Chapter 1

# **INTRODUCTION**

Mobile phone [3] is an electronic wireless portable scientific device, which is used to receive and send voice and messages. Mobile devices have become a necessity in today’s world as it helps us a lot in our daily life, and almost everyone has a mobile device or access to one. They have made the world accessible and communication convenient, all at the touch of a button. The advantage of having a mobile phone is that we communicate with others no matter where we are. Mobile phones have made communication easier with quick placing of calls and SMS. Mobile phones are of great help at the time of emergencies when we need to call the police, ambulance or other emergency services.

Information technology has improved the use of mobile phones to a great level. The advent of smart-phones has given mobile devices a new meaning. They are not only means of communication, but are also the sources of information and entertainment. They can also be depicted as sources for storing important data. Mobile phones are equipped with important features like documents, alarm, calendars, memo, stop-watch, etc. which help us become more organized.

With the help of internet, we are able to access various social media sites and apps that help us stay connected with our friends, family and the entire world. Various aspects of people’s lives can be represented by an application on a mobile device. Mobile apps have become an integral part of our lives and we rely on them in more than one way. These applications range from the meal plans to the scheduled appointments. These gadgets have also necessitated communication with friends, family and colleagues. It has even become possible to have access to your office on a mobile device. Mobile money transfer has been made possible by smart phones. Such uses of mobile phones are extremely helpful as they help us save our time and money.

App store [1] data analysis and predicting the various usage pattern of users from it will support the nascent app development industry, providing insights into the added value of features under consideration for new products and next releases. For this prediction, concepts of predictive analytics can be applied.

Predictive analytics [2] is the use of data, statistical algorithms and machine learning techniques to identify the likelihood of future outcomes based on historical data. The goal is to go beyond knowing what has happened to providing a best assessment of what will happen in the future. It uses many techniques from data mining, statistics, modeling, machine learning, and artificial intelligence to analyze current data to make predictions about future. It uses a number of data mining, predictive modeling and analytical techniques to bring together information technology and modeling business process to make prediction about the future. The app usage pattern of users are predicted based on their previous search patterns, routines, etc. Based on this prediction, apps that are suitable for them are predicted and suggested to them. So different algorithms of predictive analytics are used to predict the app usage pattern of users.

Personalised notifications can enhance the experience of using mobile apps. With personalised notifications, the main purpose is to suggest the apps as per the interest of the users.

### **PROBLEM STATEMENT**

To develop an intelligent mobile user classification system, the “ADAPTIVE FILTERED MINER” to enhance the experience of using mobile internet with the help of personalized notifications. Intelligent techniques are used to deal with the user profile and ranking information. Artificial Neural Networks (ANN) is used here for building the classification model. Labeling of the attributes is necessary to group similar user behavior together. The labeled information together with the corresponding input variables form the training data is used for establishing the classification model. The testing set is used to test the performance of the proposed intelligent mobile personalization model and to predict the content items correctly.

### **PROJECT OBJECTIVE**

The main objective of the system is to provide an automated platform to efficiently utilize the data collected from the app stores to predict the characteristic user behaviour. A further objective is to present a recommendation system that can accurately provide users with personalized notifications regarding other apps which can potentially satisfy the user requirements.

### **DESIGN AND IMPLEMENTATION CONSTRAINTS**

The primary design constraint is the mobile platform. Since the application is designated for mobile handsets, limited screen size and resolution will be a major design consideration. Other constraints such as limited memory and processing power are also worth considering. Availability of dataset to enhance the user behavioral study is very scarce thus causing a limitation.

Chapter 02

# **LITERATURE SURVEY**

**2.1. FRAMEWORK FOR PREDICTIVE ANALYTICS AS A SERVICE USING ENSEMBLE MODEL**

Dr. S. Vasavi [1] says that Cloud computing offers service delivery models that facilitate users during development, execution and deployment of workflows. In this Big-data era, Organizations require value out of big data. For this they need not have to deploy complex infrastructure, but can use services that provide value. As such there is a need for a flexible and scalable service called Predictive Analytics as a Service (PAaS).

Predictive analytics can forecast trends, determines statistical probabilities and to act upon fraud and security threats for big data applications such as business trading, fraud detection, crime investigation, banking, insurance, enterprise security, government, healthcare, e-commerce, and telecommunications Prediction algorithms can be supervised or unsupervised with different configurations, and the optimal one may be different for each kind of data.

This paper summarizes existing service frameworks for big data and proposes PAaS framework that can be used by business to deal with prediction in big data. This proposed framework is based upon ensemble model that uses best out of prediction algorithms such as Artificial Neural Networks (ANN), Auto Regression algorithm (ARX) and Gaussian process (GP).

**2.2. MOBILE CONTENT PERSONALISATION USING INTELLIGENT USER PROFILE APPROACH**

Worapat Paireekreng [2] says that the use of mobile devices for mobile internet is getting popular these day. However, there are still some limitations when using mobile internet on mobile phone. The limitations are small screen displays, limited input capabilities and information overload. To improve the experience of using mobile internet on mobile phone, one possible way is to implement some forms of personalization. However, there are still some limitations when using mobile internet on mobile phone. The limitations are small screen displays, limited input capabilities and information overload.

To improve the experience of using mobile internet on mobile phone, one possible way is to implement some forms of personalization. In this paper, she proposed the use of intelligent techniques to deal with the user profile and ranking information. The user data used for this paper contains user's ranking on the desire information as well as the demographic information. In order to categorize the users into appropriate behaviour classes, classification models need to be established. In this paper, first they introduced a clustering technique to cluster all the users from the set of collected data.

After the data have been clustered, labelling is performed to label the different clusters obtained. In this paper they used intelligent techniques to deal with the user profile and ranking information. The user data used for this paper contains user's ranking on the desire information as well as the demographic information. . Experiment results show that the proposed intelligent technique can provide reasonable results when used to predict up to 2 items of the desired content for the users.

