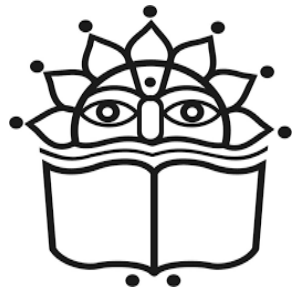
Developing a Unified Customer Segmentation Framework using Multi-Industry Behavioral Data

Anish Date (2347005)
Vipashyana Jawale (2347003)
Yashraj Devrat (2347075)
Shubham Keskar (2347031)



Guide : Dr. Pradip Paithane

Department of Artificial Intelligence & Data Science
Vidya Pratishthan's Kamalnayan Bajaj Institute of Engineering and Technology
Vidyanagari, Baramati-413133

Contents

Motivation

- RFM segmentation, we got an opportunity to unlock the true potential of e-commerce business using RFM modeling.
- RFM analysis, soars to new heights, fueled by a deep understanding of customers and their needs.
- Develop successful strategies adaptable to each cluster, optimize advantages and establish a win-win scenario for the organization and its customers.

Literature Survey

Table 1: Literature Survey

Sr.no	Author	Technique Used	Advantages	Gaps
1	Juan Liao et.al. [2022]	Two Statistical methods entropy and superiority chart method	Application utilization and improve targeted promotion	Performance Evaluation Measures are not elaborated
2	A.Joy Christy et.al [2018]	Repetative K-Means Algorithm	Proposed Algorithm has good complexity	RM K-Means problem with clusters

3	A. Syaifudin et.al [2023]	Fuzzy C-Means clustering, Genetic Programming to optimize FCM	GP overcomes local minimum issue in FCM.	Comparison with other clustering algorithms not shown.
4	Hanaa Hachimi et.al [2023].	statistical clustering method	Improved Customer Segmentation considering diversity.	Insufficient explanation of CLV factor calculation in RFM-D model

Problem Statement

To enhance the previous RFM modeling technique to integrate different customer behaivours and implement effective marketing strategy.

Objectives

Resolve Problem of Literature Survey Papers

- To divide customers into homogeneous clusters based on their RFM values, identify distinct customer segments with specific characteristics and properties.
- Gain insights into customer preferences and needs and understand product development and innovation.
- Develop targeted acquisition strategies to attract new customers who align with existing segments.
- To compare the performance of traditional K-means clustering, Fuzzy C-Means clustering, RM Kmeans, DBSCAN for customer segmentation.

