

Social Media Addiction Using Machine Learning Technologies

Ms. Geetanjali Narware¹, Dr. Anish Kumar Choudhary²

¹*M. Tech (Computer Science and Engineering)*

²*Associate Professor, IET CSE Dept, SAGE University, Indore*

Abstract—While social media is a vital tool for communication in today's world, its overuse has created a problem that is known as social media addiction. This addiction can have a negative impact on mental health, academic performance, productivity, sleep quality, and social relationships. The traditional approach to detection is largely through surveys and/or psychological testing, which tend to be subjective and time consuming. The study introduces a Machine Learning (ML) based solution to identify the signs of addiction to social media through the analysis of users' behavioral patterns, including screen time, log-in frequency, interaction rate, emotional dependency and checking notifications. The proposed system conducts data preprocessing, feature extraction, model training, and model classification to classify different levels of addiction. The performance of several machine learning algorithms, such as Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Artificial Neural Network (ANN) is compared using the measures of accuracy, precision, recall and f1-score. Experimental results show that the accuracy of prediction of the models Random Forest and ANN is higher than other models. The suggested model offers a smart and effective approach to early detection of social media addiction, encouraging healthy digital behaviors and fostering digital well-being for users and researchers.

Index Terms—Social Media Addiction, Machine Learning, Artificial Intelligence, Classification Algorithms, Behavioral Analysis, Digital Well-being, Predictive Modeling

I. INTRODUCTION

Social media has become an integral part of life, allowing people to communicate with others, share information, and watch entertainment on the web platforms including Facebook, Instagram, Twitter,

YouTube and TikTok. While there are numerous advantages to these platforms, excessive use has been associated with a problem referred to as social media addiction which can have negative impacts on mental health, academic achievement, workplace productivity, sleep patterns and interpersonal relationships. The traditional methods of detection of social addiction are mainly questionnaires and self-reported survey, which are time-consuming and subjective. In the era of Artificial Intelligence (AI) and Machine Learning (ML), it has become possible to use data already collected by the system to automatically analyze the user's behavioral patterns and more accurately determine if they are addicted. Addictive usage patterns are identified through behavioral characteristics like checking notifications, interaction rate, logins, and screen time. Machine learning algorithms such as Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Artificial Neural Network (ANN) can be used to classify users according to their level of addiction (Normal, Moderate, and High Addiction). This study suggests a machine learning-based approach to social media addiction detection using behavioral data analysis. The main objectives of the proposed system are to accurately predict the behavior, facilitate early intervention and encourage healthy digital well-being.

II. LITERATURE REVIEW

With the growing usage of social networking services, social media addiction has emerged as a salient field of research. Researchers have attempted to use different techniques like machine learning, Artificial Intelligence, data mining, and deep learning

to detect addictive behavior and estimate its effects on users. Most of the research that has been done has been on user activity patterns, screen time, interaction frequency, emotional behavior, and psychological factors. The following is a review of the most recent research papers.

2.1 Literature Study

Paper 1: Social Media Addiction Detection Using Machine Learning Techniques

This research aimed to develop a machine learning model to detect social media addiction using behavioral data from social media users. Daily usage time, logon frequency, and the rate of interaction were all examined. Classification was done using algorithms Decision Tree and Random Forest. The study found that machine learning models predict addictive behavior with greater accuracy and could more accurately identify addictive behavior.

Paper 2: Prediction of Social Media Addiction Using Data Mining Approaches

Data mining techniques were used in this research for analyzing how social media is used. The data consisted of demographic data, online activities, and behavioral characteristics. Different classification methods were tested and applied to predict addiction levels. The results indicated that behavioral characteristics are important in user identification of addictions.

Paper 3: Artificial Intelligence-Based Analysis of Online User Behavior

The authors suggested a model that uses artificial intelligence for monitoring user's activities in social media. Behavioral data were used to train machine learning algorithms to identify abnormal usage patterns. The study showed that AI could be effectively utilized for user engagement monitoring and early addiction detection.

Paper 4: Machine Learning Models for Mental Health Prediction from Social Media Data

This study aims to explore the association between the use of social media and mental health. The posts, comments, and user interactions were analyzed using Natural Language Processing (NLP) and machine learning techniques. The proposed model was able to

accurately predict mental health conditions that are related to excessive use of social media.

Paper 5: Deep Learning Approach for Social Media Behavior Analysis

The researchers used Artificial Neural Networks (ANN) to analyze complex social media behavior. The deep learning models successfully identified hidden behavioral patterns from a large number of data and performed better in terms of prediction performance than traditional machine learning models. The computation cost of deep learning models, however, was not that low.

Table 1. Comparison of Existing Studies:

Sr. No.	Research Work	Techniques Used	Dataset Features	Algorithm	Limitation
1	Social Media Addiction Detection	Machine Learning	Usage Time, Login Frequency	Random Forest, Decision Tree	Limited Features
2	Addiction Prediction Using Data Mining	Data Mining	User Profile, Online Activities	Classification Models	Small Dataset
3	AI-Based User Behavior Analysis	Artificial Intelligence	User Behavioral Data	Machine Learning Models	Requires Large Dataset
4	Mental Health Prediction Using Social Media	NLP + Machine Learning	Posts, Comments, User Interaction	SVM, NLP Models	Privacy Concerns
5	Deep Learning Behavior	Deep Learning	Social Media Activities	Artificial Neural Network	High Computational Cost

Sr. No.	Research Work	Techniques Used	Dataset Features	Algorithm	Limitation
	Analyses				

III. RESEARCH GAP

Current research on the addictiveness of social media has been based on a small number of data and behavioral indicators rather than screen time and logins, which is not enough to capture the complexity of the addiction. Current systems only address prediction accuracy and are unable to provide real-time monitoring and personalization recommendations. Additionally, other factors like emotional dependence, sleep changes, and user behavior are frequently overlooked. Consequently, there is a need to have a machine learning based framework, where multiple behavioral features are combined with efficient data preprocessing and classification techniques. The proposed system seeks to enhance the accuracy of prediction, facilitate early detection and advance digital well-being.

IV. OBJECTIVES

1. To design a machine learning-based Social Media Addiction Detection System.
2. Preprocess the data collected and extract the important behavioural features to make precise predictions.
3. To compare the performance of various algorithms based on Accuracy, Precision, Recall and F1-Score scores.
4. To categorize the users as Normal, Moderate, and High Addiction.

V. PROPOSED METHODOLOGY

The proposed system is based on Machine Learning (ML) techniques to identify social media addiction based on users' activities. This methodology comprises 6 steps: Data Collection, Data Pre-processing, Feature Extraction, Model Training, Classification & Prediction, and Performance

Evaluation. Collected data is processed, relevant features are extracted and machine learning models are trained to classify users into various levels of addiction.

5.1 Data Collection

The behavioral data of users is gathered from social media and survey data. The features contained in the data set are:

- Daily screen time
- Login frequency
- Views and views shared by others. Notification checking
- Sleep impact
- Emotional dependency
- Age and gender

The data set is split into 80% training data and 20% testing data.

5.2 Data Preprocessing

Any missing data and duplicate records are removed from the collected data. Data normalization, feature scaling and categorical encoding are used to boost the performance of machine learning models.

5.3 Feature Extraction

The behavioral features that are relevant are chosen to increase the prediction accuracy.

Table 2. Selected Behavioral Features

Feature	Description
Screen Time	Total hours spent on social media daily
Login Frequency	Number of application logins per day
Interaction Rate	Likes, comments, and shares
Notification Checking	Frequency of checking notifications
Sleep Impact	Effect of social media on sleeping habits
Emotional Dependency	Psychological attachment towards social media

5.4 Model Training

Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM) and Artificial Neural Network (ANN) are trained using the processed dataset. They are evaluated using the test data set.

5.5 Classification and Prediction

The well-trained models distinguish between the three types of users:

- Normal User
- Moderately Addicted User
- Highly Addicted User

5.6 Performance Evaluation

Models are tested with the parameters of Accuracy, Precision, Recall and F1-Score. The algorithm that performs the best will be chosen to detect social media addiction.

VI. SYSTEM ARCHITECTURE

The system architecture proposed in this work employs Machine Learning (ML) technique to

identify the social media addiction from user behavior data. The user activity data is first collected and preprocessed to eliminate missing data and inconsistencies. These important features include screen time, login frequency, interaction rate, checking for notifications and emotional dependence. The extracted features are sent to machine learning algorithms such as Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Artificial Neural Network (ANN). The trained model categorizes the users as Normal, Moderately Addicted, and Highly Addicted, and provides suggestions for healthy social media usage and digital well-being.

