THE TURKISH JOURNAL ON ADDICTIONS
www.addicta.com.tr

ORIGINAL ARTICLE

Pilot Study to Predict Smartphone Addiction Through Usage Pattern of Installed Android Applications and to Derive Correlations Between Addiction and Phone Usage Behavior

Swati Aggarwal*, Shubhi Gupta*, Simrat Pal Singh Satia*, Shivin Saluja, Varshika Gambhir

Netaji Subhas University of Technology, New Delhi, Delhi, India

ORCID IDs of the authors: S.A. 0000-0001-5986-6915; S.G. 0000-0003-0264-4718; S.P.S.S. 0000-0001-6986-6601; S.S. 0000-0002-7999-1227; V.G. 0000-0001-9960-5496.

Main Points

- The research has successfully built a machine-learning-based prediction model that can predict smartphone addiction with an accuracy of 75%.
- The study demonstrates a correlation between smartphone addiction and phone usage behavior, which includes screen-on duration, frequency of messages and calls, and duration of call, along with categorization of apps installed in the user's phone.
- The authors have categorized the various apps into five buckets based on the associated tags on Google Play Store, and used them as input features to predict smartphone addiction.
- Smartphone addiction is positively associated with the usage of apps belonging to the "social," "gaming," and "shopping/food and drinks" buckets which have a positive correlation of 0.208, 0.18, and 0.201.
- Buckets comprised of utility applications primarily consist of applications like Google Maps which negatively correlate to smartphone addiction, suggesting that work does not contribute to addiction.

Abstract

Mobile phone technology has been completely revolutionized in recent years. From being used for calls and text messages, the usage pattern has now shifted, and involves heavy applications (apps) based on platforms such as Android and iOS. Therefore, people have become exceedingly dependent on smartphones, with most of them suffering from a smartphone addiction disorder. In this study, the authors collect app usage patterns and categorize all apps into five buckets: social, entertainment, utility, gaming, and shopping/food and beverage. The data were collected through a self-developed app, named Activity Tracker, installed in the smartphones of individuals participating in the study. Activity Tracker also collects additional phone usage features such as call duration, number of texts/calls, and screen-on duration. The authors predict smartphone addiction using supervised machine-learning algorithms with good accuracy. This is a promising attempt to uncover multi-dimensional aspects of smartphone addiction by studying the usage patterns of installed Android applications. The study also includes gender-wise variations of smartphone usage patterns and explicit parallelism between usage patterns of addicted and non-addicted users.

Keywords: Android application, smartphone addiction, smartphone usage pattern, supervised machine-learning algorithms

*Equal and major contribution.

Corresponding Author: Swati Aggarwal
E-mail: swati1178@gmail.com

Received: July 13, 2021
Accepted: November 11, 2021

Introduction

Smartphones are an indispensable part of our everyday lives, and provide the functionalities of media player, digital camera, web browser, gaming, navigation, and high-speed internet access, all with a single touch. They have replaced cellphones and serve as a viable substitute for laptops, personal computers, and

Cite this article as: Aggarwal, S., Gupta, S., Satia, S. P. S., Saluja, S., & Gambhir, V. (2022). Pilot study to predict smartphone addiction through usage pattern of installed android applications and to derive correlations between addiction and phone usage behavior. *Addicta: The Turkish Journal on Addictions*, 9(1), 63-74.

many other devices. According to a study, the number of smartphone users in the world is likely to cross 3.8 billion people by 2021 (Takahashi, 2018). India is likely to have 829 million users by 2022 (ind, 2021). The significant increase can be attributed to user-friendly applications (apps) available on the regularly upgraded platforms such as Android and iOS.

There is a pertinent concern around the misuse of this privilege. In this study, the authors primarily focus on participants of Indian origin. An average Indian now uses around 1GB of data per day, compared to the earlier level of 4GB a month, a 650% increase in data consumption, with 50% of the time spent on chat, browsing, video streaming, social networking, and image apps (Nielsen, 2018). Smartphone use often has harmful effects on users, who neglect their well-being and fail to care for their surroundings. A study revealed that using smartphones while driving reduces driving performance and can increase the risk of an accident. Since smartphones are closely related to internet, they are significant perpetrators of internet addiction, accompanied by various psychological and physical problems. Excessive engagement with smartphones leads to poor academic performance, irregular sleep patterns, decreased satisfaction in life, anxiety, stress, and decreased overall mental well-being (Samaha & Hawi, 2016; Thomée et al., 2011).

Smartphones are now so ingrained in our lives that their excessive use is deemed as an addiction. The aspects of smartphone addiction have been proven to be very similar to that of drug and substance abuse (Lin et al., 2014). The evident growing use, overuse, and misuse of smartphones are significant topics of research and need to be explored thoroughly.

Literature Review

The growing use of smartphones has led to a greater degree of problematic usage, leading to smartphone addiction. In this section, the authors discuss the work done in the past on smartphone usage behavior and how it is related to addiction and problematic usage.

In 2005, Bianchi and Phillips predicted phone usage, considering extraversion, self-esteem, neuroticism, gender, and age as the potential predictors (Bianchi & Phillips, 2005). They derived a correlation between mobile phone usage and the psychological factors stated above. The Mobile Problem Usage Scale (MPPUS) was developed to detect problematic use of smartphones. This questionnaire primarily consists of current insights on behavior and technological addiction and information about the kind of relationship people have with their smartphones.

In 2018, Nahas and Hlais conducted a study focused on adolescents and university students to determine the extent of problematic smartphone use among adults aged 18 – 65 (Nahas et al., 2018). The study showed that chatting was the most used smartphone function and prevalent among younger and unmarried users, and those with mobile data subscriptions. In 2017, a representative sample of the population was taken to estimate the prevalence of problematic cell phone use in Spain (De-Sola et al., 2017). The users were categorized into four types – – casual, regular, at-risk, and problematic – – using age, gender, level of education, and daily cell phone use as potential indicators. The results, based on multiple criteria, show that such problematic use shares the features of recognized addictions, affecting not only adolescents but large segments of the population. Both the studies stated above use the MPPUS scale to estimate the extent of problematic cell phone usage.

In 2007, to address the claims that phone usage is addictive, a study was undertaken by Hooper and Zhou (2007) to categorize phone usage behavior based on the underlying motivation. Six categories were identified: addictive, compulsive, dependent, habitual, voluntary, and mandatory. This categorization process is helpful in understanding smartphone usage behavior and motivation. In order to know how people relate to these categories, a survey of 184 students was conducted. The study revealed that the behavior could not be conclusively categorized as any specific type. However, more robust support for mobile phone usage was categorized as dependent, voluntary, or mandatory behavior rather than as addictive, compulsive, or habitual. Furthermore, according to another study conducted on youngsters in Pakistan, most of them could use their cell phones within reasonable limits and did not tend toward extreme behaviors that lead to addictive cell phone usage (Ahmed & Perji, 2011).

In 2010, Oulasvirta et al. (2012) suggested that compulsive "checking behavior" comprises a significant portion of smartphone addiction. There are three types of reward value associated with its usage, namely, information, interaction, and awareness. Mads Bødker et al. (2009) used the theory of computational value to evaluate smartphone user experience and realized how the conditional, functional, social, emotional, and epistemic value of the device varies over time with each user.

Shin and Dey (2013) collected a wide range of smartphone usage data from smartphones and identified several features, namely the number of apps used per day, the ratio of SMSs to calls, number of event-initiated sessions, and the length of non-event-initiated sessions, which are useful for detecting problematic usage. They built detection models based on AdaBoost, with machine-learning algorithms to automatically detect problematic use. They inferred that the above features are useful for detecting problematic usage.

Lin et al. (2015) identified smartphone addiction using a mobile app. A novel empirical mode decomposition was used to observe the trend in smartphone use. The study was conducted for one month. Daily use duration and frequency were the parameters used to quantify excess usage and other features such as the relationship between the tolerance symptoms and the trend for the median duration of a use epoch. The study was further empowered by psychiatrists' assistance in determining self-reported phone usage time, which turned out to be significantly less than the time recorded through the app. The difference in usage time was positively correlated with actual smartphone use.

