Credit Card User Frequent Buying Predicton Analysis using Cluster Methods

N.Akshaya, Sundar Santhoshkumar, E.Ramaraj

Abstract: *Today the world becomes more digital. The cashless transactions are increased in all sectors. The large amounts of data in digital form are generated every day. The companies need to analyze the existing transactions, to predict the user requirements in the future. The payment during the purchase can be done in different modes by the user. In this work, the credit card transactions are analyzed. There are many data mining techniques are used to predict the frequent sets of items during purchase. Data clustering in one of the familiar and widely used technique to identify a similar set of items in a group or dataset. In this work, the two familiar existing techniques k-means and k-mediods are compared with the same datasets. The results show the best clustering algorithm.*

Index Terms: Data Mining, Clustering, K-means, kmediods.

I. INTRODUCTION

The increase of digital transactions leads to the accountability of digital data which is an asset of companies. The user profile and interests are most wanted buzz words for all business companies. The value of user data is getting more enterprise and become more investment field. Most of the investors undergo predictive analysis before investment. The financial transactions help to find the economic stability, investment interests and income of a particular user. The credit transaction is highly used in digital mode rather less in traditional purchases. There are many data mining techniques are used to extract frequent transaction and itemsets used in transactions. Clustering is one of the oldest and consistent techniques used in a similar set of attributes in a set. The credit card transactions encourage the customer to buy a product with its relevant product. When credit card transactions are analyzed one can able to predict the frequently occurred itemsets in a transaction. The prediction of frequently occurred itemset with credit card support may help the user and investor to predict the frequent selection of items sets.

FOR PAPER SUBMISSION

A. Review Stage

Submit your manuscript electronically for review.

B. Final Stage

When you submit your final version, after your paper has been accepted, prepare it in two-column format, including figures and tables.

II. DATA MINING

Data Mining is the most used technique for Data analysis. There are significant and result oriented techniques like clustering and classification algorithms are used as a problem-solving technique and later they are used as preliminary techniques for advanced techniques like soft computing etc., knowledge discovery in transactional databases is one of the important key factors in business processes.

A. K- Means Algorithm

K-means is a partition based method which is used to cluster the similar set of items as a group called cluster with elements N. The number of cluster K is defined by the user. It gives the local maximum.

K Mediods Algorithm

k-medoid is also a classical partitioning technique of clustering, which clusters the data set of n objects into k clusters. Where ask medoids uses data points as centroids.

K Means vs Kmediods

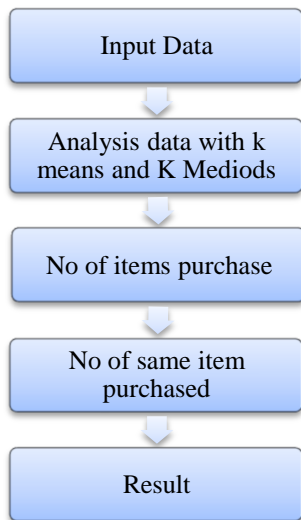| K-Means	K-Medoids
Multifaceted nature is O(ikn)	Complexity is (i(k(n-k)^2)
All the more efficient	Comparatively less effective
Touchy to outliers	Not delicate to exceptions
Curved shape is required	Convex shape is not must
Number of groups should be determined in advance	Number of bunches should be indicated ahead of time
Effective for isolated clusters	Efficient for isolated groups and little dataset

B. About Data

In order to find the frequent itemset prediction, the credit card user transaction datasets are taken for experimentation. There are one lakh records with different attributes (product items) are used. The online supermarket shopping data sets are taken from online data repositories.

C. Methodology

The credit card based online shopping data set is given input to both Kmeans and medoids algorithm. The frequently purchased items sets are predicted and given as output.

Credit Card User Frequent Buying Predicton Analysis using Cluster Methods

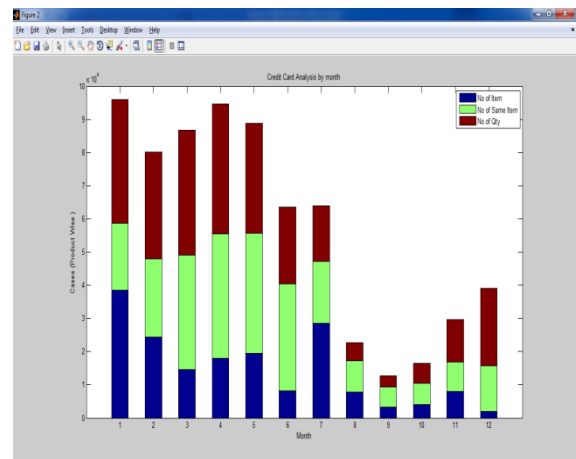## III. EXPERIMENTAL STUDY

A. Load Dataset

The shopping cart dataset is used in the experimental study. The dataset is loaded in experimentation model.

