that allows it to find a separating hyperplane. SVM can be mathematically represented by the equation (4) [28].

$$w^T \cdot x + b = 0 \tag{4}$$

Where:

- w represents the weight vector that determines the orientation or the direction of the hyperplane.
- x input feature vector.
- b offset or the distance of the hyperplane from the origin.

Decision Tree

These are supervised algorithms used for both classification and regression task, although primarily more common for classification problems. It works by recursively splitting the dataset into subsets based on the values of the input features which then forms a hierarchical or inverse tree structure composed of an uppermost node called the root node followed by the internal and leaf nodes. This is repeated until it reaches the leaf node where the final prediction is made [30], [31].

Decision Trees can be expressed mathematically by their impurity measures and Information Gain which helps guide the splitting of the nodes. Entropy is the disorder or the impurity in a dataset [31], while Information Gain measures the reduction in Entropy when a particular attribute has been split. The goal is to get the attribute with the highest Information Gain and continuously select it for the split as it results in the most homogeneous subset [32]. This can be represented in equations (5) and (6).

$$Entropy(S) = -\sum_{i=1}^{c} p_i \log_2 p_i \tag{5}$$

Where:

- S represents the entropy of the dataset.
- c are the count of unique classes.
- p_i is the probability of class *i* in the dataset.

$$IG(S,a) = Entropy(S) - \sum_{v \in Values(a)} \frac{|S_v|}{|S|} Entropy(S_v) \tag{6}$$

Where:

- a is the attribute being evaluated.
- $Entropy(S)$ represents the entropy of S.
- $|S_v|/|S|$ is the values of S_v in proportion to the values in the set.

Random Forest (RF)

Random Forest is a supervised ML algorithm that combines the prediction of multiple Decision Trees to improve accuracy and reduce overfitting [33]. Each tree is then trained on random parts of the dataset using bootstrap sampling and each split only considers a random subset of features. When used for classification tasks, the predictions from each of the sub-trees are combined in what is called majority voting where the class with the most votes is chosen as the final prediction. One of the key benefits of using RF is that it can be utilized in either classification and regression tasks and still maintain accurate results. Additional benefits include its ability to handle missing values while maintaining high accuracy and can even list the characteristics that are the most suitable for making predictions [34].

As RF is an compilation of multiple decision trees, there is no single formula to represent it mathematically. However, in classification problems the algorithm uses the Gini Index to decide which branch is *”better”* and can be expressed by the formula (7) or by its entropy as shown in (5) [35].

$$Gini(S) = 1 - \sum_{i=1}^{c} (p_i)^2 \tag{7}$$

Where:

- $Gini(S)$ represents the dataset being evaluated.
- c is the total number of classes.
- p_i is the frequency of the class.

E. Training Procedure

Two models were initially trained for each algorithm: one using all the available features that were cleaned and preprocessed and another using only selected features that are not in the scope of the demographic data and less relevant attributes. Features such as age, gender, and educational level were excluded based on findings from prior studies that identified them as insignificant predictors of social media addiction [17], [18]. Other features, including country, most used platform, relationship status were excluded from the selected features model per algorithm due to low correlation as seen in Figure 1.

