that the entire cluster must be find in subspaces of the original space.

4) Ability to function in an incremental manner: In certain cases, e.g., data warehouses, the underlying data used for the original clustering which can change over time. If the clustering algorithm can incrementally handle the addition of new data or the deletion of old data, then this is usually much more efficient than re-running the algorithm on the new data set. The proposed works have an ability to handle the new data.

6. Result Analysis

The works regarding the topic have been proposed by using the online shopping scenario. And the expected results according to the proposed work are as follows:

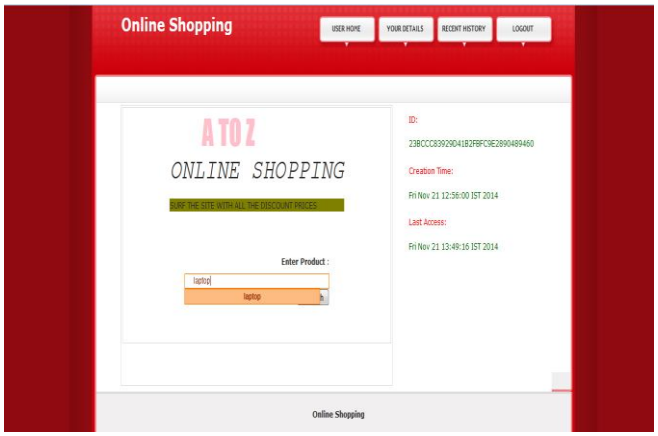

Figure 5: Searching a product (Laptop here)

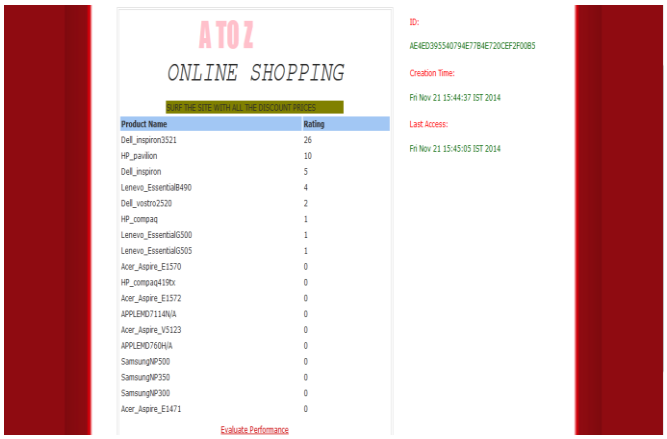

Figure 6: Suggetion Window for searching Product

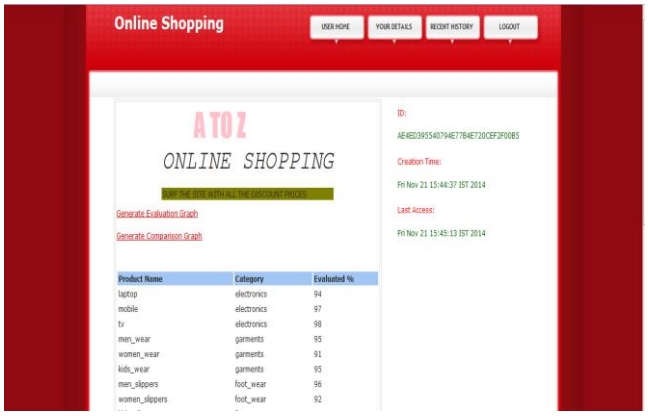

Figure 7: Performance Evaluation in%

Figure 8: General Result Analysis: Evaluation Graph

As shown in figure 5 very first the expected product will searched. As a result of this we will get all the product related to the searchable product along with their rating factors. While searching a particular product individual items in a database are consider as a separate cluster and then based on the distance criteria the most closest product related to the searching one are again put together for clustering purpose and hence we will get all the related products which are most closest to the searching one will get in next window i.e. in suggestion window as shown in fig 6. The performance of all the searched product is then evaluated in percentage value to achieve the market-basket analysis task based on the rating factor got in suggestion window as shown in fig 7. Figure 8 shows the graphical representation of general evaluation of the proposed work. In this diagram, the X-axis of a graph belongs to the no of items on which clustering have to be perform while Y-axis belongs to the efficiency of the proposed work in terms of some threshold values. The red line in graph represents the complete efficiency of a proposed work with respect to the items present in the data set. As shown in graph, for the 30 items in a data set the complete efficiency of a product is almost more than 90%. i.e. we can get at least 90% exact cluster by using the proposed work to achieve the Purpose of Market-Basket Analysis.

7. Conclusion

Market-basket analysis is an integral part of today's business world. Customer satisfaction is at the center point in Market-Basket Analysis and to achieve this it is necessary to find the main interest of customer in a particular product. The

agglomerative hierarchical clustering used in the proposed work creates the clusters by considering each item or product as a individual cluster from its starting with which the retailers can easily identify which products are frequently purchased by the customer from a huge dataset. The clustering of retail items with this technique gives more efficient and reliable result than other techniques as here clustering of item start from a individual element The placement of product in retail with the help of such clustering will not only effective and impressive but also helpful to achieve the goal of market-Basket Analysis. The technique presented is useful in the area of failure classification in retail stores or in supermarket, since the current failure classification methods do not have a definitive way to determine the number of clusters into which a set of program executions should be divided.