Smartphone addiction also has several negative impacts on youngsters' lives, such as sleep deprivation and attention deficit. Lee et al. developed the SmartLogger software to log a variety of application events such as power on/off, touch inputs, and phone events (call/SMS). They concluded that at-risk groups exhibit a highly skewed usage pattern and are more addicted to mobile instant messaging apps, followed by voice calls, SMS, and emails (Lee et al., 2014). Since the research was more oriented toward smartphone addiction, Ding et al. (2016) performed a correlation analysis between app usage features and app addiction scores to reveal that compulsive open times and usage time are good indicators of app addiction.

Through these studies, it is evident that usage of smartphones is becoming increasingly problematic. Past work has focused more on finding behavioral factors that influence smartphone usage: impulsiveness, self-esteem, self-monitoring, and loneliness. Since the assessments were self-reported, they were less reliable. A user's smartphone usage behavior can change over time. The assessment mechanisms from past work cannot be applied continuously as they are unlikely to promptly detect an individual's problematic use. This dynamic behavior of smartphone usage generates results that become inconsistent with time. Lastly, as smartphones continue to acquire various new apps with improved functionalities, there is a need to monitor the usage pattern of applications employed by users daily.

To address these shortcomings, the authors analyze smartphone usage by collecting a wide range of phone usage information from smartphones, using the Activity Tracker. This approach provides us with objective data regarding smartphone usage, which can further be used to predict problematic behavior.

Motivation

Owing to the staggering advancement in technology, mobile phones have become cheaper and are no longer used for the sole purpose of making phone calls or sending/receiving text messages. Their usage pattern has grown significantly different over time. Increased storage capacity, faster processing speed, and robust platforms such as Android/iOS provide convenient apps. These apps have come to redefine our smartphone usage pattern in recent years, from WhatsApp/Instagram (Facebook Inc., Menlo Park, CA, USA) – – used extensively for social communication – – to Netflix (Netflix, Los Gatos, CA, USA)/Amazon Prime (Amazon.com, Inc., Bellevue, WA, USA) – – used for entertainment purposes – – to ordering food through Uber Eats (Uber Technologies Inc., San Francisco, CA, USA)/Swiggy (Bundl Technologies Pvt. Ltd., Bangalore, India).

The app-driven phone usage has led users to develop a compulsive behavior that affects their school, work, and even personal relationships. Users now succumb to nomophobia (fear of being without a phone) (Bhattacharya et al., 2019) with their online compulsive behavior. Users can switch between apps with ease, and new updates/bugfixes enhance their user experience.

Previous studies that reported phone usage patterns used self-reported assessments (Elhai et al., 2016) or screen recording (Ferreira et al., 2014), which can be erroneous or inaccurate, and hence unreliable. There are no app-specific studies reported to date to predict smartphone addiction using machine-learning algorithms. In this pilot study, the authors were motivated to analyze problematic smartphone usage, the degree of usage of various apps, and their correlation with smartphone addiction. The authors developed an app named Activity Tracker to dynamically collect information regarding phone usage, thereby enhancing the collected data's reliability.

Moreover, it is the type of smartphone usage that should be taken into account while determining the degree of addiction. Considering screen time spent as the sole parameter is not enough, as the phone may be consumed passively (e.g., Google Maps) or for educational purposes (e.g., Unacademy) or content creation (e.g., photography/videography). Hence, categorization of usage based on the different kinds of applications can provide us with a wholesome picture.

The association of specific lifestyle apps with smartphone addiction has been studied (Noë et al., 2019). However, usage of other categories of applications, like shopping/eatables (Tang & Koh, 2017), apart from social functionalities, also contribute significantly to smartphone addiction.

To overcome the limitations of previous research in this field, the users were asked to install the aforementioned Android application (Activity Tracker) for a period of 7 days and monitor their smartphone usage activity in terms of number of messages, call duration, and on-screen time, and to fill in their self-esteem score (cite paper/questionnaire), and most notably, track the usage pattern of various Android applications they use. The novelty is to categorize these apps into specific buckets (as described in Table 1), and ask the users to fill the Smartphone Addiction Questionnaire (SAS-SV); using this data, we derive insightful correlations between the addiction and the smartphone usage pattern by employing various machine-learning algorithms. This also helps us derive meaningful comparisons between the usage of different apps and their contribution to smartphone addiction.

Interestingly, we came across research which determined that males are more likely to be addicted to the smartphone (Bisen & Deshpande, 2016), while at the same time, some advocated the

Table 1.
Description of Buckets for Each Android Application

S. No.	Bucket	Tags	Example	Label
1.	Social	Social/Communication	WhatsApp/Facebook/Instagram	Soc
2.	Entertainment	Entertainment/Music/Video Player	Netflix/Wynk/Saavn/Amazon Prime	Ent
3.	Utility	Maps/Navigation/Photography/Education/News	Google Docs/Google Maps/Inshorts/Duolingo	Uti
4.	Gaming	Gaming/Adventure/Sports/Strategy/Adventure	PokemonGo/PUBG/Clash Royal	Game
5.	Shopping/Food and Beverage	Shopping/Food and Beverage	Zomato/Myntra/Swiggy/Amazon	SFD

tendency of females to be more inclined toward the addiction (Tang & Koh, 2017). This also motivated us to draw a gender-based parallel on smartphone addiction.

Proposed Approach

In this study, the authors developed the Android app named Activity Tracker to study each user's smartphone usage pattern. The modus operandi is as below:

- **Step 1:** Install Activity Tracker in the user's smartphone.
- **Step 2:** Before commencing the process of data collection, the user is supposed to submit his/her demographic details such as name, age, and gender in the Activity Tracker itself.
- **Step 3:** In order to determine the ground truth regarding the user's smartphone addiction, the user also submits the Smartphone Addiction Questionnaire (SAS-SV), present in Activity Tracker 4.1.
- **Step 4:** The user also submits the Single-Item Self-Esteem Questionnaire, which is used as a feature in our machine-learning classification model to predict smartphone addiction.
- **Step 5:** Once the questionnaires have been successfully submitted, Activity Tracker runs in the background and sends information regarding the user's smartphone usage pattern to the back-end server, dynamically.
- **Step 6:** Data collected through Activity Tracker are processed to form the input features of the machine-learning classification model, as illustrated in Table 2.
- **Step 7:** The machine-learning model is trained based on input features.
- **Step 8:** The machine-learning model is used to predict likelihood if a person is addicted to the smartphone using machine-learning models.

This entire process of data collection lasted for seven days. It has been briefly summarized in Figure 1.

Standard of Questionnaires Used

Smartphone Addiction Scale – Short Version (SAS-SV)

Min Kwon et al. revised the Smartphone Addiction Scale (SAS) and developed a shorter version consisting of 10 items. Each answer's response could vary from "Strongly Agree" to "Strongly Disagree," with a score between 1 and 6 associated with each response (Kwon et al., 2013). The final result was obtained by adding the score corresponding to each question. The scores of participants involved in this study range from 10 to 53. Min Kwon et al. suggested a cut-off value of 31 for boys and a cut-off value of 33 for girls.

Single-Item Self-Esteem Scale

The single-item self-esteem scale is used to compute the self-esteem of a participant, by self-analysis and rating of their self-esteem on a Likert scale of 1 (very low) to 7 (very high) (Robins et al., 2001).

Classification Using Supervised Machine-Learning Algorithms

Supervised learning is a learning technique to teach the machine using data which are well labeled. We have input variables and an output variable and use an algorithm to learn the mapping function from the input to the output. The goal is to create an accurate map so that if we have new input data, we can correctly predict the output for that data. Supervised learning can be further divided into two categories of algorithms:

1. **Classification:** A classification problem is when the output variable is a category, such as red or blue; disease or no disease.
2. **Regression:** A regression problem is when the output variable is a real value, such as dollars or weight.