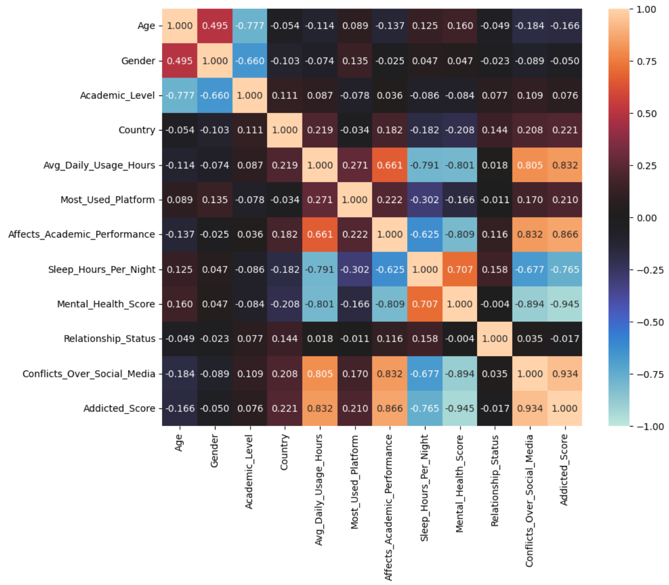

Fig. 1. Corelation Matrix

For each pair of models per algorithm, whichever yielded the better results was then optimized further with the use of Grid Search Cross Validation (GridSearchCV) for hyperparameter tuning. The GridSearchCV method performed a 5-fold-cross-validation to ensure robustness of the model, reduce overfitting, and obtain the average performance metrics across all folds.

F. Evaluation Metrics

As no single evaluation metric tells the whole story or the correctness of a model, four primary metrics were used: accuracy, precision, recall, and F1-score. These metrics were used as they provide different perspectives on the performance of the model.

a) Accuracy: Tells the percentage of the correct predictions the model made. However, utilizing accuracy as the only metric can lead to biases when the dataset is imbalanced. Accuracy can be calculated with (8).

$$Accuracy = \frac{\text{Num of Correct Predictions}}{\text{Total Num of Predictions}} \tag{8}$$

b) Precision: This indicated how many positive predictions made by the model are actually correct. Unlike accuracy, precision focuses on the quality of correct predictions rather the count. Precision is calculated as seen in (9).

$$Precision = \frac{\text{True Positive}}{\text{True Positive + False Positive}} \tag{9}$$

c) Recall: This measures the number of predicted positive cases were actually correctly identified by the model. Recall is calculated with (10).

$$Recall = \frac{\text{True Positive}}{\text{True Positive + False Negative}} \tag{10}$$

d) F1-score: This is the harmonic mean of precision and recall which means it provides a single measure for both metrics. This can be mathematically represented by (11).

$$F1Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{11}$$

These metrics were computed using the predictions from the test set after training each model. The results were then compared to identify which algorithm achieved the highest performance across all metrics. Comparisons were also made between the baseline models and their respective optimized version to determine the effect hyperparameter tuning has on the classification models' performance. This approach ensured a fair and comprehensive evaluation of how each model performed in predicting the level of social media addiction.

G. Baseline and Comparative Models

For each algorithm, initial models using either all available features or a selected subset of features were used as the baseline comparative models. These models served as the reference for the performance of the optimized models for each algorithm. The optimized models were then evaluated against these baselines. Across most algorithms, the optimized models either showed improvements or maintained comparable performance with the baseline. Models like Logistic Regression and Naïve Bayes showed a more consistent predictive accuracy from the hyperparameter tuning, while Random Forest retained the same results even after optimization. However, despite retaining the same results, the Random Forest remained the most accurate model among all models tested.

IV. Results and Discussion

This section presents the findings from the evaluation of six machine learning classification algorithms for predicting social media addiction levels among students. The results are presented systematically, beginning with baseline model performance, followed by cross-validation results, optimized model comparisons, and a comprehensive discussion of the implications and limitations of these findings.

A. Key Findings

The evaluation of six machine learning algorithms revealed significant variations in predictive performance across different model configurations. The primary findings indicate that ensemble methods, particularly Random Forest, consistently outperformed individual classifiers in predicting social media addiction levels among students.

B. Baseline Model Performance

Initial experiments compared models trained on all available features against those trained on selected features (excluding demographic variables and low-correlation attributes). Table IV presents the comprehensive results for all baseline configurations.

The results reveal clear patterns. For KNN, increasing (k) from 1 to 5 with selected features raised accuracy from 0.8369 to 0.8794, indicating that using more neighbors helps reduce noise. In most algorithms, models using all features performed better than those with selected features except for SVM, which dropped from 85.82% to 28.37% accuracy. This decline suggests that SVM struggled with high dimensionality and irrelevant features in the full dataset.

