# **LITERATURE REVIEW**

### **Customer Segmentation**

R. Gardener.et.al[1]The business world has become increasingly competitive over time, therefore businesses must increase their earnings and business by meeting client wants and luring in new customers in accordance with their requirements. It takes a lot of work to identify clients and meet their individual needs. This is due to the possibility of diverse clientele depending on their needs, wants, tastes, and other factors. Consumer segmentation divides the customer population into groups that share similar traits or behaviours rather than taking a “one size fits all” approach. The strategy of segmenting the market into uniform groups is known as customer segmentation.

Abeyasekera, S. 2005.et.al. [2]The Information employed in the customer segmentation technique, which divides the customers into groups, depends on a number of variables, including demographics, regional data, economic situations, and behavioural patterns. By better utilising their marketing budgets, businesses can obtain a competitive advantage over their rivals by proving that they understand their customers’ needs. This is done by using the customer segmentation strategy. Additionally, it aids a company in planning the marketing budget, detecting new market prospects, improving brand strategy, and identifying customer retention.

D. P. Yash Kushwaha.et.al[3]To segment customers, decision-makers consider a variety of factors. Age, gender, family, education level, and income are the most basic and widely used demographic segmentation factors. The other main factors that are employed for segmentation include socio-cultural, geographic, psychographic, and behavioural variables. Presented various clustering algorithms while taking into account Big Data’s size, noise, dimensionality, algorithm calculations, and cluster shape. A brief overview of the various clustering algorithms was provided, which were categorised as partitioning, hierarchical, density, grid-based, and model-based algorithms was also provided.

Alexander, C.E., Wilson, C.A. & Foley, D.H. 2005.et.al. [4]Investigated the need for client segmentation using clustering methods, a fundamental CRM feature. The benefits and drawbacks of the most popular K-Means and Hierarchical Clustering algorithms were examined. Finally, the notion of developing a hybrid strategy is addressed by combining the aforementioned two tactics with the potential to outperform the individual ideas. Customers in the steel business were sorted into two clusters using a combination of fuzzy c-means and genetic algorithms for clustering.

Customer Segmentation is the process of division of customer base into several groups of individuals that share a similarity in different ways that are relevant to marketing such as gender, age, interests, and miscellaneous spending habitsB.

P. E. Shreya Tripath.et.al[5]Customer segmentation is a strategy used by businesses who believe that each customer has unique needs that must be catered to through targeted marketing. Businesses strive to develop a more thorough understanding of their target market. Because of this, their goal must be clear and should be customised to meet the needs of each and every single customer. Additionally, by analysing the data gathered, businesses can better understand client preferences and the criteria for identifying profitable market categories. In doing so, they may more effectively plan their marketing strategies and reduce the risk associated with their investment.

Babbie, E. 2011.et.al.[6]A number of crucial differentiators that separate customers into targetable categories are what drive the customer segmentation technique. The company’s strategy for addressing the various segments is heavily influenced by data on demographics, location, economic position, and behavioural trends.

Market segmentation is a useful tool for identifying and satisfying customer needs. Mall consumer analysis is carried out using unsupervised machine learning techniques, the K-Means clustering algorithm, and the DBSCAN clustering method. A mall consumer analysis is conducted to identify the target clients who can converge most readily. To enable the marketing team to develop a marketing plan for new products that will appeal to their target market’s interests.

E.A. Onur DOĞAN1.et.al[7]Management and upkeep of client relationships have always been crucial in helping businesses develop, manage, and cultivate lucrative long-term customer relationships. In the modern era, it is becoming more and more critical to recognise customers as an organization’s core asset. The development of customer acquisition, maintenance, and development plans is something that businesses are interested in funding. Business intelligence is crucial in enabling businesses to leverage technical know-how to improve customer understanding and outreach initiatives.

Borgatti, S.P. 2005.et.al.[8]Customers with comparable means are grouped together using clustering techniques like k-means. The marketing team uses customer segmentation to identify and highlight several customer groups that have distinctive buying habits and ways of thinking. Customer segmentation aids in identifying customers with a range of preferences, expectations, desires, and other characteristics. The fundamental goal of customer segmentation is to put people in groups based on their shared interests so that the marketing team can come up with an efficient marketing strategy. Clustering is an iterative procedure for extracting knowledge from enormous amounts of unstructured and raw data. In several fields, including machine learning, classification, and pattern recognition, clustering is a sort of exploratory data mining.