Classification is a technique to categorize our data into a desired and distinct number of classes, with a label assigned to each class. A classification model attempts to draw some conclusion from observed values and is used for predicting discrete responses. For example, when provided with a dataset about houses, a classification algorithm can try to predict whether the houses will sell for a higher or lower price than the recommended retail price. Here, the houses will be classified according to whether their prices fall into two discrete categories: above or below the said price. Classifiers can be of two types:

1. **Binary Classifiers:** A classification problem with only two distinct classes or with only two possible outcomes. For example: male and female; spam email and non-spam email.

Table 2.
Input Features for Smartphone Addiction Prediction Machine-Learning Classification Model

S. No.	Input Features	Description (Cumulative Sum)	Label
1.	Total screen-on duration	Sum of screen-on duration	SOD
2.	Total number of messages	Sum of number of messages sent/received	NOM
3.	Total number of calls	Sum of number of calls dialed/received	NOC
4.	Total duration of calls	Sum of the duration of calls dialed/received	DOC
5.	Social	Sum of usage duration of all apps in 'Soc' bucket	Soc
6.	Entertainment	Sum of usage duration of all applications in 'Ent' bucket	Ent
7.	Utility	Sum of usage duration of all apps in 'Uti' bucket	Uti
8.	Gaming	Sum of usage duration of all apps in 'Game' bucket	Game
9.	Shopping/food and beverage	Sum of usage duration of all apps in 'SFD' bucket	SFD
10.	Esteem score	Single integer ranging from 1 to 7 to indicate self-esteem of each participant	Esteem_score

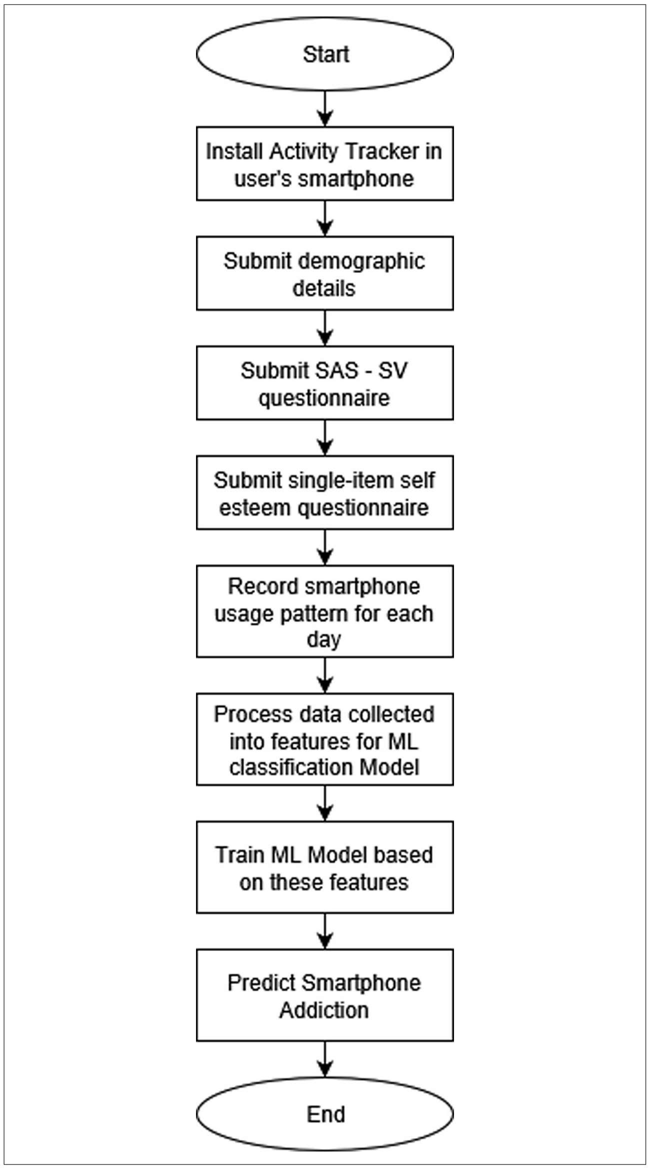

Figure 1. Proposed Approach.

2. **Multi-Class Classifiers:** A classification problem with more than two distinct classes. For example: classification of types of soil; classification of types of crops.

The different classification algorithms are:

Decision Tree

A decision tree is a supervised machine-learning algorithm that is used extensively for classification and regression problems. It has a tree-based structure with internal nodes representing the various features provided in the dataset, branches representing the decisions (yes/no), and finally, the leaf nodes denoting the desired outcome. This algorithm mimics the thinking of a human brain, and is therefore relatively easier to understand.

Decision Tree with AdaBoost

The primary difference between the decision tree and the decision tree with AdaBoost is that the weightage of certain decision trees is greater than the rest. Increased "say" or weightage is given to those that were able to perform best during the previous iterations, that is, with the least number of miscalculations. Hence, the distinguishing feature is its ability to learn from the previous iterations, so that the next one is built on the misclassification error of the last one.

Logistic Regression

Logistic regression is another supervised machine-learning algorithm which uses an equation (often sigmoid function) for its representation. $P(y = 1)$ is represented as a function of x by defining proper weights to the input values and modeling them into binary values of 0 or 1.

For example: $y = e\hat{}(b0 + b1*x)/(1 + e\hat{}(b0 + b1*x))$

where y is the desired output, $b0$ is the bias or intercept term, and $b1$ is the coefficient for the single input value (x).

Support Vector Machine

The support vector machine (SVM) is a representation of the training data as points in space separated into categories by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall. The SVMs are effective in high dimensional spaces and use a subset of training points in the decision function, and are thus also memory-efficient.

Naive Bayes Classifier

Naive Bayes is a probabilistic classifier inspired by the Bayes theorem, under a simple assumption that the attributes are conditionally independent. The classification is conducted by deriving the maximum posterior which is the maximal P(Ci|X) with the above assumption applying to the Bayes theorem. This assumption greatly reduces the computational cost by only counting the class distribution. Naive Bayes is a very simple algorithm to implement and good results have been obtained in most cases. It can be easily scalable to larger datasets since it takes linear time, rather than by expensive iterative approximation as used for many other types of classifiers. However, naive Bayes can suffer from the zero probability problem. When the conditional probability is zero for a particular attribute, it fails to give a valid prediction. This needs to be fixed explicitly using a Laplacian estimator.

k-Nearest Neighbor

The k-nearest neighbor (KNN) algorithm is based on the hypothesis that similar things exist in close proximity, thereby classifying the new incoming data closely with the existing stored data based on resemblance in the features. It is also called a lazy learner algorithm because rather than learning from the data immediately, it classifies the data into a specific category. This algorithm would take the KNNs based on the Euclidean distance, count the data points in each category, and then assign the new data point into the category with maximum neighbors.

k-fold Cross-Validation

k-fold cross-validation is a technique used to gauge the performance of a machine-learning algorithm by dividing the dataset into k subsets. $k-1$ subsets are used to train the dataset and the remaining set is used to check the performance of the trained model. Hence, we perform k such iterations, treating each fold as a validation set. This validation method has less bias and optimism in checking the accuracy of the model and is hence preferred over the train/test split validation technique.

