Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 7, July : 2023

Targeted marketing campaigns using RFM and Machine Learning

[1*] Dr. B.V. Ramana, Professor, Information Technology,
Aditya Institute of Technology & Management, Tekkali, India,
ramana.bendi@gmail.com.
[2] Dr. Kavitha, Assistant Professor, Information Technology,
Aditya Institute of Technology & Management, Tekkali, India,
kavitha.akki@adityatekkali.edu.in
[3] Smita Rani Sahu, Assistant Professor, Department of Information Technology,
Aditya Institute of Technology & Management, Tekkali, India.
smitharanisahu.it@adityatekkali.edu.in
[4] B. Manideep, Computer Science and Engineering,
Aditya Institute of Technology & Management, Tekkali, India,
bendimanideep@gmail.com
[5] Jayavardhanarao Sahukaru, Assistant Professor, Computer Science and Engineering
Aditya Institute of Technology & Management, Tekkali, India,
jayavardhanarao.mca@gmail.com

1. Abstract.

Target marketing plays a major role for the growth of any business. This paper proposes a customer segmentation Model using k-Means Clustering and RFM analysis. Essentially It is based on customer-centric marketing strategy that is analyzing the customer demographics and purchase behavior. The k-means clustering algorithm Acquires more Accuracy compared to K-Medoids and DBSCAN method and thereby it improves customer retention and loyalty.

2. Introduction

Segmentation is an important so that the company can create profitable segments and react to the selected segment based on its competitive advantages. This study is to investigate whether customer segmentation can be achieved better by using RFM analysis integrated with cluster analysis. Customer segmentation is categorized into Demographic Segmentation, Behavioral Segmentation, Geographic segmentation and Psychographic Segmentation.

Demographic segmentation assumes that customers with similar demographic profiles will exhibit similar purchasing patterns. Behavioural segmentation divides consumers into groups according to their observed behaviours. Geographic segmentation divides markets according to geographic criteria. Psychographic segmentation is measured by studying the activities, interests, and opinions of customers.

3. Related Work

Rahul shirole et.al., Proposed a methodology for extracting RFM parameters for each customer and form clusters[1]. Saurabh Patil et.al., Proposed to establish the relation between marketing campaign and customer segmentation along with the enhancement using the RFM approach [2]. Mussadiq Abdul Rahim et.al., Proposed a new clustering method using item preference based on RFM for recommendation system [3]. N.Vijayalaksmi et.al., Proposed the best approach to customer segmentation based on the weights of RFM attributes [4].

Hoang thi ha et.al., Investigated the influence of problems with data accuracy using RFM analysis for customer segmentation for two real-life direct marketing data sets [5]. Dr.Y.Suresh et.el., declared to segment hotel customers corresponding to their RFM score [6]. Kankamol Chairuangc et.al., Proposed a two-stage clustering- classification model integrated the RFM attribute and K-means algorithm for clustering the patients and optimizing health care services [7].

Cormac Dullaghan et.al., Proposed a model using the RFM approach to select attributes to cluster customers into different groups [8]. A.Mohammed et.al., Suggested a approach of customer classification. RFM model to analyze and estimate customer behavior using clustering algorithms and data mining techniques [9]. Monireh Hosseini et.al., developed an algorithm for generating all RFM patterns from customers for managerial decision- making [10].

4.Methodology:

In this analysis retail dataset from UCI machine learning repository was considered. In the data pre-processing process, customer data and transaction data need to separate for further analysis. The appropriate variables such as customer name, business size, industry, and company type are selected to be used as customer demographics. According to the RFM analysis model, purchasing transactions need to be transformed into three essential variables, that are Recency, Frequency and Monetary. In processing phase, RFM data are used in the first phase of segmentation as the input of K-means clustering. In the second phase, each cluster resulted from a previous phase is again clustered into new clusters using customer demographics data. Finally, each clusters need to be interpreted and identified the main characteristics. The entire framework is represented in fig 1.

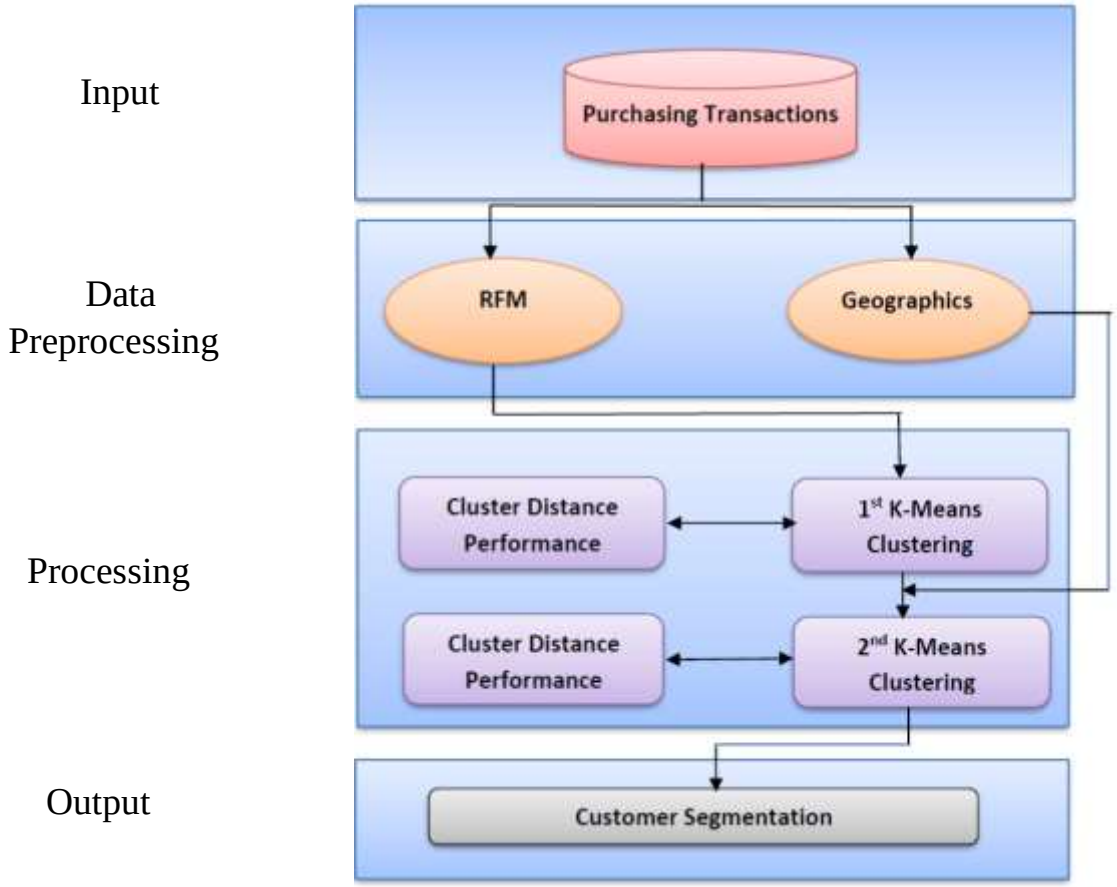

Fig 1: Conceptual Framework

RFM analytic model represents customers' consumption behaviours based on the transaction database. It has three attributes. Recency states "the period of time starting from last purchasing and present". Frequency states that "the number of transactions in a particular period". Monetary states "the total consumption money amount in a particular period".

5. Dataset

In this study online retail dataset was considered from UCI repository [11]. This dataset includes all the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online retail. It has 8 attributes and 541909 instances. The number of instances were considered 433527and 108381 for training and testing respectively. Fig 2 shows sample dataset and it is the dataset of a store whose customers are coming from all over the world. It includes information such as invoice number, invoice date, customer id, stock code, description of the product, purchased quantity and country where the customer lives.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 7, July : 2023

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
0	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	12/1/2010 8:26	2.55	17850	United Kingdom
1	536365	71053	WHITE METAL LANTERN	6	12/1/2010 8:26	3.39	17850	United Kingdom
2	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	12/1/2010 8:26	2.75	17850	United Kingdom
3	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	12/1/2010 8:26	3.39	17850	United Kingdom
4	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	12/1/2010 8:26	3.39	17850	United Kingdom

Fig. 2. Sample Online Retail dataset

The fraction of the RFM scores is represented in fig. 3.

	CustomerID	Recency	Frequency	Monetary
0	12346.0	325	1	77183.60
1	12747.0	22	103	4196.01
2	12748.0	4	4596	33719.73
3	12749.0	22	199	4090.88
4	12820.0	44	59	942.34

Fig. 3. Fraction of RFM scores

6. Results:

RFM scores are represented w.r.t Recency, Frequency and Monetary in fig 4.