Among all baseline configurations, Random Forest with all features achieved the highest performance across all metrics (95.04% accuracy, 95.28% precision, 95.04% recall, and 95.01% F1-score), followed by Decision Tree with all features (92.91% accuracy). The superior performance of ensemble methods aligns with their theoretical advantages in reducing overfitting and capturing complex non-linear relationships in the data.

C. Best Model Comparison Before Optimization

Table V presents the best performing configuration for each algorithm before hyperparameter optimization.

This comparison reveals a clear performance hierarchy, with tree-based methods (Random Forest and Decision Tree) outperforming other models (Logistic Regression, Naïve Bayes, KNN, and SVM). The performance gap between Random Forest (95.04%) and the lowest-performing model, Naïve Bayes (85.11%), amounts to approximately 10%, indicating significant differences in the algorithms' ability to capture the underlying patterns in social media addiction data.

D. Cross-Validation Results

To assess model generalization and robustness, 5-fold cross-validation was performed on the best configuration of each algorithm. Table VI presents the averaged performance metrics across all folds.

Cross-validation results reveal interesting shifts in relative performance. While Random Forest maintained its position as the top performer with 86.95% accuracy, the performance gap between models narrowed considerably compared to single train-test split results. Notably, Logistic Regression's accuracy decreased by approximately 12% (from 90.78% to 78.44%), suggesting that its initial high performance may have been partially due to favorable characteristics of the specific train-test split. Conversely, Random Forest showed a smaller decline (from 95.04% to 86.95%), demonstrating better generalization across different data subsets. The substantial performance drop observed across most models during cross-validation indicates potential overfitting in the baseline configurations, particularly for Logistic Regression and Decision Tree. This finding emphasizes the importance of cross-validation in providing realistic estimates of model performance and guided the subsequent hyperparameter optimization phase.

E. Optimized Model Performance

Following hyperparameter tuning using GridSearchCV with 5-fold cross-validation, each algorithm was re-evaluated on the test set. Table VII presents the final optimized results.

Hyperparameter optimization produced varied improvements across algorithms. Naïve Bayes demonstrated the most substantial improvement, increasing from 85.11% to 91.49% accuracy, a gain of 6.38%. This tuning allowed the model to better capture the underlying probability distributions.

Interestingly, Random Forest maintained identical performance after optimization (95.04% accuracy), indicating that the initial hyperparameters (100 estimators, random state of 5) were already near-optimal for this classification task. Decision Tree showed minimal change, with accuracy remaining at 92.91% while F1-score increased slightly from 0.9288 to 0.9320. Similarly, Logistic Regression improved from 90.78% to 91.49% accuracy, though the improvement was more modest at 0.71 percentage points.

KNN and SVM showed minimal or no improvement through optimization, with both algorithms achieving 85.82% accuracy in their optimized forms. For KNN, this suggests that the selected feature set and distance metric were more influential than the number of neighbors, while for SVM, the challenge of navigating the high-dimensional feature space persisted despite kernel and regularization parameter tuning.

F. Overall Performance Ranking

The final ranking establishes Random Forest as the superior classifier for this problem, achieving balanced performance across all metrics. Decision Tree secured second place with 92.91% accuracy, while Naïve Bayes and Logistic Regression tied for third place at 91.49% accuracy. The relatively close performance of Naïve Bayes and Logistic Regression suggests that while the relationship between features and addiction levels contains non-linear components best captured by tree-based methods, linear assumptions still provide reasonable approximations.

V. Conclusion

Social media addiction has become an increasingly pervasive problem that affects millions of individuals worldwide, particularly among younger and more impressionable generations. Excessive use has been repeatedly shown to contribute to serious mental health issues such as anxiety, depression, and social dysfunction. As such, the primary objectives of this research were to determine which classification algorithm yields the best results and to identify key features that can serve as indicators for preventative measures or early signs of social media addiction. Overall, the research aimed to train a reliable predictive model that can accurately detect the severity of social media addiction among students.