System Architecture

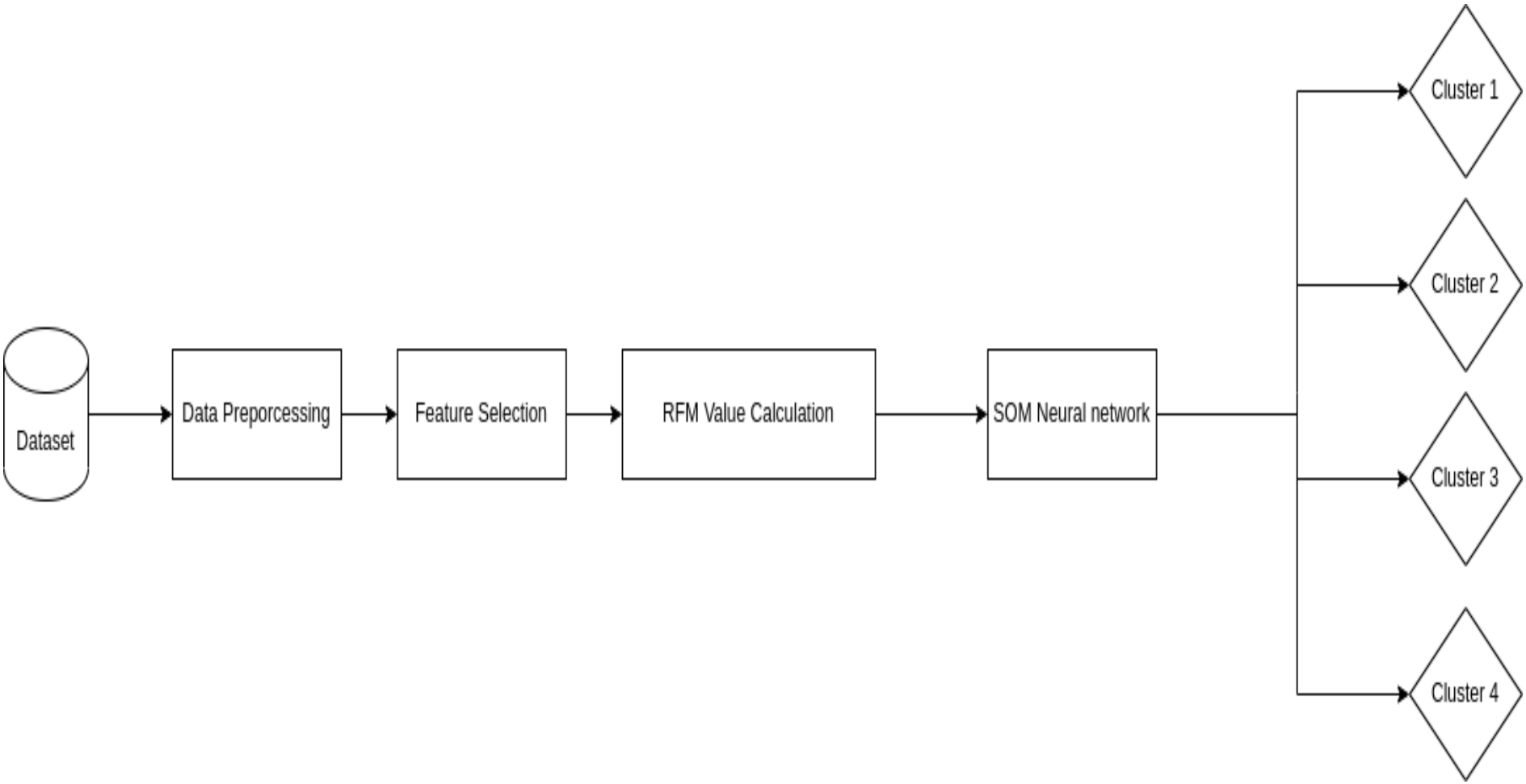


Figure: Architecture

Software Requirement Specification

Software Requirements :

- Operating System: Linux
- OS Type: 64-bit
- Python Version: 3.11.4
- Tools: Google Colaboratory / Jupyter Notebook.

Hardware Requirements :

- Disk Space: 200 GB
- Processor: 11th Gen Intel Core i5
- GPU: NVDIA Corporation.
- RAM: 8 GB