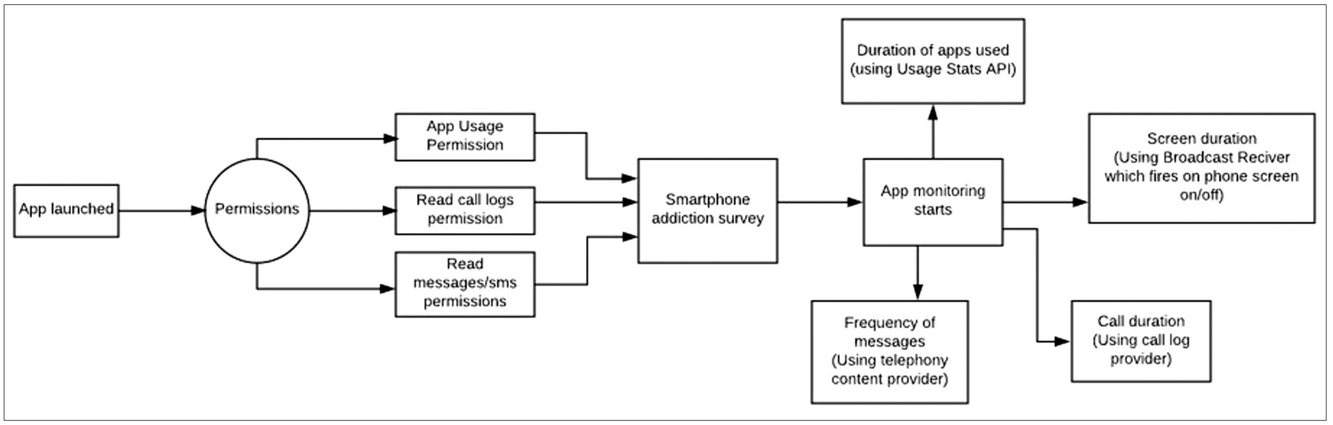

Figure 2. Working of the App on the Client Side.

Activity Tracker App Development

Activity Tracker was developed on Android Studio IDE using JAVA. It is compatible with 99% of Android devices. The working of Activity Tracker is depicted in Figure 2.

Step 1: Activity Tracker collects the data daily from each participant. The data collected through the app are as follows:

1. **Screen Duration:** The time (in milliseconds) for which the phone was used, the total screen-on duration was recorded for every participant each day. This was implemented using a *never-ending service*, running in the background. A *broadcast receiver* was used to trigger an intent when the participant switches the smartphone's screen-on/off. The interval between each on and off action was computed to give the total screen-on duration of each day, and the sum of screen-on duration for all the days then gave the total smartphone usage feature.
2. **App Usage:** The time (in milliseconds) for which individual apps installed in the smartphone were fetched for each day was determined using *UsageStats* API. The *UsageStatsManger* queries to fetch the usage history of all the apps of a fixed time interval daily.
3. **Call Logs:** The *CallLog* provider contains information about placed and received calls. This provider uses *content URI* to access call log entries. The call history is fetched daily for a week.
4. **Messages:** The *Telephony* provider contains all sent and received text messages in the SMS app. This provider uses content *URI* to access SMS/text messages. The frequency of the messages sent and received is fetched daily for a week.

Step 2: The data collected from each individual's smartphone is uploaded on the server every 24 hours using an *AsyncTask*. It allows performance of long-lasting tasks/background operations and shows the result on the user interface thread without affecting the main thread.

Step 3: The data sent from Activity Tracker are fetched as a *JSON* Object and then stored on the cloud as a *JSON* Array. The cloud service platform used is *Heroku* (Salesforce, San Francisco, CA, USA) which supports several languages, helping developers to build, run, and operate an app entirely in the cloud.

Step 4: The back-end of the Activity Tracker is supported by *MongoDB* (MongoDB Inc., New York City, USA) using *Atlas*. It is a fully-managed secure cloud database that handles all the complexity of deploying, managing, and healing the deployments on the cloud service provider. Figure 3 depicts the basic integration of the back-end process with Android smartphone and server.

Step 5: A *.json* file is extracted from the server which is further converted into *CSV* format. This *CSV* file is used as input to the data cleaning phase where the data are cleaned, as discussed in Section 6.2.

Dataset Creation

Raw Data

Ninety-six participants were involved in the study; 47 were males and 49 females, their mean age being 24.04. Activity Tracker was installed in the smartphones of all these participants. It collects data daily from each participant. Age and gender are recorded as the demographic details of the participants. Self-esteem scores and smartphone addiction scores are computed through questionnaires filled by the participants. The screen-on duration, number of messages, number of calls, duration of calls, list of apps, and usage duration of apps are recorded over a period of 7 days. The data collected through the app are as follows:

(1) **Age:** The age of the participant, in years.
(2) **Gender:** Whether the participant is a male or a female.

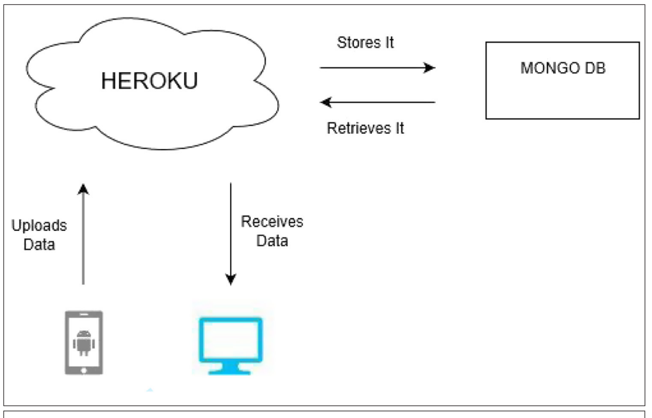

Figure 3. Storage and Retrieval of Smartphone Usage Features from the Cloud Service Platform.

(3) **Screen-on Duration**: The total amount of time spent on the smartphone each day with the screen turned on.
(4) **Number of Messages**: The total number of text messages sent/received by the participant each day.
(5) **Number of Calls**: The total number of calls made/received by the participant each day.
(6) **Duration of Calls**: The total duration of calls made/received by the participant each day.
(7) **List of Apps**: Day-wise list of apps that the participant used.
(8) **Usage Duration of Application**: Day-wise list of usage duration of each app that the participant used.
(9) **Esteem Score**: The self-esteem of a participant on a Likert scale of 1 – 7 (Robins et al., 2001).
(10) **Smartphone Addiction Questionnaire Score**: The resultant score that was computed through the Smartphone Addiction Scale – Short Version (SAS-SV) (Kwon et al., 2013).

Data Cleaning

The process of data cleaning was performed in Python language using libraries such as Pandas, NumPy, and Sklearn. Snapshots of raw data collected and processed data have been alluded as Figure 4 and Figure 5. The following steps were performed to clean the raw data retrieved from Activity Tracker:

Step 1: Clean Android app's name to remove redundant words such as com and Android. For example, "com.android.truecaller" renamed "Truecaller."

Step 2: Create a list of different apps that all participants used.

Step 3: Create a list of all participants present in the study.

Step 4: Record demographic details such as name, age, and gender of each participant.

Step 5: Record the result of Smartphone Addiction Questionnaire and Self-Esteem Questionnaire for each participant.

Step 6: Record phone usage pattern (screen-on duration, number of calls/texts, and duration of calls) for each participant.

Step 7: Record app usage duration of each app for each participant.

Step 8: Remove those apps with usage duration less than the defined threshold of 5 minutes a day. The assumption is that for any meaningful/purposeful interaction with any app, a user needs to spend more than 5 minutes a day on that app. Hence, a usage duration of fewer than 5 minutes should not contribute to smartphone addiction. Perhaps, the study can be further extended to perform a detailed analysis of this threshold value.

Step 9: Categorize the remaining apps into five buckets. The buckets are formed using tags associated with each app on Google Play Store, as per Table 1.

Step 10: Find the cumulative sum of the features over a period of 7 days.

Step 11: Normalize the values of all the features. The process of normalization was performed using the preprocessing.MinMaxScaler. fit_transform function present in Python to restrict all values within the range of 0 to 1 using the minimum and maximum values for a particular feature.

Step 12: Determine the output label, computed using Smartphone Addiction Questionnaire score and gender as follows:

- For males, the threshold value of the score was 31, above which they were classified as addicted.
- For females, the threshold value of the score was 33, above which they were classified as addicted (Kwon et al., 2013). The output labels are depicted in Table 3.

Step 13: The final dataset is thereby built. It is wholly annotated with output labels 0 and 1 and includes ten features to be used as input to the machine-learning classification models, as mentioned in Table 2.

Experiments and Analysis

Experiment 1: Checking the accuracy of various machine-learning-based classifiers to predict smartphone addiction.