**2.3. APP STORE MINING AND ANALYSIS: MSR FOR APP STORES**

Mark Harman [3] proposed that app stores provide a rich source of information about apps concerning their customer, business and technically focused attributes. Customer information is available concerning the ratings that were given to apps by the users who have downloaded them. This provides both qualitative and quantitative data about the customer perception of the apps. Business information is available concerning the downloads and price of apps. Technical information is also available in the descriptions of apps, but it is in free text format, so data mining is required to extract the technical details required for analysis. This approach was applied to the 32,108 non-zero priced apps available in the Blackberry app store in September 2011.

This approach to app store analysis consisted of the four phases. The first phase extracts raw data from the app store. In the second phase parse the raw data extracted in the first phase to retrieve all of the available attributes of each app relating to price, ratings and textual descriptions of the app itself. The third phase uses data mining to extract feature information from the textual descriptions and the final phase computes metrics concerning the technical, business and customer information extracted.

The feature extraction algorithm was used to extract different features and then these features were compared to find the common pattern among the users and hence common features were grouped using the greedy clustering algorithm. With this algorithm, redundant data were effectively handled.

**2.4. STUDYING SMARTPHONE USAGE: LESSONS FROM A FOUR-MONTH FIELD STUDY**

Ahmad Rahmati [4] conducted a study that took place in Pecan Park, an underserved community in Houston, Texas, where the average household income is below the poverty line. 14 teenage participants were recruited for the study. They were between 15 and 18 years old, either attending or had just finished high school. The participants had little or no prior experience with smartphones, providing a relatively clean slate for studying the usage evolution of smartphone-based services. All participants had PC-based Internet access at school and a good command of Internet knowledge. An experimental HTC Wizard phone was given to each participant.

The logging software reduces the standby battery lifetime of the smartphones from about five to three days. The participants were informed of the battery lifetime from the start. The logging software records the battery level, charging status, and display status (on/off) every minute. It also records visible Wi-Fi access points and their signal strength every 5 minutes. They employed the Wi-Fi traces to cluster the most visited access points into areas according to their proximity. Each cluster corresponds to a unique physical area, enabling us to study the location dependence of phone usage with minimal disclosure of location information.

**2.5. ROUTINE BASED ANALYSIS FOR USER CLASSIFICATION AND LOCATION PREDICTION**

Yibing Xiong [5] says that a person’s daily routine is a valuable piece of data, used to reveal user identity and state as well as to imply habits and interests. At present, there are two common ways to analyze daily routines. The first is classification, which focuses on a group of people and delineates the similarities and differences that they exhibit, in order to locate existing communities within the group or to recommend friends among group members.

The second is a prediction, which focuses instead on a single individual and analyzes his historical trajectories in order to predict his future routines. These predictions can then be applied to recommend attractions or even to assess the likely validity of alibis. In general, both means of analysis are committed to knowing more about the users and their requirements through their trajectories, so as to offer them personalized services. To address the problems, a Routine Based Classification (RBC) method was developed, which takes routine itself as the standard for routine analysis. Unlike probability methods, this partition based prediction method is more practical.

The whole process can be concluded as following steps: 1.Transform user routines into Nearest Relative Distance sequence; 2. Mine patterns in the training set; 3. Divide patterns into states and generate prediction rules; 4. Cover routines with states and predict the NRDs; 5. Predict the routines. Compared with other location-based methods of analysis, the main advantage of this method is its practicality. Using only limited and high-granularity routine information to successfully classify users and predict their locations.

Chapter 03

# **PROPOSAL FOR DISSERTATION**

Through Adaptive Filtered Miner, the aim is to develop and manage an automated system that can perform the functionalities of online recommendation. In the current scenario, a wide range of mobile applications are made available to mobile users to enhance the potential use of internet and to make the world more accessible. But the users are quite unaware of the opportunities ahead and are pinned to the conventional and clichéd applications. Prolonging to this accustomed strategy, the users would never experience the significance of mobile internet. This automated system aims at every user and advances them with information about the latest applications, with an intention to proffer the potential customers the knowledge about the recent and latest technological inventions that may serve to be beneficial for them.

Chapter 04

# **SYSTEM OVERVIEW AND REQUIREMENTS SPECIFICATION**

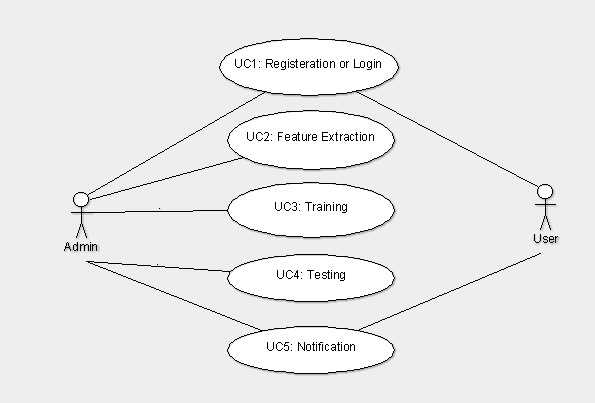
## **4.1. SYSTEM OVERVIEW**

The “Adaptive Filtered Miner” is an intelligent online recommendation system which is used to enhance the experience of using mobile internet. The system ensures that the platform is capable of generating personalized notifications that can be delivered to the user. Intelligent techniques are used to deal with the user profile and thereby to study the user behaviour. Artificial Neural Networks (ANN) are used here for building the classification model.

Labelling of the attributes is necessary to group similar user behaviour together. The labelled information together with the corresponding input variables form the training data and is used for establishing the classification model. The testing set is used to test the performance of the proposed intelligent mobile personalization model and to predict the content items correctly.

## 

## **4.2. FUNCTIONAL REQUIREMENTS**

  
**Figure 4.2.1:** Use case Diagram

USECASE 1: REGISTRATION AND LOGIN

* Use Case Description

The user has to access the dataset that has been collected form where the required features has to be selected. Also the user has to login using a forum to fill in the user related details.

* Flow of Event

The admin has to login in order to access the dataset for further proceedings. For user, he has to login in to a forum in order to provide with necessary details.

USECASE 2: FEATURE EXTRACTION

* Use Case Description

The required features must be selected and extracted manually from the dataset.