TABLE IV
BASELINE MODEL PERFORMANCE ACROSS ALL CONFIGURATIONS

Algorithm	Model Configuration	Accuracy	Precision	Recall	F1-Score
KNN	Model 1: All Features (K=1)	0.8582	0.8685	0.8582	0.8607
	Model 2: All Features (K=3)	0.8227	0.8289	0.8227	0.8232
	Model 3: All Features (K=5)	0.8369	0.8415	0.8369	0.8368
	Model 4: Selected Features (K=1)	0.8369	0.8427	0.8369	0.8388
	Model 5: Selected Features (K=3)	0.8582	0.8595	0.8582	0.8572
	Model 6: Selected Features (K=5)	**0.8794**	**0.8803**	**0.8794**	**0.8764**
Logistic Regression	Model 1: All Features	**0.9078**	**0.9158**	**0.9078**	**0.8993**
	Model 2: Selected Features	0.8652	0.8792	0.8652	0.8516
SVM	Model 1: All Features	0.2837	0.0805	0.2837	0.1254
	Model 2: Selected Features	**0.8582**	**0.8398**	**0.8582**	**0.8373**
Naïve Bayes	Model 1: All Features	**0.8511**	**0.7947**	**0.8511**	**0.8171**
	Model 2: Selected Features	0.7376	0.6477	0.7376	0.6788
Decision Tree	Model 1: All Features	**0.9291**	**0.9317**	**0.9291**	**0.9288**
	Model 2: Selected Features	0.8440	0.8539	0.8440	0.8454
Random Forest	Model 1: All Features	**0.9504**	**0.9528**	**0.9504**	**0.9501**
	Model 2: Selected Features	0.8794	0.8779	0.8794	0.8761

TABLE V
COMPARISON OF BEST BASELINE MODELS PER ALGORITHM

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	**0.9504**	**0.9528**	**0.9504**	**0.9501**
Decision Tree	0.9291	0.9317	0.9291	0.9288
Logistic Regression	0.9078	0.9158	0.9078	0.8993
KNN (K=5)	0.8794	0.8803	0.8794	0.8764
SVM	0.8582	0.8398	0.8582	0.8373
Naïve Bayes	0.8511	0.7947	0.8511	0.8171

TABLE VI
CROSS-VALIDATION RESULTS (5-FOLD) FOR BEST MODELS

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	**0.8695**	**0.8867**	**0.8695**	**0.8553**
SVM	0.8298	0.8016	0.8298	0.8055
Decision Tree	0.8213	0.8326	0.8213	0.8174
KNN (K=5)	0.7943	0.7970	0.7943	0.7786
Logistic Regression	0.7844	0.8132	0.7844	0.7701
Naïve Bayes	0.7348	0.7185	0.7348	0.6979

Among all models evaluated, the Random Forest consistently produced the most reliable and accurate predictions. The fact that its performance did not improve after tuning indicates that the model was already operating efficiently under its initial settings. Moreover, the results from cross-validation show that the Random Forest maintained consistent accuracy when tested on different portions of the data, demonstrating its capability to generalize well to unseen cases. Overall, these findings suggest that Random Forest is a robust and dependable model for predicting social media addiction levels.

The main limitation of this study is the unequal representation of classes within the dataset. Some of the 110 countries had only one participant, making cross-country comparisons less reliable. The addiction scores were also unevenly distributed, with most data concentrated at level 7 and no representation for levels 1 and 10, which may have introduced bias and reduced the model's ability to generalize. Future studies are encouraged to apply data balancing techniques such as SMOTE or collect larger, more diverse datasets to improve model performance. Ensuring even class distribution would promote fair representation and lead to more balanced and reliable predictions.

