

Approaches to Clustering in Customer Segmentation

Shreya Tripathi^{1*}, Aditya Bhardwaj², Poovammal E³

^{1,2}Studies in the Department of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India

³Faculty of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India

*Corresponding Author Email: ¹shreyatripathi0610@gmail.com, ²bh_aditya@outlook.com, ³poovammal.e@ktr.srmuniv.ac.in

Abstract

Customer Relationship Management(CRM) has always played a crucial role as a market strategy for providing organizations with the quintessential business intelligence for building, managing and developing valuable long-term customer relationships. A number of business enterprises have come to realize the significance of CRM and the application of technical expertise to achieve competitive advantage. This study explores the importance of Customer Segmentation as a core function of CRM as well as the various models for segmenting customers using clustering techniques. The available clustering models for customer segmentation, in general, and the major models of K-Means and Hierarchical Clustering, in particular, are studied and the virtues and vices of the techniques are pointed out. Finally, the possibility of developing a hybrid solution by the combination of the above two techniques, having the ability to outperform the individual models, is discussed.

Keywords: Customer Relationship Management, Customer Segmentation, Clustering, K-Means Clustering, Hierarchical Clustering

1. Introduction

In the contemporary day and age, the importance of treating customers as the principal asset of an organization is increasing in value. Organizations are rapidly investing in developing strategies for better customer acquisition, maintenance and development. The concept of business intelligence has a crucial role to play in making it possible for organizations to use technical expertise for acquiring better customer insight for outreach programs. In this scenario, the concept of CRM garners much attention since it is a comprehensive process of acquiring and retaining customers, using business intelligence, to maximize the customer value for a business enterprise [1].

One of the two most important objectives of CRM is customer development through customer insight. This objective of CRM entails the usage of an analytical approach in order to correctly assess customer information and analysis of the value of customers for better customer insight. Keeping up with the changing times, organizations are modifying their business flow models by employing systems engineering as well as change management and designing information technology(IT) solutions that aid them in acquiring new customers, help retain the present customer base and boost the customer's lifelong value [2]. According to [3], due to the diverse range of products and services available in the market as well as the intense competition among organizations, customer relationship management has come to play a significant role in the identification and analysis of a company's best customers and the adoption of best marketing strategies to achieve and sustain

competitive advantage.

One of the most useful techniques in business analytics for the analysis of consumer behavior and categorization is customer segmentation. By using clustering techniques, customers with similar means, end and behavior are grouped together into homogeneous clusters[4]. Customer Segmentation helps organizations in identifying or revealing distinct groups of customers who think and function differently and follow varied approaches in their spending and purchasing habits.

Clustering techniques reveal internally homogeneous and externally heterogeneous groups. Customers vary in terms of behavior, needs, wants and characteristics and the main goal of clustering techniques is to identify different customer types and segment the customer base into clusters of similar profiles so that the process of target marketing can be executed more efficiently. This study aims to explore the avenues of using customer segmentation, as a business intelligence tool within the CRM framework as well as the use of clustering techniques for helping organizations redeem a clearer picture of the valuable customer base. The concepts of customer relationship management, customer segmentation as a core function of CRM as well as the approach of segmenting customers using clustering techniques are discussed.

The available clustering models for business analysis in the context of customer segmentation, the advantages and disadvantages of the two main models chosen for our study- K-Means and Hierarchical Clustering, as well as the possibility of developing a hybrid model which can outperform the individual models is surveyed.

The paper is organized as follows. Section 2 provides an overview of customer segmentation and related concepts. The role of

clustering in the context of customer segmentation and the major models of K-Means and Hierarchical Clustering are described in Section 3. Finally, a general discussion and concluding remarks are provided in Section 4.

2. Related Work

Key concepts such as CRM, customer segmentation and the usability of customer segmentation are reviewed and discussed in this section. The significance of these concepts to businesses and organizations is also highlighted.