The authors employed a *k*-fold cross-validation methodology to check machine-learning models' accuracy to predict smartphone

no_of_calls	no_of_messages	Gender	age	call_duration	ques_score	esteem_score	date	doa/0/_id	doa/0/ap	doa/0/du	doa/1/_id	doa/1/ap	doa/1/du	doa/2/_id	doa/2/ap	doa/2/du
42	3	F	21	4587	37	7	7-May-19	5cd80522	com.googl	277.302	5cd80522	com.truec	53.266	5cd80522	com.googl	17.996
6	0	F	21	434	37	7	8-May-19	5cd80523	com.googl	673.311	5cd80523	com.truec	17.346	5cd80523	com.googl	5.116
9	0	F	21	1763	37	7	9-May-19	5cd80524	com.googl	6.197	5cd80524	com.truec	3.025	5cd80524	com.googl	1.383
22	1	F	21	3553	37	7	10-May-19	5cd80525	com.truec	19.575	5cd80525	com.googl	0.803	5cd80525	com.what	5594.306
15	2	F	21	3806	37	7	11-May-19	5cd80526	com.googl	118.599	5cd80526	com.truec	10.707	5cd80526	com.googl	2.822
14	12	F	21	419	37	7	12-May-19	5cd81a81	com.googl	287.354	5cd81a81	com.googl	1.567	5cd81a81	org.telegra	1.409
8	13	F	21	138	37	7	13-May-19	5cd81a82	com.googl	3663.044	5cd81a82	com.googl	3.88	5cd81a82	com.what	14202
33	10	M	21	2145	44	6	7-May-19	5cd81a83	com.googl	104.049	5cd81a83	com.googl	0.628	5cd81a83	org.telegra	12.208
23	22	M	21	2197	44	6	8-May-19	5cd81a83	com.googl	1288.898	5cd81a83	com.googl	0.004	5cd81a83	org.telegra	34.445
5	20	M	21	1551	44	6	9-May-19	5cd81a84	com.googl	237.406	5cd81a84	com.what	11413.25	5cd81a84	com.miui.	14.377
10	12	M	22	245	44	6	10-May-19	5cd82974	com.googl	4342.89	5cd82974	com.tesla	0.18	5cd82974	com.tesla	472.073
20	18	M	22	1789	44	6	11-May-19	5cd82975	com.googl	10262.94	5cd82975	com.googl	319.185	5cd82975	com.sams	23.272
5	17	M	22	340	44	6	12-May-19	5cd82976	com.googl	4135.357	5cd82976	com.googl	381.778	5cd82976	com.sams	8.027
6	18	M	22	200	44	6	13-May-19	5cd82977	com.googl	3477.306	5cd82977	com.googl	139.192	5cd82977	com.tesla	0.223
4	21	F	22	450	41	7	7-May-19	5cd82978	com.googl	4471.9	5cd82978	com.googl	66.059	5cd82978	com.sams	12.333
10	4	F	22	11	41	7	8-May-19	5cd82cc7b	com.googl	14.533	5cd82cc7b	com.truec	137.008	5cd82cc7b	com.a0sof	50.322
10	8	F	22	144	41	7	9-May-19	5cd82cc8b	com.googl	3699.746	5cd82cc8b	com.truec	41.178	5cd82cc8b	com.onep	56.788
8	11	F	22	43	41	7	10-May-19	5cd82cc8b	com.truec	6.554	5cd82cc8b	com.quora	681.571	5cd82cc8b	com.what	6083.715
4	8	F	22	0	41	7	11-May-19	5cd82cc9b	com.googl	449.902	5cd82cc9b	com.truec	23.102	5cd82cc9b	com.quora	2159.265
0	8	F	22	0	41	7	12-May-19	5cd82ccab	com.quora	1178.511	5cd82ccab	com.what	8724.09	5cd82ccab	com.onep	1175.492
7	2	F	21	385	41	7	13-May-19	5cd830feb	com.googl	3205.12	5cd830feb	com.onep	223.153	5cd830feb	com.a0sof	39.307
12	3	M	21	310	37	6	7-May-19	5cd830feb	com.googl	2181.686	5cd830feb	com.googl	140.246	5cd830feb	com.starm	569.711

Figure 4. Data Collected.

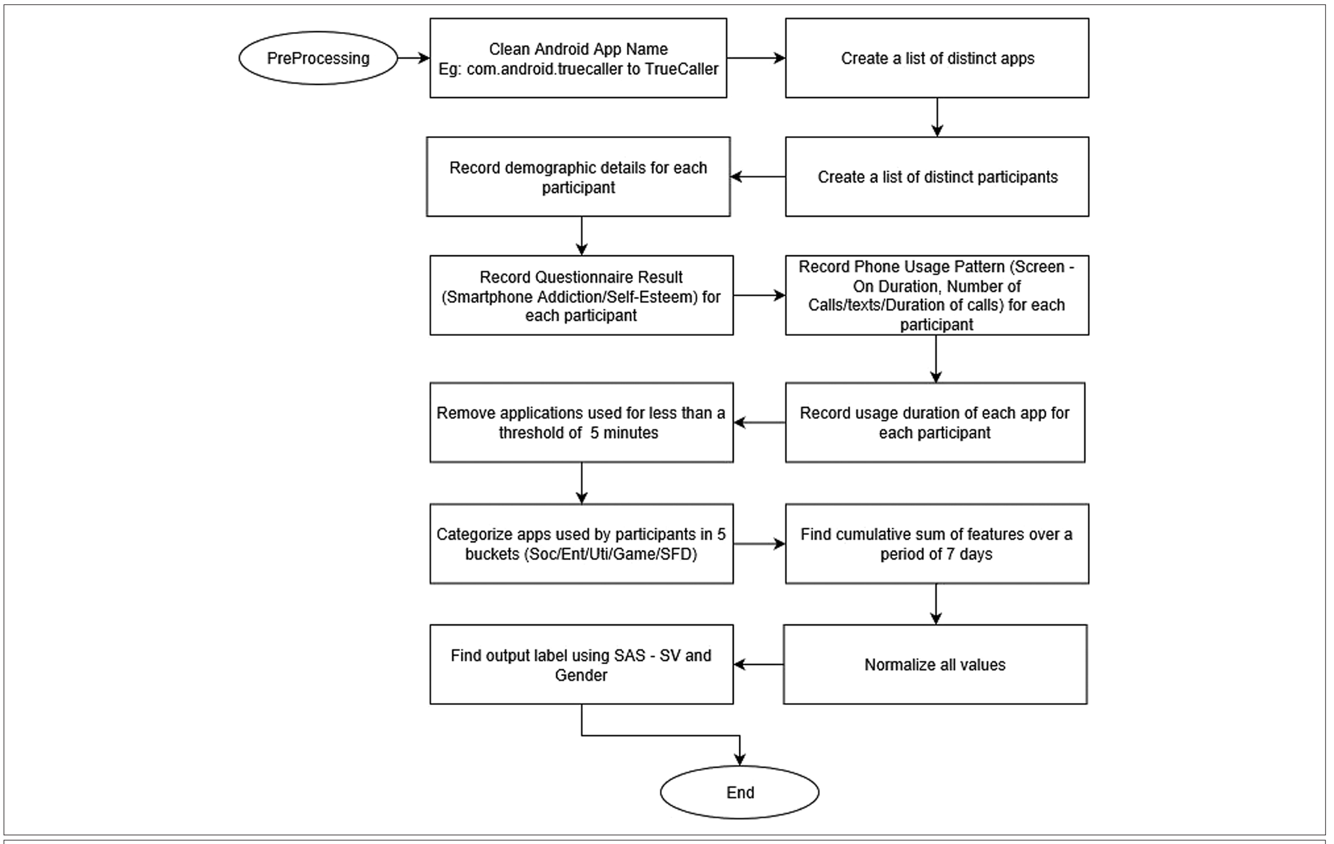

Figure 5. Data Pre-processing.

addiction. In this approach, the data set is randomly divided into *k* groups, or folds, of approximately equal size. The first fold is treated as a validation set, and the method fits on the remaining *k*– *1* folds. The machine-learning classification models employed are decision tree, decision tree with AdaBoost, logistic regression, KNNs, and naive Bayes. This task was performed on Python.