SRS

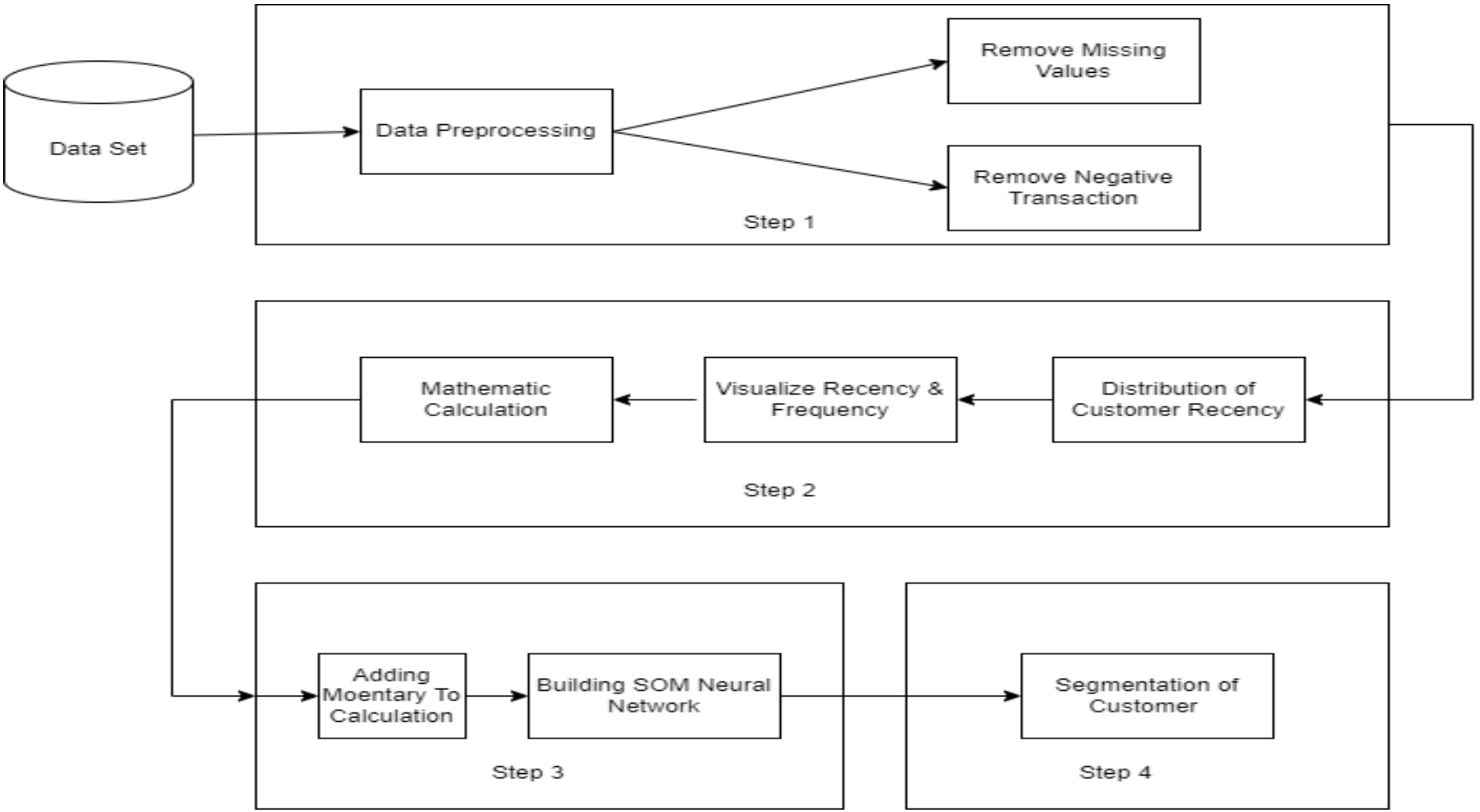


Figure: Data Flow Diagram

Data Flow Diagram

1. DFD Level 0

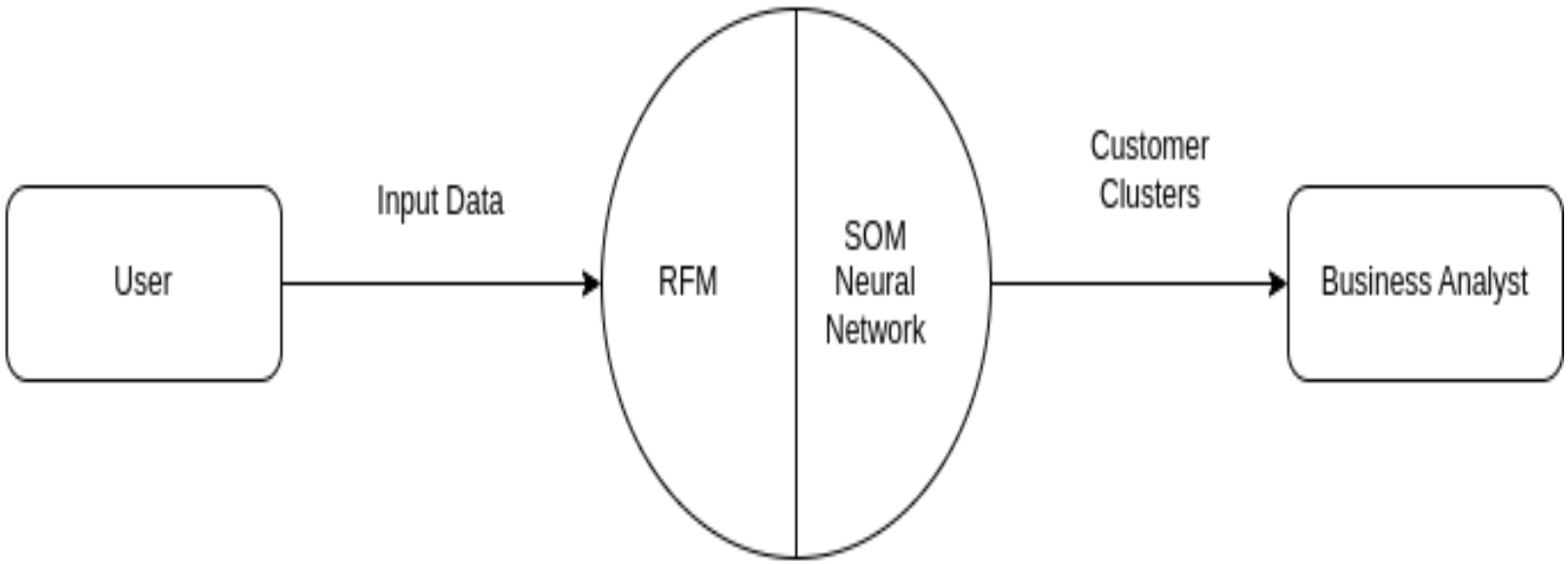


