

## Application of Classification technique to super market using mobile devices

Dr. R. R. Shelke<sup>1</sup>, Ms. Archana. B<sup>2</sup>, Ms S.A. Deshpande

<sup>1</sup> H.V.P.M. COET, Amravati,

<sup>2</sup> Reva University, Bangalore

<sup>3</sup> MIT, School of Engineering, Pune

<sup>1</sup>rajeshrisheshe@rediffmail.com, <sup>2</sup>archanab@reva.edu.in ,

<sup>3</sup>sonali.deshpande@mituniversity.edu.in

### **Abstract**

*The fast expansion and enhancements in information technology permit data to be accumulated sooner and in larger quantities. Due to vast new information resources, people like scientists, engineers, and businessmen need efficient systematic techniques to extract useful information and effectively valuable knowledge patterns. With the assistance of mobile data processing technique, user are going to be ready to visualise classification leads to mobile device any time anywhere. Classification of large dataset for mobile device represents a really favourable area for users and professionals that require to research data where users, resource and applications are mobile.*

**Keywords:** data mining, classification, supermarket dataset, mobile.

### **1. Introduction**

Classification could be a supervised learning technique. It maps the information into predefined groups. it's accustomed develop a model for classifying the population of records at large level. Classification algorithm needs that the classes be distinct supported the information attribute value. It defines these classes per the characteristics of the information that's already known and belong to the classes. Classification consists of two steps, training step and therefore the testing step. within the training step, the model is built. And within the testing step classifiers are tested to indicate the accuracy of the classifiers. Various kinds of classification algorithm are Support Vector Machine (SVM), Neural Network, Decision tree, Naive Bayesian.

Data mining techniques using mobile devices, especially Smartphone is an emerging research area. Mobile devices have recently expanded in several communities like governmental agencies, enterprises, social work providers (e.g., insurance, Police, fire departments), healthcare, education, and engineering organizations. However, despite of significant improvement in mobile computing capabilities, still computing requirements of mobile users, especially marketing users, isn't achieved. To realize results of knowledge mining on mobile device, one can perform data mining techniques on separate machine and visualise results on mobile device.

### **2. Literature Review**

Most of the favoured data mining algorithms are designed to figure for centralized data and that they often don't concentrate to the resource constraints of distributed and mobile environments [1-2]. In support of the third generation of knowledge mining systems on distributed and large data, an efficient distributed and mobile algorithm has been developed by Frank Wang and Nu Helian for global association rule mining[3], which doesn't have to ship all of local data to one site thereby not causing excessive network communication cost [4-5]. Classification techniques are compared by many

investigators [6-9]. in keeping with Megha Gupta and Naveen Aggarwal, data mining is that the knowledge discovery from the big amounts of knowledge stored in databases[10]. it's a knowledge discovery from databases (KDD), the automated or suitable extraction of patterns representing knowledge implicitly stored in large databases. Data processing tools forecast future trends and behaviours, allowing businesses to form practical, knowledge-driven decisions. Classification techniques are generally utilized in data mining to categorize data among various classes. Classification techniques are getting used in many industries to simply identify the kind and group to which a specific tuple belongs.

One of the foremost challenging tasks within the whole KDD process is to choose the proper data mining technique. The commercial software tools provide more and more possibilities together and therefore the decision requires more and more expertise on the practical point of view. Aditi Goel and Saurabh Kr. Srivastava found that SVM is that the best classifier amongst all the classifiers. They used learning algorithms with the historical dataset to train the classifier and therefore the test samples are accustomed validate the correctness of the classifier[11]. Sharon Carl, Glaston D'souza and Linet Varghese focused on Implementation of Classification Algorithms and their Comparison for Educational Dataset[12]. T. Rajesh Kumar et. Al. used data mining techniques for student's Performance [13]. Adelaja Oluwaseun Adebayo and Mani Shanker Chaubey applied data mining classification techniques on the analysis of student's performance[14]. Chitra Jalota and Rashmi Agrawal performed analysis of educational data mining using Classification[15].

### **3. Problem Statement**

The use of information mining techniques and rapidly development in mobile technology it's encouraging to introduce techniques to utilize these emerging technologies. The utilization of classification technique of data mining using mobile devices is one in every of the challenging task, the shifting of pricy data processing task to mobile devices is challenge by the constraints carry these devices. Also, mobile devices have limited processing capabilities and memory capacities. Another important issue during data transformation between mobile device and data provider over low bandwidth wireless network isn't feasible. Keeping visible the importance of data mining and also the employment of mobile devices, the demand is to effectively visualize results of information mining on mobile devices in such the way to produce information very fast.