References

[1] Ashok Kumar D and Loraine Charlet Annie M.C, "Market Basket Analysis for a Supermarket Based on Frequent Itemset Minng", IJCSI, Vol. 9. Issue 5, No.3,September 2012.
[2] Aastha Joshi, Rajneet Kaur, "Comparative study of Clustering Techniques in Data mining", IJARCSSE, 2012.
[3] Berry, M.J.A., Linoff, G.S.: Data Mining Techniques: for Marketing, Sales and Customer Relationship Management (second edition), Hungry Minds Inc., 2004.
[4] "Cluster analysis" in http://en.wikipedia.org/ wiki/Cluster_Analysis
[5] Chen, Y.-L., Tang, K., Shen, R.-J., Hu, Y.-H.: "Market basket analysis in a multiple store environment, Decision Support Systems", 2004.
[6] Erik Buchmann, Leonardo Weiss Ferreira Chaves and Klemens Bohm, "Finding misplaced items in retail by clustering RFID data", EDBT 2010,March 22-26,2010, Lausanne, Switzerland.
[7] Er. Arpit Gupta 1,Er.Ankit Gupta 2,Er. Amit Mishra 3, " Research Paper On Cluster Techniques Of Data Variations", International Journal of Advance Technology & Engineering Research (IJATER).
[8] Fionn Murtagh and Pedro Contreras, "Methods of Hierarchical Clustering",arXiv:1105.021v1 [CS:IR], 30th April 2011.
[9] Introduction to Data Mining and Knowledge Discovery, Third Edition ISBN: 1-892095-02-5, Two Crows Corporation, 10500 Falls Road, Potomac, MD 20854 (U.S.A.), 1999.
[10] Kalyani M Raval, " Data Mining Techniques", IJARCSE, Volume 2, Issue 10, October 2012
[11] Manish Verma, Mauly Srivastava, Neha Chack, Atul Kumar Diswar, Nidhi Gupta, "A Comparative Study of Various Clustering Algorithms in Data Mining",International Journal of Engineering Reserch and Applications (IJERA), Vol. 2, Issue 3, pp.1379-1384, 2012
[12] MR ILANGO, Dr V MOHAN, "A Survey of Grid Based Clustering Algorithms", International Journal of Engineering Science and Technology, pp.3441-3446, 2010.
[13] "Measuring Association d12 Between Clusters 1 and 2" in http://www.stat.psu.edu/online/courses/stat505/18_cluste r/0 5_cluster_between.html
[14] Neha Soni, Amit Ganatra, — "Categorization of Several Clustering Algorithms from Different Perspective: A Reviewl", International Journal of Advanced Research in Computer Science and Software Engineering, vol. 2,no.8,pp.63-68,Aug. 2012.
[15] P. Berkhin. (2001) — "Survey of Clustering Data Mining Techniques" [Online]. Available:http://www.accure.com/products/rp_cluster_re view.pdf
[16] Rui Xu, Donald C. Wunsch II, — "Survey of Clustering Algorithmsl", IEEE Transactions on neural Networks, vol. 16, pp. 645-678, May 2005.
[17] Rahmat Widia Sembiring, JAsni Mohamad Zain, Abdullah Embong, "A Comparative Agglomeartive Hierarchical Clustering Method to Cluster Implemented Course", journal of Computing, volume 2,Issue 12, December 2010.
[18] S.M. Savaresi and D. Boley, "On the performance of bisecting K-means and PDDP",. Proc. SIAM Data Mining Conf, 2001.
[19] W. Bishop, "Documenting the value of merchandising. Technical report, National Association for Retail Merchandising Service", 2000.

Author Profile

Ms. Rujata Saraf pursuing M.E. Degree in Computer Science and Engineering from North Maharashtra University, Jalgaon and received the B.E., degree from Mumbai Univ. in 2011. After degree worked as an assistant professor (from January 2012) in the Dept. of Information Technology, the G.H.Raisoni Institute of Engineering and Management, Jalgaon., and now working as a Lecturer (from January 2015) at ST. Francis Institute of Technology, Borivali in Mumbai University

Prof. Sonal P. Patil received the Diploma and B.E.. degrees, from MSBTE and NMU Univ. in 2005 and 2008, respectively. She received the M.Tech degree from Bhopal Univ. in 2013. She have been Appreciated as Best Outgoing Student during Diploma & Degree College. Also Appreciated by North Maharashtra University, Jalgaon for active involvement and management of YUVARANG 2009. She worked as an assistant professor (from 2009 to 2014) in the Dept. of Information Technology, in G.H.Raisoni Institute of Engineering and Management, Jalgaon and now (from July 2014) working as a HOD of Information Technology department in Same college. She has successfully published book having a title "Computer Organization" for Second Year CSE & IT Students of Engineering in January 2014. Her research interest includes Data Mining and its Application. She is a member of CSI, CAP process of North Maharashtra University from Dec 2010 Exam, Syllabus setting committee member for the Computer & I.T. department subjects and Certified & lifetime member of ISTE.