R. Azarnoush Ansari.et.al.[9]In order to segment clients, a business must collect specific information about them (data) and analyse it to find patterns that can be utilised to define segments. Job title, location, and products purchased, for example, are some of the information that may be learned from purchasing data. Some of it might be discovered by looking at the customer’s system entry. An online marketer using an opt-in email list may divide marketing communications into different categories based on the opt-in offer that drew the client, for instance. However, other data, such as consumer demographics like age and marital status, will have to be gathered in different methods.

Canever, M.D., Van Trijp, H. & Van der Lans, I. 2007.et.al.[10]Common traits among consumer groups can help a business choose which items or services to advertise to which segments and how to market to each one. Based on segment expertise that tells them younger musicians have less disposable cash than their older counterparts, a small business selling hand-made guitars might choose to push lower-priced products to younger guitarists and higher-priced quality guitars to elderly musicians. Similar to this, a meal-delivery business may focus on convenience for millennial clients and “tastes-just-like-mom-used-to-make” advantages for baby boomers.

Christensen, C.M. & Raynor, M.E. 2003.et.al.[11]Customer segmentation can be practiced by all businesses regardless of size or industry and whether they sell online or in person. It begins with gathering and analysing data and ends with acting on the information gathered in a way that is appropriate and effective.

Common characteristics in customer segments can guide how a company markets to individual segments and what products or services it promotes to them. A small business selling hand-made guitars, for example, might decide to promote lower-priced products to younger guitarists and higher-priced premium guitars to older musicians based on segment knowledge that tells them that younger musicians have less disposable income than their older counterparts. Similarly, a meals-by-mail service might emphasize convenience to millennial customers and “tastes-like-mother-used-to-make” benefits to baby boomers.

Cornelissen, J. 2002.et.al.[12]Customer segmentation can be practiced by all businesses regardless of size or industry and whether they sell online or in person. It begins with gathering and analysing data and ends with acting on the information gathered in a way that is appropriate and effective.

Clustering is one of the most common methods used in exploring data to obtain a clear understanding of the data structure. It can be characterized as the task of finding the subtitles and subgroups in the complete dataset. Similar data is clustered in many subgroups. A cluster refers to a collection of aggregated data points due to some similarities. Clustering is used in Market basket analysis used to segment the customers based on their behaviours and transactions.

Denscombe, M. 2007.et.al.[13]K Means Clustering is the most common and simplest Machine learning algorithm and it follows an iterative approach which attempts to partition the dataset into different “k” number of predefined and non-overlapping subgroups where each data point belongs to only one subgroup according to their similar qualities.

Unlike Supervised Learning, Unsupervised Learning has only independent variables and no corresponding target variable. The data is unlabelled. The aim of unsupervised learning is to model the underlying structure or distribution in the data in order to learn more about the data.

1. We are going to examine a dataset that is about mall visitors for segmentation. There is no any feature about label of customers. That is to say, we don’t have information about customer’s characteristics. We are going to try clustering clients through identifying similarities with machine learning algorithms. Segmentation of customers has a pretty significant position for companies in new marketing disciplines. Firms must reach to the right target audiences with right approaches because of costsI.

C. M. S. R. a. K. V. N. T. SajSaj.et al.[14]K Means Clustering is the most common and simplest Machine learning algorithm and it follows an iterative approach which attempts to partition the dataset into different “k” number of predefined and non-overlapping subgroups where each data point belongs to only one subgroup according to their similar qualities.

It is the distance or radius around each object.

The DBSCAN will process each and every object/points in this fashion and at the end it will obtain categorization of all the points as either core, border or noise points. Once the categorization of the points is obtained, the next step is to use them to construct the clusters. DBSCAN take up a core point and then look at the points which are inside its Epsilon radius circle and assign a Cluster label to those points, So the key idea is to give the same label to all the points inside the circle of a core point.

Multiple iterations will be run for different core points to assign Cluster label, please note algorithm will not assign new Cluster label to those points which have already be considered in earlier iteration.