* Flow of Event

The features have to be extracted from the dataset manually. These features act as the input for the classifier.

USECASE 3: TRAINING

* Use Case Description

The features which has been extracted from the dataset is given has the input layers .These input layers are separated into different stages according to weight age .These input layers are used for training to get an output used for prediction.

* Flow of Event

The extracted features act as the input layers to the classifier .These inputs are used for training.

USECASE 4: TESTING

* Use Case Description

The user has to enter their details in a forum that has been provided.This details are compared with the app details in the admin panel inorder to provide the predicted output.

* Flow of Event

The details from the forum and admin are compared, along with the output obtained from the classifier.

USECASE 5: NOTIFICATION

* Use Case Description

The output from the classifier, the details from the admin panel and the users forum are compared together and if they meet a threshold, then notifications are send.

* Flow of Event

The output from classifier and other details are compared inorder to get a predicted output.

## **4.3. SOFTWARE AND HARDWARE REQUIREMENTS**

Software Specification

Front End : PHP

Platform : Android

Software : Android Studio

Back End : MySQL

Operating System : Windows

Hardware Specification

Processor : Pentum IV or above

Primary Memory : 1 GB RAM

Storage : 40 GB Hard Disk

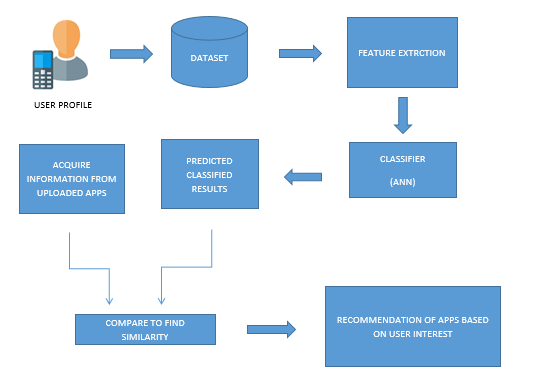
Display : VGA Color Monitor

Keyboard : Windows compatible

Mouse : Windows compatible

Chapter 05

# **SYSTEM ARCHITECTURE**



**Figure 5.1:** System Architecture of “Adaptive Filtered Miner”

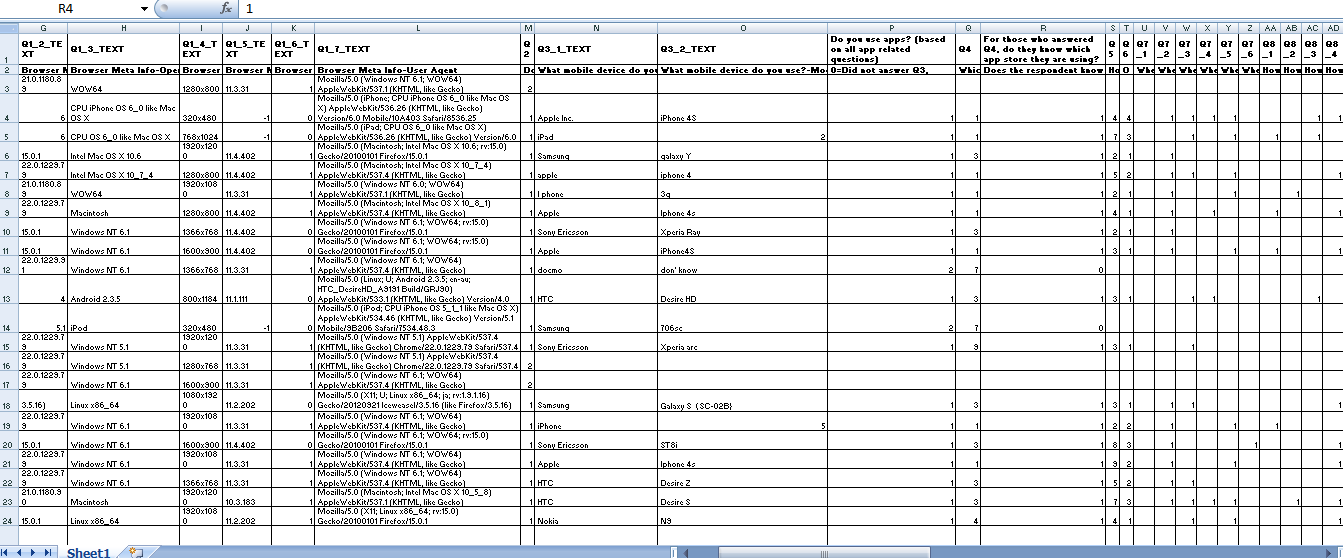
## **5.1. MOBILE USER PROFILE**

User profile is an important attribute in order to facilitate the content personalization process. It is difficult to predict the needs of the users as accurately as possible with raw user profile. Some form of categorizing the users based on their user profile could be useful. The attributes like age, gender, income and marital status have some impact in the ways the users use wireless device. Therefore it is suggested that with appropriate user profile information, personalization can be improved.

## **5.2. DATASET**

Dataset includes the behavioral features of users. Datasets from different locations are collected for analysis. The required dataset was collected from the site:

<https://dataverse.harvard.edu/dataverse.xhtml?alias=worldwide_app_user_behavior>



**Figure 5.2.1:** Dataset collected

## **5.3. FEATURE EXTRACTION**

In this step, the accurate features which can potentially perform content personalization are extracted manually. Feature patterns are identified and noise words are filtered out from the feature patterns.

## **5.4. CLASSIFIER**

The artificial neural network performs the operation of classification. An artificial neural network (ANN) is a computational model based on the structure and functions of biological neural networks. Information that flows through the network affects the structure of the ANN because a neural network changes - or learns, in a sense - based on that input and output. Initially random weights are assigned to all the linkages (input to hidden layer node). Using the inputs and the linkages activation rate of hidden nodes are recognized.

## **5.5. PREDICTED CLASSIFIED OUTPUT**

At the end of the entire procedure, content personalization is acquired that can enhance the experience of using mobile internet.