Figure: DFD-0

Data Flow Diagram

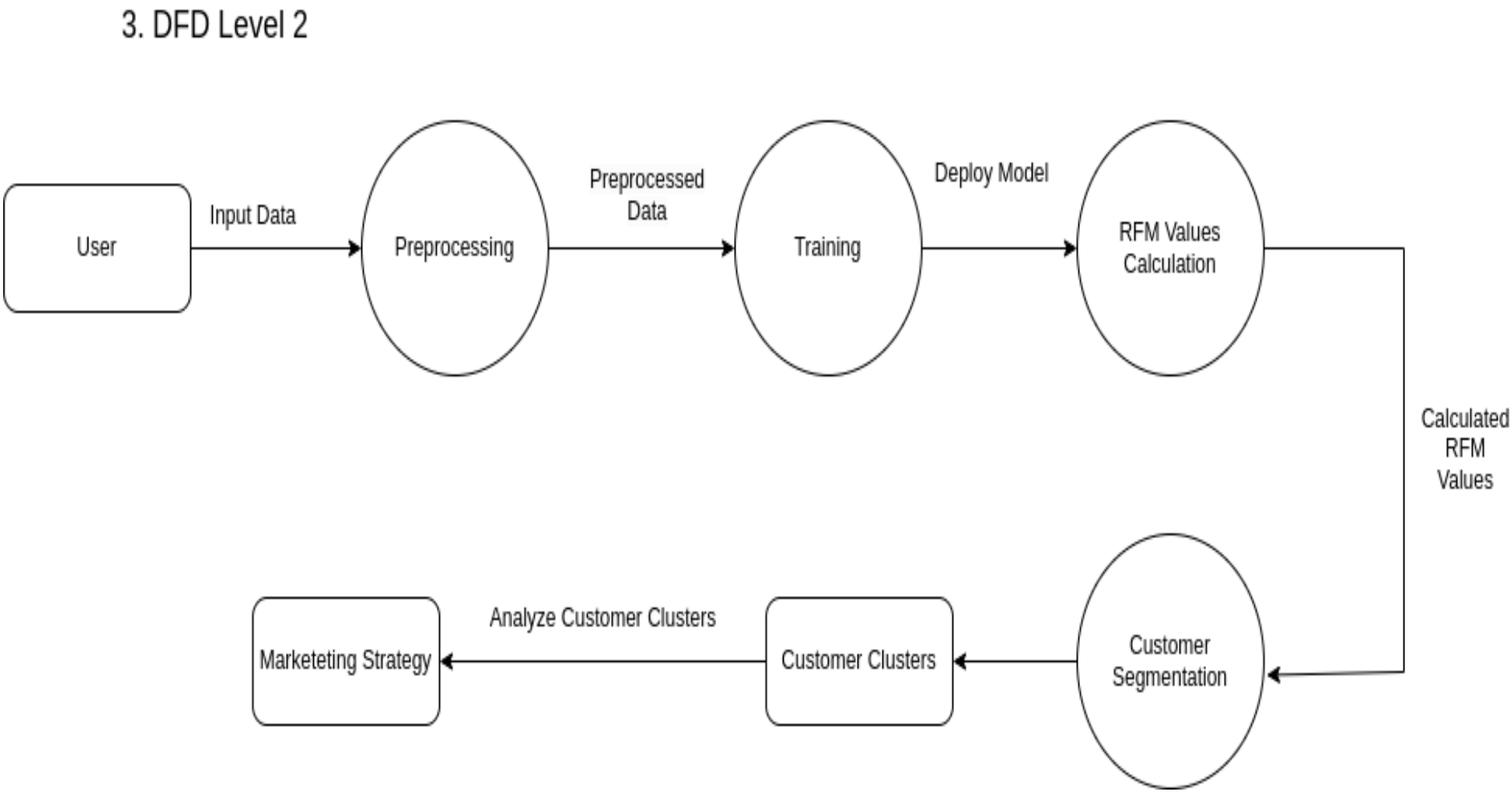


Figure: DFD-1

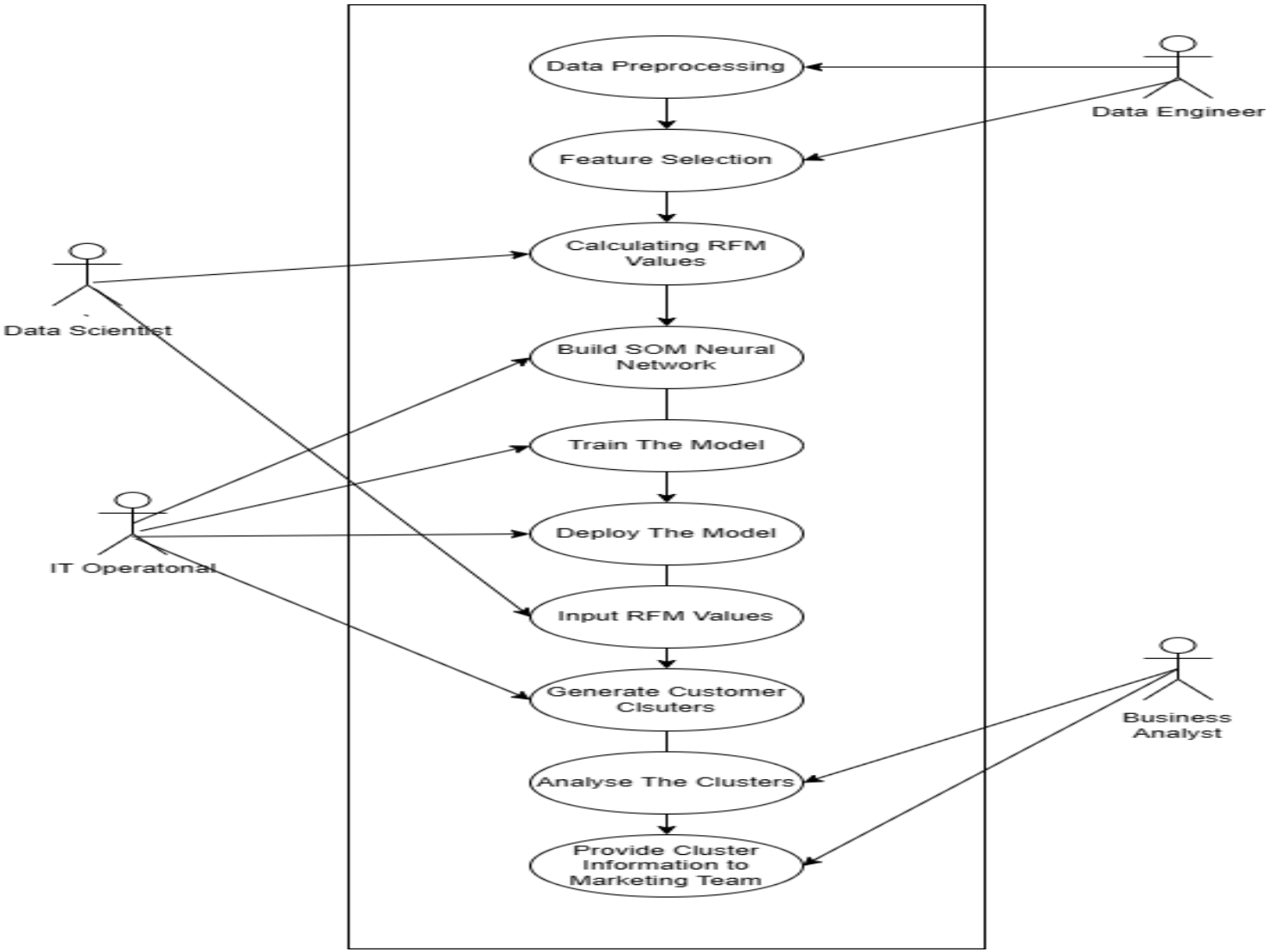