A. Customer Relationship Management

Customer Relationship Management is an important business approach for developing and securing steady, long-term customer associations. The modern marketing approach promotes the usage of CRM as part of the organization's business strategy for enhancing customer service satisfaction[5]. CRM enables business enterprises in customer value analysis as well as the targeting of those customers that prove of greater value. It also helps business organizations in developing high-quality and long-term customer-company relationships that increase loyalty and profits. An accurate evaluation of customer profitability and the targeting of highvalue customers are important factors that contribute to the success of CRM [6].

CRM being a customer-centric strategy, it is important for business organizations to be familiar with their customer base in terms of its characteristics and behavior. These insights into the customer data can then become useful when it is employed in such Information Technology(IT) solutions that provide valuable outputs for better targeting the profitable customers [7].

CRM plays a major role in targeting the customer base, once it is identified using essential segmentation strategies. According to [8] and [9], the CRM strategy is a closed circular structure with four dimensions; Customer identification, customer attraction, customer retention, and customer development. Thus, customer identification which lays the foundation of this structure clearly implies that the act of grouping or segmenting customers according to their behavior and characteristics-customer segmentation, emerges as a core function of CRM.

B. Customer Segmentation

As the market is widening, the rate of competition between all business entities is rapidly growing. Hence, these business enterprises are increasing their expenditure on their marketing strategies to achieve competitive advantage [10], [11]. In this context, the significance of employing Information Technology(IT) solutions to marketing campaigns emerges as a pivotal step in a modern approach to business. Customer Segmentation is a popular technique of partitioning the customer base into externally distinct and internally uniform groups in order to create varied marketing strategies for targeting each group according to its characteristics. Generally speaking, it is defined as the process whereby the consumers of a business enterprise are divided into groups according to their preferences, characteristics and purchasing behavior [12].

By studying and analyzing large volumes of collected customer data, businesses can improve their marketing decisions based on the customers' preferences [13]. According to [14], maximum profits can be generated for any business entity if the resources are utilized judiciously in order to cultivate the most loyal and useful group of customers once customer segmentation and clustering have enabled the allocation of customers to such groups. The total customer set can be divided and grouped into clusters based on their buying behavior, frequency, demographics etc. Hence, instead of studying each customer individually, firms can group

similar customers together so that their needs can be better understood [15].

C. Use of Customer Segmentation

Target Marketing and Customer Segmentation are so closely related that they are often used interchangeably. According to [16] and [17], target marketing refers to the grouping of buyers based on certain characteristics which the firms intend to serve. It has been referred to as a personal branding strategy in the context of a specific customer. There are three steps that ought to be followed in order to devise segmentation-based marketing strategies. First, customers in the selected market are segmented into different groups based on their characteristics. Secondly, the segments are studied for their properties and the different ways in which marketing tactics can be applied to that specific group. At last, required comparisons on the competing brands and studies about customer behavior to their products can be completed. Hence, a segmentation model which is useful will be able to effectively increase the profitability and competitive value for a company. The next section delves into the concept of using clustering for customer segmentation and the various algorithms involved.

3. Using Clustering Techniques for Customer Segmentation

Clustering, which is also called as cluster analysis, is an important subject in data mining [18]. The data set is partitioned into some groups and the data points in each group, that is, the cluster are more related to each other than to those in other clusters. These data points are clubbed together by detecting correspondences according to the attributes found in raw data [19], however the main purpose is to find the suitable number of clusters which are relevant as well as insightful for analysis purposes. This process is a repeatable and iterative task where vast amounts of raw data are scanned for similarities and patterns [20]. The unorganized data is searched for knowledge that is important and then data points are assigned. For getting favorable results, a specific clustering algorithm along with certain parameters may be best suited in a market domain [21].

According to [22], clustering is a type of data mining technique used in a number of applications, involving areas such as machine learning, classification and pattern recognition. There are various clustering algorithms which differ from each other in terms of the approach they follow in order to do the grouping of the objects according to their characteristics.