### **4. Objectives Of The Research**

1. Study of Data mining techniques.
2. Study of methodologies and algorithms of data mining for mobile devices.
3. Use of an efficient approach of data mining through Web services.
4. To study classification techniques for large data.
5. Simulation testing and performance execution.

### **5. Classification In Present System**

Classification is a data mining task that assigns items in a group to target categories or classes. The objective of classification is to predict the target class for each case in the data accurately. A classification technique begins with a data set in which the class targets are known. Classification is a technique of supervised learning. It maps the data into predefined groups. It is used to grow a model that can classify the population of records at large level.

Classification algorithm requires that the classes to be defined based on the value of data attribute. Classification has been applied on four tables of Supermarket dataset.

Algorithm used in proposed work is as follows:

## Variables

These are the variables and array needed:

Count - containing the count of all is the current iteration

res - is the result set displaying data from database

entities - containing all distinct values for product, customers, sales, promotions

ids - contains all id's (for customer, product, sales, promotions tables), is the index of the target input data vector in the input data set

list - is a target input data vector

idcnt - is the counter containing list of all entities related to particular category

iteration gives condition for how much times the loop will rotate res\_array is two dimensional array for training data\_set train\_dataset – recursive function for generating result T\_net – class that gives result after number of iterations rotations.

## Algorithm

Step 1: Retrieve all entities class id from main categories

```
select distinct entity_class_id from different entities tables (one at a time)
```

```
while (res.next()) { idcnt++;  
    list.add (entity_class_id);  
} ids = new String[idcnt];
```

Step 2: Retrieve the different parameters, their count and entity\_class\_id from distinct entity table (one at a time)

```
for(i=0; i<ids.length; i++)  
{ select distinct parameters (entities), count(*), entity_class_id from distinct entity table.  
    if (res.next()) { Get Name of attributes  
        Geneate combolist[i] & combolist2 [i]  
    } }
```

Step 3: Generate two dimensional array list for providing input to classifier for actual classification

```
res_array = [idcnt][2];  
for (i=0; i<idcnt; i++) { for (j=0;j<2;j++)  
    {  
        res_array [i][j] = 0.0;
```

}}

Step 4: Grouping of category, count of items in the category and number of sub categories for each entity category.

Step 5: Sub category count will indicate the result and gives classified data

## 6. Application of Classification Technique to Super Market Dataset

Requirement of Classification algorithm is that on the basis of data attribute value classes must be defined . Depending on the characteristics of the data classes are described which is already known to classes. Classification has been applied on four tables of Supermarket dataset. Classification has been applied to tables of data set. Algorithm for classification of product table is as follows:

1. all main categories from product\_category table are retrieved
2. all product class id from main categories are retrieved
3. sub category names from the class id from product category table are retrieved
4. each class count for sub categories with the product class count from the products table are matched.
5. If the class id product\_category table is matched with class id of products table, increment the count for each class id
6. array of all counts of class id is created.
7. Grouping of category, count of items in the category and number of sub categories for each product category
8. Average for each sub category count by the total category count is calculated.

Similarly, classification can be applied to Customer, Sales and promotions.

## 7. Experimental Results of Classification Technique

Data analysis is a very important process that is needed to facilitate data interpretations and aided indirectly the decision makers by providing need knowledge for analysis. Graphics and visualization are used in data analysis to facilitate data understanding. Data analysis of Product table has been given in following fig 1.

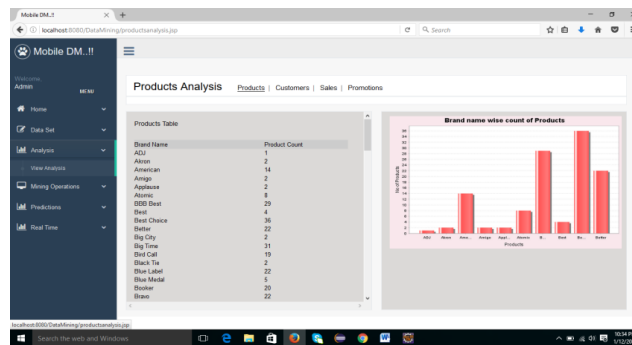


Fig. 1. Analysis of Product

Proposed work has focused on classification of supermarket dataset. Data of supermarket has been divided into suitable classes in order to enhance business decision making process. An algorithm for classification has been applied and tested for different tables of dataset. The classification results for Product table of dataset is given in fig. 2