## **5.6. ACQUIRING INFORMATION FROM USER**

A forum is used to acquire information about the user which includes the user name, email id, age category he/she belongs to, gender etc.

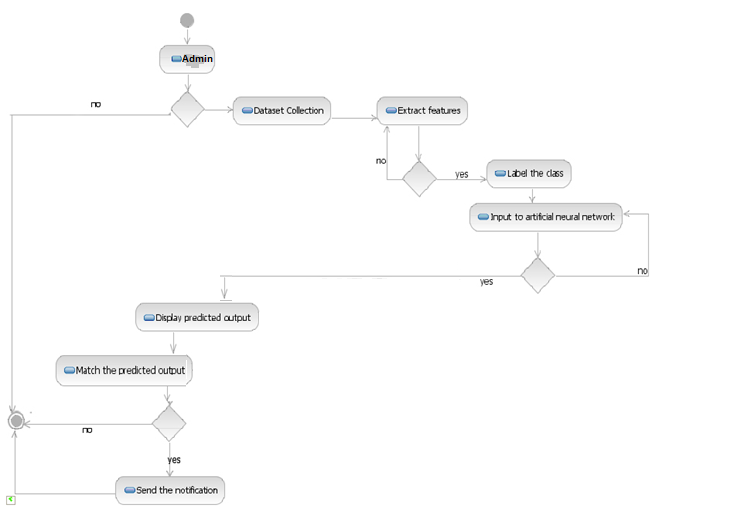
## **5.7. MATCHING THE OUTPUTS**

Later the information acquired from the user is matched with the predicted output to obtain the desired output and the notification is send to the relevant users accordingly

Chapter 06

# **DETAILED SYSTEM DESIGN**

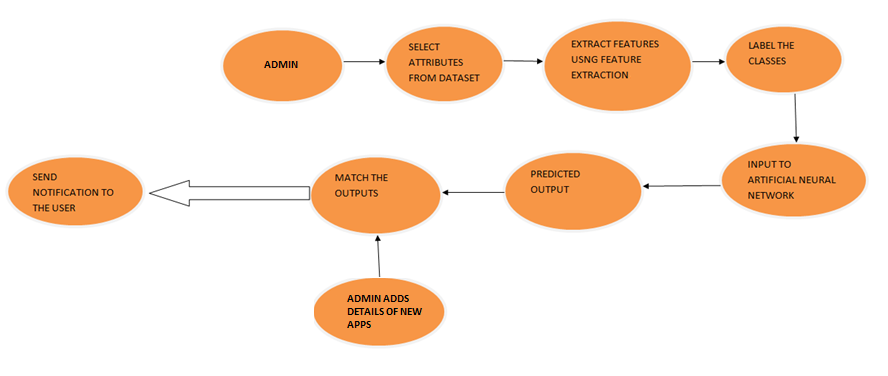
## **6.1. ACTIVITY DIAGRAM**

****

**Figure 6.1.1:** Activity Diagram

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. An activity diagram shows the overall flow of control. Arrows run from the start towards the end and represent the order in which activities happen.The activity diagram shows the flow of activities in each functions in the project. Here the administrator collects the dataset and this was used to acquire information about various users. Then the dataset is analysed and potential features are extracted from it manually. Classes are identified and are labelled accordingly. The labelled classes are then given as the input to the artificial neural network. The output of the artificial neural network is the predicted output. The administrator also adds the details of the new applications that come up in the application store. The predicted output and the details of the apps are then compared and if they match a notification is sent to the user.

## **6.2. WORKFLOW DIAGRAM**

****

**Figure 6.2.1:** Work flow diagram

A workflow diagram is a visual representation of work flow, usually done through a flowchart. It uses standardized symbols to describe the exact steps needed to complete a process. A workflow analysis can help in defining, standardizing and identifying critical areas of a process. The admin selects few important attributes from the dataset and extract features manually. Then the classes are labelled and is given as the input to the artificial neural network. The predicted result and the details of the new apps entered by the admin are analysed and if they match, the app will be recommended to the user.

## **6.3. SEQUENCE DIAGRAM**

****

**Figure 6.3.1:** Sequence Diagram

A sequence diagram is an interaction diagram that shows how objects operate with one another and in what order. It is a construct of a message sequence chart. A sequence diagram shows object interactions arranged in time sequence. Sequence diagrams are sometimes called event diagrams or event scenarios. It shows, as parallel vertical lines, different processes or objects that live simultaneously, and, as horizontal arrows, the messages exchanged between them in the order in which they occur. The user logs in. The admin extracts features, label the classes and give it as the input to the classifier which is the artificial neural network. The output obtained from the classifier and the features of new apps are compared and the result is sent to the admin. User fills the forum and the output is predicted.

## **MODULES OF THE SYSTEM**

* + 1. Feature Extraction from Dataset

The features which can potentially perform content personalization are extracted manually. Feature patterns are identified and noise words are filtered out from the feature patterns. A feature pattern consists of three parts: the phrases that signify the start of a feature list, the feature list itself and closing phrase that signifies the end of the feature list. From the feature list, the noise words, which are determined to be those from the English language are filtered out.

The filtering is done to make more accurate predictions. The features that were extracted were age, gender, marital status and income. The main features that form the input layers are age, gender, marital status, whether the user takes reviews of an app to download it and if the user downloads paid app or not .

* + 1. Training Classifier

The training classifier used here is an artificial neural network. Artificial Neural Network (ANN) is an efficient computing system whose central theme is borrowed from the analogy of biological neural networks. The artificial neural network performs the operation of classification. ANNs are also named as “artificial neural systems,” or “parallel distributed processing systems,” or “connectionist systems.” ANN acquires a large collection of units that are interconnected in some pattern to allow communication between the units. These units, also referred to as nodes or neurons, are simple processors which operate in parallel. Every neuron is connected with other neuron through a connection link. Each connection link is associated with a weight that has information about the input signal. This is the most useful information for neurons to solve a particular problem because the weight usually excites or inhibits the signal that is being communicated. Each neuron has an internal state, which is called an activation signal. Output signals, which are produced after combining the input signals and activation rule, may be sent to other units.