These are stated as follows [23] –

In Partitioned Based Clustering all the data points are taken as a single cluster in the beginning. These data points are then separated into clusters by iteratively positioning these objects between the clusters. Some of the partitioning algorithms are K-Means, K-Medoids and K-Modes. For Hierarchical Clustering, one of the two present approaches can be followed for implementation [24]. One is Agglomerative (bottom-up) approach and another one is Divisive (top-down) approach. In the agglomerative process, each observation begins in its own cluster or segment and then pairs of these formed segments are combined when moving up the hierarchy. In the divisive approach, all the observations begin in one cluster and then are repeatedly split into different clusters. The results are in the form of dendrograms. In Density Based Clustering, the clusters are defined as regions of higher density than the rest of the dataset. Objects are differentiated as core, noise and border points. Grid Based Clustering algorithms partition the data set into grid structures containing a number of cells. Grid algorithm uses subspace and hierarchical clustering techniques. STING, CLIQUE are some of the grid-based clustering techniques. For implementing Model Based Clustering, the data points are grouped together on the

basis of various techniques like statistical methods, conceptual and robust clustering methods. Some of the well-known algorithms are SOM, SLINK and COBWEB.

A. Clustering for Segmentation Purposes

Clustering techniques reveal internally homogeneous and externally heterogeneous groups. Customers vary in terms of behavior, needs, wants and characteristics and the main goal of clustering techniques is to identify different customer types and segment the customer base into clusters of similar profiles so that the process of target marketing can be executed more efficiently. Both, hierarchical and non-hierarchical clustering algorithms are widely used in customer segmentation, most prominent among them being K-Means and Agglomerative Hierarchical Clustering. In [25], K-Means has been used as part of their clustering approach. [26] also implemented K-Means for customer segmentation on their dataset. Although, hierarchical clustering algorithm seems unsuitable to many, [27] have used it for intelligent customer segmentation for their research and [28] have made use of it for applying clustering algorithms on the transaction data from a supermarket. K-means and Hierarchical Clustering algorithms are useful for clustering data and find extensive usage in customer segmentation. Hence, they will be our main focus of interest.

B. K-Means Clustering

K-Means [29] is one of the most widely used clustering algorithms, and is simple and efficient. The aim of K-Means algorithm is to divide M points in N dimensions into K clusters (assume k centroids) fixed a priori. These centroids should be placed in a wise fashion so that the results are optimal which otherwise can differ if locations of the centroids change. So, they should be placed as far as possible from each other. Each data point is then taken and associated with the nearest centroid until no data points are pending. This way an early grouping is done and at this point, k new centroids have to be recalculated as these will be the centers of the clusters formed earlier. After having calculated these centroids, the data points are then allocated to the clusters to the nearest centroids. In this iteration, the centroids change their position stepwise until no further modifications have to be done and the location of the centroids remain intact.

The K-Means algorithm is relatively simple. The 'K' cluster points, which will be the centroids, are placed in the space among the data points. Each data point is assigned to the centroid for which the distance is the least. After each data object has been assigned, centroids of the new groups are re-calculated. The above two steps are repeated until the movement of the centroid ceases. This means that the objective function of having the least squared error is completed and it cannot be improved further. Hence, we get K clusters as a result.

K-Means algorithm aims at minimizing an objective function, which here, is the squared-error. It is an indicator of the distance of the data points from their respective cluster centers. The process in this algorithm always terminates but the relevance or the optimal configuration cannot be guaranteed even when the condition on the objective function is met. The algorithm is also sensitive to the selection of the initial random cluster centers. That is why it runs multiple times to reduce this effect but for a large number of data points, it tends to perform very well even though it is iterative. A variety of cluster validity indices are used, major of them are Dunn, Davies Bouldin, Silhouette, Sum of Squares within Cluster (SSWC), C, Calinski-Harabasz [30]. SSWC is simple and the most widely used criterion to gauge the validity of the clusters. Smaller values of SSWC mean better clusters [31]. Here we apply K-Means Clustering algorithm on a relatively small dataset and the results are depicted. The dataset is based on customer information for a mall and has 5 attributes named

Customer, Genre, Age, Annual Income and Spending Score. It consists of 200 observations, each of which refers to a unique customer and the spending scores are decided and calculated by the company, based on their spending habits. Hence annual income and spending scores are the key indicators in this data. The age attribute of the customers can also be experimented with, to analyze which age group works best for a business. Any business would always keep the monetary values of any customer as top indicators. Thus, the annual income and spending scores of the customers will be best suited for clustering.