Figure: Use Case Diagram

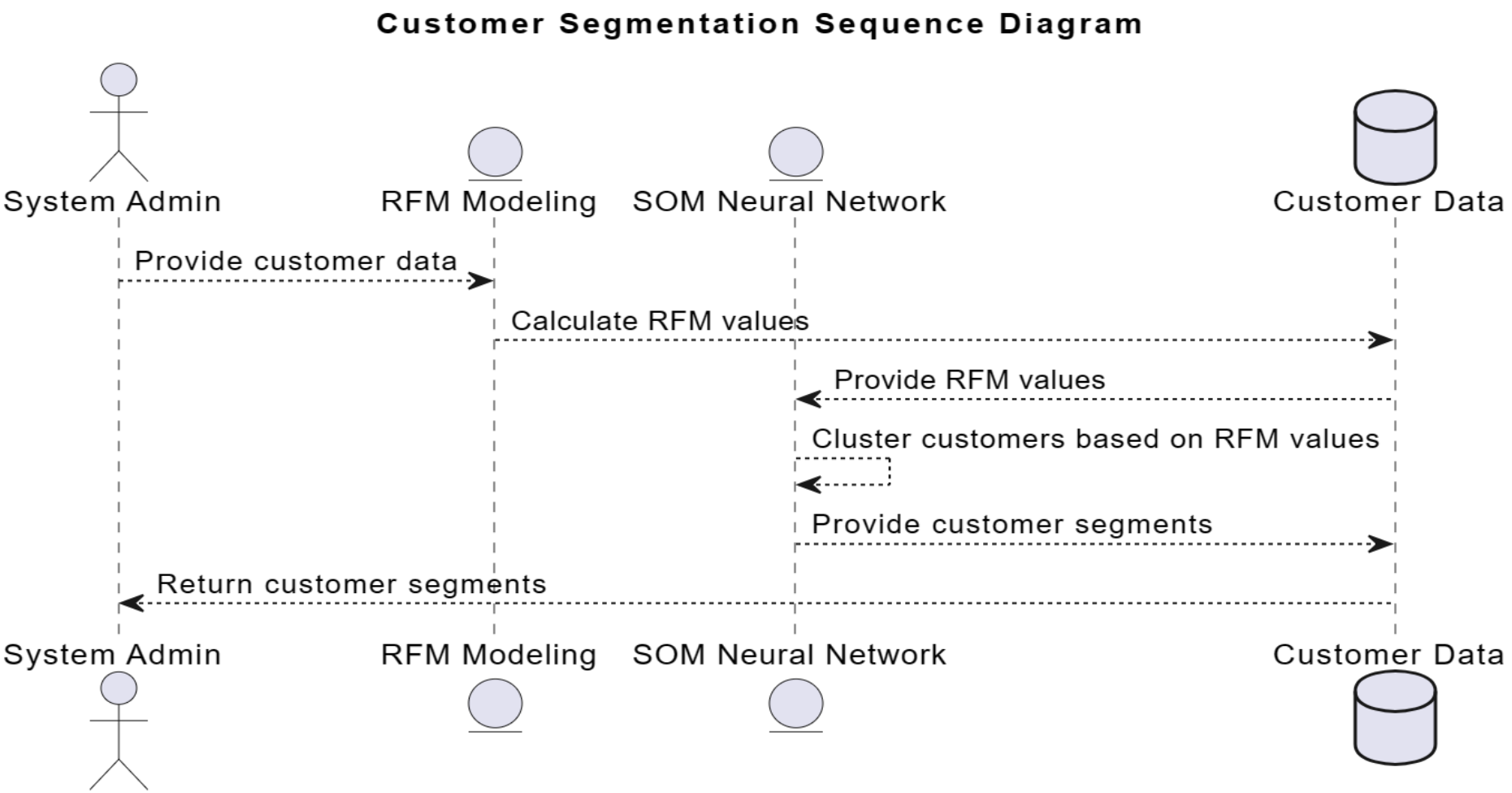


Figure: Sequence Diagram

Algorithms And Techniques

Recurrent Neural Networks (RNN):

- For time-series behavioral data, RNNs like LSTMs capture sequential patterns influencing customer segments.

Deep Embedding Clustering (DEC):

- Jointly learn feature embeddings and cluster assignments using deep networks in an unsupervised manner.