Conclusion: The data set was divided into four folds by choosing the value of *k* as 4. The authors were successfully able to predict smartphone addiction using various features with good accuracy of 75% of decision tree and decision tree with AdaBoost, as mentioned in Table 4, along with the accuracy of other classification models used. Similar models could henceforth be employed for these kinds of investigations related to the prediction of smartphone addiction.

Experiment 2: Establishing correlations between smartphone usage pattern and addiction using Pearson coefficient.

Also referred to as Pearson's *r*, the Pearson product-moment correlation coefficient or the bivariate correlation compares the linear correlation between two variables, *X* and *Y*. Its value ranges from +1.0 to -1.0. A positive correlation depicts that as the value of one parameter increases, so does the corresponding parameter's value. A negative correlation depicts that as the value of one parameter increases, the corresponding parameter value decreases. A zero correlation coefficient depicts no association between the parameters [Wikipedia, 2021].

Conclusion: Table 5 depicts Pearson's correlation between smartphone usage patterns used as input features and smartphone addiction. It is observed that as the total duration of screen time and number of messages increases, the likelihood of the person being addicted to a smartphone also increases. The maximum effect can be seen through the usage of apps present in the "social" bucket such as Facebook, WhatsApp, and Instagram, with a positive correlation of 0.208. Apps present in gaming and shopping/food and beverages also show a positive correlation. In contrast, apps used for a utility such as maps/navigation, document reader, or photography show a negative correlation to smartphone addiction as they are often used for productive utilization of time rather than contributing to addiction. The total

Table 3.
Smartphone Addiction Model Output Labels

S. No.	Result	Label
1.	Addicted	1
2.	Non-addicted	0

Table 4.
Accuracy of Various Machine-Learning Classification Models

S. No.	Model	Cross-Validation Accuracy
1.	Decision tree	75%
2.	Decision tree with AdaBoost	75%
3.	Logistic regression	64.58%
4.	k-Nearest neighbors	68.75%
5.	Naive Bayes	66.66%

Table 5.
Pearson Correlation Between Input Features and Smart Phone Addiction

S. No.	Feature	Smartphone Addiction
1.	Total screen-on duration	0.119
2.	Total number of messages	0.111
3.	Total number of calls	−0.040
4.	Total duration of calls	−0.106
5.	Social	0.208
6.	Entertainment	0.012
7.	Utility	−0.19
8.	Gaming	0.18
9.	Shopping/food and beverage	0.201
10.	Esteem score	0.169

number of calls and apps present in the "entertainment" bucket shows a negligible correlation to the addiction levels.

Experiment 3: To find gender-based variation between addicted and non-addicted participants.

A study was conducted on 96 participants, comprising an almost equal number of males and females. The trend of smartphone addiction in males and females was analyzed.

Conclusion: According to the study, 35 males in the group are addicted to smartphones while the remaining 12 are non-addicted. Twenty-six females are smartphone-addicted, while 23 females are not. Figure 6 shows the pictorial representation of the gender-wise distribution of smartphone addiction among the participants. The study suggests that males are more susceptible to smartphone addiction when compared to females.

Experiment 4: To determine the degree of usage of various smartphone usage parameters for addicted and non-addicted participants.

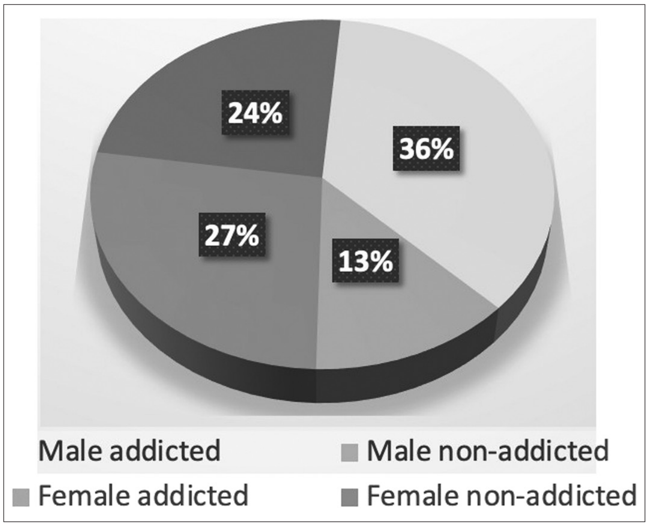

Figure 6. Gender-wise Distribution of Smartphone Addiction.

The various smartphone usage parameters for addicted and non-addicted participants are comparatively analyzed. The degree of usage of each parameter for addicted and non-addicted participants is shown in Figure 7. This gives a significant insight into the smartphone usage trend among addicted and non-addicted participants.

Conclusion: As shown in Figure 7, the screen-on usage parameter shows each participant's total interaction with the phone. Smartphone-addicted participants have a higher usage of social and entertainment applications. They are likely to send/receive fewer messages and place/receive fewer calls than non-addicted users. The usage of utility apps is higher in non-addicted users, suggesting that their smartphone usage for work does not contribute to addiction.

Experiment 5: To observe the trend of usage of various kinds of applications.

There is a staggering difference between the kinds of apps that users employ nowadays. Hence, the authors analyze the general trend of app usage and make comparisons based on usage duration. All participants' mean time on applications belonging to each bucket was computed and then normalized to restrict the values to a range of −1 to +1, using *preprocessing.MinMaxScaler().fit_transform*, present in Python 3.0.

Conclusion: Figure 8 analyzes the application usage pattern. It is observed that the participants spend their maximum time on social applications, followed by apps related to entertainment, and then by gaming. They spend their least time on apps related to shopping/food and beverage.

Discussion and Conclusion

Mobile phones, earlier restricted to mere communication, are now used for online shopping, gaming, social networking, and banking. All tasks that required manual operation or personal presence can now be accomplished using smartphones. However, there are major challenges in overcoming smartphone addiction. First, smartphones are widely and socially accepted due to their ease of access. Secondly, the user may begin using the smartphone as a one-stop solution to a number of functionalities such

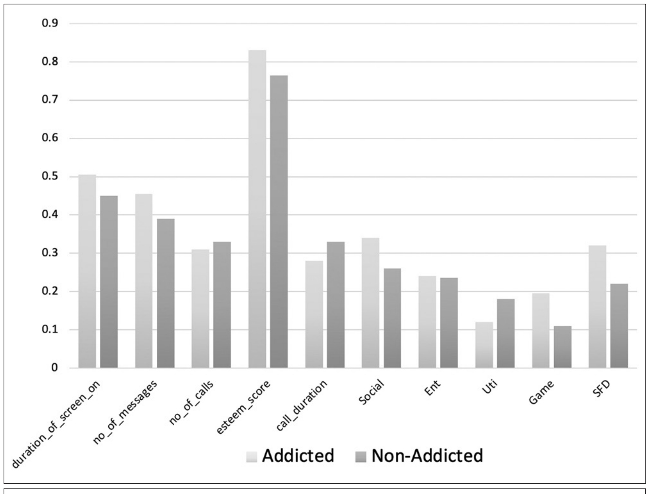

Figure 7. Comparison Between Normalized Degree of Usage and Types of Applications.

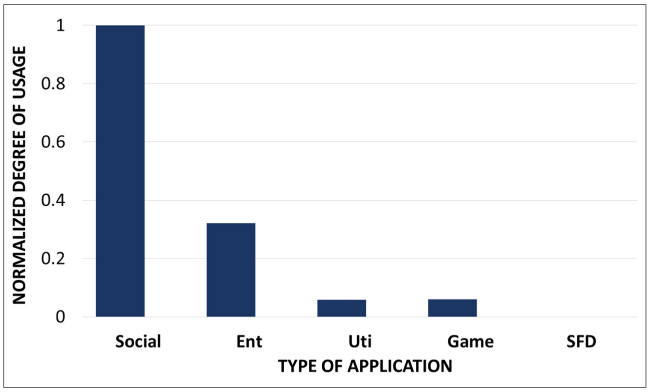

Figure 8. Comparison Between Various Buckets on the Basis of Degree of Usage.

as inbuilt watch, FitBit, GPS, camera, or calculator, but unregulated usage of other applications like social networking and gaming apps may cause the user to become addicted to smartphones. Thirdly, apps are designed to make users prolong their usage, with an infinite scrolling option that encourages users to continue online activity and attracts them to return using notifications or the daily rewards option. This prolonged usage leads to high levels of smartphone addiction and becomes challenging to control.