As K-Means algorithm requires the number of clusters as input, below we will use the elbow method to get the optimal number of clusters which can be formed [32]. It works on the principal that after a certain number of 'K' clusters, the difference in SSE (Sum of Squared Errors) starts to decrease and diminishes gradually. Here, the WCSS (Within-Cluster-Sum-of-Squared-errors) metric is used as an indicator of the same. Hence, the 'K' value, specifies the number of clusters.

In Figure 1, it can be observed that an elbow point occurs at $K=5$. After $K=5$, the difference in WCSS is not so visible. Hence, we will choose to have 5 clusters and provide the same as input to the K-Means algorithm.

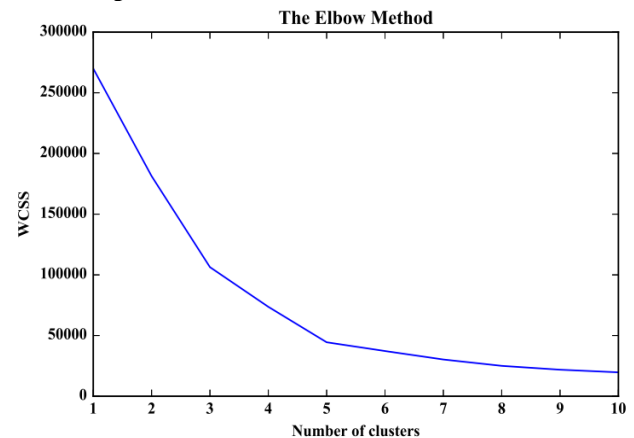


Fig. 1: Finding the optimal number of clusters using Elbow method for K-Means Clustering

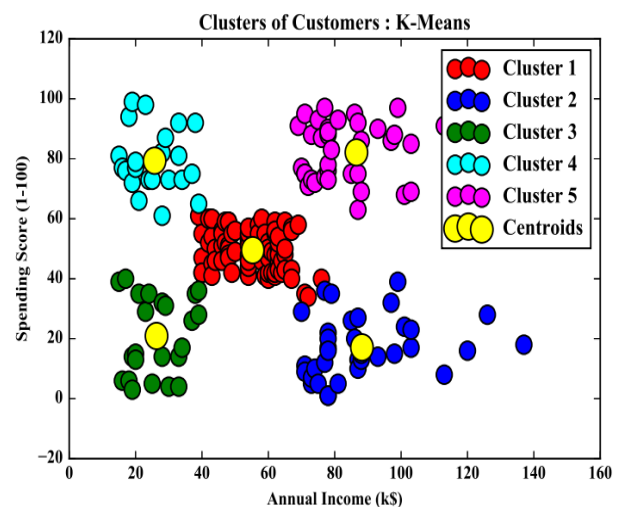


Fig. 2: Clusters formed as a result of applying K-Means Clustering on the dataset taken for study

As shown in Figure 2, the scatter plot of the clusters is created with Annual income plotted against X-axis and Spending Score against Y-axis. The data points under each cluster are represented using distinct colors and the centroids are also highlighted, as shown above.

K-Means has been used in all the major domains; [33] have used it in their research about finding clusters in retail industry, [28]

have taken a dataset of the transactions in a supermarket in their comprehensive analysis of data clustering algorithms. [34] Have also applied K-means algorithm when doing customer segmentation. [31] have also chosen K-means as the partitioning algorithm for the customer data. [35] also chose a form of the K-means algorithm as a part of their clustering technique.

Hence K-means is widely accepted and largely used in customer segmentation. One of the major problems in cluster analysis is the selection of the number of clusters 'K'. Hence, this algorithm does not give the optimal number of clusters but instead this has to be given as input while performing clustering [31]. The ideal number of clusters are the ones in which variation is minimum among the clusters and maximum between the clusters. One major advantage of K-Means clustering is that the computational speed of this algorithm is higher than other hierarchical methods of clustering [36] and it is also easy to implement. According to [37], when the differences between the categories is small and the data is scaled, this algorithm tends to perform better but one major downside is that it depends upon the initial cluster centers. So, if these are chosen incorrectly, then the results tend to be inaccurate and unstable. The time complexity of K-Means is given by $O(nkl)$, where n represents the number of data points, k gives the number of clusters and l is the number of iterations undergone before the algorithm converges. The space complexity is given by $O(k+n)$ [38].