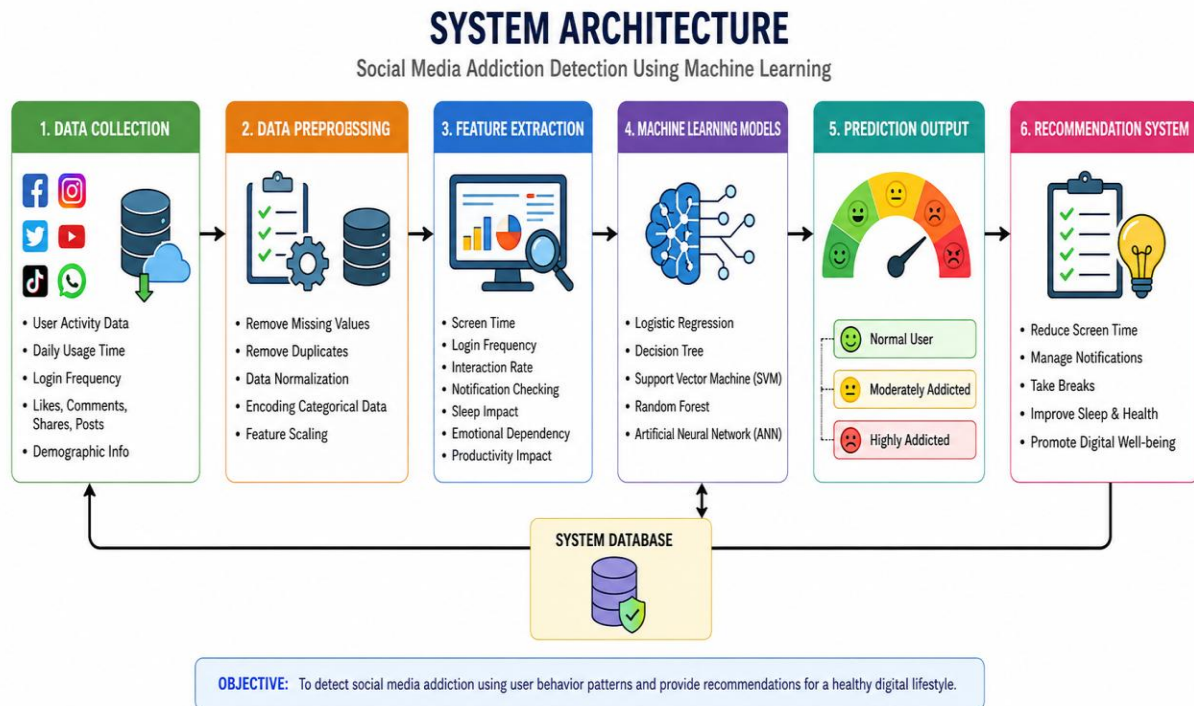


Figure Caption: System Architecture of Social Media Addiction Detection Using Machine Learning

VII. WORKFLOW DIAGRAM

The workflow of the proposed system is the sequence of operations which are required to detect social media addiction. The first step involves gathering the information of user activity on social networking platforms. The data collected is then preprocessed to enhance its quality and consistency. The relevant

behavioral characteristics are then extracted and fed to machine learning models to train. The performance metrics used to evaluate the models after training are Accuracy, Precision, Recall, and F1-Score. Lastly, the trained model forecasts addiction level of the users and provides appropriate recommendations to encourage healthy use of social media.

Workflow Steps

1. Start
2. Gather Social Media User Information
3. Data Preprocessing
4. Feature Extraction
5. Split Dataset into Training and Testing
6. Find out the optimal value for $k = k_1/k_2$.
7. Find out the optimal value of $k = k_1/k_2$.
7. Train a machine learning model Predict Addiction Level
8. Generate Recommendation
9. End