This research has successfully built a machine-learning-based prediction model that can predict smartphone addiction with an accuracy of 75% with the decision tree and decision tree with AdaBoost classification algorithms. The study demonstrates a correlation between smartphone addiction and phone usage behavior, which includes screen-on duration, frequency of messages and calls, and duration of call, along with categorization of apps installed in the user's phone. This sets the ground for researchers to determine smartphone addiction levels using the aforementioned model.

Unlike the previous studies, the authors have categorized the various apps into five buckets based on the associated tags on Google Play Store, and used them as input features to predict smartphone addiction. The authors inferred that smartphone addiction is positively associated with the usage of apps belonging to the "social," "gaming," and "food and shopping" buckets. These apps are the major contributors to predict smartphone addiction. The results proved that the "social," "gaming," and "food and shopping" buckets have a positive correlation of 0.208, 0.18, and 0.201. This supports the significant correlation of social media and internet gaming with smartphone addiction. However, the findings are not limited to any particular app, but rather a broad range of Android applications. This range of Android applications can be explored further to measure safe limits of application usage and enhance human – computer interaction while ensuring the mental well-being of users. In particular, buckets comprised of utility applications primarily consist of applications like Google Maps, Video Recorder, HealthifyMe, Unacademy, etc., which negatively correlate to smartphone addiction, suggesting that work does not contribute to addiction (Noë et al., 2019). Researchers can leverage the individuals' usage time for utility applications to measure effective time spent to enhance productivity and personal growth in comparison to other buckets which are positively correlated to smartphone addiction. The findings also suggest that unhealthy food intake and shopping addiction have physiological and behavioral disorders like depression and anxiety which lead to increased smartphone addiction (Tang & Koh, 2017).

Another interesting experiment was conducted to resolve the perplexing gender-based correlation between addicted and non-addicted participants. In our study, it was found that males are more prone to smartphone addiction than females.

This is in accordance with previous study conducted among 100 participants (50 males and 50 females), to measure gender differences in smartphone addiction (Bisen & Deshpande, 2016).

duration_	no_of_messages	no_of_calls	esteem_score	call_duration	Social	Ent	Uti	Game	SFD	Result
0.336069	0.057692308	0.303225806	1	0.401083319	0.219994779	0.010971098	0.057883073	0.24787237	0.016576549	1
0.813918	0.740384615	0.267741935	0.833333333	0.182916454	0.630602155	0.107762542	0.042676898	0	0.33705549	1
0.448994	0.826923077	0	1	0	0.193951475	0.363262262	0.022152346	0	0.120764643	1
0.585425	0.375	0.087096774	0.833333333	0.00561511	0.28181474	0.417975795	0.060603608	0	4.31E-05	1
0.771027	0.144230769	0.15483871	0.833333333	0.147410811	0.504566709	0.295383401	0.082652336	0	0.000581123	1
0.44543	0.769230769	0.329032258	0.666666667	0.275310533	0.283817559	0.084984916	0.11672976	0	0.530554812	1
0.672613	0.846153846	0.335483871	1	0.408059668	0.492039589	0.146639673	0.096381393	0	0.035658482	0
0.134293	0.028846154	0.225806452	0.833333333	0.698769213	0.093084017	0.053001505	0.019022346	0	0.028720218	0
0.677757	0.528846154	0.306451613	0.833333333	0.306533946	0.308291954	0.285785426	0.054130352	0.432145489	0.257819749	1
0.474553	0.826923077	0.474193548	1	0.295615677	0.168580135	0.173093051	0.150681616	0.476242488	0.027399223	1
0.500227	0.663461538	0.167741935	1	0.073903919	0.150106474	0.467963894	0.112133768	0	0.204271637	0
0.17558	0.471153846	0.329032258	0.833333333	0.224065566	0.120829716	0.061675694	0.027635691	0	0.046576997	1
0.540073	0.740384615	0.309677419	1	0.130877432	0.356574098	0.076080839	0.095778927	0.140277575	0.336821523	1
0	0.25	0.087096774	0.833333333	0.128750496	0	0.014803155	0	0	0.047468624	0
0.785909	0.346153846	0.393548387	0.666666667	0.654982701	0.316748762	0.625047139	0.150683139	0	0.04069063	1
0.311532	0.432692308	0.180645161	1	0.614514208	0.241974448	0.031207387	0.084081288	0	0.023180527	1
0.300996	0.269230769	0.522580645	0.833333333	0.457801599	0.029107365	0.427317212	0.045528174	0	0.032365985	1
0.287572	0.173076923	0.1	1	0.04154614	0.246388375	0.020292649	0.022555516	0	0	0
0.406626	0	0.174193548	0.5	0.537008678	0.208435473	0.156704559	0.051200526	0	0.82502367	1
0.478462	0.076923077	0.087096774	0.666666667	0.015030344	0.415701482	0	0.054320388	0	0	1
0.630668	0.903846154	0.55483871	0.666666667	0.53332199	0.439496193	0.03345847	0.36113868	0	0.03713351	0
0.689044	0.211538462	0.161290323	0.833333333	0.10620498	0.288810814	0.533202182	0.015008716	0.099392978	0.134121516	1

Figure 9. Dataset After Pre-processing.

The previous studies, which determined phone usage patterns solely through self-reported assessments, were prone to inaccuracies and hence unreliable. Therefore, in the current study, data were collected objectively with Activity Tracker, a self-developed app which dynamically records information regarding phone usage. The authors have successfully predicted smartphone addiction with good accuracy of 75%, using features collected through Activity Tracker. As there has been no app-specific study undertaken in the past, this study includes the effect of app usage patterns on smartphone addiction. The authors proposed categorizing the various apps into five buckets based on the associated tags on Google Play Store and used them as input features for the machine-learning classification model. The effects of apps on smartphone addiction were thereby investigated extensively. The study includes explicit parallels between addicted and non-addicted users based on normalized usage factors such as the number of calls/texts, duration of calls, and apps used.

The authors were successfully able to apply various machine-learning-based classifiers to predict smartphone addiction.

This pilot study also found that the maximum contribution to addiction is through the usage of apps present in the "social" bucket such as Facebook, WhatsApp, and Instagram, with a positive correlation of 0.208. Apps present in gaming and shopping/food and beverage also show a positive correlation. In contrast, apps used for a utility such as maps/navigation, document reader, or photography show a negative correlation to smartphone addiction. The total number of calls and apps present in the "entertainment" bucket shows a negligible correlation to the addiction levels.

Another interesting finding of the present pilot study is that males are much more addicted than females. However, this aspect needs to be studied with a more extensive database for generalizing this finding.

This research also attempted to study the degree of usage with various smartphone usage parameters for addicted and non-addicted participants. It was found that smartphone-addicted participants have higher usage of social and entertainment applications and negative correlation with the usage of utility apps. This study also correlated the trend of usage of various applications, and it is observed that the participants spend their maximum time on social applications, followed by apps related to entertainment, and then by gaming. Users tend to spend their least time on apps related to shopping/food and beverage.

This study promises to uncover multi-dimensional aspects of smartphone addiction by studying usage patterns of installed Android applications, and hence paves the path for further detailed investigations of smartphone addictions by analyzing installed app usage.

Future Scope

In this study, each application was classified in a limited number of buckets to predict smartphone addiction. This classification can be narrowed down into more buckets, resulting in more input features to the classification model, thereby providing a better understanding of the app usage pattern. Since this was a pilot study performed on a limited number of participants, further studies can involve a larger group of volunteers. Furthermore, the defined threshold of 5 minutes of usage per day of an app has not contributed to addiction level. Researchers can further investigate this threshold to determine a more appropriate value for app usage duration, beyond which it might contribute to addiction. Various other features can also be included to predict smartphone addiction, such as app-initiated events, the number of notifications from each app, and more detailed investigations.

