CUSTOMER SEGMENTATION AND CROSS-SELLING RECOMMENDATION ENGINE FOR THE TELECOM INDUSTRY: A MACHINE LEARNING APPROACH

Janadhi Dilmini Nanayakkara

219445C

Master of Science in Telecommunications

Department of Electronics & Telecommunications Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

July 2025

CUSTOMER SEGMENTATION AND CROSS-SELLING RECOMMENDATION ENGINE FOR THE TELECOM INDUSTRY: A MACHINE LEARNING APPROACH

Janadhi Dilmini Nanayakkara

219445C

Dissertation submitted in partial fulfillment of the requirements for the degree
Master of Science in Telecommunications

Department of Electronics & Telecommunications Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

July 2025

DECLARATION

I declare that this is my own work and this Dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date: 18/07/2025

The supervisor should certify the Dissertation with the following declaration.

The above candidate has carried out research for the Master of Science in Telecommunications Dissertation under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of Supervisor: Prof. (Mrs) Dileeka Dias

Signature of the Supervisor: Date:

DEDICATION

To my parents, for their endless sacrifices and unwavering belief in me, and to all those who stood by me with patience and encouragement - this achievement belongs as much to you as it does to me.

ACKNOWLEDGEMENT

This research would not have been a success without the kind support and help of many individuals and organizations. I would like to extend my sincere thanks to all of them.

First of all, I wish to express my heartfelt gratitude to my academic supervisor, Prof. (Mrs.) Dileeka Dias, for her continuous guidance, insightful feedback, and invaluable encouragement throughout the course of this research. Her guidance and dedication were crucial for this study.

I would also like to extend my special thanks to Dr. Kasun Hemachandra, the course coordinator of the PG Diploma/ MSc in Telecommunications program, for his expert advice and dedication towards mentoring students and inspiring them to achieve their full potential.

A special note of appreciation goes to all individuals who offered their assistance in gathering the required datasets and insights which contributed significantly to the successful completion of this research.

To my friends and colleagues, thank you for your encouragement, thoughtful discussions, and moral support throughout this journey.

Finally, and most importantly, I am thankful to my family for the moral support, encouragement and guidance extended in many ways at all stages of my life. This accomplishment would not have been possible without their constant support.

ABSTRACT

In the telecommunication industry, personalized service delivery and effective cross-selling strategies are vital for enhancing customer engagement and driving revenue growth. This research presents a data-driven approach to customer segmentation and personalized recommendation, aiming to promote the transition from Double-Play (Voice and Internet) to Triple-Play (Voice, Internet and IPTV) service adoption.

The study applies unsupervised machine learning techniques, namely Gaussian Mixture Models (GMM) and K-Means clustering to segment customers based on demographic, billing, and service usage attributes. Multiple internal evaluation metrics, including the Silhouette Score, Calinski-Harabasz Index, Davies-Bouldin Index and the Elbow method are employed to determine the optimal clustering structure and validate the segmentation quality.

To generate personalized service recommendations, a collaborative filtering approach based on Singular Value Decomposition (SVD) is utilized. The model predicts the most suitable IPTV package for Double-Play customers by learning from existing customer interaction patterns. The performance of the recommendation system is evaluated using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE).

The implementation leverages Python-based tools and libraries for data preprocessing, modeling and visualization. The research offers a practical and scalable framework for targeted marketing and intelligent service personalization, contributing to the advancement of data-driven strategies in the telecommunications domain.

Keywords: Machine Learning, Telecommunication, Customer segmentation, Recommendation Systems

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
CH	Calinski-Harabasz
DBI	Davies-Bouldin Index
DP	Double-Play
DQ	Data Quality
GMM	Gaussian Mixture Model
LTV	Lifetime Value
ML	Machine Learning
RFM	Recency, Frequency, Monetary
RMSE	Root Mean Squared Error
SVD	Singular Value Decomposition
TP	Triple-Play
VAS	Value-Added Services

LIST OF APPENDICES