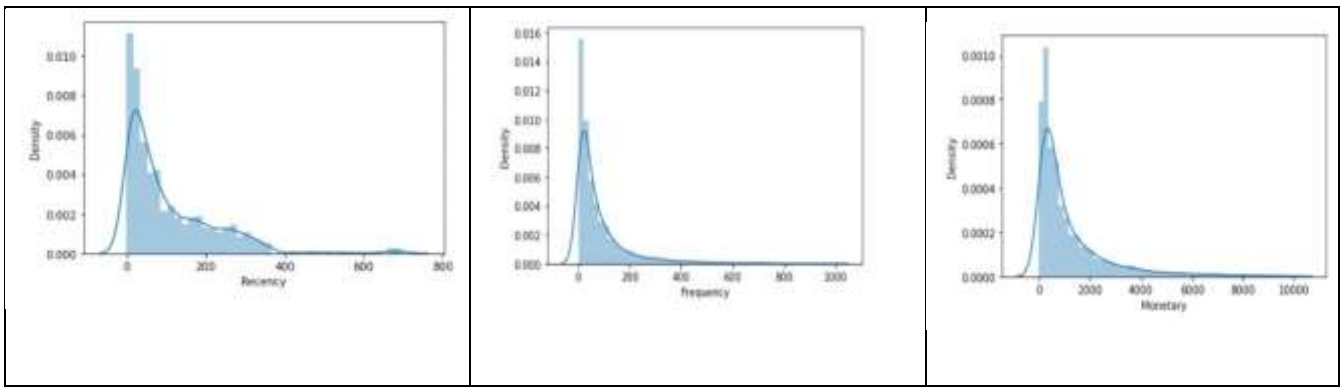

Fig.4. RFM scores

Customers are grouped based on loyalty and are recognized as platinum, gold, silver and bronze. Fig 5 shows the sample customer data with loyalty level and represented graphically in fig 6.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 7, July : 2023

	CustomerID	Recency	Frequency	Monetary	R	F	M	RFMGroup	RFMScore	RFM_Loyalty_Level
0	12346.0	325	1	77183.60	4	4	1	441	9	Silver
1	12747.0	22	103	4196.01	2	1	1	211	4	Platinum
2	12748.0	4	4596	33719.73	1	1	1	111	3	Platinum
3	12749.0	22	199	4090.88	2	1	1	211	4	Platinum
4	12820.0	44	59	942.34	2	2	2	222	6	Gold

Fig.5. Loyalty level of customer

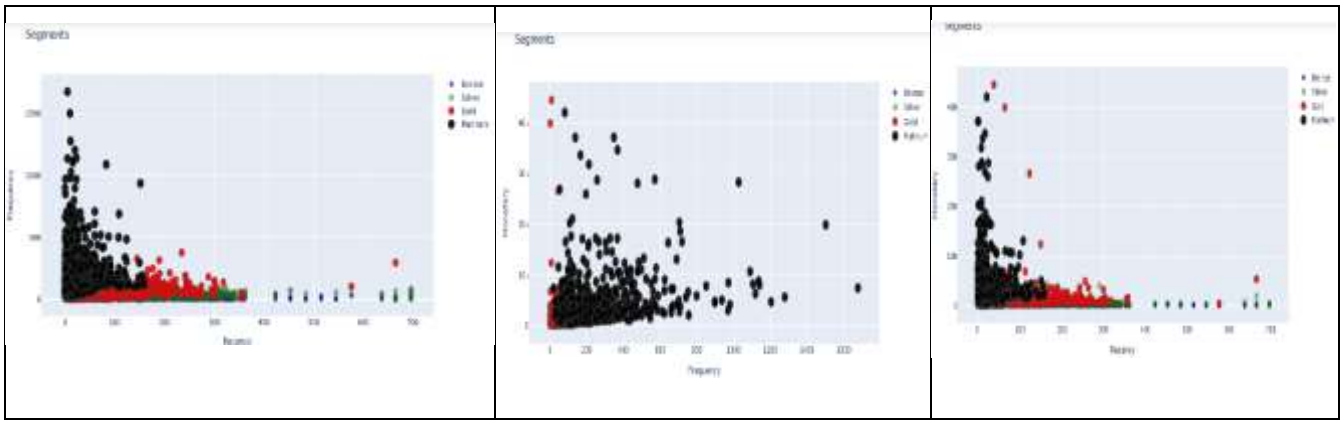

Fig.6. Loyalty segment frequency for RFM

K-Means Clustering Algorithm utilized on customer loyalty data for improving the performance of customer segmentation and represented in fig 7 and represented graphically in fig 8.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 7, July : 2023

CustomerID	Recency	Frequency	Monetary	R	F	M	RFMGroup	RFMScore	RFM_Loyalty_Level	Cluster
12346.0	325	1	77183.60	4	4	1	441	9	Silver	0
12747.0	22	103	4196.01	2	1	1	211	4	Platinum	1
12748.0	4	4596	33719.73	1	1	1	111	3	Platinum	1
12749.0	22	199	4090.88	2	1	1	211	4	Platinum	1
12820.0	44	59	942.34	2	2	2	222	6	Gold	0