An artificial neural network (ANN) is considered to be a computational model based on the structure and functions of biological neural networks. Information that flows through the network affects the structure of the ANN because a neural network changes or learns, in a sense - based on that input and output. Initially random weights are assigned to all the linkages (input to hidden layer node).Using the inputs and the linkages, activation rate of hidden nodes are recognized.

The activation rate of hidden nodes and linkages to output, the activation rate of output nodes is computed. Then the error rate at the output node was found and recalibrate all the linkages between the hidden nodes and output node. This is done using the weights and error found at output nodes. The error is cascaded down to hidden nodes to recalibrate the weights between hidden node and the input node. This process is repeated till the conversion criterion is met. Different weights are assigned to the features. Here the features that were extracted from the dataset were given as the input to the artificial neural network. The maximum weight was given to age, gender and marital status. The output was obtained from the classifier.

* + 1. Admin Panel

The Admin Panel also known as the Back-end, Administrator application or Control Panel, is the interface where administrators and other site officials with appropriate privileges can manipulate the data. There are many tasks which can be done with the administrator interface. Here the admin panel was created to enter the details about several applications.

The details entered by the admin panel include name of the application, a brief description about the app, the url, features of the app, rating out of five and whether the app is paid or not. The user has to register by filling a forum. The fields of the forum are name, email id, gender, age, marital status, password, income and review.

* + 1. Recommendation or Notification

The last or final module is to send notifications or recommendations to the user. For this, the required inputs are the information entered by the user using the forum, the details entered by the admin about the apps and the output from the classifier obtained after training. This is obtained from the previous modules.

The basic idea behind this is behaviour matching. Here it is the synchronization of the features extracted and the details entered by the admin panel.

For sending notifications, recommendation algorithms are used. The use of efficient and accurate recommendation techniques is very important for a system that will provide good and useful recommendation to its individual users. Content-based filtering technique, Collaborative filtering technique and Hybrid filtering technique are the three different types. Collaborative filtering technique is used in this case.

Collaborative filtering is a domain-independent prediction technique for content that cannot easily and adequately be described by metadata such as movies and music. Collaborative filtering technique works by building a database of preferences for items by users. It then matches users with relevant interest and preferences by calculating similarities between their profiles to make recommendations. The learning algorithms used in model-based recommender systems are Association rule, Clustering, Decision tree, Artificial Neural network, Link analysis, Regression, Bayesian Classifiers.

So here Artificial Neural Networks are used for predicting the output. When a new application is uploaded in the application store, that is when the details of a new application is entered by the admin and if they match with the predicted data and the forum filled by the user, then a notification has to be send to the user as recommendation.

Chapter 07

# **IMPLEMENTATION**

Artificial neural networks systems are inspired by the biological neural networks similar to a human brain. They constitute different layers such as input layer, hidden layers and output layer. The inputs are given to the input layer and the hidden layers and the required output is obtained at the output layer. These neural connections have different weights that will be adjusted as the training of the inputs progresses. The weights might vary for each input parameter depending on the strength of the parameter .Once the training is done, testing has to be performed to apply the obtained weights for future use.

The system has been implemented using PHP and Android. The goal of the system ‘Adaptive Filtered Miner’ is to predict the type of mobile app that a user is likely to use. The user behavior is observed based on a dataset issued by the UCI. The required features has been extracted manually from the dataset. The features selected are age, gender, marital status, whether the user uses paid apps and if the user downloads an app on the basis of the review about the app.

The initial phase is the training phase. The concept of Artificial Neural Network is applied here. The extracted features form the input layer of the neural network. They are weighted according to the priority set to each feature. At the end of training, a threshold value that determines whether the user is likely to use that particular app or not was obtained.

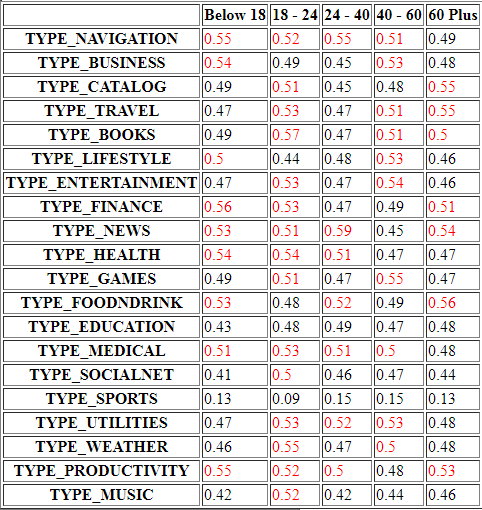
Based on the analysis a table is obtained showing the ratio of each features that were given as an input. It is based on this ratios the final threshold value is computed.

Tables 7.1, 7.2 and 7.3 that are given below are the output table that shows the ratio of the features of mobile apps by features that has been extracted manually from the database.

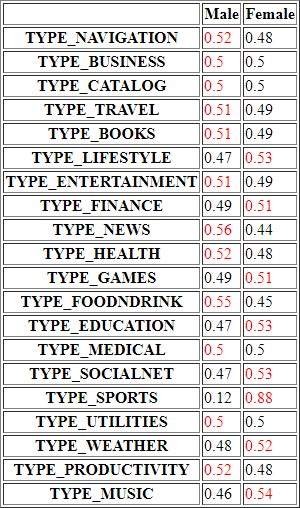
From trial and error methods a threshold value 0.5 has been set. The ratios above 0.5 shows that those group of users are likely to use the particular featured app.

The main features that form the input layers are age, gender, marital status, whether the user takes reviews of an app to download it and if the user downloads paid app or not .

