iJRASET
International Journal For Research in Applied Science and Engineering Technology



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 12 **Issue:** I **Month of publication:** January 2024

**DOI:** https://doi.org/10.22214/ijraset.2024.58144

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 12 Issue I Jan 2024- Available at www.ijraset.com*

# Customer Segmentation

Ishita Shah[1], Sanghavi Kedar[2], Gracy Wagh[3], Prof. Deepali Srikhande[4]
*Vidyalankar Institute of Technology*

***Abstract:*** ***Effective marketing involves targeting specific customer groups with personalized products, services, and campaigns, making customer segmentation a crucial strategy in modern business. This paper introduces a pioneering customer segmentation method that utilizes machine learning techniques to accurately and efficiently segment customers based on their behaviors, demographics, and transaction history. By combining transfer learning, Rfm (recency, frequency, monetary) modeling, and clustering algorithms like K-means, our approach generates meaningful customer segments, offering valuable insights for personalized marketing and improved customer experiences. We showcase the positive impact of our method on a real-world dataset, displaying noteworthy enhancements in marketing effectiveness and customer satisfaction.***
***Keywords: machine learning, marketing, transfer learning, rfm model.***

## I. INTRODUCTION

In today's fast-paced business world, the needs and wants of customers are constantly shifting, making it a vital task for businesses to stay up-to-date and effectively connect with a diverse range of clients.

The concept of customer segmentation, which involves organizing customers into distinct categories based on common qualities, has become a crucial strategy for tackling this challenge. From marketing tactics tailored to specific groups to improving overall customer satisfaction, segmentation has proven to be a valuable tool in the ever-increasing world of data-driven business. As companies navigate through the overwhelming amount of information on customer behavior, segmentation has emerged as a strategic approach to interpreting and utilizing this valuable data.

In the current climate, a one-size-fits-all strategy for customer engagement is no longer effective. Nowadays, customers demand personalized interactions, tailored content, and products that speak to their unique needs. Therefore, businesses must pivot and utilize advanced data analysis techniques to deconstruct their customer demographics and identify what sets each segment apart. This not only streamlines marketing efforts, but also allows companies to adapt to evolving market trends, foster stronger connections with customers, and ultimately gain a competitive edge.

This paper seeks to enhance the ongoing discourse on customer segmentation by presenting cutting-edge techniques and methodologies that harness advanced analytics, machine learning, and data-driven strategies. By examining demographics, behaviors, and transaction records, we delve into the complexities of segmentation.

Through our research, we aim to tackle the obstacles and potential benefits of customer segmentation, providing a holistic understanding of how this practice can enable businesses to better cater to their customers and excel in today's data-driven landscape.

## II. LITERATURE SURVEY

*1)* Authors uMahmoud SalahEldin Kasema , Mohamed Hamadab , Islam Taj-Eddinc says in the paper title " Customer Profiling, Segmentation, and Sales Prediction using AI in Direct Marketing" that Since most of the purchases are made online, the online market has increased a lot in recent years, and the data that is collected there has also increased tremendously. This data can be used to find out what customers want, what kind of customers buy their product, and what kind of customers to focus on.

*2)* Authors Asif Iqbal, Rajeev Ranjan Pandey, Subhraneel Bagchi, Saikat Ruj, Sujata Dawn says in the paper title "Customer Segmentation using Machine Learning with a Coupon Generator GUI" that The rise of numerous competitors and entrepreneurs which has led to a great deal of competition among businesses, compelling them to seek out new customers while retaining their existing ones.

*3)* Authors Professor Harshad Kubade, Pratik Jitendra Gharde, Tushar Dilip Fulbandhe , Aman Amit Pandey, Kiran Sharad Rehpade, Sanskruti Ratnakar Hedaoo says in the paper title " Customer Segmentation: Types of Models and Clustering Techniques" that Customer segmentation is a powerful tool that allows companies to identify and target specific customer groups based on their needs and characteristics.

©IJRASET: All Rights are Reserved 

*4)* Author Chenguang Wang says in the paper title "Efficient customer segmentation in digital marketing using deep learning with swarm intelligence approach" that Successful firm marketing is to offer the right product to the right person at the right time. To market the product to the customer successfully, it is needed to segment the customer by finding their behavioural patterns. The customer behaviours and their purchasing patterns are used to generate profit for the company.