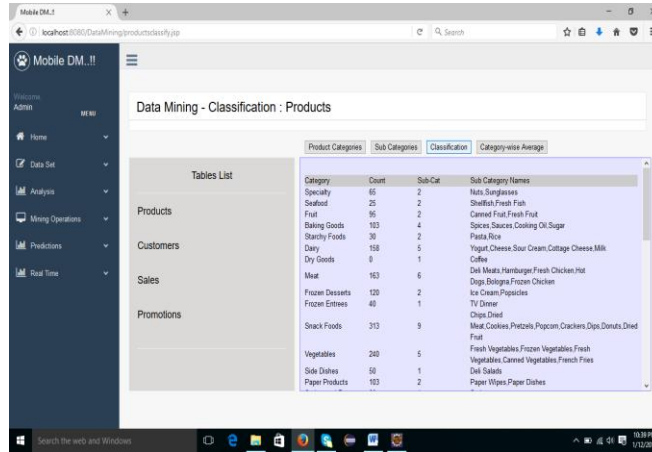


Fig 2. Classification of Products

Classification is used to find out in which group each data instance is related within a given dataset. It is used for classifying data into different classes according to some constrains. In proposed system classification has been performed in server and results are given to mobile devices with the help of web services. Results of classification on mobile device are as shown in fig 3.

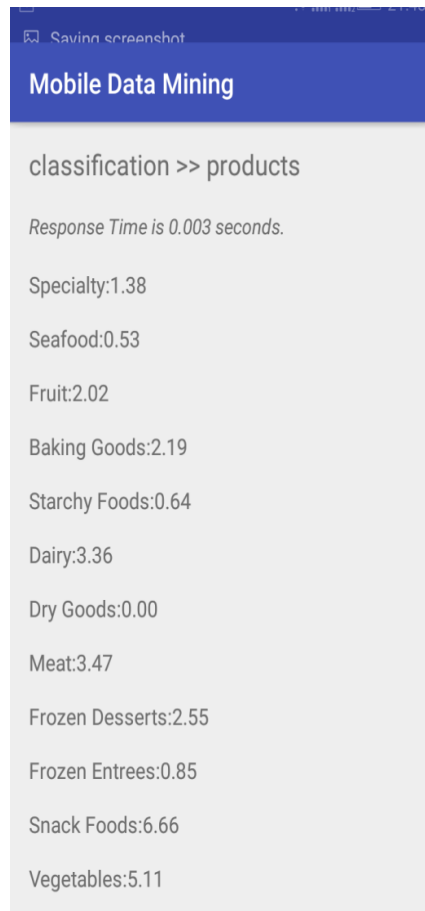


Fig 3. Classification of Products

### 7.1 Comparison of Classification technique in web and mobile applications

Sr. No	Operation type	Time required for web application in seconds	Time required for Mobile application in seconds
1	Classification of Product	2.873	0.236
2	Classification of Customer	0.266	0.071
3	Classification of Sales	66.332	0.344
4	Classification of Promotions	2.827	0.045

Table1. Time required for web and mobile application

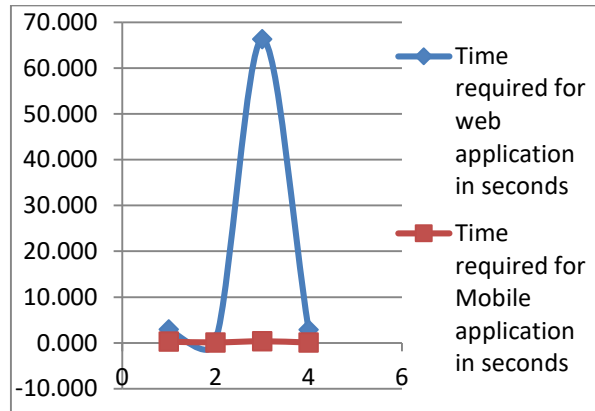


Fig 4. Comparison of time required for web and mobile application.

