PREDICT THE USAGE OF WHATSAPP USING MACHINE LEARNING

Adithi N, Anusha H,
Anusha N and Athmika H R
Computer Science and Engineering,
K S School of Engineering and Management,
Bangalore, India

Prof. Meena G
Assistant Professor,
Computer Science and Engineering,
K S School of Engineering and Management,
Bangalore, India

Abstract— WhatsApp is one of the most popular messaging applications used to share texts, images, videos and audios. Many people getting addicted to the application because of its features. WhatsApp addiction may emerge as one of the most important behavioral addictions having large negative consequences, we cannot completely restrict ourselves but we can reduce using. This project aims to classify WhatsApp users as addicted, moderate and low users based on their name, gender, age, number of hours spent on the application, number of checks and send them mails also alert the addicted users by sending certain measures. WhatsApp usage data is collected from the users through google form survey. Then the data is preprocessed to make it suitable for Machine learning model. The classification models used are Decision Tree, K nearest Neighbor, Random Forest, Naïve Bayes, Logistic regression, Support Vector Machine. Among this classification models, the accuracy is found and the accurate model is Decision Tree. Prediction of next week's usage is also done by taking previous week's usage. The work expected results were obtained and the predictions was able to show the level of addiction of the various WhatsApp users and also predict next week's usage.

Keywords— WhatsApp, Machine learning, classification models .

I INTRODUCTION

The model classifies the WhatsApp users as addicted, moderate and low users and sends mail to the users based on their WhatsApp usage, age and gender. The prediction of next week's usage is done by analyzing the previous week's data. The process of gathering and measuring information that is data collection was done by each individual through google form survey by checking their screen time and times opened in the digital wellbeing. Digital wellbeing is an inbuilt application which helps users to track daily usage of different applications. The collected data was pre-processed to make it suitable for machine learning model. The pre-processed data was trained by using machine learning models. The report is sent to the respective users based on their WhatsApp usage using SMTP Library. It is user interface developed to reduce the WhatsApp addiction. It helps users to know their WhatsApp addiction level and if the user falls under the addicted category, they can follow certain measures to reduce WhatsApp usage. Prediction of WhatsApp usage of next week is done by using previous week's data. Users can get to know how much time they use WhatsApp.

II. LITERATURE SURVEY

Social media addiction is one of the biggest problems these days. Due to the outburst of various social media applications such as WhatsApp, Instagram, snapchat etc, people are getting addicted to social media. The present research is designed to help people to realize their addiction to social media and also to provide constructive results according to the inputs given. The application uses machine learning to predict the result based on past data taken from surveys[1]. The study uses statistical and data analytic tools in analysis. According to the study different genders and age demographics had significantly different usage habits in almost all message and group attributes[2].

III. HOW DOES THE SYSTEM WORK?

The survey is conducted for collecting data of WhatsApp users through Google form. The responses are stored in .csv file. Firstly, the data is pre-processed to find and remove null value columns. Then the data is divided for training and testing. 80% data is used for training and remaining 20% is kept for testing. The model is trained using various classification algorithms like logistic Regression, decision tree, random forest, support vector mechanism, k nearest neighbour and naïve bayes. The most accurate model is chosen based on the accuracy, sensitivity, precision and F score. The new data is fed into the model for testing from the GUI, the learned model predicts the output. Age, Gender, Time used, number of checks is fed into the model to predict whether the person is status of the usage i.e., addicted user, low users and normal users. if the user is addicted then the alert message is sent through the mail. In addition to this we are sending the report of usage in the form of pdf to the users

IV. SYSTEM OVERVIEW
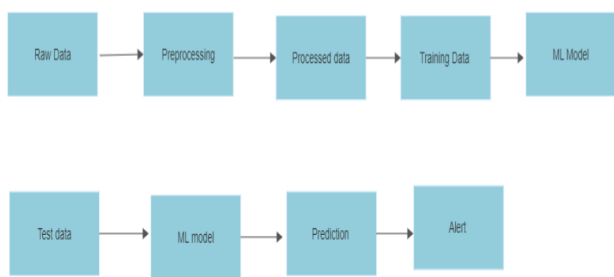
WhatsApp usage data is collected from the users through google form survey. Then the data is preprocessed to make it suitable for Machine learning model. The classification models used are Decision Tree, K nearest Neighbor, Random Forest, Naïve Bayes, Logistic regression, Support Vector Machine.