*5)* This Authors Kayalvily Tabianan, Shubashini Velu, Vinayakumar Ravi says in the paper title "K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data" that The purpose of customer segmentation is to determine how to deal with customers in each category in order to increase the profit of each customer to the business. Segmenting the customers assists business to identify their profitable customer to satisfy their needs by optimising the services and products.

*6)* Authors Srinivas Dileep, M. Seshashayee says in the paper title "Customer Segmentation using machine learning" that The emergence of the many competitors and entrepreneurs has caused tons of tension mong competing businesses to seek out new buyers and keep the old ones.

*7)* Authors Mrs.P.Rajeshwari, G. Shravya , Jeet S Patel , K.Dinesh says in the paper title "Customer Segmentation using machine learning" that The main purpose for any organization and enterprise is to recognize their centered clients, How their purchasers perform and use their offerings. Every customer may also use a organization's offerings differently.

*8)* Authors Gali Venkata Durga Ayyappa Babu, M Durga Sathish says in the paper title "E-Commerce Customer Segmentation Using Machine Learning" that In order to optimise each customer\'s worth to business, segmenting clients is done in order to select how to interact among them. Client segmentation belongs to the specific strategy used by businesses to target particular, smaller groups about customers among appropriate messages certain would encourage them to make a purchase.

*9)* Authors Sushilkumar Chavhan, R. C. Dharmik, Sachin Jain, Ketan Kamble says in the paper title "RFM ANALYSIS OF CUSTOMER SEGMENTATION USING MACHINE LEARNING: A SURVEY OF A DECADE OF RESEARCH" that Based on RFM values (Recent, Frequency, and Cost) of customers, the successful classification of company customers is divided into groups with comparable behaviors.

*10)* Author Kushal S says in the paper title "Customer Segmentation" that Looking around and finding that most companies are now data-driven. They make strategicdecisions based on data analysis, enabling them to examine and organize their data for betterservice.

*11)* Authors Disha Maini, Muskan Aggarwal, Prof. Ankur Jain says in the paper title "Customer Segmentation" Similarly, clustering means putting things together in such a way that similar types of things remain in the same group. In this study, a machine learning (ML) algorithm is implemented in the python programming language to perform customer segmentation on credit card data sets to determine the appropriate marketing strategies.

## III. PROPOSED SYSTEM

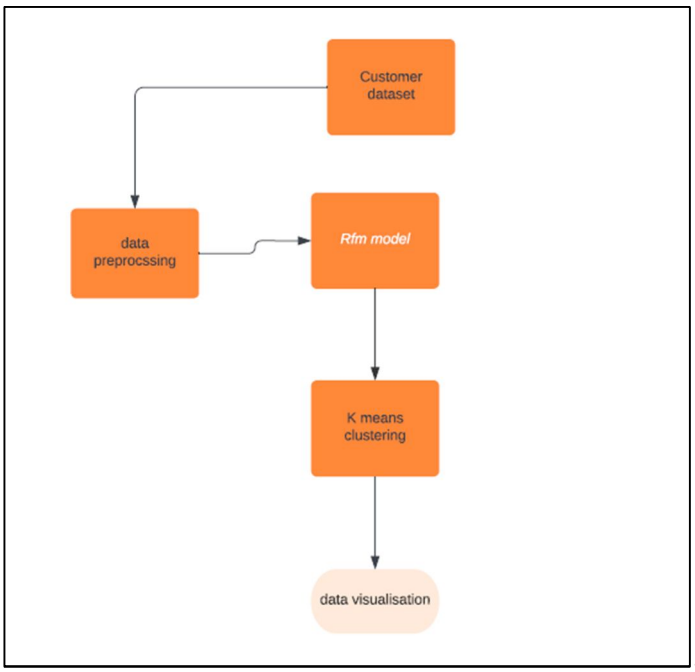
In this paper, we present a strong and thorough framework for customer segmentation - a critical aspect of modern marketing and business strategies.

Our innovative approach utilizes state-of-the-art machine learning techniques and a diverse set of data, including customer interactions, demographic information, and transaction history. The system follows a structured process, starting with data collection and integration from a range of sources, both structured and unstructured. From there, data preprocessing ensures the accuracy and relevance of the data.