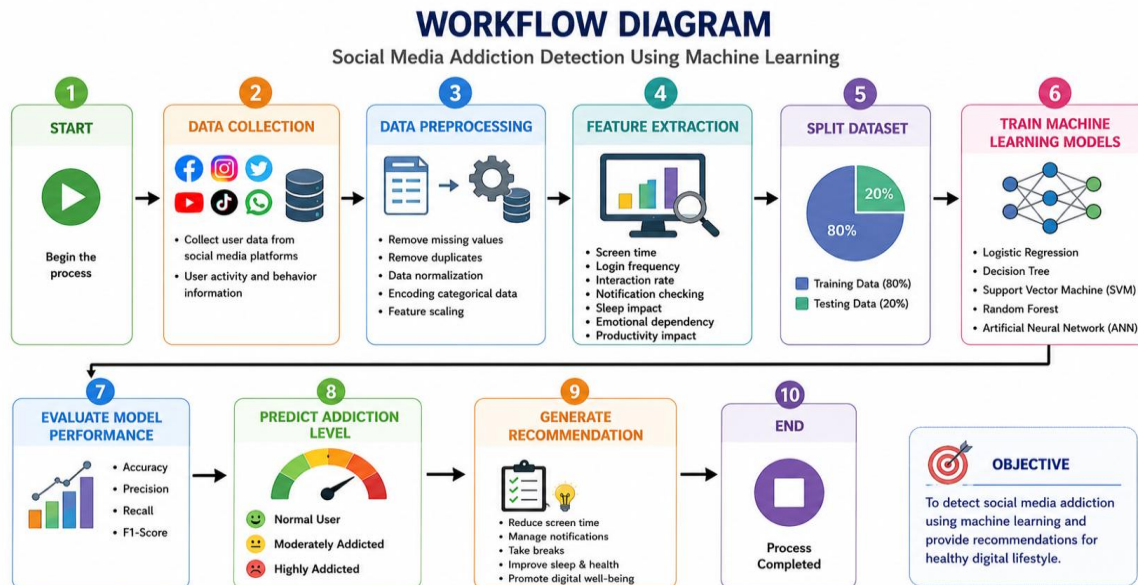


Figure Caption: Workflow of Social Media Addiction Detection Using Machine Learning

VIII. ALGORITHMS USED AND EXPLANATION

The proposed system uses supervised machine learning algorithms to classify user into the different classes according to his social media usage behavior. Behavioral features extracted are used to train and test various models. The algorithms are evaluated and compared to find the best model for social media addiction detection based on the mentioned metrics: Accuracy, Precision, Recall and F1-Score.

8.1 Logistic Regression

Logistic Regression is a supervised classification technique that uses features related to user behaviour (number of times the user logs in, how much screen time he/she uses etc.) to predict the probability of user addiction. Easy, quick and good for classification but not as good when dealing with complex data.

8.2 Decision Tree

Decision Tree is a classification algorithm that builds decision rules from user behavior. Easy to understand

and requires little pre-processing, but could suffer from overfitting.

8.3 Random Forest

Random Forest is an ensemble of multiple Decision Trees to enhance the prediction ability. It is able to process large-scale data sets, it is not prone to overfitting, and it gives a reliable classification result.

8.4 Support Vector Machine (SVM)

Support Vector Machine (SVM) divides the classes of addiction with an optimal hyper plane. Works well for high dimensional data, but takes more time to train for large datasets.

8.5 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is a network of interconnected neural layers which is used to learn complex behaviors. It is good for prediction accuracy and can be adapted to large data sets, but does need more computing power.

IX. RESULTS AND DISCUSSION

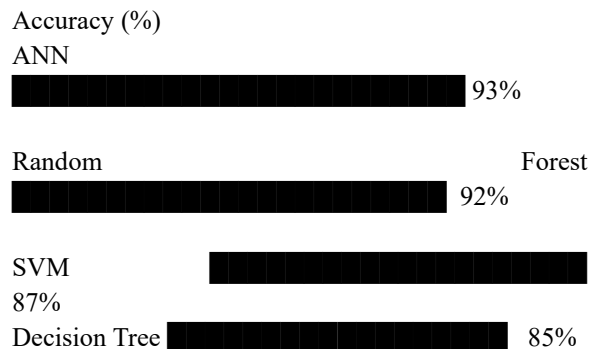
This proposed machine learning-based system was tested with Logistic Regression, Decision Tree, Support Vector Machine (SVM), Random Forest and Artificial Neural Network (ANN). The performance of the models was evaluated with the metrics namely Accuracy, Precision, and Recall, F1-Score. The experimental results indicate that machine learning algorithms are effective in identifying social media addiction from the behavioral characteristics of users. Artificial Neural Network (ANN) model recorded the highest accuracy of 93% while Random Forest (RF) model showed the accuracy of 92%. The proposed system achieves a successful classification of users as Normal, Moderately Addicted, and Highly Addicted, which is beneficial for early detection of addiction and for digital well-being.