V. ACTIVITY DIAGRAM
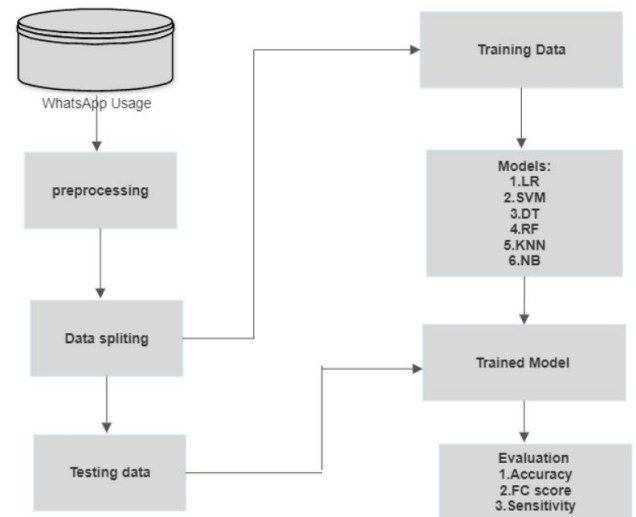
Among this classification models, the accuracy is found and the accurate model is Decision Tree. The considered factors are F1-Score, Precision, recall, Support. The alert mail is sent to addictive users based on their usage levels and it contains certain measures to reduce the WhatsApp Usage.

VI. COMPARISON WITH OTHER ALGORTIHMS

When deciding between algorithms, Logistic Regression, Support vector machine, Decision tree, Random Forest, K nearest neighbor and Naïve Bayes were implemented to choose the model which gives more accuracy .Among these algorithms, decision tree gave the highest accuracy of 97% for the dataset we have used. The most accurate model that is Decision tree is chosen based on the accuracy, sensitivity, precision and F score.

VII. FUTURE WORK

As enhancement to this project, the following can be implemented:

● New UI improvement offers can be done.
● Enhance the user experience of this application.
● Dataset size can be increased to get more accuracy.

CONCLUSION

The prediction of the WhatsApp usage is done by using the machine learning models(classification models) Decision Tree, K nearest Neighbor, Random Forest, Naïve Bayes, Logistic regression, Support Vector Machine. This work was able to discuss the WhatsApp usage based on name, age, gender, number of hours users use the application, number of times they check and classify them based on the usage levels that is addicted users, moderate and low users. The prediction of next week's WhatsApp usage is done by having the previous week's usage. Among this classification models, the accuracy is found and the accurate model is Decision Tree. At the end of the work expected results were obtained and the prediction was able to show the level of addiction of the various WhatsApp users and also predict next week's usage. On serious note this system has the ability to predict the WhatsApp usage.

REFERENCES

[1] Nandakishor Valakunde ,Srinath Ravikumar "Prediction of Addiction to Social Media" Department of Computer Engineering Vishvakarma Institute of Technology, Pune, India.

[2] Worawat Lawanont and Masahiro Inoue, "A Development of Classification Model for Smartphone Addiction Recognition System Based on Smartphone Usage Data"Shibaura Institute of Technology, Graduate School of Engineering and Science, Saitama, Japan.

[3] Avi Rosenfeld1 , Sigal Sina2 , David Sarne2 , Or Avidov2 , Sarit Kraus2,3 "WhatsApp Usage Patterns and Prediction Models" Institute for Advanced Computer Studies, University of Maryland

[4] Harshil Kanakia, Pooja Raundal,"Analysis of Social Networks using Naive Bayes" Sardar Patel Institute of Technology Mumbai, India, IEEE xplore 2019.

[5] Majid Altuwairiqi, "Problematic Attachment to Social Media, the Psychological States vs Usage Styles" 1 Bournemouth University, UK ,IEEE xplore 2019