Ethics Committee Approval: N/A.

Informed Consent: Electronic consent was taken from all the participants as a pre-requisite to take part in the study. The participation of all the participants in the study was kept anonymous.

Peer-review: Externally peer-reviewed.

Author Contributions: SA conceptualized and supervised the study, SG contributed in building the activity tracker, SPSS did data collection and both SG and SPSS were involved in detailed analysis, SS and VG were involved in data cleaning and coordinating with the participant in cases where they faced some issue. All the authors were involved in writing manuscript and final review of the manuscript. All the authors approved the final version of the manuscript.

Declaration of Interests: The authors have no conflicts of interest to declare.

Funding: The authors declared that this study has received no financial support.

References

Ahmed, I., & Perji K. A. (2011). Mobile phone to youngsters: Necessity or addiction. *Information Management and Business Review*, *2*(5), 229 – 238. [CrossRef]

Bhattacharya, S., Bashar, M. A., Srivastava, A., & Singh, A. (2019). Nomophobia: No mobile phone phobia. *Journal of Family Medicine and Primary Care*, *8*(4), 1297 – 1300. [CrossRef]

Bianchi, A., & Phillips, J. G. (2005). Psychological predictors of problem mobile phone use. *Cyberpsychology and Behavior* , *8*(1), 39 – 51. [CrossRef]

Bisen, S., & Deshpande, Y. (2016). An analytical study of smartphone addiction among engineering students: A gender differences. *International Journal of Indian Psychology*, *4*(1), 70 – 83.

Bødker, M., Gimpel, G., & Hedman, J. (2009). The user experience of smart phones: a consumption values approach. 8th global mobility roundtable, GMR.

De-Sola, J., Talledo, H., Rodríguez de Fonseca, F., & Rubio, G. (2017). Prevalence of problematic cell phone use in an adult population in Spain as assessed by the mobile phone problem use scale (mppus). *PLoS One*, *12*(8), e0181184. [CrossRef]

Ding, X., Xu, J., Chen, G., & Xu, C. (2016). Beyond smartphone overuse: Identifying addictive mobile apps. In *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems* (pp. 2821 – 2828). San Jose, CA, USA: CHI '16.

Elhai, J. D., Levine, J. C., Dvorak, R. D., & Hall, B. J. (2016). Fear of missing out, need for touch, anxiety and depression are related to problematic smartphone use. *Computers in Human Behavior*, *63*, 509 – 516. [CrossRef]

Ferreira, D., J., Kostakos, V., Barkhuus, L., & Dey, A. K. (2014). Contextual experience sampling of mobile application micro-usage. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services* (pp. 91 – 100). Toronto, Ontario, Canada: MobileHCI '14.

Hooper, V., & Zhou, Y. (2007). Addictive, dependent, compulsive? A study of mobile phone usage. *BLED 2007 Proceedings*, p. 38.

Bureau, E. (2020). India to have over 800 million smartphone users by 2022: Cisco study. *The Economic Times*.

Kwon, M., Kim, D. J., Cho, H., & Yang, S. (2013). The smartphone addiction scale: Development and validation of a short version for adolescents. *PLoS One*, *8*(12), e83558. [CrossRef]

Lee, U., Lee, J., Ko, M., Lee, C., Kim, Y., Yang, S., Yatani, K., Gweon, G., Chung, K.-M., & Song, J. (2014). Hooked on smartphones: An exploratory study on smartphone overuse among college students. In *Proceedings of the 32nd annual ACM conference on human factors in computing systems* (pp. 2327 – 2336). ACM.

Lin, Y. H., Chang, L. R., Lee, Y. H., Tseng, H. W., Kuo, T. B., & Chen, S. H. (2014). Development and validation of the smartphone addiction inventory (spai). *PLoS One*, *9*(6), e98312. [CrossRef]

Lin, Y. H., Lin, Y. C., Lee, Y. H., Lin, P. H., Lin, S. H., Chang, L. R., Tseng, H. W., Yen, L. Y., Yang, C. C., & Kuo, T. B. (2015). Time distortion associated with smartphone addiction: Identifying smartphone addiction via a mobile application (app). *Journal of Psychiatric Research*, *65*, 139 – 145. [CrossRef]

Nahas, M., Hlais, S., Saberian, C., & Antoun, J. (2018). Problematic smartphone use among Lebanese adults aged 18 – 65 years using mppus-10. *Computers in Human Behavior*, *87*, 348 – 353. [CrossRef]

Nielsen (2021). Average Indian smartphone user spends 4x time on online activities as compared to offline activities. *Nielsen*. Available from: https://www.nielsen.com/in/en/press-releases/2018/average-indian-smartphone-user-spends-4x-time-on-online-activities-as-compared-to-offline-activities/

Noë, B., Turner, L. D., Linden, D. E. J., Allen, S. M., Winkens, B., & Whitaker, R. M. (2019). Identifying indicators of smartphone addiction through user-app interaction. *Computers in Human Behavior*, *99*, 56 – 65. [CrossRef]

Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing*, *16*(1), 105 – 114. [CrossRef]

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, *27*(2), 151 – 161. [CrossRef]

Samaha, M., & Hawi, N. S. (2016). Relationships among smartphone addiction, stress, academic performance, and satisfaction with life. *Computers in Human Behavior*, *57*, 321 – 325. [CrossRef]

Shin, C., & Dey, A. K. (2013). Automatically detecting problematic use of smartphones. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing. Academic Medicine*, 335 – 344. Zurich, Switzerland: ACM.

Takahashi, D. (2018). Smartphone users will top 3 billion in 2018, hit 3.8 billion by 2021. *Venturebeat*. Available from: https://venturebeat.com/2018/09/11/newzoo-smartphone-users-will-top-3-billion-in-2018-hit-3-8-billion-by-2021/

Tang, C. S.-k., & Koh, Y. Y. W. (2017). Online social networking addiction among college students in Singapore: Comorbidity with behavioral addiction and affective disorder. *Asian Journal of Psychiatry*, *25*, 175 – 178. [CrossRef]

Thomée, S., Härenstam, A., & Hagberg, M. (2011). Mobile phone use and stress, sleep disturbances, and symptoms of depression among young adults: A prospective cohort study. *BMC Public Health*, *11*(1), 66. [CrossRef]

Wikipedia. (2021). Pearson correlation coefficient. *Wikipedia*.

A Appendix 1: Smartphone addiction survey in Android Application

The survey was a part of the Android application circulated among the participants. It collected each participant's demographic information and required them to fill in specific measures. The survey was composed of the following items:

A.1 Demographic Details

1. Age
2. Gender
3. Name

A.2 Questionnaire 1 (Smartphone Addiction Scale – Short Version)

1. Missing planned work due to smartphone use?
2. Having a hard time concentrating in class, while doing assignments, or while working due to smartphone use?
3. Feeling pain in the wrists or at the back of the neck while using a smartphone?
4. Won't be able to stand not having a smartphone?
5. Feeling impatient and fretful when I am not holding my smartphone?
6. Having my smartphone in my mind even when I am not using it?
7. I will never give up using my smartphone even when my daily life is already greatly affected by it?
8. Constantly checking my smartphone so as not to miss conversations between other people on Twitter or Facebook?
9. Using my smartphone longer than I had intended?
10. The people around me tell me that I use my smartphone too much?

A.3 Questionnaire 2 (Single-Item Self-Esteem Scale)

1. Do you have high self-esteem?

B Appendix 2: Data pre-processing and collection

The steps followed to pre-process data collected through Activity Tracker are shown in Figure 9.

The raw data are then pre-processed and normalized as shown in Figure 8. Normalization is a technique to change the values of numeric columns in the data set to use a standard scale, without distorting differences in the ranges of values or losing information.