C. Hierarchical Clustering

Hierarchical clustering is a method of cluster analysis which builds a hierarchy of data points as they move into a cluster or out of it. Strategies for this algorithm generally fall into two categories [24].

1. Agglomerative - This is a bottom-up approach where each observation begins as an initial cluster and then merges into clusters as they move up the hierarchy. Divisive technique is a top-down approach where there is only one cluster initially and is then split into finer cluster groups as they move down the hierarchy. This merging and splitting of clusters takes place in a greedy manner and the hierarchical algorithm yields a dendrogram which represents the nested grouping of patterns and the levels at which groupings change.

2. Divisive - These are quite rarely used in market research and hence the agglomerative approach is the one that is widely followed by the practitioners. Here in each step, the two closest clusters are merged based on a specific linkage criterion. The linkage criterion defines the distance between the two clusters. The grouping of data points depends very much on the choice of the linkage criterion. Some of them are complete, single, ward or average. The time complexity of linkage metrics-based hierarchical clustering is high and it is generally given by $O(n^3)$. It is sometimes reduced to $O(n^2 \log n)$ or $O(n^2)$ for some algorithms like CLINK [39] and SLINK[40]. The algorithm for the agglomerative hierarchical clustering approach proceeds by taking each observation in a cluster of its own. A pair of clusters with the shortest distance between them is chosen. The above two clusters are replaced with a new cluster by merging the original clusters in the previous step. Previous two steps are repeated until only one cluster remains and that cluster will contain all the observations. One major advantage of Hierarchical clustering is that we do not need to know the exact number of clusters beforehand and we can choose the formation of the clusters as they merge. The dataset containing 200 mall customers and their information is taken, same as used for the KMeans algorithm, and agglomerative hierarchical clustering is applied. We can visualize the cluster formation in the form of a tree-diagram called a dendrogram, which allows us to illustrate the hierarchical organization of entities just like in Figure 3.

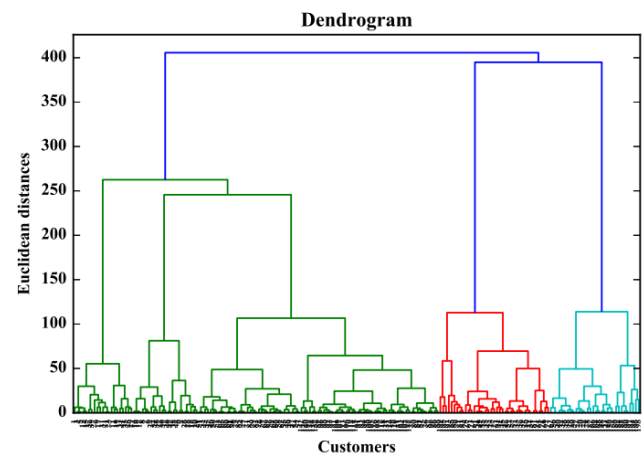


Fig. 3: Visualization of the formation of clusters in the studied dataset with the help of a dendrogram

We have the option of choosing the required number of clusters from the dendrogram itself by selecting the range of maximum distance and then placing a cut-off line at that position. It simply indicates that the distance between the formed clusters is maximum and distinction can be made among them. Hence, according to Figure 3, for satisfactory results, we can choose five clusters ($K=5$). The clusters are depicted as follows in Figure 4.