Our core system is built upon advanced machine learning models, such as clustering algorithms like k-means, transfer learning, and Recency, frequency, and monetary that is Rfm model, allowing for precise segmentation based on customer behavior, demographics, and preferences.

Our extensive and meticulous approach ensures a comprehensive understanding of customer patterns and behaviors. Our model can be easily deployed into production environments and seamlessly integrated into marketing and customer relationship systems, allowing for its real-world applicability. Through real-time data processing, the segments remain up-to-date, while our continuous feedback loop enables dynamic adjustments. These customer segments are then further profiled, providing valuable insights for personalized marketing, product development, and customer engagement strategies. Our adaptable and scalable system is the perfect solution for organizations looking to improve their marketing efforts and enhance customer satisfaction in today's data-driven and customer-centric market.

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 12 Issue I Jan 2024- Available at www.ijraset.com*

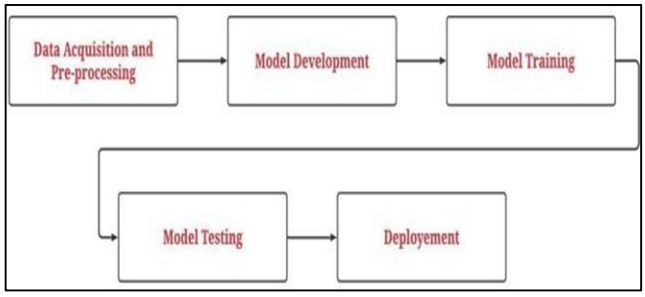



Flowchart of the model

The above figure is the design of our project's model which uses various algorithms on a dataset of customers' transaction history. This model shows the steps in which the system flow works.

*A. Methodology of the System*

Steps in methodology: The Methodology of developing and deploying a machine learning model, includes the following steps: Data acquisition and preprocessing, model development, model training, model testing, and deployment.




*1) Data Acquisition and Preprocessing*

The first step is to determine the problem at hand and the specific type of data needed to resolve it.

Then, we must gather the required data from a range of sources, including databases and real-time markets.

Next, we must carefully sift through the data and eliminate any irrelevant or inconsistent entries, as well as address any missing values or outliers. Finally, we must preprocess the data by performing tasks such as normalization, feature scaling, feature engineering, and encoding categorical variables, to prepare it for further analysis.

*2) Model Development*

Selecting the appropriate machine learning algorithm or a combination of algorithms depends on the nature of the problem, whether it is a classification, regression, or clustering problem.

After preprocessing the data, it is essential to divide it into training and validation sets. The training set is utilized to train the model, while the validation set is used to fine-tune its parameters and measure its performance during development.

The model's architecture, including the number and type of layers or the algorithm's configuration, should be carefully defined.

©IJRASET: All Rights are Reserved 

To control the learning process, it is crucial to set hyperparameters, such as the learning rate, regularization strength, and batch size. The final step involves implementing the model using a suitable programming language or a machine learning library/framework, ensuring smooth and effective execution.

*3) Model Training*

Introduce the training data into the model and utilize an optimization strategy, such as gradient descent, to progressively adjust the model's parameters. Throughout the training process, carefully monitor the model's progress on the validation set to identify any potential overfitting or underfitting. Adjust the hyperparameters accordingly to optimize performance.
Continue training until the model achieves desirable results on the validation set or reaches a stable state.

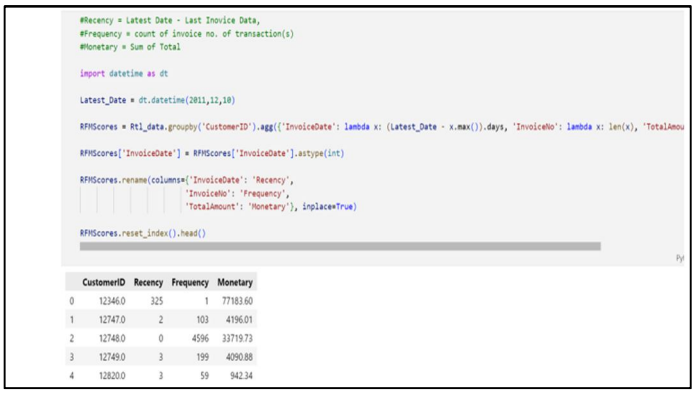
*4) Model Testing*

Once the training is finalized, conduct a thorough evaluation of the model's performance on a new, unseen test set.
Utilize various evaluation metrics, such as accuracy, precision, recall, and score, to assess the model's suitability for the specific problem. Analyze the model's performance and make necessary tweaks or enhancements based on the evaluation results.