El-Adly, M.H. 2007.et.al. [15]As clustering is unsupervised learning, need to analyse each cluster and have a definition with respect to business data because Clustering is always guided by some business rules. Once clusters are close to business rules, model will make sense.

For identifying, prioritizing, and targeting your best current customer segments, simply following it does not guarantee success. To be effective, you must prepare and plan for the various challenges and hurdles that each step may present, and always make sure to adapt your process to any new information or feedback that might change its output.

Additionally, you cannot force feed this process on your business. If the key stakeholders that will be impacted by the best current customers segmentation process do not fully buy-in, then the outputs produced from it will be relatively meaningless.

Forza, C. 2002.et.al. [16] If you properly manage the best current customer segmentation process, however, the impact it can have on every part of your organization — sales, marketing, product development, customer service, etc. — is immense. Your business will possess stronger customer focus and market clarity, allowing it to scale in a far more predictable and efficient manner.

Ultimately, that means no longer needing to take on every customer that is willing to pay for your product or service, which will allow you to instead hone in on a specific subset of customers that present the most profitable opportunities and efficient use of resources. That is critical for every business, of course, but at the expansion stage, it can often be the difference between incredible success and certain failure.

#### **Clustering and K-Means Algorithm**

Clustering algorithms generates clusters such that within the clusters are similar based on some characteristics. Similarity is defined in terms of how close the objects are in space.

K-means algorithm in one of the most popular centroid based algorithms. Suppose data set, D, contains n objects in space. Partitioning methods distribute the objects in D into k clusters, C1,……Ck , that is, Ci ⊂ D and Ci ∩Cj = ∅ for (1 ≤ i, j ≤ k). A centroid-based partitioning technique uses the centroid of a cluster, Ci, to represent that cluster. Conceptually, the centroid of a cluster is its centre point. The difference between an object p ∈ Ci and ci, the representative of the cluster, is measured by dist(p,ci), where dist(x,y) is the Euclidean distance between two points x and y.

**Algorithm:** The k-means algorithm for partitioning, where each cluster’s centre is represented by the mean value of the objects in the cluster. Input: k: the number of clusters, D: a data set containing n objects. Output: A set of k clusters. Method: Arbitrarily choose k objects from D as the initial cluster centres; repeat (re)assigns each object to the cluster to which the object is the most similar, based on the mean value of the objects in the cluster; update the cluster means, that is, calculate the mean value of the objects for each cluster; until no change.

# **REFERENCES**

1.Dr. R. Gardener “The Essential R Reference” (2014),

2.Abeyasekera, S. 2005. Multivariate methods for index construction. Department of Economic and Social Affairs Statistics Division.

3.Concepts of customer segmentation [http://www.business-science.io](http://www.business-science.io/)

4.Alexander, C.E., Wilson, C.A. & Foley, D.H. 2005. Agricultural input market

Segments: who is buying what? Journal of Agribusiness, 23(2):113-132

5. <https://labs.openviewpartners.com/customer-segmentation/>

6.Babbie, E. 2011. The basics of social research. 5th ed. Belmont, USA:

Wadsworth.

7. Source data related to our analysis has been collected from

8.Borgatti, S.P. 2005. Thinking theoretically. [Online] Available from:

<http://www.analytictech.com/mb870/handouts/theorizing.htm> [Accessed: 2012-

08-08] <https://github.com/mdancho84/orderSimulatoR/tree/master/data>

9. <https://www.kaggle.com/>

10..Canever, M.D., Van Trijp, H. & Van der Lans, I. 2007. Benefit-feature

Segmentation: A tool for the design of supply-chain strategy. Marketing

Intelligence & Planning, 25(5):511-533

11. <https://www.r-project.org/>

12. D. P. Yash Kushwaha, “Customer Segmentation using K-Means Algorithm,” 8th Semester Student of B.tech in Computer Science and Engineering

13. E. A. Onur DOĞAN1, “CUSTOMER SEGMENTATION BY USING RFM”.

14. M. S. R. a. K. V. N. T. Sajana, “A Survey on,” in Indian Journal of Science and Technology, Volume 9, Issue 3, Jan 2016.

15.El-Adly, M.H. 2007. Shopping malls attractiveness: A segmentation approach.

International Journal of Retail & Distribution Management, 35(11):936-

950.

16.Forza, C. 2002. Survey research in operations management: A process-based