**Table 7.1:** Classification of category based on age



**Table 7.2:** Classification of category based on gender.



**Table 7.3:** Classification of category based on marital status.

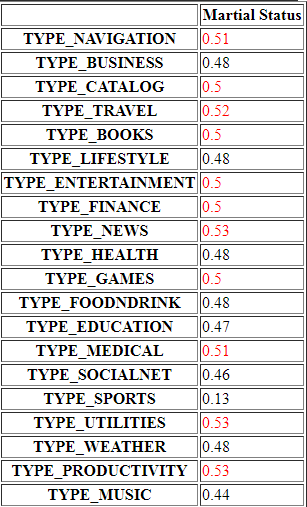
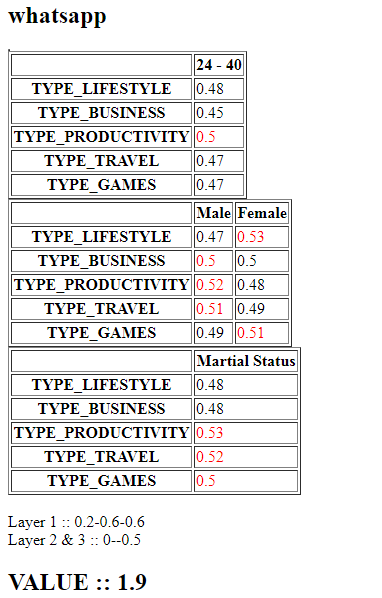
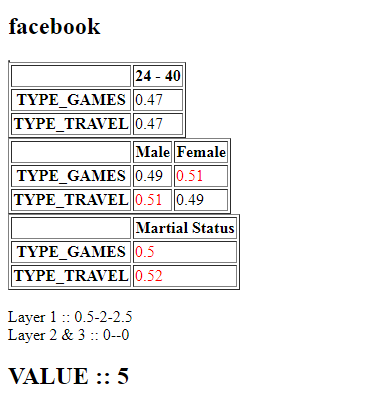


Figure 7.1 is the table for the app WhatsApp showing the features of the app and the threshold values of the users of the age group



**Figure 7.1:** WhatsApp

Figure 7.2 is the table for the app Facebook showing the features of the app and the threshold values of the users of the age group



**Figure 7.2:** Facebook

The features age, gender and marital status forms the layer 1 of the neural network. From figures 4 and 5, the ratios of these features can be taken. Layer 2 constitutes the features whether the user downloads an app based on the review and whether a user downloads paid app or not. After several trial and error method a threshold value 2 has been set.

The system has an admin panel that has several apps that has been added along with their app details. Once the training phase has been completed, the next phase is the testing phase. The front end of the system is a mobile app where the user can enter his/her details. This information is compared with the threshold value obtained. Our aim is to send a notification of the most likely apps that a user may use.

Once the result is obtained and if it goes above the threshold, then a notification will be sent to that particular user about the different types of apps they are likely to use.

Chapter 08

# **TESTING AND TEST RESULTS**

Software Testing is the evaluation of software against requirements gathered from users and system specifications. Testing is conducted at the phase level in software development life cycle or at module level in program code.

Software testing comprises of Validation and Verification, where validation is process of examining whether or not the software satisfies the user requirements. It is carried out at the end of the software development life cycle. If the software matches requirements for which it was made, it is validated.

Validation ensures the product under development is as per the user requirements and it emphasizes on user requirements. Verification is the process of confirming if the software is meeting the business requirements, and is developed adhering to the proper specifications and methodologies. Verification ensures the product being developed is according to design specifications and it concentrates on the esign and system specifications.

Testing Levels

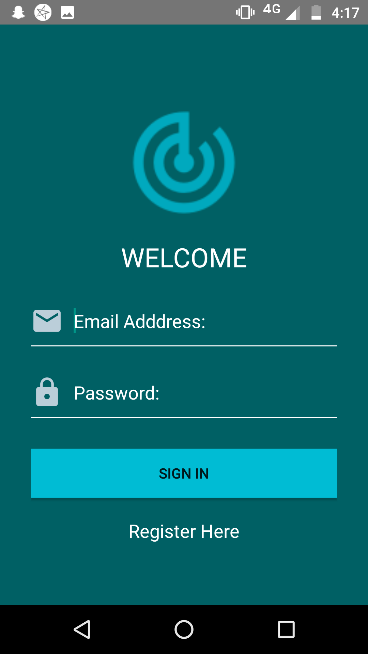
Testing itself may be defined at various levels of software development life cycle. The testing process runs parallel to software development. Before jumping on the next stage, a stage is tested, validated and verified. Testing separately is done just to make sure that there are no hidden bugs or issues left in the software.

Software is tested on various levels -Unit Testing, Integration Testing, System Testing, Acceptance Testing, and Regression Testing.

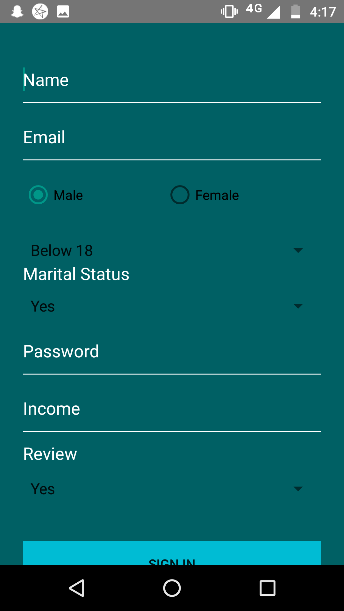
**Unit testing** is a level of software testing where individual units/ components of software are tested. The purpose is to validate that each unit of the software performs as designed.

Adaptive filtered miner utilizes the technique of unit testing which analysis the user inputs given to the system. A user interface is used to collect information about the users including their name, age category they belongs to, income and if they download apps based on review etc. Unit testing is applied here in order to validate the data entered by the users and then use it as testing data for the system.

The system runs successfully if the data is validated and verified properly, otherwise would not proffer the desired output. The testing data is then compared against the trained data in order deliver the predictions to the users.