*5) Deployment*

Before deploying the model, it is essential to package it in a suitable format for the desired deployment environment. This could involve creating a serialized model file or containerizing the application. Next, carefully integrate the model into the target system, which may entail code writing for managing input/output, data preprocessing, and communication with other system components.
To ensure seamless functioning in the production environment, thoroughly test the deployed model.
It is crucial to monitor its performance and gather real-world feedback to continuously enhance and refine the model as needed."

## IV. IMPLEMENTATION



```
#Recency = Latest Date - Last Inovice Data,
#Frequency = count of invoice no. of transaction(s)
#Monetary = Sum of Total

import datetime as dt

Latest_Date = dt.datetime(2011,12,10)

RFMScores = Rtl_data.groupby('CustomerID').agg({'InvoiceDate': lambda x: (Latest_Date - x.max()).days, 'InvoiceNo': lambda x: len(x), 'TotalAmou

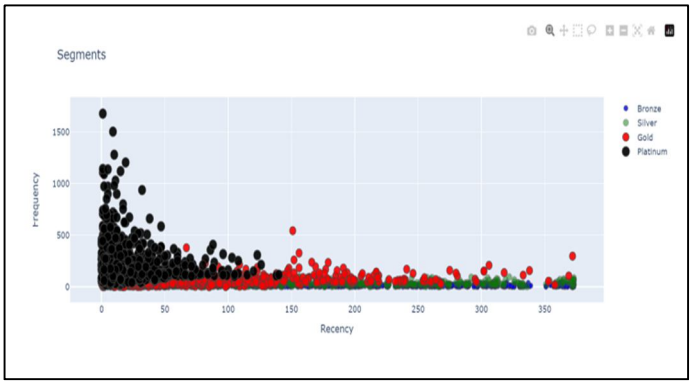
RFMScores['InvoiceDate'] = RFMScores['InvoiceDate'].astype(int)

RFMScores.rename(columns={'InvoiceDate': 'Recency',
                          'InvoiceNo': 'Frequency',
                          'TotalAmount': 'Monetary'}, inplace=True)

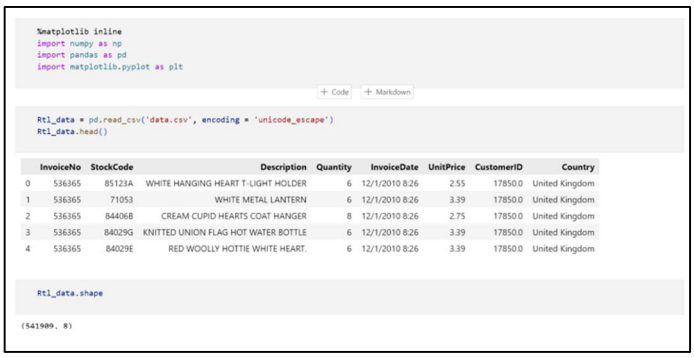
RFMScores.reset_index().head()
```