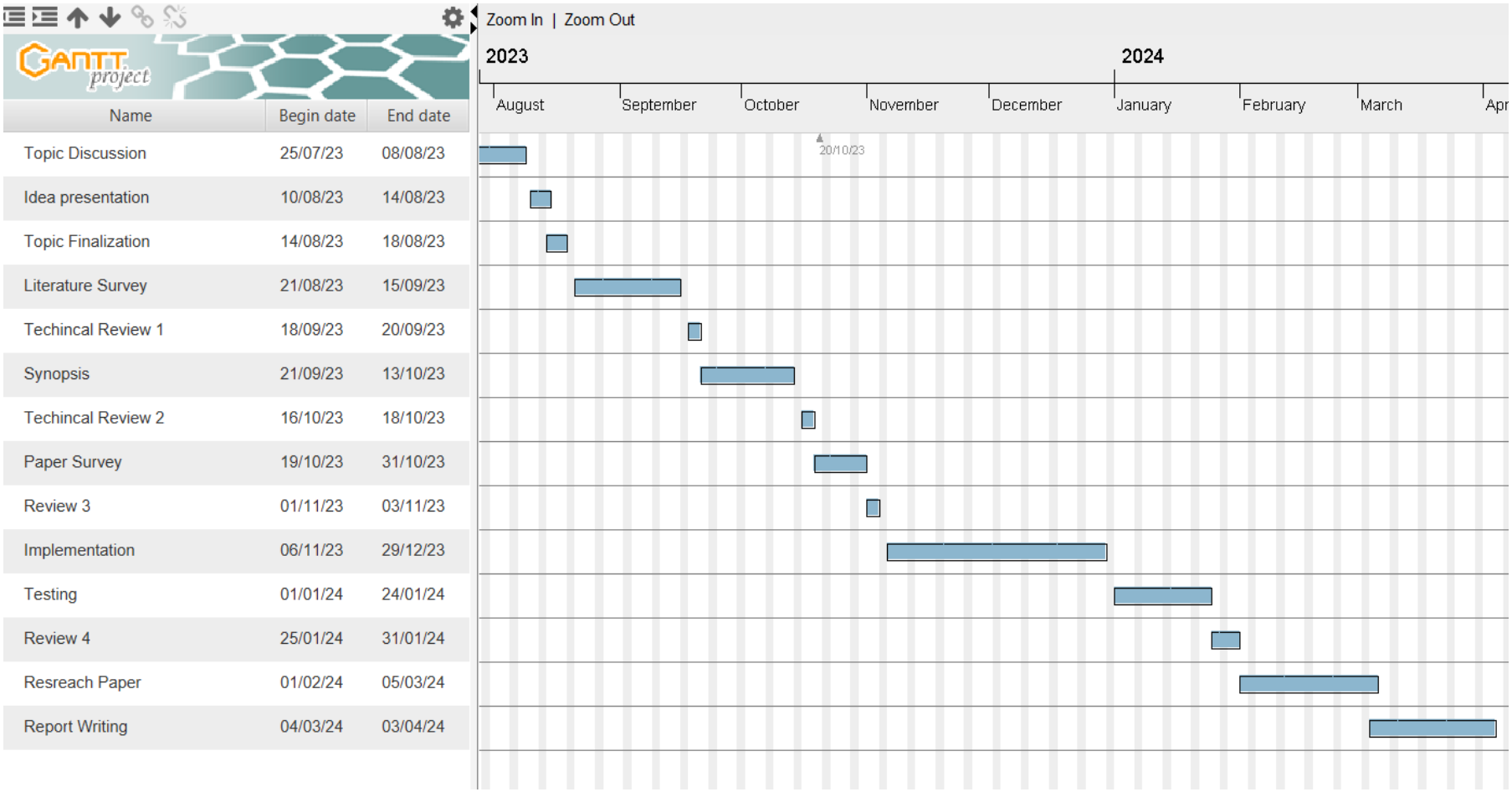
Module split-up

- Module 1 : Data Gathering
  1. Dataset
  2. Data Preprocessing
- Module 2 : Recency and Frequency Value Calculation
  1. Superiority Chart Method
  2. Entropy Value Method
- Module 3 : Calculate Omega Value
  1. Add Monetary Value to Calculated Recency and Frequency Values .

Module split-up

- Module 4 : Neural Network Architecture
  1. Model Building Using SOM Neural Network
  2. Training model
  3. Testing model
- Module 5 : Results and Experiments
  1. Evaluation Metrics
     1. Accuracy
     2. Recall
     3. Precision
     4. F1 score
- Modue 6 : Clusters as Output
  1. Segmentation of Customers based on various properties .

Project plan 2.0



Name	Begin date	End date
Topic Discussion	25/07/23	08/08/23
Idea presentation	10/08/23	14/08/23
Topic Finalization	14/08/23	18/08/23
Literature Survey	21/08/23	15/09/23
Techincal Review 1	18/09/23	20/09/23
Synopsis	21/09/23	13/10/23
Techincal Review 2	16/10/23	18/10/23
Paper Survey	19/10/23	31/10/23
Review 3	01/11/23	03/11/23
Implementation	06/11/23	29/12/23
Testing	01/01/24	24/01/24
Review 4	25/01/24	31/01/24
Resreach Paper	01/02/24	05/03/24
Report Writing	04/03/24	03/04/24

Figure: Gantt Chart

Conclusion

Customer segmentation is critical for business to leverage the power of data analysis for improving profitability . Various clustering techniques divides data into categories of clusters and patterns gives idea for complementary marketing strategies .

References

- RFM ranking – An effective approach to customer segmentation, A. Joy Christy a, A. Umamakeswari a, L. Priyatharsini b, A. Neyaa 2018
- New RFM-D classification model for improving customer analysis and response prediction, Moulay Youssef SMAIL, Hanaa HACHIMI,2023
- Customer Segmentation Using Fuzzy-AHP and RFM Mode, Anu Gupta Aggarwal, Shweta Yadav, 2020.
- Customer Segmentation with RFM Model using Fuzzy C-Means and Genetic Programming, Anas Syaifudin , Purwanto , Heribertus Himawan , M. Arief Soelema, 2023.
- Customer Segmentation Based on RFM Model Using K-Means, K-Medoids, and DBSCAN Methods, Rahma Wati Br Sembiring Berahmana, Fahd Agodzo Mohammed, Kankamol Chairuang, 2020

Thank You