Fig.7. Cluster identification for Loyalty segment data

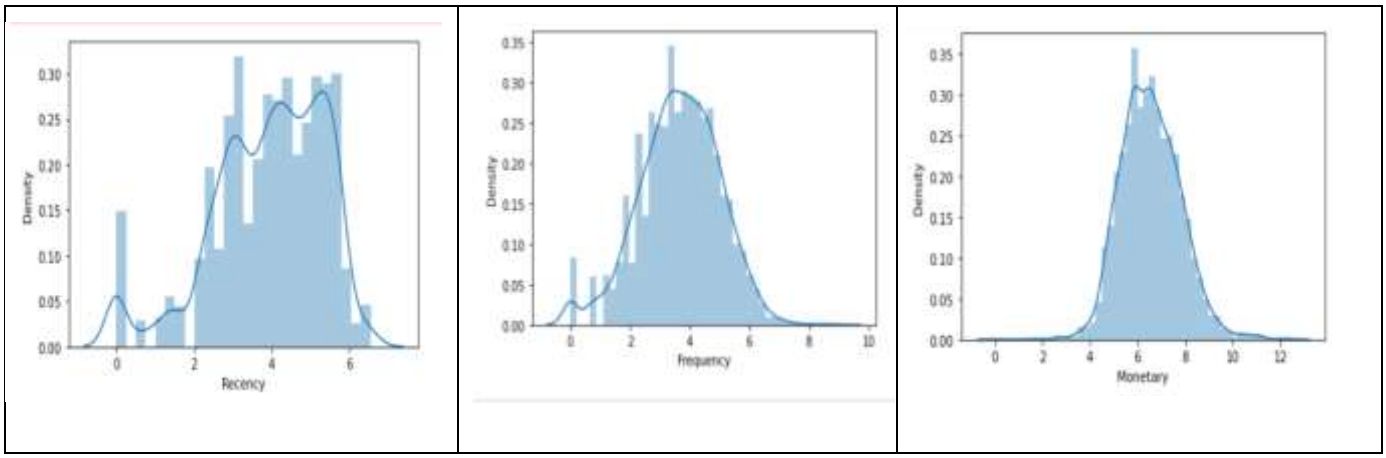

Fig.8. Cluster identification of RFM for Loyalty segment data

7.Conclusion

Organizations should better understand their components. Especially, it is essential for businesses, they should have detailed understanding about their customers' characteristics, behaviors, demographics, etc. with this Motivation, this Project is Implemented by RFM modelling and K-Means clustering algorithm. The k-means clustering algorithm Acquires more Accuracy compared to K-Medoids and DBSCAN method. The goal of K-Means is to group data points into distinct non-overlapping sub groups. One of the major application of k means clustering is segmentation of customers to get a better understanding of them which in turn could be used to increase the revenue of the company. In this implementation the Project acquires 96.7% Accuracy.

7. References

[1].Rahul shirole, Laxmiputra salokhe, and Saraswati Jadhav .(2021).” Customer Segmentation using RFM Model and K-Means Clustering”. IJSRST. Volume 8, Issue 03.

[2].Saurabh Patil, Hasnath khan, Sachin Mehta and Prof Umakant Mandawkar. (2021).” Study of Customer Segmentation using K-Means clustering and RFM modelling”. Journal of engineering sciences. Volume 12, Issue 06.

[3]. Mussadiq Abdul Rahim, Muhammad Mushafiq, Salabat khan and Zulfiqar Ali arain.”RFM based repurchase behavior for customer classification and segmentation”. Journal of Retailing and consumer services 61 (2021) 102566.

[4].K.Keerthi, G.Lakshmi tirupathamma, N.Vijayalaksmi, D.Aparna, U.Vineela.(2021).”Customer Segmentation Analysis and visualization”. IJSRCSEIT. Volume 7, Issue 01.

[5].Musthofa galih Pradana, Hoang thi ha.(2021).”Maximizing Strategy Improvement in mall customer segmentation using K-Means Clustering”. Journal of Applied Data sciences. Volume 02, NO 1, pp.19-25.

[6].Dr.Y.suresh,Dr.J.senthilkumar,Dr.V.Mohanraj,Dr.Selvarajkesavan.(2020).”Customer Segmentation using machine learning in python”. Turkish Journal of physiotherapy and rehabilitation.

[7]. Rahma Wati Br Sembiring Berahmana, Fahd Agodzo Mohammed, Kankamol Chairuangc.(2020).”Customer segmentation based on RFM model using K-Medoids and DBSCAN method”. Volume 11, No.1.

[8].Cormac Dullaghan, and Eleni Rozaki.(2017).”Customer segmentation based on RFM model and K-Means Clustering Technique”. IJDKP. Volume 7,No 1.

[9].J. tu toit, R.Davimes, A.Mohammed, K.patel and J.M. Nye.(2016).”Customer segmentation using unsupervised learning on Daily energy load profiles”. Journal of Advances in Information Technology. Volume 7, No.6.

[10].Monireh Hosseini and Mostafa Shabani.(2016).”New Approach to Customer Segmentation based on changes in customer value”. Journal of Marketing Analytics. Volume 3, No.3.

[11]. http://archive.ics.uci.edu/ml/datasets/online+retail