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

RFM Model




RFM Graph



```
%matplotlib inline
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

+ Code + Markdown

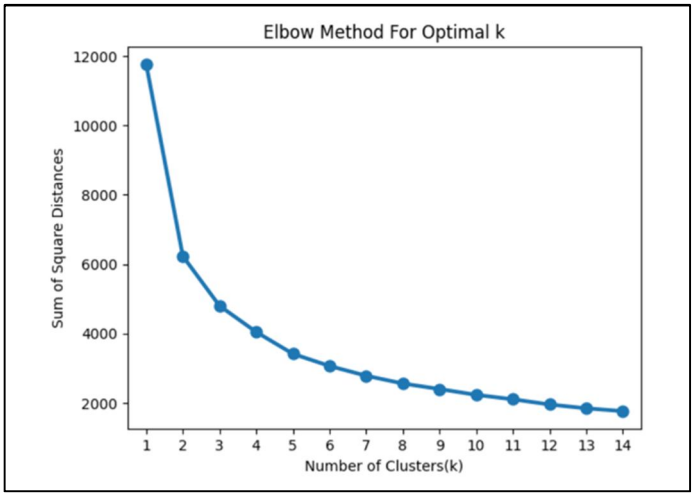
```
Rtl_data = pd.read_csv('data.csv', encoding = 'unicode_escape')
Rtl_data.head()
```

| | 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 | 17850.0 | United Kingdom |
| 1 | 536365 | 71053 | WHITE METAL LANTERN | 6 | 12/1/2010 8:26 | 3.39 | 17850.0 | United Kingdom |
| 2 | 536365 | 84406B | CREAM CUPID HEARTS COAT HANGER | 8 | 12/1/2010 8:26 | 2.75 | 17850.0 | United Kingdom |
| 3 | 536365 | 84029G | KNITTED UNION FLAG HOT WATER BOTTLE | 6 | 12/1/2010 8:26 | 3.39 | 17850.0 | United Kingdom |
| 4 | 536365 | 84029E | RED WOOLLY HOTTIE WHITE HEART. | 6 | 12/1/2010 8:26 | 3.39 | 17850.0 | United Kingdom |