## 8. Conclusion

The internet and web services usage on mobile devices are continuously and rapidly growing. Therefore the demand is to own efficient mobile interface that may effectively display information and efficiently utilize the little size mobile screen, low bandwidth and unreliable connection etc. Also business organizations including supermarket sectors yields tones of records and sale every day. The most suitable mechanisms that can handle such vast growth of data set and information is data mining techniques like Classification. Data mining is known as the process of monitoring new and innovative information from vast amount of data sets by viewing structured of dataset table and actual data of supermarket. If mobile devices are used for classification then user can view data mining results anytime anywhere. But because of some limitations of mobile devices, it is better to perform classification on server and visualize results on mobile device. So that, people will be able to make sense of huge amount of data as well as come up with new, creative ways to extract useful information from it.

## 9. References

- [1] Domenico Taliayand Paolo Trunfioz, "Mobile Data Mining on Small Devices Through Web Services." <http://grid.deis.unical.it/papers/pdf/MI2010.pdf> Viewed on 25-03-2011
- [2] Dijiang Huang, Xinwen Zhang, Myong Kang, Jim Luo, "MobiCloud: Building Secure Cloud Framework for Mobile Computing And Communication.", [www.dj.eas.edu/courses/CSE\\_548/2010Spring/papers/mobile\\_%20cloud.pdf](http://www.dj.eas.edu/courses/CSE_548/2010Spring/papers/mobile_%20cloud.pdf), Viewed on 15-03-2011.
- [3] Frank Wang and Nu Helian, "A Distributed and Mobile Data Mining System", 2003 IEEE.
- [4] Sherif Rashad, Mehmed Kantardzic, and Anup Kumar, "Mobile Data Mining for Radio Resource Management in Wireless Mobile Networks",
- [5] Mohamed Medhat Gaber, "Foundations Of Adaptive Data Stream Mining For Mobile And Embedded Applications", Proceedings of the 2008 IEEE, CIBEC'08
- [6] Debajyoti Mukhopadhyay, Pradipta Biswas, Young-Chon Kim, "A Syntactic Classification based Web Page Ranking Algorithm", 6th International Workshop on MSPT Proceedings 2006, pp 83-92.

- [7] Jeffrey Xu Yu, Yuming Ou, Chengqi Zhang and Shichao Zhang, “Identifying Interesting Visitors through Web Log Classification”, 1541-1672/05/\$20.00 © 2005 IEEE
- [8] Media Anugerah Ayua, Siti Aisyah Ismaila, Ahmad Faridi Abdul Matina, Teddy Mantorob,” A comparison study of classifier algorithms for mobile-phone’s accelerometer based activity recognition”, International Symposium on Robotics and Intelligent Sensors 2012 (IRIS 2012)[ELSVIER]
- [9] Ben D. Fulcher and Nick S. Jones, “Highly Comparative Feature-Based Time-Series Classification”, IEEE Transactions On Knowledge And Data Engineering, Vol. 26, No. 12, December 2014.
- [10] Megha Gupta, Naveen Aggarwal, “ Classification Techniques Analysis”, NCCI 2010 -National Conference on Computational Instrumentation CSIO Chandigarh, INDIA, 19-20 March 2010.
- [11] Aditi Goel, Saurabh Kr. Srivastava, “Study of Data Mining Algorithms in the Context of Performance Enhancement of Classification ”, International Journal of Computer Applications (0975 – 8887) Volume 134 – No.9, January 2016
- [12] Sharon Carl, Glaston D’souza and Linet Varghese, “ Implementation of Classification Algorithms and their Comparison for Educational Dataset”, International Journal of Innovative Science, Engineering & Technology(IJISSET ), Vol. 3 Issue 3, March 2016. www.ijiset.com ISSN 2348 – 7968
- [13] T. Rajesh Kumar et. Al., “Students Performance Prediction Using Data Mining Techniques”, published in proceeding of International Conference on Intelligent Sustainable Systems (ICISS)-2019 by IEEEExplore Digital Library
- [14] Adelaja Oluwaseun Adebayo , Mani Shanker Chaubey , “Data mining classification techniques on the analysis of student’s performance”, Global Scientific Journals, Volume 7, Issue 4, April 2019 ISSN 2320-9186.
- [15] Chitra Jalota, Rashmi Agrawal, “Analysis of Educational Data Mining using Classification”, published in proceeding of International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)- 2019 by IEEEExplore Digital Library