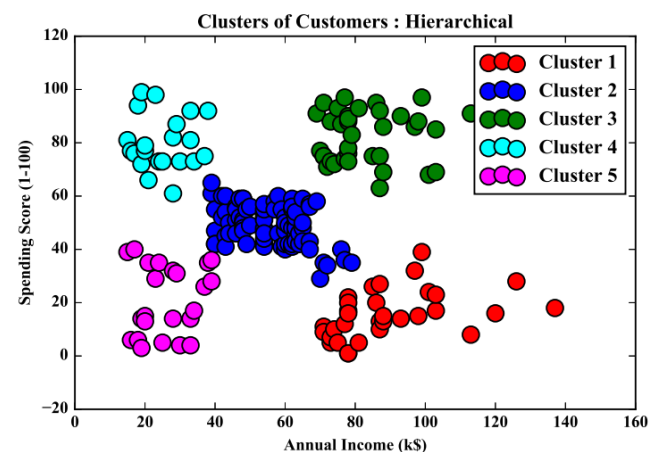


Fig. 4: Clusters formed as a result of applying Hierarchical Clustering on the dataset taken for study

Each color represents a different cluster and data points are plotted against Annual Income on the X-axis and the Spending Score on Y-axis.

Hierarchical clustering has been extensively used for segmentation purposes due to its ability to produce results in a visual way [24]. It helps in determining the number of clusters for any analysis. It can be used for varied datasets like categorical, spatial and time series with numerical data set being the most common as it consists of data as just real numbers [28]. It was also used and compared with other clustering algorithms in bank customer segmentation [34]. [41] and [42] have used Chameleon in their customer segmentation process, which is a hierarchical clustering algorithm as well. [43] have made use of a hierarchical pattern-based clustering. The main advantage of Hierarchical clustering is that the output is in the form of a hierarchy(dendrogram) which tells us exactly at which point the clusters merged or split. Hence it is easy to choose and decide on the number of clusters that we wish to take by looking at the dendrogram. However, for a large number of observations its computational speed is very low as compared to the nonhierarchical methods of clustering [23]. Hence, the size and order of the data have an impact on the final results obtained. Nevertheless, one does not need to plug the number of clusters as the input to the algorithm and hence we can

have different partitioning groups, the choice of which completely depends on the end user.

4. Discussion and Conclusion

Due to increasing commercialization, consumer data is increasing exponentially. When dealing with this large magnitude of data, organizations need to make use of more efficient clustering algorithms for customer segmentation. These clustering models need to possess the capability to process this enormous data effectively. Each of the above discussed clustering algorithms come with their own set of merits and demerits. The computational speed of K-Means clustering algorithm is relatively better as compared to the hierarchical clustering algorithms as the latter require the calculation of the full proximity matrix after each iteration [36]. K-Means clustering gives better performance for a large number of observations while hierarchical clustering has the ability to handle fewer data points.

The major hindrance produces itself in the form of selecting the numbers of clusters 'K' for the K-Means process, which have to be provided as an input to this non-hierarchical clustering algorithm. This limitation does not exist in the case of hierarchical clustering since it does not require any cluster centers as input. It depends on the user to choose the cluster groups as well as their number. Hierarchical clustering also gives better results as compared to K-Means when a random dataset is used. The output or results obtained when using hierarchical clustering are in the form of dendrograms but the output of K-Means consists of flat-structured clusters which may be difficult to analyze. As the value of k increases, the quality(accuracy) of hierarchical clustering improves when compared to K-Means clustering. As such, partitioning algorithms like K-Means are suitable for large datasets while hierarchical clustering algorithms are more suitable for small datasets [38].

Both K-Means and Hierarchical clustering have drawbacks that make them unsuitable when used individually. For business use, data visualization forms a major part of efficient data analysis and hierarchical clustering aids in doing so. Furthermore, when the performance aspect is taken into account, K-Means tends to deliver better results. With the vices and virtues of the two techniques pointed out, it comes to light that an amalgam of the best of these algorithms could outperform the individual models. In summary, different clustering algorithms, owing to their properties towards different kinds of data can be used in succession such that the advantages of these techniques could be harnessed in full. However, the selection process of these suitable techniques as well as their judicious implementation could require a considerable time investment in studying and processing the data along with an adequate understanding of the goals and requirements.