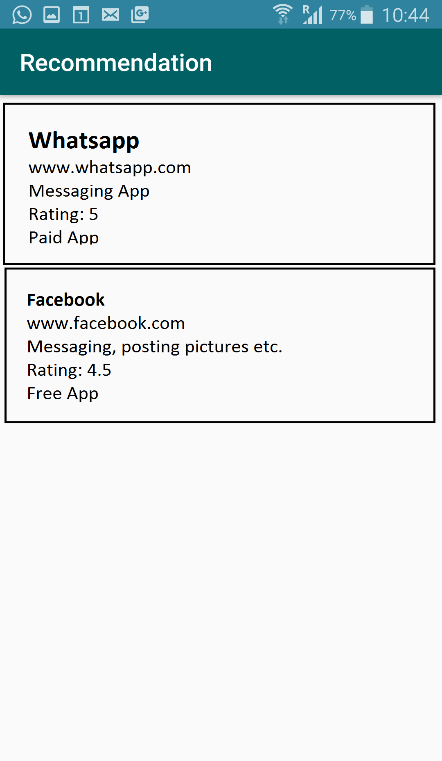
****

**Figure 8.1:** Login page of Adaptive Filtered Miner

****

**Figure 8.2:** Registeration form

The dataset consists of 10,000 records. The results were obtained and it was observed that the output was content personalized. So this can be converted into a product that will expand the use of several applications deployed in the application store. The test was conducted for different age groups ranging from 18-60 respectively. The system can also be expanded to a dataset with more entries.



**Figure 8.3:** Recommendation given to the user after succesful login

Integration testing allows individuals the opportunity to combine all of the units within a program and test them as a group. This testing level is designed to **find interface defects between the modules/functions**. This is particularly beneficial because it determines how efficiently the units are running together.

Adaptive filtered miner utilizes the technique of integration testing by combining the users that belongs to a particular age group and checks if the data required output has been obtained or not.

Functional testing is a software testing process used within software development in which software is tested to ensure that it conforms to all requirements. Functional testing is a way of checking software to ensure that it has all the required functionality that's specified within its functional requirements.

Adaptive filtered miner utilizes the technique of functional testing inorder to ensure that the obtained output conforms to all the functional requirements such as feature extraction, efficient training of data etc.

Chapter 09

# **CONCLUSION**

In this project, with the help of a dataset parameters of the users were classified. These parameters are used as an input for feature extraction process. They are fed as an input to the artificial neural network where they are trained.

Each stage undergoes proper testing. Then the trained algorithm is tested with another new input and the output, which the predicted values are obtained. With the help of the web crawler the information and details of apps deployed in the play store are extracted.

The pulled information and the predicted output are then matched to see if they satisfy the required conditions. If they satisfy then notification is send to the concerned user.

Chapter 10

# **FUTURE ENHANCEMENT**

With the help of Adaptive Filtered Miner, every user would be updated with the knowledge of the recent innovations in the field of mobile applications. It provides the user with online recommendations of new applications based on their miscellaneous features and the user behaviour. The system proves itself to be a reliable, fast and convenient system that can potentially predict the behaviour of its users and equip them with suggestions that can intensify the utilization of mobile internet.

This system diminishes the chances of users being oblivious about the contemporary applications convenient for them. While the system currently yields customized suggestions based on the age, gender, marital status and income of each user, an ability to pull information about various applications from an app store would result in developing a system which is more efficient in providing recommendations for the user in the near future.

# **REFERENCES**

1. Dr. S.Vasavi, “Framework for Predictive Analytics as a Service using Ensemble Model”, 7th IEEE International Advanced Computing Conference(IACC), 2017.
2. Worapat Paireekreng, Kok Wai Wong, “Mobile Content Personalisation using Intelligent User Profile Approach”, 3rd International Conference on Knowledge Discovery and Data Mining, IEEE 2010.
3. Mark Harman, Yue Jia, Yuanyuan, “App store mining and analysis: MSR for app stores”, IEEE 2012.
4. Ahmad Rahmati, Lin Zhong, “Studying Smartphone Usage: Lessons from a four-month field study”, IEEE transactions on Mobile Computing, Vol. 12, no. 7, July 2013.
5. Yibing Xiong, “Routine Based Analysis for User Classification and Location Prediction”, 9th International Conference on Ubiquitous Intelligence and Computing and 9th International Conference on Autonomic and Trusted Computing, 2012.
6. Hwaiyu Geng," Data Analytics and Predictive Analysis in the era of big data", Internet of Things and Data Analytics Handbook, 2017.
7. Xuanzhe Liu, Huoran li, Xuan lu, Tao xie, “Understanding diverse usage patterns from large-scale appstore-service profiles”, IEEE transactions on software engineering , Vol. xx, no. x, Feb 2017.
8. Philip s. yu, University of Illinois at Chicago, IEEE member, “On the feature discovery for app usage prediction in smartphones”, September 2013.
9. Danilo Hernane Spatti, Ivan Nunes da Silva, Luisa Helena Bartocci Liboni &quot; Artificial Neural Networks: A Practical Course&quot;, Springer International Publishing,2016
10. Vijay Kotu, Bala Deshpande ,&quot;Predictive Analytics and Data Mining: Concepts and Practice with RapidMiner &quot;,British Library Catalogue,2015
11. ABC-CLIO, Brill, CRC Press, De Gruyter, Elsevier, Sage, Taylor &amp; Francis, Wiley,&quot; Predictive Analytics and Algorithms&quot;,2012
12. Ronald .J.Brachman and Thomas G.D.&quot;Synthesis lectures on artificial intelligence and support vector machine&quot;,Series Editors,2008.
13. Hwaiyu Geng,&quot; Data Analytics and Predictive Analysis in the era of big data&quot;,Internet of Things and Data Analytics Handbook, 2017.
14. <https://www.worthofread.com/advantages-and-disadvantages-of-mobile-phones-for-students/>
15. <https://www.sas.com/en_us/insights/analytics/predictive-analytics.html>

# **APPENDICES**

**APPENDIX A – PROGRAM CODE FOR ANALYSIS**

<style>

.red

{

color:#F00;

}

</style>

<?php

//$app\_feature="TYPE\_LIFESTYLE,TYPE\_MEDICAL,TYPE\_PRODUCTIVITY";

$val2=0;

$age="24 - 40";

$gender="Male";

$martial\_status="Yes";

$age\_group=array($age);

$review="Yes";

$employ="No";

mysql\_connect("localhost","root","");

mysql\_select\_db("smart\_phone");