```
Rtl_data.shape
```

```
(541909, 8)
```

Data Visualization




Elbow Method

## V. CONCLUSION

Our study introduces a cutting-edge method for customer segmentation, merging the RFM model with K-Means clustering to elevate marketing initiatives and tailor experiences for customers. Our findings demonstrate the success of this technique in categorizing customers into unique clusters according to their recent transactions, level of engagement, and monetary contributions. By utilizing the K-Means algorithm, customers were efficiently grouped revealing crucial insights into their behaviors and preferences. The practical application of our research has notably bolstered marketing efficiency and customer interaction, by providing personalized experiences and delivering more focused campaigns.

As businesses increasingly seek to harness the power of data-driven strategies, the RFM and K-Means model emerges as a valuable tool for gaining a comprehensive understanding of customer behavior. This research contributes to the growing body of knowledge in customer segmentation and encourages further exploration of innovative data-driven techniques, which will play a crucial role in shaping the future of personalized marketing and customer-centric business practices.

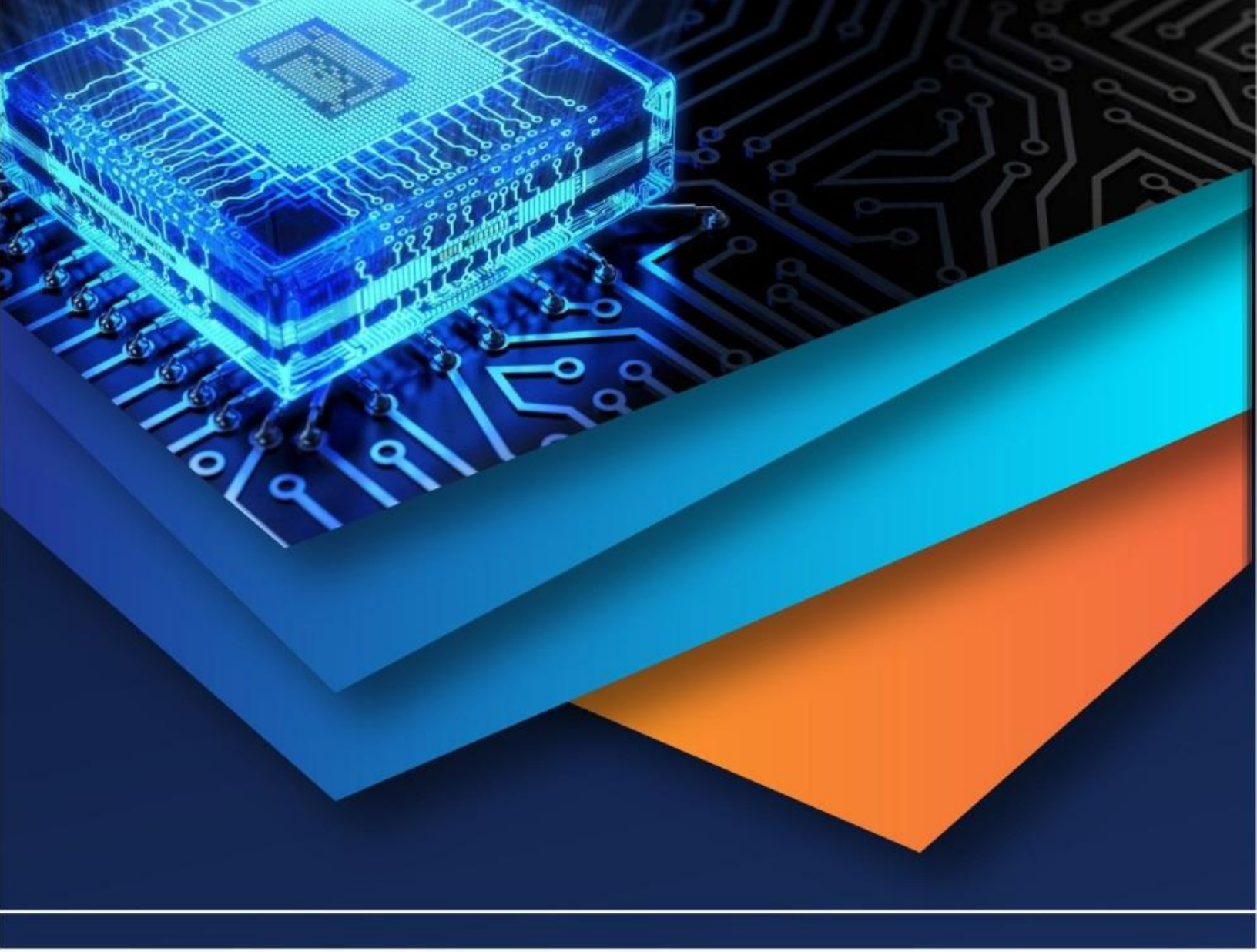
## VI. FUTURE SCOPE

Our research opens the door to future studies on enhancing and expanding customer segmentation techniques. To continue pushing the boundaries of this field, scholars can delve into incorporating supplemental data sources, utilizing machine learning methods, and implementing personalization tactics. Furthermore, the adaptability of our approach to various industries and larger datasets presents exciting opportunities for further investigation. The fusion of RFM-based customer segmentation and K-Means clustering offers a powerful tool for gaining deeper insights into customer behavior. It will empower data-driven decision-making for businesses and elevate customer experiences.

## REFERENCES

[1] uMahmoud SalahEldin Kasema , Mohamed Hamadab , Islam Taj-Eddinc "Customer Profiling, Segmentation, and Sales Prediction using AI in Direct Marketing"
[2] Asif Iqbal, Rajeev Ranjan Pandey, Subhraneel Bagchi, Saikat Ruj, Sujata Dawn "Customer Segmentation using Machine Learning with a Coupon Generator GUI"
[3] Professor Harshad Kubade, Pratik Jitendra Gharde, Tushar Dilip Fulbandhe , Aman Amit Pandey, Kiran Sharad Rehpade, Sanskruti Ratnakar Hedaoo "Customer Segmentation: Types of Models and Clustering Techniques"
[4] Chenguang Wang "Efficient customer segmentation in digital marketing using deep learning with swarm intelligence approach"
[5] Kayalvily Tabianan, Shubashini Velu, Vinayakumar Ravi "K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data"
[6] Srinivas Dileep, M. Seshashayee "Customer Segmentation using machine learning"
[7] Mrs.P.Rajeshwari, G. Shravya , Jeet S Patel , K. Dinesh "Customer Segmentation using machine learning"
[8] Gali Venkata Durga Ayyappa Babu, M Durga Sathish "E-Commerce Customer Segmentation Using Machine Learning"
[9] Sushilkumar Chavhan, R. C. Dharmik, Sachin Jain, Ketan Kamble "RFM ANALYSIS OF CUSTOMER SEGMENTATION USING MACHINE LEARNING: A SURVEY OF A DECADE OF RESEARCH"
[10] Kushal S "Customer segmentation"
[11] Disha Maini, Muskan Aggarwal, Prof. Ankur Jain "Customer Segmentation"
[12] Prof. Nikhil Patankar, Soham Dixit, Akshay Bhamare Ashutosh Darpel and Ritik Raina "Customer Segmentation using machine learning"