TABLE VII
OPTIMIZED MODEL PERFORMANCE AFTER HYPERPARAMETER TUNING

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	**0.9504**	**0.9528**	**0.9504**	**0.9501**
Decision Tree	0.9291	0.9377	0.9291	0.9320
Naïve Bayes	0.9149	0.9144	0.9149	0.9093
Logistic Regression	0.9149	0.9213	0.9149	0.9091
SVM	0.8582	0.8620	0.8582	0.8534
KNN	0.8582	0.8595	0.8582	0.8572

The findings of this study highlight the potential of machine learning, particularly Random Forest, in predicting and understanding social media addiction among students. By demonstrating that reliable predictions can be achieved even with minimal model tuning, this research emphasizes the practicality of data-driven approaches in addressing behavioral and mental health challenges. In the long term, such predictive models can contribute to early detection and intervention efforts, supporting the development of more informed strategies to promote healthier and more balanced social media usage among students.

REFERENCES

[1] Lotus Behavioral Health, "Social media addiction: Statistics, facts, and how to break free," *Lotus Behavioral Health Report*, April 2024. [Online]. Available: https://www.lotusbh.org/post/social-media-addiction-statistics-and-facts

[2] World Health Organization Regional Office for Europe, "Teens, screens and mental health: New data reveals sharp rise in problematic social media use," WHO Europe, Tech. Rep., September 2024. [Online]. Available: https://www.who.int/europe/news/item/25-09-2024-teens--screens-and-mental-health

[3] CloudWards Research Team, "21+ eye-opening social media addiction statistics in 2025," *CloudWards Digital Health Report*, April 2025. [Online]. Available: https://www.cloudwards.net/social-media-addiction-statistics/

[4] M. Mahat, "Prediction and prevention of addiction to social media using machine learning," in *2021 International Conference on Computing and Communications*. IEEE, 2021, pp. 1–8.

[5] Lanier Law Firm Research Division, "Social media addiction statistics: Understanding the digital mental health crisis," *Digital Health and Safety Report*, August 2025. [Online]. Available: https://www.lanierlawfirm.com/social-media-addiction/statistics/

[6] Stanford Law School, "Social media addiction and mental health: The growing concern for youth well-being," https://law.stanford.edu/2024/05/20/social-media-addiction-and-mental-health-the-growing-concern, May 2024, accessed: 15 October, 2025.

[7] A. Rahman, S. Kumar, and L. Chen, "Machine learning for detecting social media addiction patterns: Analyzing user behavior and mental health data," *International Journal of Innovations in Science & Technology*, vol. 7, no. 1, pp. 112–128, January 2025.

[8] D. Tullett-Prado, V. Stavropoulos, R. Gomez, and J. Doley, "Social media use and abuse: Different profiles of users and their associations with addictive behaviours," *Addictive Behaviors Reports*, vol. 17, p. 100479, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352853223000019

[9] S. R. Chamberlain and J. E. Grant, *Behavioral Addictions*. Cambridge University Press, 2019, p. 401–412.

[10] V. Starcevic, "Behavioural addictions: A challenge for psychopathology and psychiatric nosology," *Australian & New Zealand Journal of Psychiatry*, vol. 50, no. 8, pp. 721–725, 2016, pMID: 27357713. [Online]. Available: https://doi.org/10.1177/0004867416654009

[11] H. Pallathadka, L. K. Pallathadka, and S. K. Devi, "Reasons of social media addiction among youth: An exploratory study," *Integrated Journal for Research in Arts and Humanities*, vol. 2, no. 6, pp. 137–143, 2022.

[12] M. E. Aksoy, "A qualitative study on the reasons for social media addiction," *European Journal of Educational Research*, vol. volume–7–2018, no. volume7-issue4.html, p. 861–865, Oct 2018.

[13] A. Larson, "What makes social media so addictive, and how does it affect our mental health?" *Psychology Student Publications*, Oct 2024. [Online]. Available: https://commons.und.edu/cgi/viewcontent.cgi?article=1062&context=psych-stu

[14] T. Subair, S. ADEBOLA, and D. YAHYA, "Social media: Usage and influence on undergraduate studies," *International Journal of Education and Development using ICT*, vol. 15, no. 3, 2019.

[15] G. J. Hjetland, J. C. Skogen, M. Hysing, and B. Sivertsen, "The association between self-reported screen time, social media addiction, and sleep among norwegian university students," *Frontiers in public health*, vol. 9, p. 794307, 2021.