References

- [1] E. Ngai, L. Xiu and D. Chau, "Application of data mining techniques in customer relationship management: A literature review and classification", *Expert Systems with Applications*, vol. 36, no. 2, pp. 2592-2602, 2009.
- [2] J. Peppard, "Customer Relationship Management (CRM) in financial services", *European Management Journal*, vol. 18, no. 3, pp. 312-327, 2000.
- [3] A. Ansari and A. Riasi, "Taxonomy of marketing strategies using bank customers clustering", *International Journal of Business and Management*, vol. 11, no. 7, pp. 106-119, 2016.
- [4] M. Ghzanfari, et al., "Customer segmentation in clothing exports based on clustering algorithm", *Iranian Journal of Trade Studies*, vol. 14, no. 56, pp. 59-86, 2010.
- [5] C. Rygielski, J. Wang and D. Yen, "Data mining techniques for customer relationship management", *Technology in Society*, vol. 24, no. 4, pp. 483-502, 2002.
- [6] J. Lee and S. Park, "Intelligent profitable customers segmentation system based on business intelligence tools", *Expert Systems with Applications*, vol. 29, no. 1, pp. 145-152, 2005.
- [7] D. A. Kandeil, A. A. Saad and S. M. Youssef, "A two-phase clustering analysis for B2B customer segmentation", in *International Conference on Intelligent Networking and Collaborative Systems*, Salerno, 2014, pp. 221-228.
- [8] R. Swift, *Accelerating Customer Relationships: Using CRM and Relationship Technologies*, 1st ed. Upper Saddle River, N.J.: Prentice Hall PTR, 2000.
- [9] A. Parvatiyar and J. N. Sheth, "Customer relationship management: Emerging practice, process and discipline", *Journal of Economic and Social Research*, vol. 3, no. 2, pp. 1-34, 2001.
- [10] A. Riasi, "Barriers to international supply chain management in Iranian flower industry", *Management Science Letters*, vol. 5, no. 4, pp. 363-368, 2015.
- [11] A. Riasi, "Competitive advantages of shadow banking industry: An analysis using porter diamond model", *Business Management and Strategy*, vol. 6, no. 2, pp. 15-27, 2015.
- [12] J. Qian and C. Gao, "The application of data mining in CRM", in *2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC)*, Deng Leng, 2011, pp. 5202-5206.
- [13] D. Liu and Y. Shih, "Integrating AHP and data mining for product recommendation based on customer lifetime value", *Information & Management*, vol. 42, no. 3, pp. 387-400, 2005.
- [14] V. Zeithaml, R. Rust and K. Lemon, "The customer pyramid: Creating and serving profitable customers", *California Management Review*, vol. 43, no. 4, pp. 118-142, 2001.
- [15] A. Ansari and A. Riasi, "Customer clustering using a combination of fuzzy c-means and genetic algorithms", *International Journal of Business and Management*, vol. 11, no. 7, pp. 59-66, 2016.
- [16] P. Kotler and G. Armstrong, *Principles of marketing*. Englewood Cliffs, N.J.: Prentice Hall, 1991.
- [17] J. Aaker, A. Brumbaugh and S. Grier, "Nontarget Markets and Viewer Distinctiveness: The Impact of Target Marketing on Advertising Attitudes", *Journal of Consumer Psychology*, vol. 9, no. 3, pp. 127-140, 2000.
- [18] Q. Zhao and P. Franti, "Centroid Ratio for a Pairwise Random Swap Clustering Algorithm", *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 5, pp. 1090-1101, 2014.
- [19] I. Dhillon and D. Modha, "Concept decompositions for large sparse text data using clustering", *Machine Learning*, vol. 42, no. 1/2, pp. 143-175, 2001.
- [20] T. Kanungo, et al., "An efficient k-means clustering algorithm: analysis and implementation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 881-892, 2002.
- [21] D. MacKay, "An example inference task: Clustering", in *Information theory, inference, and learning algorithms*. Cambridge, UK: Cambridge University Press, 2003, pp. 284-292.
- [22] M. Inaba, N. Katoh and H. Imai, "Applications of weighted Voronoi diagrams and randomization to variance-based k-clustering", in *Proceedings of the tenth annual symposium on Computational geometry*, New York, 1994, pp. 332-339.
- [23] T. Sajana, C. Sheela Rani and K. Narayana, "A Survey on Clustering Techniques for Big Data Mining", *Indian Journal of Science and Technology*, vol. 9, no. 3, 2016.
- [24] O. Maimon and L. Rokach, "Clustering methods", in *Data Mining and Knowledge Discovery Handbook*. Boston: Springer US, 2005, pp. 321-352.
- [25] Y. Chen, et al., "Identifying patients in target customer segments using a two-stage clustering-classification approach: A hospital-based assessment", *Computers in Biology and Medicine*, vol. 42, no. 2, pp. 213-221, 2012.
- [26] G. Lafait and T. Kechadi, "Customer segmentation architecture based on clustering techniques", in *Fourth International Conference on Digital Society*, Sint Maarten, 2010, pp. 243-248.
- [27] M. Namvar, M. Gholamian and S. KhakAbi, "A two-phase clustering method for intelligent customer segmentation", in *International Conference on Intelligent Systems, Modelling and Simulation*, Liverpool, 2010, pp. 215-219.
- [28] D. Gaur and S. Gaur, "Comprehensive analysis of data clustering algorithms", in *Future Information Communication*

