INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 5, Issue 1)
Available online at: www.ijariit.com

Application usage behaviour prediction in Android OS using improved unsupervised learning algorithms

Praveen Thirumurugan
thirupraveen18@gmail.com
Sri Krishna College of Engineering and Technology, Coimbatore, Tamil Nadu

Keerthana Ravichandran
keerthiravi2198@gmail.com
Sri Krishna College of Engineering and Technology, Coimbatore, Tamil Nadu

Poojaa Krishnamoorthy
15eucs089@skcet.ac.in
Sri Krishna College of Engineering and Technology, Coimbatore, Tamil Nadu

ABSTRACT

The paper is about 'Application Usage Behaviour Prediction in Android OS using Improved Unsupervised Learning Algorithms'. This algorithm works with the input dataset collected from the usage of various applications by a user in an Android Device for a continuous period of time regularly. The algorithm which has the ability to adapt and tune itself to the user's behaviour and makes predictions which are in turn communicated to the user via notifications to improve his/her usage productivity.

Keywords— *Machine learning, Android, Tensor flow, Apps, Behavior, Prediction*

1. INTRODUCTION

Mobile phone apps are no longer just an application. It is just like a track which runs parallel along with us towards our destiny. The reign of mobile applications started and we get our work done in minutes of time. We live in a world where life without a smartphone is absolutely not possible. These smartphones allow us to install mobile applications which solve most of our life problems. These apps keep the user engaged in the different form of connectivity, with feeds, with the social network, instant messengers and this list is endless.

But with over adoption of smartphone culture, the number of people who are using smartphones gets increased every day. Hence, there are also various reasons and concerns which should be considered when the usage of smartphones get negative impacts in life. This might be due to unforeseen reasons and there must be an active tracking set up to alert the user in case the productivity drops using the phone the wrong way. Manual tracking is difficult since not all parameters and historical data is considered every time. Due to these observations, a Machine Learning algorithm based on unsupervised learning is needed to be developed. In this paper, we propose the algorithm which can be proof positive to eliminate the wrong usage of smartphones in a few days of observation of the user.

2. EXISTING SYSTEM

Applications are always developed keeping the people in mind and also, to adapt to the current environmental needs. This significantly delivers instant access to any type of information in any form. At this moment, there is an almost infinite number of mobile applications which eases out the user's time in just a matter of 'tap'. We have an app for everything, form apparel shopping to grocery delivery, booking movie tickets to booking travel tickets and this list goes on.

Various apps are available in the Google Play Store to track the usage of the applications in the device, but none of them gives predictions to the user or appears to be using Machine Learning algorithms to compute the calculations. They use the mere observations acquired and report them using data visualization.

3. INPUT DATASET PARAMETERS

The unsupervised learning algorithm proposed uses various data points to make the machine learning work and all the data is collected via the services setup while the installation takes place on the device. The various data points collected are,

3.1 Screen Time

The screen time of the applications used every day is collected from the Android OS using a call to the usage android library every 5 minutes automatically in the background.

3.2 Application Names

The names of the apps that are used are collected to classify the data from those apps. Both the App ID and the App Name is collected during this process. This data can be used to find the applications that are reported spam or helpful by the fellow users of the algorithm.

3.3 Application Category

Each application has a category associated with it like Photography, Productivity, Entertainment, etc. Normally a user's patterns look like using many categories of apps every day. This data gives a view of the user's taste of usage.

3.4 IBFC Rating

The IBFC Rating gives the age restrictions of every application and this will help us to detect the user's age and compare with the application's restrictions. For instance, while checking the apps against a child, usage of the app with IBFC rating of 16+ is not advisable. This can be communicated to the user to prevent abuse or mental problems.

3.5 Historical Observations & System Parameters

Various other parameters like Battery Status, Total Screen Time, System Uptime, etc, are also collected to match with various internal calculations like night time usage, most used time, etc. The historical observations from the output of our algorithm are also considered every time so that the prediction can be more accurate with more data and previous accuracy.

With these set of parameters, the machine learning algorithm based on unsupervised learning computes the prediction using various cluster creation.

4. ALGORITHM WORKING

4.1 Data collection from user phone

```
PROCEDURE dataCollectionBackgroundService()
        appList <- From System Data
        appMetaData <- appList.MetaData
        IF appList > 0
                START ASYNC TASK
                        appMetaData -> PUSH Device Data to Cloud
                END ASYNC TASK
        END IF
END PROCEDURE
```

4.2 Machine learning in Google Cloud Platform

```
PROCEDURE tensorProcessing()
        rawData <- PULL New Data from Cloud Storage
        processedData <- Observations Inferred from Past ML
        IF rawData IS NOT NULL
                ML_MODEL <- rawData + processedData + Tensor Flow
                processedData <- Observations in ML_MODEL
                processed data -> Google Cloud Storage
        END IF
END PROCEDURE

PROCEDURE ML_MODEL(data)
        IF data IS NOT NULL
                timeHabit <- data
                futurePredictions <- data
                anamolyPredictions <- data
                suggestionPredictions <- data
                categoryHabit <- data
                productivityLevel <- data
        End if
END PROCEDURE
```

4.3 Notifications in user device

```
PROCEDURE notifyUser()
        START ASYNC TASK
                observations <- PULL ML Data from Cloud Storage
        END ASYNC TASK
        LocalStorage <- STORE observations
        IF (APP) IS NOT OPENED
```



```
                PUBLISH SYSTEM NOTIFICATIONS
            END IF
            START DataCollection() ACTIVITY
        END PROCEDURE
```

There a sequence of steps involved before the prediction actually works. It starts with the collection of data points from the sources mentioned above and is classified accordingly based on how it can be fed into the system. This includes classification based on application usage, name and categories. These 3 are termed as separate clusters and will be uploaded to the cloud storage every day once.

The data is securely sent over an SSL/HTTPS channel so that the privacy of the user is ensured to be protected and no intruder can access that data anytime. The security of the cloud storage is also made tight using Google's proprietary algorithms and several tier security walls both physically and virtually.

The processed data is clustered and are set to the cloud storage like Google Cloud Storage Buckets and a trigger is invoked which will alert the cloud systems that some data has arrived and the computation needs to be performed in due course.

Then the unsupervised learning takes place using TensorFlow and using a high-performance TPU also known as Tensor Processing Unit on Google Cloud Platform. This service can be used to perform highly intense calculations specialized exclusively for generating Machine Learning Models.

Step 1: Processing the collected Data

Step 2: Uploading the Data to the Cloud Storage

Step 3: Machine Learning in Cloud using TensorFlow

Step 4: Prediction of Usage and Alerts in the Mobile

Fig. 1: Sequence of steps

5. CONCLUSION

The question about whether, and how, Machine Learning and data science can improve and modify usage decision making remains an empirical question. This algorithm will be useful in order to classify a user's productivity needs and improve his phone habits day by day. At a broader scale, the role and meaning of such tools in a policy and governance context requires further critical attention. This will be made more privacy enabled in the future so that our algorithm will not turn the other side of the proposed usage.

6. REFERENCES

[1] Self-Adaptive Machine Learning - http://www.academia.edu/Documents/in/Machine_Learning
[2] TensorFlow Usage Tutorials - https://www.tensorflow.org/tutorials/
[3] Google Cloud Tensor Processing Unit - https://cloud.google.com/tpu/