[16] N. Ergün, Z. Özkan, and M. D. Griffiths, "Social media addiction and poor mental health: Examining the mediating roles of internet addiction and phubbing," *Psychological reports*, vol. 128, no. 2, pp. 723–743, 2025.

[17] L. Braghieri, R. Levy, and A. Makarin, "Social media and mental health," *American Economic Review*, vol. 112, no. 11, pp. 3660–3693, 2022.

[18] A. Simsek, K. Elciyar, and T. Kizilhan, "A comparative study on social media addiction of high school and university students," *Contemporary educational technology*, vol. 10, no. 2, pp. 106–119, 2019.

[19] O. A. Abdullahi, M. Bahari, S. Miskon, and M. Abu Yazid, "Social media addiction and academic performance: A bibliometric analysis approach," *International Journal of Academic Research in Business and Social Sciences*, vol. 14, no. 1, pp. 1217–1228, 2024.

[20] M. N. Mim, M. Firoz, M. M. Islam, M. Hasan, and T. Habib, "A study on social media addiction analysis on the people of bangladesh using machine learning algorithms," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 5, Jul 2024. [Online]. Available: https://beei.org/index.php/EEI/article/view/5680

[21] A. T. B. S. Tegar, H. Hasanah, and I. Oktaviani, "Predicting social media addiction using machine learning and interactive visualization with streamlit," *bit-Tech*, vol. 8, no. 1, pp. 778–788, 2025.

[22] GeeksforGeeks, "K-nearest neighbor (knn) algorithm," https://www.geeksforgeeks.org/machine-learning/k-nearest-neighbours/, last updated: 23 August 2025. Accessed: 15 October 2025.

[23] IBM, "What is the k-nearest neighbors (knn) algorithm?" https://www.ibm.com/think/topics/knn, accessed: 15 October, 2025.

[24] ——, "What is logistic regression?" https://www.ibm.com/think/topics/logistic-regression, accessed: 15 October 2025.

[25] Google Developers, "Logistic regression: Calculating a probability with the sigmoid function," https://developers.google.com/machine-learning/crash-course/logistic-regression/sigmoid-function, last updated: 22 September 2025, Accessed: 15 October 2025.

[26] GeeksforGeeks, "Naive bayes classifiers," https://www.geeksforgeeks.org/naive-bayes-classifiers/, last updated: 25 August 2025. Accessed: 15 October 2025.

[27] IBM, "What are naïve bayes classifiers?" https://www.ibm.com/think/topics/naive-bayes, accessed: 15 October 2025.

[28] GeeksforGeeks, "Support vector machine (svm) algorithm," https://www.geeksforgeeks.org/machine-learning/support-vector-machine-algorithm/, last updated: 07 August 2025. Accessed: 15 October 2025.

[29] IBM, "Support vector machine (svm)," https://www.ibm.com/think/topics/support-vector-machine, accessed: 15 October 2025.

[30] GeeksforGeeks, "Decision tree," 2025, accessed: 15 October 2025. [Online]. Available: https://www.geeksforgeeks.org/machine-learning/decision-tree/

[31] IBM, "What is a decision tree?" 2025, accessed: 15 October 2025. [Online]. Available: https://www.ibm.com/think/topics/decision-trees

[32] ScienceDirect, "Information gain - an overview," 2025, accessed: 15 October 2025. [Online]. Available: https://www.sciencedirect.com/topics/computer-science/information-gain

[33] IBM, "What is random forest?" 2025, accessed: 15 October 2025. [Online]. Available: https://www.ibm.com/think/topics/random-forest

[34] GeeksforGeeks, "Random forest algorithm in machine learning," 2025, accessed: 15 October 2025. [Online]. Available: https://www.geeksforgeeks.org/machine-learning/random-forest-algorithm-in-machine-learning/

[35] M. Schott, "Random forest algorithm for machine learning," 2019, accessed: 2025-10-15. [Online]. Available: https://medium.com/capital-one-tech/random-forest-algorithm-for-machine-learning-c4b2c8cc9feb