$node1=0;

$node21=0;

$node22=0;

$node3=0;

$i=0;

$flag=0;

$buy=0;

$rev=0;

$data = array();

$table="smart\_phone";

$threshold=.41;

mysql\_query("UPDATE application SET status='0' ");

$result22= mysql\_query("select count(\*) as cc from $table ");

while($row22 = mysql\_fetch\_array($result22))

{

$tot=$row22['cc'];

}

echo "Total $tot <br>";

//$fields=array("TYPE\_NAVIGATION","TYPE\_BUSINESS","TYPE\_CATALOG","TYPE\_TRAVEL","TYPE\_BOOKS","TYPE\_LIFESTYLE","TYPE\_ENTERTAINMENT","TYPE\_FINANCE","TYPE\_NEWS","TYPE\_HEALTH","TYPE\_GAMES","TYPE\_FOODNDRINK","TYPE\_EDUCATION","TYPE\_MEDICAL","TYPE\_SOCIALNET","TYPE\_SPORTS","TYPE\_UTILITIES","TYPE\_WEATHER","TYPE\_PRODUCTIVITY","TYPE\_MUSIC");

$result= mysql\_query("select \* from application ");

while($row = mysql\_fetch\_array($result))

{

echo "<h2>".$row['app\_name']."</h2>";

$appreview=$row['review'];

$paid=$row['paid'];

$y=0;

$result22= mysql\_query("select \* from app\_features where app\_id='$row[id]' ");

while($row22 = mysql\_fetch\_array($result22))

{

//echo "ff<br>";

//$fields=array("TYPE\_LIFESTYLE","TYPE\_BUSINESS","TYPE\_PRODUCTIVITY","TYPE\_TRAVEL","TYPE\_GAMES");

$fields[$y]=$row22['feature'];

$y++;

}

$array\_count=count($fields);

echo "<table border=1 >";

//$condition=array("Below 18","18 - 24","24 - 40", "40 - 60","60 Plus");

$condition=$age\_group;

echo "<table border=1 ><tr><th></th>";

foreach ($condition as $cond) {

echo "<th>$cond</th>";

}

echo "</tr>";

foreach ($fields as $field) {

//$condition=array("Below 18","18 - 24","24 - 40", "40 - 60","60 Plus");

//$condition=array("Below 18");

echo "<tr><th>$field</th>";

foreach ($condition as $cond) {

$result22= mysql\_query("select count(\*) as cc from $table where age='$cond' ");

while($row22 = mysql\_fetch\_array($result22))

{

$val=$row22['cc'];

}

$result22= mysql\_query("select count(\*) as cc from $table where age='$cond' and $field=1");

while($row22 = mysql\_fetch\_array($result22))

{

$ratio=round($row22['cc'] / $val,2);

if($ratio>=.5)

{

$node1++;

echo "<td class='red'> $ratio</td>";

}

else

echo "<td > $ratio</td>";

}

}

echo "</tr>";

}

echo "</table>";

echo "<table border=1 >";

echo "<tr><th></th><th>Male</th><th>Female</th></tr>";

foreach ($fields as $field) {

echo "<tr><th>$field</th>";

$result22= mysql\_query("select count(\*) as cc from $table where GENDER='Male' ");

while($row22 = mysql\_fetch\_array($result22))

{

$val=$row22['cc'];

}

$result22= mysql\_query("select count(\*) as cc from $table where GENDER='Male' and $field=1");

while($row22 = mysql\_fetch\_array($result22))

{

$ratio=round($row22['cc'] / $val,2);

$rat=abs(1-$ratio);

if($ratio>=.5)

{

$node21++;

echo "<td class='red'> $ratio</td><td > $rat</td>";

}

else

{

$node22++;

echo "<td > $ratio</td><td class='red'> $rat</td>";

}

}

echo "</tr>";

}

echo "</table>";

echo "<table border=1 >";

echo "<tr><th></th><th>Martial Status</th></tr>";

foreach ($fields as $field) {

echo "<tr><th>$field</th>";

$result22= mysql\_query("select count(\*) as cc from $table where MARITAL\_STATUS='$martial\_status' ");

while($row22 = mysql\_fetch\_array($result22))

{

$val=$row22['cc'];

}

$result22= mysql\_query("select count(\*) as cc from $table where MARITAL\_STATUS='Yes' and $field=1");

while($row22 = mysql\_fetch\_array($result22))

{

$ratio=round($row22['cc'] / $val,2);

$rat=abs(1-$ratio);

if($ratio>=.5)

{

$node3++;

echo "<td class='red'> $ratio</td>";

}

else

echo "<td > $ratio</td>";

}

echo "</tr>";

}

echo "</table>";

if($gender=="Male")

$node2=$node21;

else

$node2=$node22;

echo "<br> Layer 1 :: ".$node1/$array\_count."-".$node2/$array\_count."-".$node3/$array\_count;

if($employ=="Yes" && $paid=="paid")

$buy=.5;

else

$buy=0;

if($paid=="free")

$buy=.5;

if($review=="Yes" && $appreview>3)

$rev=.5;

echo "<br> Layer 2 & 3 :: ".$buy."--".$rev;

$val2=($node1/$array\_count)+($node2/$array\_count)+($node3/$array\_count);

$val2=$val2+$buy+$rev;

echo "<h2>VALUE :: $val2 </h2>";

if($val2>2)

{

mysql\_query("UPDATE application SET status='1' WHERE id='$row[id]'");

}

$buy=0;

$rev=0;

unset($fields);

}

$sql = "select \* from application where status='1'";

$result = mysql\_query( $sql) or die("Error in Selecting " . mysqli\_error($connection));

//create an array

$emparray = array();

// $emparray['status']['success']="Error";

// $emparray['status']['error'] ="no";

while($row =mysql\_fetch\_assoc($result))

{

$emparray['data'][] = $row;

}

$emparray['status']['success']="true";

//$emparray['status']['error'] ="";

echo json\_encode($emparray);

?>

**APPENDIX B – PRESENTATION SLIDES**