- Technology and Applications. Dordrecht: Springer Netherlands, 2013, pp. 753-762.
- [29] J. MacQueen, "Some methods for classification and analysis of multivariate observations", in Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, 1967, pp. 281-297.
- [30] E. Rendon, et al., "A comparison of internal and external cluster validation indexes", in American Conference on Applied Mathematics and The Fifth WSEAS International Conference on Computer Engineering and Applications, Puerto Morelos, 2011, pp. 158-163.
- [31] H. Guccdemir and H. Selim, "Integrating multi-criteria decision making and clustering for business customer segmentation", *Industrial Management & Data Systems*, vol. 115, no. 6, pp. 1022-1040, 2015.
- [32] T. Kodinariya and P. Makwana, "Review on determining number of cluster in K-means clustering", *International Journal of Advance Research in Computer Science and Management Studies*, vol. 1, no. 6, pp. 90-95, 2013.
- [33] S. Salvador and P. Chan, "Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms", in 16th IEEE International Conference on Tools with Artificial Intelligence, Boca Raton, 2004, pp. 576-584.
- [34] D. Zakrzewska and J. Murlewski, "Clustering algorithms for bank customer segmentation", in 5th International Conference on Intelligent Systems Design and Applications (ISDA'05), Warsaw, 2005, pp. 197-202.
- [35] P. Brito, et al., "Customer segmentation in a large database of an online customized fashion business", *Robotics and Computer-Integrated Manufacturing*, vol. 36, pp. 93-100, 2015.
- [36] P. Badase, G. Deshbhratar and A. Bhagat, "Classification and analysis of clustering algorithms for large datasets", in International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, 2015, pp. 1-5.
- [37] C. Xiong, et al., "An Improved K-means Text Clustering Algorithm by Optimizing Initial Cluster Centres", in 7th International Conference on Cloud Computing and Big Data (CCBD), Macau, 2016, pp. 265-268.
- [38] O. A. Abbas, "Comparisons between data clustering algorithms", *International Arab Journal of Information Technology*, vol. 5, no. 3, pp. 320-325, 2008.
- [39] D. Defays, "An efficient algorithm for a complete link method", *The Computer Journal*, vol. 20, no. 4, pp. 364-366, 1977.
- [40] R. Sibson, "SLINK: An optimally efficient algorithm for the singlelink cluster method", *The Computer Journal*, vol. 16, no. 1, pp. 30-34, 1973.
- [41] S. Yoon, et al., "A data partitioning approach for hierarchical clustering", in 7th International Conference on Ubiquitous Information Management and Communication, Kota Kinabalu, 2013, p. 4.
- [42] J. Li, K. Wang and L. Xu, "Chameleon based on clustering feature tree and its application in customer segmentation", *Annals of Operations Research*, vol. 168, no. 1, pp. 225-245, 2008.
- [43] D. Suib and M. Deris, "An efficient hierarchical clustering model for grouping web transactions", *International Journal of Business Intelligence and Data Mining*, vol. 3, no. 2, pp. 147-157, 2008.