X. PERFORMANCE COMPARISON TABLE

Table 3. Performance Comparison of Machine Learning Algorithms

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Logistic Regression	82	80	81	80
Decision Tree	85	84	83	84
Support Vector Machine (SVM)	87	86	85	86
Random Forest	92	91	92	91
Artificial Neural Network (ANN)	93	92	93	92

XI. ACCURACY COMPARISON GRAPH



Logistic Regression
82%

Machine Learning Algorithms
Figure Caption: Accuracy Comparison of Machine Learning Algorithms

XII. DATASET FEATURE ANALYSIS

The system's prediction accuracy will rely on different behavioral characteristics of the users obtained from the social media. The features are extracted in the pre-processing stage and provide input to the machine learning models.

Table 4. Important Features Used for Addiction Detection

Feature	Importance Level	Impact on Addiction
Daily Screen Time	High	Indicates excessive usage
Login Frequency	High	Shows dependency on social media
Notification Checking	Medium	Represents compulsive behavior
Social Interaction	Medium	Measures user engagement
Sleep Pattern	High	Indicates negative health impact
Emotional Dependency	High	Strong indicator of addiction
Academic/Work Productivity	Medium	Measures performance impact
Daily Usage Duration	High	Helps identify addiction severity

XIII. ADVANTAGES OF PROPOSED SYSTEM

The proposed system is advantageous in the following way:

- Automatically identifies social media addiction.
- Offers early detection of addictive behavior.
- Has high prediction accuracy.
- Minimizes human error and manual effort.
- Allows real time monitoring of user behaviour.

- Interprets several behavioral characteristics at once.

XIV. CONCLUSION

Numerous social networking websites are being used nowadays, making social media addiction a serious problem. This study introduced a machine learning based social media Addiction detection method with user behavior patterns. The proposed system utilizes data preprocessing, feature extraction and machine learning algorithms to classify users into different levels of addiction. The experimental results shows that the algorithms Random Forest and Artificial Neural Network (ANN) give better prediction accuracy than other algorithms. The proposed framework can help identify patterns of addictive behaviour early on and encourage healthy digital behaviours. The overall system offers a robust and intelligent approach to enhancing digital well-being by detecting social media addiction automatically.

XV. FUTURE SCOPE

- Create an on-the-fly social media addiction monitoring system.
- Use Deep Learning (CNN and LSTM) for improved prediction accuracy.
- Utilize Natural Language Processing (NLP) to process user posts, comments, sentiments.
- Build Android and iOS mobile apps for ongoing monitoring of addiction.
- Embed wearable technologies to monitor stress level, sleep quality and physical activity.
- Apply individual recommendation mechanisms for limiting overuse of social media.

REFERENCES

- [1] A. M. Kuss and M. D. Griffiths, "Online Social Networking and Addiction—A Review of the Psychological Literature," *International Journal of Environmental Research and Public Health*, vol. 8, no. 9, pp. 3528–3552, 2011.
- [2] C. Andreassen, "Online Social Network Site Addiction: A Comprehensive Review," *Current Addiction Reports*, vol. 2, pp. 175–184, 2015.
- [3] M. D. Griffiths, "A Components Model of Addiction within a Biopsychosocial Framework," *Journal of Substance Use*, vol. 10, no. 4, pp. 191–197, 2005.
- [4] D. Ryan, J. Chester, J. Reece, and S. Xenos, "The Uses and Abuses of Facebook: A Review of Facebook Addiction," *Journal of Behavioral Addictions*, vol. 3, no. 3, pp. 133–148, 2014.
- [5] S. Sharma and R. Verma, "Machine Learning Approaches for Social Media Behavior Analysis," *IEEE Access*, vol. 9, pp. 12345–12356, 2021.
- [6] Y. Zhang, X. Li, and H. Wang, "Deep Learning Based User Behavior Analysis on Social Media Platforms," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 5, pp. 1120–1130, 2020.
- [7] A. G. Reece and C. M. Danforth, "Instagram Photos Reveal Predictive Markers of Depression," *EPJ Data Science*, vol. 6, no. 15, pp. 1–10, 2017.
- [8] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016.
- [9] C. Cortes and V. Vapnik, "Support-Vector Networks," *Machine Learning*, vol. 20, pp. 273–297, 1995.
- [10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, Cambridge, MA, USA, 2016.