The data set contains a list of products available and set items purchased by the credit card users. B. Frequent Itemset Prediction The dataset is clustered using the application model with both Kmeans and Kmediods algorithms. The application allows for selecting the purchased item. Once the attribute is selected the frequently purchased items and its frequent itemsets are identified. The process involves the following steps.
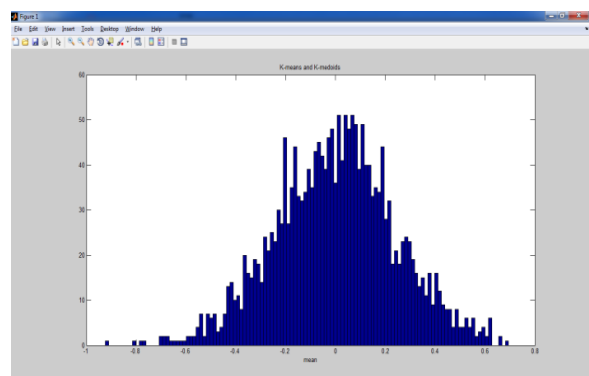
- Item or product selection
- Transaction quantity
- Comparison between K means and Kmediods Algorithms
- Result

C. Comparision

The data set is given as input to both algorithms and the execution time and clustering accuracy are analyzed with respect to its generic functionality. The product wise with occurrence and similarity is predicted.

The result is then compared with respect to the number of similarities in the dataset.

The following figure shows the comparative study of K-means and Kmediods algorithm with credit card shopping dataset. Based on the attribute selection and similarity index Kmedios give better performance than the k-means algorithm.

CONCLUSION

Credit card based user frequent purchase prediction using K-means and K-medoids algorithms have experimented in this work. This a preliminary study hwere online shopping dataset is taken form analysis. The result shows that Kmediod obtained better results compare with the k-means algorithm. Hence this is a budding research work, the existing clustering are compared without any modification. As future work, the existing techniques will be modfied with respect to data used for prediction.

APPENDIX

It is optional. Appendixes, if needed, appear before the acknowledgment.

ACKNOWLEDGMENT

REFERENCES

1. J. Han, M. Kamber, Data Mining Concepts and Techniques, Morgan Kaufmann, 2006 (2nd Edition), http://www.cs.sfu.ca/~han/dmbook
2. I.H. Witten, E. Frank, Data Mining: Practical Machine Learning Tools and Techniques, Second Edition, Morgan, Kaufmann, 2005,http://www.cs.waikato.ac.nz/~ml/weka/book.htm l
3. Pang-Ning Tan, Michael Steinbach, and Vipin Kumar, Introduction to Data Mining (First Edition), Addison Wesley, (May 2, 2005).
4. Margaret H. Dunham, Data Mining: Introductory and Advanced Topics, Prentice Hall, 2003,http://lyle.smu.edu/~mhd/book
5. A K-means-like Algorithm for K-Medoids Clustering and Its Performance, Hae-Sang Park, Jong-Seok Lee, and Chi-Hyuck Jun, Department of Industrial and Management Engineering, POSTECH, South Korea, shoo359@postech.ac.kr, jongseok@postech.ac.kr, chjun@postech.ac.kr
6. Computational Complexity between K-Means and KMedoids Clustering algorithms for Normal and Uniform Distributions of Data Points T. Velmurugan and T. Santhanam, Department of Computer Science, DG Vaishnav College, Chennai, India
7. Best of Both: A Hybridized Centroid-Medoid Clustering Heuristic, NizarGriragrira@nii.ac.jp, Michael E.Houlemeh@nii.ac.jp, National Institute of Informatics, Japan
8. K-Medoids: CUDA Implementation, Douglas Roberts, May 21, 2009
9. International Journal of Database Management Systems (IJDMS), Vol.3, No.1, February 2011 [10] Top 10 algorithms in data mining, XindongWu • Vipin Kumar • J. Ross Quinlan.

AUTHORS PROFILE

N. Akshaya is a M.Phil research scholar of computer Science. Her research interests are Data Mining and Big Data Analytics.
Email: akshayanarayanan1995@gmail.com

Dr.S.Santhoshkumar is Assistant Professor in Computer Science, Alagappa University, Karaikudi, Tamil Nadu. His research interests are Machine Intelligence and Computational Biology.
Email: santhoshkumars@alagappauniversity.ac.in.

Dr.E.Ramaraj is Professor and Head in computer science, Alagappa University, Karaikudi, Tamil Nadu. This research interests are multimedia, IoT and data mining.
Email: eramaraj@rediffmail.com

Retrieval Number: I8145078919/19©BEIESP
DOI:10.35940/ijitee.I8145.078919

2225

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication