Customer Data Processing in Internet of Things Environments Using RFM Analysis and K-means Clustering

V. Jai Kumar[1,*], Y.V.Bhaskar Reddy[2], V.S.Nishok[3], Vivek K. [4]

[1,2,3] Department of Electronics and Communication Engineering, QIS College of Engineering and Technology, Ongole-523272, Andhra Pradesh, India

[4] Department of Computer Science and Engineering, QIS College of Engineering and Technology, Ongole-523272, Andhra Pradesh, India

Emails: qispublications@qiscet.edu.in; bhaskar.reddy@qiscet.edu.in ; nishok.vs@qiscet.edu.in ;vivek.kolla@qiscet.edu.in

Corresponding Author Mail: qispublications@qiscet.edu.in

Abstract

Businesses, shops, banks, and other types of organizations all engage in cutthroat competition today. The retrieval techniques are responsible for arranging, processing, and listing the documents in the corpus in response to the user query. The strategies are distinct from one another in any of the processes stated above. As a result, the body of published work is crammed full of different retrieval theories and strategies. It is possible to combine the advantageous aspects of several different strategies to improve the retrieval systems' overall performance. Once again, the success of the merging process is dependent on the careful selection of the separate schemes that will be merged together. The selection process is carried out using optimization-seeking tools. The Genetic Algorithm is going to be used for this job. Using GA as the instrument and with the intention of evaluating both the even and the odd point crossover's effects, The odd and even point crossover is primarily employed as an exploratory tool, and its influence on the Internet of Things is evaluated throughout the information retrieval process. The enormous combination that results from the fusion function retrieval strategies and their weights may be understood as follows: The investigation is the only thing that can help us find the best answer out of all of these possible permutations. As a method of investigation, we made use of both odd and even point crossing. This exploration tool has a lack of convergence, which is a setback. It is possible to get a higher convergence rate by combining the genetic algorithm with tabu search, which is the best local search. In a scenario like this one, customer segmentation may be helpful in bringing in new customers while also helping to keep the ones you already have. An effective customer segmentation strategy for a business splits consumers into groups based on the RFM (Recency, Frequency, and Monitor) values of the Monitors. These groups have behavior in common. This will be of use to us in determining the possible customers for the firm. Following the completion of an RFM analysis, we use a conventional k-means method in order to extend the scope of the research to include clusters. Maintaining positive relationships with customers makes it much easier to market effectively to certain demographics of consumers, which in turn helps bolster a company's competitive position.

Keywords: RFM; K-Means Clustering; Internet of Things; customers; customer segmentation

DOI: https://doi.org/10.54216/JISIoT.070101
Received: April 02, 2022 Accepted: October 07, 2022

1. Introduction

This proposed work makes a feeble effort to understand the fundamental ideas behind information retrieval, the Internet of Things, genetic algorithms, and Tabu search. At the very conclusion of this proposed work, you will get a summary of the issue as well as the purpose of the work that we want to do. Text, as well as multimedia, are included in digital media [1]. The textual data have been around since prehistoric times. Because of this, the field of library science evolved throughout time. The idea behind library science was appropriated and adapted for the purpose of developing an autonomous system that is capable of dealing with digital text data. Because of the emphasis placed on autonomous systems, a new field has emerged, which is known as information retrieval (IR) [2]. Information Retrieval is a procedure that involves selecting the appropriate documents from a collection (or corpus) in order to fulfill a request for them [3].

The retrieval methods are helpful to the process since they define the different steps that are involved. The first two primary steps are indexing and matching the data. The user's query is compared to the indexed terms in order to determine whether or not they are relevant. During the process of matching, a distinct two-step approach is used in order to compute the level of relevance [4]. Indexing and matching have each been elaborately outlined by a number of different IR models. The models each have their own specialized indexing method, in addition to a variety of matching and similarity measurements that they each possess. Both the models and the matching procedure come with a wide variety of benefits and cons—information Retrieval Models.

In order to efficiently carry out the process of storing information and retrieving it, a number of different models have been developed [5]. These models provide an explanation of the mechanism that is used to store the keywords or index phrases [6]. These operations convert the documents into a format that is acceptable, making it simple to store them and retrieve them when necessary [7]. As a direct consequence of this, the procedure for recognizing the papers becomes less complicated. A model is a collection of premises and an algorithm for rating documents in relation to a user query. Models are used to predict how documents will respond to a search. A more formal definition of an IR model would be the quadruple [D, Q, F, R(qi, dj)], where D and Q are a set of logical views of documents and queries, R(qi, dj) is a ranking function that associates a numeric ranking to the query qi and the document dj, and F is the framework for modeling document and query models. There is no difference between rank R and strategy or scheme (qi, dj). The term "system" refers to the actual physical implementation of an IR method, which may have a number of different modes of operation or a number of different parameter values. Therefore, a single infrared (IR) system is capable of executing several distinct IR schemes, provided the appropriate adjustments are made.

A large number of companies strive to set themselves apart from the competition in the industry by delivering better goods and services. However, in terms of providing a respectable experience and an improved system, they should mean that customers do not leave in search of rival businesses because the enterprise exists because of them, and they assist the company in expanding the market, which generates revenue and profits. This is because customers are the reason the enterprise exists.

When we speak about our customers, we talk about their ages, where they come from, and their mindsets. However, the purchasing habits of our customers are the most important component to consider. As a consequence of this, each sector of the market needs goods at a range of different price points. Because of this, every segment of customers needs its own unique marketing strategy. In order to do this, we process procedures for customers in related industries.

They organize customers into groups or clusters depending on whether they are clustered indirectly or directly. Client data may include information such as exposure to several social media sites, transaction data, and the number of hours spent on a particular post. Nevertheless, this article describes some past business dealings carried out by a consumer in the UK who shops online. In order to solve this issue, the K-means clustering method, also known as unsupervised learning, is often recommended by data analysts to use in conjunction with the RFM model. RFM is an abbreviation that refers to a customer's recency, frequency, and monetary values.

This article will focus on recognizing the nature of customers (big customers, medium customers, and bottom customers), as well as their value so that business owners can evaluate which categories of customers create wholesome earnings and which don't, as well as what innovative marketing

DOI: https://doi.org/10.54216/JISIoT.070101
Received: April 02, 2022 Accepted: October 07, 2022

plans, they can incorporate to uplift growth in revenue. The big customer is the most valuable type of customer, followed by the medium customer, and then the bottom customer.

2. Related Work

Based on the shopping habits of the consumers, the purpose of this research is to determine which customers contribute the most to the success of the superstore's retail management and which customers contribute the least to the success of a given transaction. We are able to perform customer segmentation and provide management strategy recommendations by combining the RFM and K-means algorithms. On top of this foundation, using the RFM paradigm and the K-means algorithm [8] we conduct customer value analysis and profile creation. Identifying patterns and differences among consumers, predicting their actions, and providing them with improved alternatives and opportunities have become very necessary for fostering a connection between customers and companies [9].

Any business that is capable of understanding the requirements of each of its customers will be in a better position to serve those customers by developing individualized customer service strategies and providing focused customer services. The provision of structured customer service paves the way for the achievement of this comprehension. Customers in each group have characteristics that are comparable to those of the market [10]. Machine learning comes into play because it enables improved decision-making, which is accomplished by using simple algorithms to uncover previously hidden patterns in the data. The concept of a customer segmentation process achieves the notion of which group to target by using the clustering approach as part of the process. The use of offshore, time-consuming formal market assessments, which are frequently ineffective when there are too many customers has been largely supplanted by the development of an efficient algorithm for customer profiling thanks to the traditional data processing concepts that have had a significant influence on its creation.

In addition, the research classifies clients into subgroups according to a phenomenon known as the Recency, as well as another phenomenon known as the Recency, Frequency, and Monetary model. The technique for those segments' characteristics was used by applying the clustered element in Statistical Analysis, which was carried out with the assistance of a professional Minor. The algorithm classified the datasets according to the relevance, value, and regularity values of the customers that were included in the datasets. During the training of the model, we applied both of the k-means techniques. When the training phase of the data collection for customers is over, the information needed to determine the recency, frequency, and mean monetary values of customers in the different segments may be collected [7]. In addition to this, we designed a web model that makes it possible for e-commerce companies and business analysts to evaluate their customers using the framework that we created. Consequently, using this would allow you to target consumers in a suitable manner and build your company by maintaining a strong connection with your customers [8].

In the early 1970s, Fisher was the first person to combine two Boolean searches into a single one [10]. This was the beginning of early work on an Internet of Things approach that does not utilize training data. In his approach, one search is performed on the title word, while the other search is performed on manually constructed index phrases. Both searches are carried out simultaneously. He was successful in making substantial strides in improving the efficiency of the information retrieval system. This approach effectively combines the fewest possible retrieval methods, in contrast to the linear combination method, which may successfully combine a greater number of retrieval strategies (Fox and Shaw 1994). Through the use of weights, the linear combination approach evaluates each individual strategy [11]. Fusion functions are distinct from Comb-functions in that the relevance is calculated based on the rank that is given to documents, as opposed to the relevance scores approach that is used by Comb-functions. This is in contrast to the way that relevance is determined by Comb-functions. Two of these fusion systems, known as Board fusion and Condorcet fusion, are among the few that imitate the social voting schemes. [12] have all done a significant amount of work on comb functions [13]. Lee has suggested many new justifications and indications for the process of the Internet of Things. Experiments have been performed by him throughout the TREC data collection. He arrived at the conclusion that CombMNZ exhibited superior performance when compared to the other Comb-functions. The Comb-functions may combine additional techniques in a linear fashion. The Probabilistic technique is distinct from the Comb-functions in that it chooses the 33 strategies from the pool that have the greatest track record of success. The probabilistic model will only

DOI: https://doi.org/10.54216/JISIoT.070101
Received: April 02, 2022 Accepted: October 07, 2022

choose one strategy from the pool, and the rest of the strategies will be put on hold. As a result, evolutionary algorithms are used in order to choose the tactics that prove to be the most successful. Billhart came up with an Internet of Things technique that was based on heuristics. The retrieval score is combined using an algorithm that is genetic in nature. This kind of algorithm not only gives ratings to several separate strategies but also chooses the approach that has performed the best so that it may be combined with the others.

The fusion approaches combine the benefits of its individual member tactics into a single strategy to maximize their effectiveness. According to [14] when the tactics are combined, it has a tendency to leverage the following effects. i. Skimming effect ii. Effect of the Chorus, and iii. Effect of the Dark Horse The skimming effect is produced whenever the combining functions combine just the top-rank list from each of the several retrieval strategies. The Chorus Effect is produced when many retrieval strategies converge on the same conclusion that an item is relevant to a query. This kind of evidence for relevance is often more convincing than that produced by a solitary retrieval strategy. The Dark Horse Effect is a phenomenon in which one retrieval method may generate estimates of relevance that are particularly accurate (or erroneous), in comparison to the estimates that are produced by the other retrieval methods, for at least some of the items being retrieved. Researchers were able to make use of the benefits that these previously mentioned effects offered thanks to the careful design of the combination function. The success of the fusion procedures that need training data is dependent on the relevant input from the active participant. Our suggested function makes use of the benefits of the skimming effect. Therefore, the performance of these kinds of systems varies from user to user and is contingent upon that user's level of expertise in estimating the relevancy of the materials. Hence it is focusing on total user-independent fusion approaches. According to a study of the relevant literature, comb functions have been shown to have the best performance of all the functions in this category.

According to [15] the Internet of Things issue in the field of information retrieval is a multiobjective function. It is necessary to take into account more than one goal in order to come up with the best possible answer. As a result, a method that can optimize for several objectives is required. Information retrieval makes considerable use of the genetic algorithm, which is the best multi-objective function currently available [16] GA has been used for a variety of tasks, including query expansion, user profile matching, text summarising, and others [17]. The traditional IR has used up all of GA's potential to the fullest extent [18]. Because GA has restricted use in traditional IR, there is little doubt that hybrid search tools will become more popular. The challenge of the Internet of Things is more difficult than IR, and it requires an optimization tool that incorporates both hybrid and multi-objective search criteria. In addition, the hybrid GA plays a part in the integration of data.

3. Proposed Methodology

3.1 CROSSOVER

It accomplishes both exploration and exploitation via the processes of reproduction, genetic crossing, and mutation. It was determined that the crossing was the most significant of them. It is the GA's principal exploration operator and serves in that capacity. The numerous disruptions to the structure of the gene become the primary reason for the accomplishment of effective exploration. Therefore, it is only possible to envisage GA with crossover. Because the search space calls for a more specialized exploration tool, GA is moving toward a number of operators that fit this description. As a result, the fundamental crossover operator has been subjected to a number of modifications, each of which comes with its own set of benefits and drawbacks. In their most basic form, these crossover operators may be divided into two categories: single-point crossover and multipoint crossover.

Single Point Crossover

In this operator, the crossover operation has been carried out in a single position. Based on the selection, the chromosome structure gets destructed. The basic operation has been explained in the following Figure 1.

DOI: https://doi.org/10.54216/JISIoT.070101
Received: April 02, 2022 Accepted: October 07, 2022

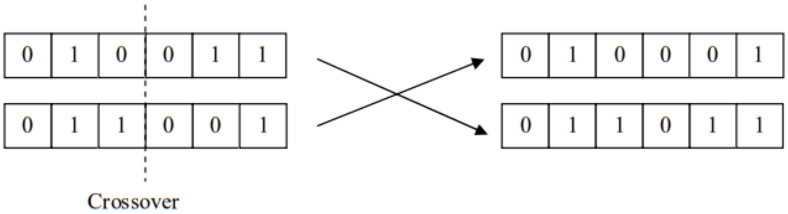

Figure 1: Crossover technique

Algorithm-

The recommended procedure may, more or less, be divided up into four distinct parts. The information that is linked with this is described as

The first step, known as "data preparation," involves doing an initial review of the data in order to extract or recognize patterns via the use of statistics or visualizations. This particular visualization configuration is helpful in locating one-of-a-kind customers, the proportion of total orders, data information, description inconsistencies, stock code checks, and null value checks. In addition, it may be used to track down and erase incorrect transactions, customer identification numbers, and other data.

Execution of the RFM Analysis is the second step. After the data has been pre-processed, it should be inspected for client spending amounts, transaction records, and transaction frequency. RFM analysis is a method that is often used in database marketing for the purpose of customer identification and segmentation. Each customer's RFM score is based on their performance across three categories.

The number of days that have passed since a consumer made their most recent purchase as of the reference date is referred to as "recency." The degree of recency between customer visits to a shop and those visits is correlated to the number of visits. The length of time that passes between a customer's successive two purchases is referred to as the customer's frequency. The more regularly customers shop at the establishment, the higher their score for Frequency will be. Financially speaking, this refers to the amount of money spent by a customer over a certain amount of time.

Step 3: Using the RFM values of the consumers as a guide, use the K-Means clustering technique to split the customers into groups. K-Means is applied twice to determine the amount gained for recent and frequent transactions so that it can be used to categorize customers according to the amount produced with recent transactions and the amount created with frequent transactions. This helps to ensure that customers are categorized correctly.

Step 4: Cluster assessment The number of clusters is denoted by "K." Use either K = 3 or K = 5 to calculate the optimal cluster size, which varies based on the value. Compare the sales after the analysis; doing so will help you identify the customers who have the largest sales volume, frequency, and recent sales.

DOI: https://doi.org/10.54216/JISIoT.070101
Received: April 02, 2022 Accepted: October 07, 2022

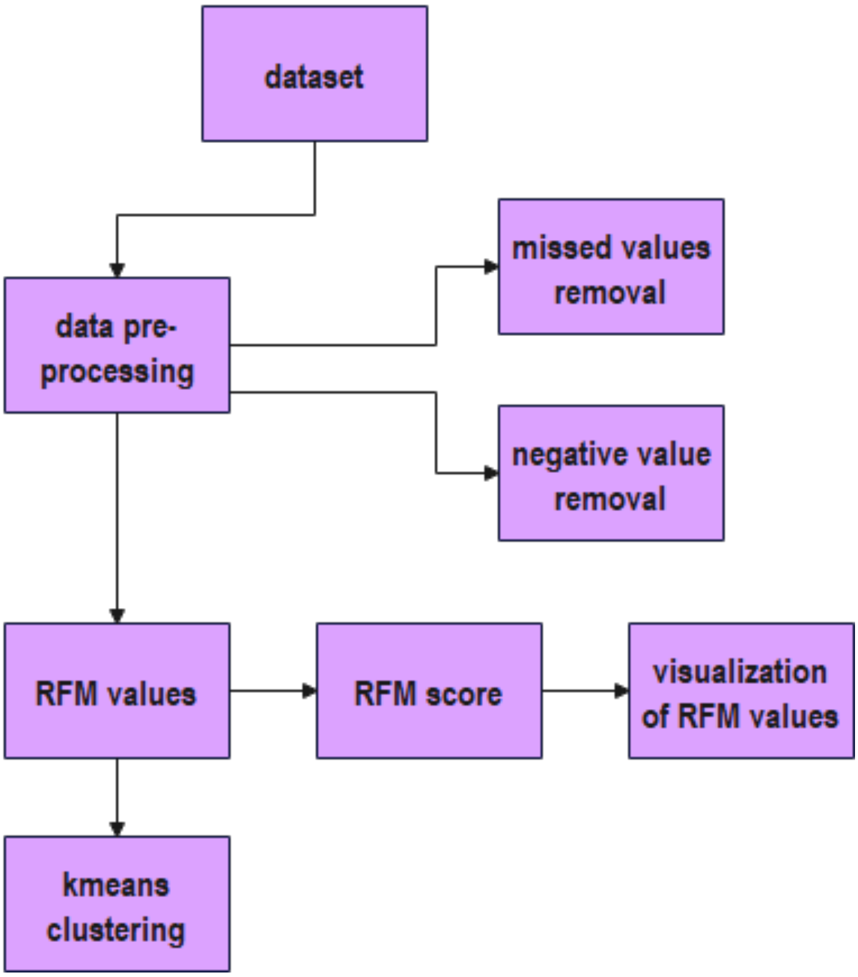

Figure 2: Flow Chart for the proposed work

In contrast to the other types of crossover operators already in existence, the odd and even point crossover is used as a powerful exploration tool. It was able to explore the search space effectively. In the preceding proposed work, both the method and the results achieved by making use of these operators are covered in detail.

The performance of the IR system is presented in the proposed work before this one, but there is no discussion of the tool attributes. This proposed work's original purpose was to investigate the convergence characteristic of the exploration tool, more specifically the odd and even point crossing. That optimization tool's convergence feature must be considered very important in the evaluation process. The convergence of a GA is often affected by three fundamental operators, namely reproduction, crossover, and mutation. There is no way to declare with certainty that these three operators each have an equal amount of impact on the convergence. Each one has an influence of its own. If the other two operators are held constant, then the impact of these operators may be investigated. In other words, one must not make use of the same two operators in order to conduct an analysis of the effect exerted by the third operator. In our situation, there is a shown interest in exploration that is solely focused on the crossover. Because of this, the reproduction and mutation mechanisms employed are identical to those seen in traditional GA. Because of this design, we were able to do the convergence analysis, which takes into account the crossover. In our situation, the crossover is a point-based crossover that alternates between odd and even.

There are an excessive number of possible solutions to the selection issue posed by the Internet of Things in information retrieval. It is important to investigate every possible combination. This demand causes periodic disruptions in the gene structure, which in turn leads to a gradual approach to convergence. The price was paid for gradual convergence rather than the requirement for robust

DOI: https://doi.org/10.54216/JISIoT.070101
Received: April 02, 2022 Accepted: October 07, 2022

research, which was necessary given the nature of our challenge. It is going to make for an excellent testing ground for the projected crossover operator. The functionality of the suggested operator was scrutinized in the prior proposed work. In this proposed work, we are going to do the crucial task of analyzing convergence, which is a trait of vital importance. The convergence property was evaluated by taking into account the two factors listed below. 1) The individual's overall level of fitness on average 2) The degree to which the average values of fitness change from one generation to the next.

3.2 Reason for Selecting this Two Parameter

In traditional GA, there are two different factors that are utilized to determine when to quit. The first factor is the predetermined number of generations, and the second factor is the shift in the average values of physical fitness that occurs between each generation. Each of these two options comes with its own set of benefits and drawbacks. In the first scenario, the investigation will be halted after a certain number of generations even if there is effective variation between succeeding generations. This is the case regardless of whether or not the variation is additive or multiplicative.

This is a signal that there are certain places that have not yet been investigated, and it is utilized in that capacity. If all regions are investigated, then the amount of genetic variation that exists between two succeeding generations will be at its lowest. When it comes to the second scenario, the precise number of generations is still being determined. The difference between the two generations that came after one another was utilized as an indicator. In the event that it is found to be inside the range, the exploration will be terminated. It comes with its own set of disadvantages. There may be very little variance at the beginning of the process. If at that time a feeble exploration tool is discovered, the GA will immediately cease its exploration while it is still in the initial condition. It indicates that there will be a significant amount of uncharted territory accessible. A hybrid termination condition is offered as a means of addressing both of these potential issues. In this hybrid termination scenario, both the number of generations and the difference in average fitness levels between each succeeding generation have been chosen. In any one of these three scenarios, the difference between the two generations that have come before plays a significant role. When dealing with a predetermined amount of generations, one has no choice but to rely on the knowledge of an expert. However, the situation is quite different in the other two cases. As a result, the difference in the average fitness value between the two generations that came after the previous one has been chosen as the indicator.

3.3 Convergence Analysis

The convergence analysis has been completed using two different sections. The first one has a tendency to investigate the convergence of traditional GA. The second category typically performs an analysis of the GA that is based on odd and even point crossovers. Experiments are carried out under these two headings by altering the total number of bits that are used in the encoding process. Because of these analyses, we were able to confirm that the GA algorithm that uses odd and even point crossovers converges slowly It is worth mentioning that the coefficients subscribed by the index N are of neutrosophic values.

4. Experimental Analysis

Step 1- Preparation of Data:

Numpy carries a high multi-dimensional array of objects and manipulation capabilities. Pandas is a toolkit that gathers data formats and analytic tools that are fast and simple. Numpy is often used as a foundation.
The sklearn Standards Scaler package can be used for normalization. Seaborn is a program for visualizing graphs.

Classifier, regression, cluster, and dimension reduction are just a few machine-learning and standard statistical algorithms included in the sklearn package. I used Kaggle's Retail Sales Data, which offers several columns that would benefit the study, including Invoice-No, Invoices, Customer-Id, Retail Price, Volume, and others. The data about online retail can be found here.

DOI: https://doi.org/10.54216/JISIoT.070101
Received: April 02, 2022 Accepted: October 07, 2022

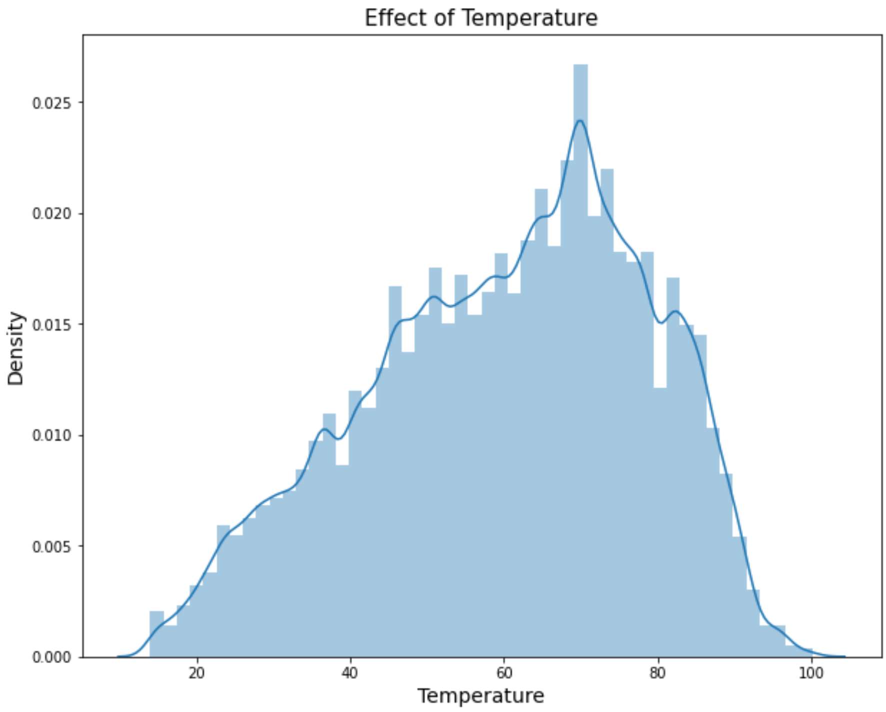

Figure 3: Imported Data

It is responsible for data pretreatment, such as determining whether any values are missing or null. The dataset's details and description will be displayed after that.

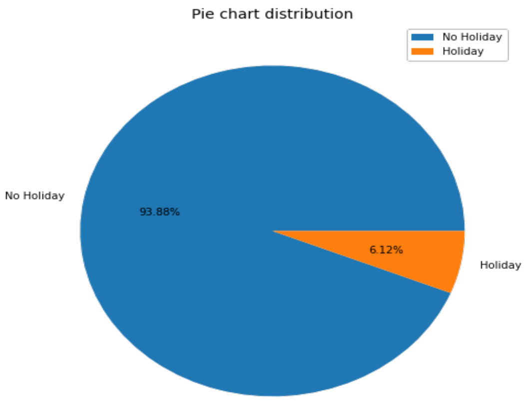

Figure 4: Data Description

It defines the score, which keeps track of rows inside the table, and even the table's average, standard deviation, variance, min and max quartiles, etc.

DOI: https://doi.org/10.54216/JISIoT.070101
Received: April 02, 2022 Accepted: October 07, 2022

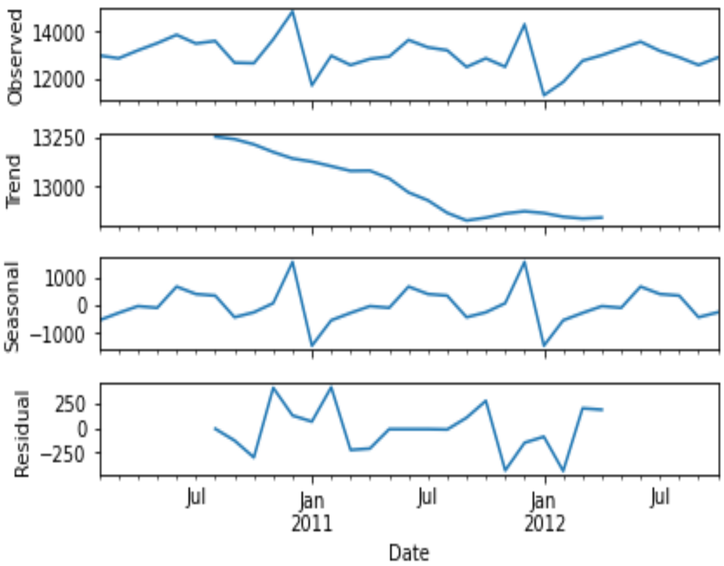

Figure 5: Null Value Checking

It denotes the absence of any null values.

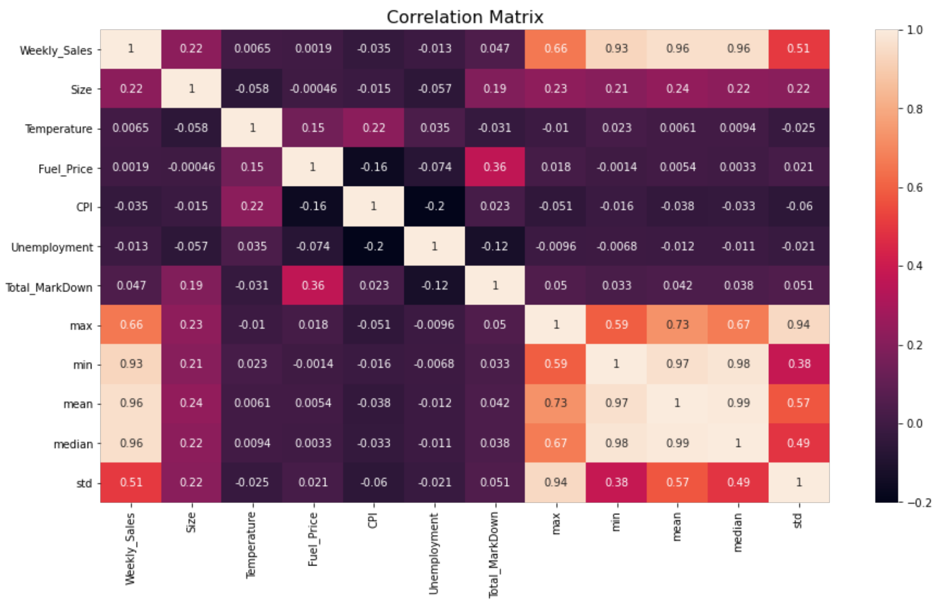

Figure 6: Prepared Data the date info has been imported, which can be used to calculate recency.

Step 2-Analysis of RFM:

Recency is calculated by the size of time between both the correct date and the last Purchased date by each customer.

Calculate the total number of trades per customer for a particular period. Calculate each consumer's total purchase price for monetary.

We'll now divide the statistics into segments using quantiles. On the order of 1 to 4, we rank RFM value. One is the best deal, while four is perhaps the worst value. To acquire a final RFM rating, add the individual R, F, and M score figures together.

DOI: https://doi.org/10.54216/JISIoT.070101
Received: April 02, 2022 Accepted: October 07, 2022

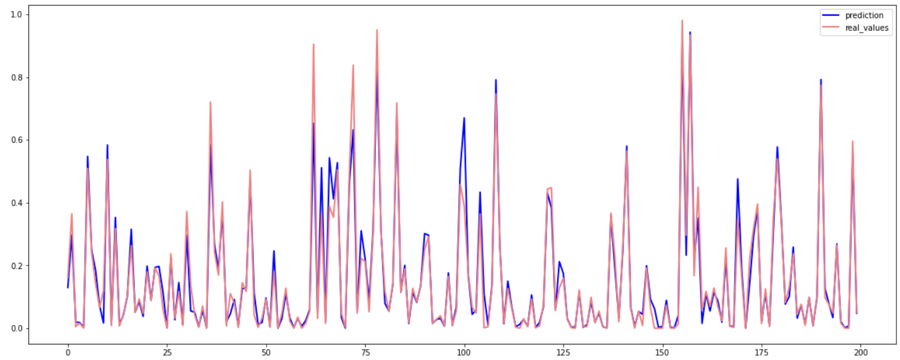

Figure 7: RFM Table

Since the data is considerably skewed, we will utilize log transforms to reduce the variability of each factor. I add a small constant since logistic regression needs all items to be positive. Analyzation Using Plotting

Fig .7: RFM Analysis Step 3-K-Means Algorithm:
It tries to keep clusters as widely apart as feasible while making sub-pieces of data as identical as practicable.
It allocates datasets to cluster with the smallest feasible summation of its euclidean difference from centroids. The main goal of the k-means clustering algorithm, which fits the algorithm, is to reduce the distance between the nodes in a cluster and their centroid.
Our data collection needs to be clustered and represented by Kmeans. fit(argument). Because fit(X, y) and predict(X) routines
should be used for everything, according to early Sklearn decisions. And due to backward compatibility, it probably won't change.

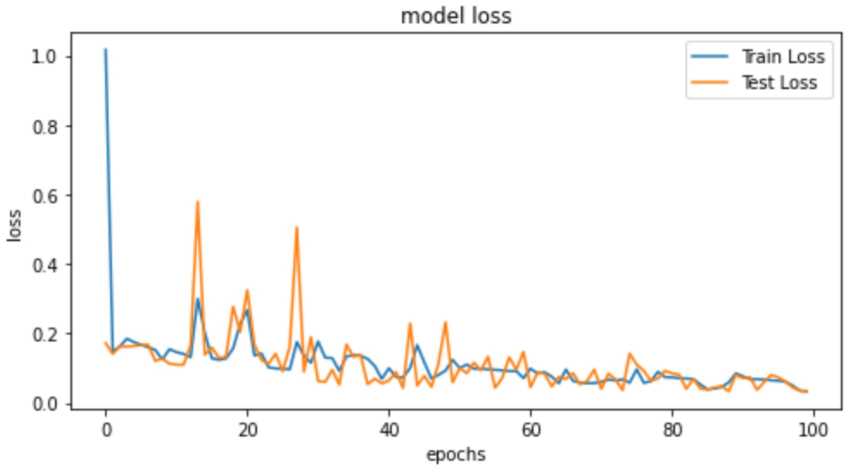

Figure 8: Elbow Graph Step 4-Evaluation of Clusters:
-We can see the clusters here. Using the k-value.

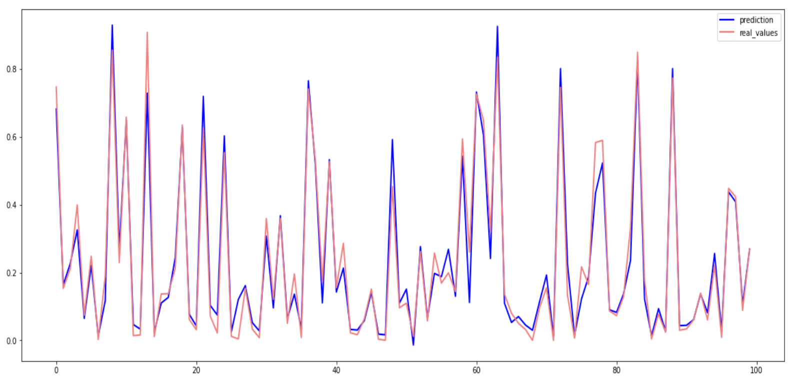

Figure 9: Visualization of cluster

DOI: https://doi.org/10.54216/JISIoT.070101
Received: April 02, 2022 Accepted: October 07, 2022

Interpretation of the clusters formed using k-means.

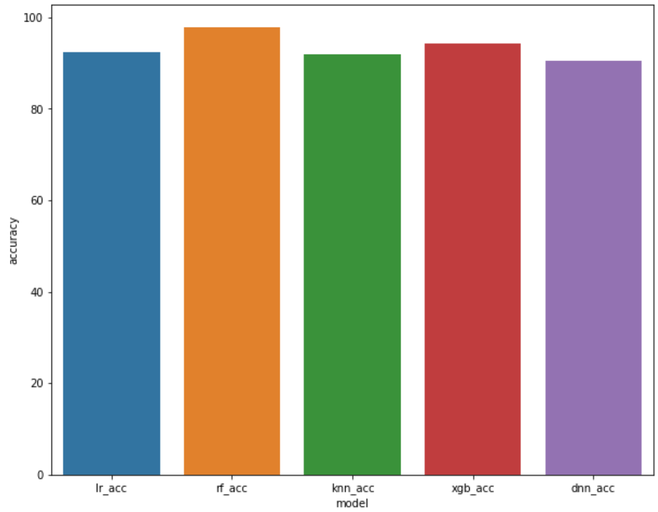

Figure 10: Final RFM Table

This aids in locating the clientele with the highest sales volume, frequency, and recent sales

6. Conclusion

We concluded that the very first cluster belongs to the portion of "Best-customers," which we already knew to be customers who had recently made a purchase (R=1), who were regulars (F=1), and who paid the most (M=1) in total. Clientele participating in the second activity is more likely to be going through the process because their most recent transaction occurred quite some time ago (R=4), they bought fewer items (F=4) and spent less money (M=4) than in the previous activity. The company needs to come up with more effective programs to assist individuals in securing permanent seats. Considering that they have not yet purchased it in the long run (R=3) but have used it to acquire and spend a great deal, the third group is more likely to be associated with the "Almost Lost" section. • The very devoted customers who spend a fortune make up the final core of the customer base.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] A. M. A. Zamil and T. G. Vasista, "Customer Segmentation Using RFM Analysis: Realizing Through Python Implementation," *Pacific Bus. Rev. Int.*, vol. 13, no. 11, pp. 24-36 WE-Emerging Sources Citation Index (ESCI), 2021.

[2] J. Wu *et al.*, "An Empirical Study on Customer Segmentation by Purchase Behaviors Using an RFM Model and K -Means Algorithm," *Math. Probl. Eng.*, vol. 2020, no. November 2017, 2020, doi: 10.1155/2020/8884227.

[3] O. Doğan, E. Ayçin, and Z. A. Bulut, "Customer Segmentation by Using RFM Model and Clustering Methods: A Case Study in Retail Industry," *Int. J. Contemp. Econ. Adm. Sci.*, vol. 8, no. 1, pp. 1–19, 2018, [Online]. Available: www.ijceas.com

[4] G. Noida, "Customer Segmentation Using," pp. 1–7.

[5] C. Pascal, S. Ozuomba, and C. Kalu, "Application of K-Means Algorithm for Efficient Customer Segmentation: A Strategy for Targeted Customer Services," *Int. J. Adv. Res. Artif. Intell.*, vol. 4, no. 10, pp. 40–44, 2015, doi: 10.14569/ijarai.2015.041007.

[6] S. Idowu, "Customer Segmentation Based on RFM Model Using K-Means, Hierarchical and Fuzzy C- Means Clustering Algorithms," no. August 2019, 2020, doi: 10.13140/RG.2.2.15379.71201.

[7] D. Devarapalli, S. Veera, V. Satya, S. Geddam, A. S. Sravya, and A. P. Devi, "Analysis of RFM Customer Segmentation Using Clustering Algorithms," no. February 2022.

DOI: https://doi.org/10.54216/JISIoT.070101
Received: April 02, 2022 Accepted: October 07, 2022

[8] R. Shirole, L. Salokhe, and S. Jadhav, "Customer Segmentation using RFM Model and K-Means Clustering," *Int. J. Sci. Res. Sci. Technol.*, pp. 591–597, 2021, doi: 10.32628/ijsrst2183118.

[9] Ibrahim, S. Jafar Ali, and M. Thangamani. "Prediction of Novel Drugs and Diseases for Hepatocellular Carcinoma Based on Multi-Source Simulated Annealing Based Random Walk." Journal of medical systems 42, no. 10 (2018): 188. https://doi.org/10.1007/s10916-018-1038-y

[10] S. Jafar Ali Ibrahim and Thangamani, M "Proliferators and Inhibitors Of Hepatocellular Ca," Sonoma," International Journal of Pure and Applied Mathematics (IJPAM) Special Issue of Mathematical Modelling of Engineering Problems Vol 119 Issue. 15. July 2018 ISSN 1311-8080, https://acadpubl.eu/hub/2018-119-16/1/94.pdf

[11] Bleiholder, J., & Naumann, F. (2009). Internet of Things. *ACM computing surveys (CSUR)*, *41*(1), 1-41.

[12] Castanedo, F. (2013). A review of Internet of Things techniques. *The scientific world journal*, *2013*.

[13] Manal Nasir ,Ahmed N. Al-Masri. (2021). Multi-source Heterogeneous Ecological Big Data Adaptive Fusion Method Based on Symmetric Encryption. Journal of Fusion: Practice and Applications, 5 (1), 08-20.

[14] Bar-Shalom, Y., Willett, P. K., & Tian, X. (2011). *Tracking and Internet of Things* (Vol. 11). Storrs, CT, USA:: YBS publishing.

[15] Hall, D. L., & Llinas, J. (1997). An introduction to multisensor Internet of Things. *Proceedings of the IEEE*, *85*(1), 6-23.

[16] Esraa Mohamed. (2020). The Relationship between Artificial Intelligence and Internet of Things: A quick review. Journal of Journal of Cybersecurity and Information Management, 1(1), 30-34.

[17] Meng, T., Jing, X., Yan, Z., & Pedrycz, W. (2020). A survey on machine learning for the Internet of Things. *Information Fusion*, *57*, 115-129.

DOI: https://doi.org/10.54216/JISIoT.070101
Received: April 02, 2022 Accepted: October 07, 2022