[13] Garima Sharma, Ankita Nainwal, Bhaskar Pant, Vikas Tripathi, Mr Akash Chauhan "Customer Segmentation using machine learning"
[14] Rahul Shirole, Laxmiputra Salokhe, Saraswati Jadhav "Customer Segmentation using RFM Model and K-Means Clustering"
[15] V Vijlesh, A.Harini, M.Hari Dharshini, R.Priyadharshini "Customer Segmentation using machine learning"
[16] Saurabh Patil, Hasnath Khan, Sachin Mehta and Prof. Umakant Mandawkar "STUDY OF CUSTOMER SEGMENTATION USING k-MEANS CLUSTERING AND RFM MODELLING"
[17] Azad Abdulhafedh "Incorporating K-means, Hierarchical Clustering and PCA in Customer Segmentation"
[18] Musthofa Galih Pradana, Hoang Thi Ha "Maximizing Strategy Improvement in Mall Customer Segmentation using K-means Clustering"
[19] Aman Banduni, Ilavendhan "Customer Segmentation using machine learning"
[20] Ritik Raina "Customer Segmentation using machine learning"
[21] Patel Monil , Patel Darshan , Rana Jecky , Chauhan Vimarsh , Prof. B. R. Bhatt "Customer Segmentation using Machine Learning"
[22] Maha Alkhayrat, Mohamad Aljnidi, Kadan Aljoumaa "A comparative dimensionality reduction study in telecom customer segmentation using deep learning and PCA"
[23] Jun Wu, Li Shi, Wen-Pin Lin, Sang-Bing Tsai, Yuanyuan Li, Liping Yang and Guangshu Xu "An Empirical Study on Customer Segmentation by Purchase Behaviors Using a RFM Model and K-Means Algorithm"
[24] Hansi Chen, Lei Zhang, Xuening Chu, Bo Yan "Smartphone customer segmentation based on the usage pattern"
[25] A. Joy Christy, A. Umamakeswari, L. Priyatharsini, A. Neyaa "RFM ranking – An effective approach to customer segmentation".
[26] Wafa Qadadeha, Sherief Abdallah "Customers Segmentation in the Insurance Company (TIC) Dataset"
[27] Mr. Abhijit Bag "Customer Segmentation & Opportunity Analysis".
[28] Dogan O., Aycin E., Bulut Z. A. "Customer Segmentation By Using RFM Model And Clustering Methods: A Case Study in Retail Industry".
[29] M.A.Ishantha "Mall Customer Segmentation Using Clustering Algorithm".
[30] Serhat Peker, Altan Kocyigit, P. Erhan Eren "LRFMP model for customer segmentation in the grocery retail industry: a case study"
[31] Diana L. Huerta-Muñoz a, Roger Z. Ríos-Mercado b, Rubén Ruiz "An iterated greedy heuristic for a market segmentation problem with multiple attributes"
[32] Cormac Dullaghan and Eleni Rozaki "Integration of Machine Learning Techniques to Evaluate Dynamic Customer Segmentation for Mobile Customers"
[33] Sari, Juni Nurma; Nugroho, Lukito Edi; Ferdiana, Ridi; Santosa, P. Insap "Review on Customer Segmentation Technique on Ecommerce"
[34] Mieczysław Pawłowski, Jarosław Banaś "The use of Clustering Methods and Machine Learning Algorithms in the Trading Enterprise for Customer Segmentation"
[35] Hülya Güçdemir, Hasan Selim "Integrating multi-criteria decision making and clustering for business customer segmentation"
[36] Jie Lu, Vahid Behbood, Peng Hao, Hua Zuo, Shan Xue, Guangquan Zhang "Transfer learning using computational intelligence: A survey"
[37] Fadly Hamka , Harry Bouwman , Mark de Reuver ,, Maarten Kroesen "Mobile customer segmentation based on smartphone measurement"
[38] Kishana R. Kashwan, Member, IACSIT, and C. M. Velu "Customer Segmentation Using Clustering and Data Mining Techniques".
[39] Payam Hanafizadeh, Neda Rastkhiz Paydar "A Data Mining Model for Risk Assessment and Customer Segmentation in the Insurance Industry"
[40] Chih-Fong Tsai, Ya-Han Hu, Chia-Sheng Hung, Yu-Feng Hsu "A comparative study of hybrid machine learning techniques for customer lifetime value prediction"

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)