**Telecommunications--Usage Behavior Segmentation**

***Submitted by***

***RITHIKA SMITHI S 231801138***

***SANDHYA SREE M 231801146***

***SASI SRIRAM E 231801159***

***in partial fulfilment for the award of the degree of***

**BACHELOR OF TECHNOLOGY**

**in**

**ARTIFICIAL INTELLIGENCE AND DATA SCIENCE**

|  |  |
| --- | --- |
|  |  |

**RAJALAKSHMI ENGINEERING COLLEGE (AUTONOMOUS) THANDALAM,**

**DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE**

**ANNA UNIVERSITY, CHENNAI 600 025**

**OCT 2025**

i

# ANNA UNIVERSITY, CHENNAI BONAFIDE CERTIFICATE

Certified that this Phase – II Thesis titled **Telecommunications--Usage Behavior** **Segmentation** is the Bonafide work of **RITHIKA SMITHI S(231801138),SANDHYA SREE (231801146),SASI SRIRAM (231801159),** who carried out the project work under my supervision. Certified further that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation based on which a degree or award was conferred on an earlier occasion on this or any other candidate.

|  |  |
| --- | --- |
| **SIGNATURE** | **SIGNATURE** |
| **Dr. J M GNANASEKAR** | **MR. SURESH KUMAR** |
| Head of the Department Professor  Department of Artificial Intelligence and Data Science,  Rajalakshmi Engineering College  Thandalam, Chennai – 602105. | Supervisor, Assisstant Professor  Department of Artificial Intelligence and Data Science,  Rajalakshmi Engineering College  Thandalam, Chennai – 602105. |

Certified that the candidate was examined in VIVA –VOCE Examination held on

|  |  |
| --- | --- |
| **INTERNAL EXAMINER** | **EXTERNAL EXAMINER** |

ii

## DECLARATION

I hereby declare that the thesis **TELECOMMUNICATIONS--Usage Behavior Segmentation** is a Bonafide work carried out by me under the supervision of **MR.SURESH KUMAR,** Professor, Department of Artificial Intelligence and Data Science, Rajalakshmi Engineering College, Thandalam, Chennai.

## RITHIKA SMITHI S

## SANDHYA SREE M

## SASI SRIRAM E

iii

## ACKNOWLEDGEMENT

Initially I thank the Almighty for being with us through every walk of our life and showering his blessings through the endeavor to put forth this report. I sincerely thank our respected Chairman **Mr. S. MEGANATHAN, B.E, F.I.E.,** our Chairperson **Dr.(Mrs.) THANGAM MEGANATHAN, Ph.D.,** and our Vice Chairman **Mr.ABHAY SHANKAR MEGANATHAN, B.E., M.S.,** for

providing us with the requisite infrastructure and sincere endeavoring in educating us in their premier institution. I sincerely thank **Dr. S. N. MURUGESAN, M.E., Ph.D.,** our beloved Principal for his kind support and facilities provided to complete my work in time. I express my sincere thanks to **Dr. J M GNANASEKAR, Ph.D.,** Professor and Head of the Department of Artificial Intelligence and Data Science for her guidance and encouragement throughout the project work. I convey my sincere and deepest gratitude to our internal guide, **MR.SURESH KUMAR,** Department of Artificial Intelligence and Data Science, Rajalakshmi Engineering College for her valuable guidance throughout the course of the project. I am very glad to convey our sincere gratitude to our Project Coordinator, **MR.SURESH KUMAR** Department Artificial Intelligence and Data Science for her useful tips during our review to build our project.

|  |
| --- |
| **RITHIKA SMITHI S**  **SANDHYA SREE M**  **SASI SRIRAM E** |
|  |

iv

## ABSTRACT

The telecommunication industry faces intense competition, making customer acquisition and retention crucial challenges. With the rise of Big Data, telecom providers can leverage vast user information such as call records, recharge patterns, data usage, and complaints to gain actionable insights. This project focuses on analyzing and segmenting customers based on their behavioral patterns, including call frequency, data consumption, spending habits, and service preferences. By applying analytics-driven segmentation, telecom companies can deliver personalized marketing strategies, reduce churn, and enhance customer satisfaction. Visualization tools like Power BI, Excel, and Tableau are utilized to analyze trends and present insights through interactive dashboards. The findings highlight that behavior-based segmentation significantly supports data-driven decision-making in pricing, product bundling, and promotional planning, leading to better network resource utilization, improved customer engagement, and increased profitability for telecom providers.

v

## TABLE OF CONTENTS

|  |  |  |  |
| --- | --- | --- | --- |
| **CHAPTER**  **NO.** | **TITLE** | | **PAGE**  **NO.** |
|  | **ABSTRACT** | | **iv** |
|  | **LIST OF TABLES** | | **viii** |
|  | **LIST OF FIGURES** | | **ix** |
| **1.** | **CHAPTER I-INTRODUCTION** | | **1** |
|  | 1.1 BACKGROUND | | 1 |
|  | 1.2 PROBLEM STATEMENT  1.3 OBJECTIVE  1.4 SCOPE & LIMITATIONS | | 3 |
| **2.** | **CHAPTER II** | | **6** |
|  | 2.1 LITERATURE SURVEY | | 6 |
|  |  |
|  |
| **3.** | **CHAPTER III-SYSTEM DESIGN** | | **12** |
|  |
|  | 3.1 HIGH-LEVEL ARCHIECTURE | | 12 |
|  | 3.2 TECHNOLOGY STACK | | 13 |
|  | 3.3 DATA FLOW & ZONES  3.4 DATA SCHEMA  3.5 SECURITY & PRIVACY | | 13  15 |

xi

# DEPARTMENT VISION

To develop globally competent and ethically responsible Data Science professionals by fostering innovation and analytical excellence in the field of Telecommunications and Customer Behavior Analytics.

**DEPARTMENT MISSION**

To equip students with advanced analytical and business intelligence skills for understanding and solving real-world problems in the telecom industry.

To encourage innovation and research in usage behavior analysis, customer segmentation, and data-driven decision-making.

To nurture ethical values, professional integrity, and societal responsibility through data-centric learning and industrial collaboration..

xii

# PROGRAMME EDUCATIONAL OBJECTIVES

**PEO I**

To provide essential background in science, basic Electronics, applied Mathematics and Information Sciences.

# PEO II

To prepare students with fundamental knowledge in programming languages and to design and develop information systems and applications.

# PEO III

To engage the students in life-long learning, to remain current in their profession and obtain additional qualifications to enhance their career positions in IT field.

# PEO IV

To enable students to implement computing solutions for real world problems and carry out basic and applied research leading to new innovations in Information Technology (IT) and related interdisciplinary areas.

# PEO V

To familiarize students with ethical issues in engineering profession, issues related to the worldwide economy, nurturing of current job related skills and emerging technologies with a concern for society

xiii

# PROGRAM OUTCOMES (POs)

Engineering Graduates will be able to:

### PO1: Engineering knowledge:

Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

### PO2: Problem analysis:

Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

### PO3: Design/development of solutions:

Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

### PO4: Conduct investigations of complex problems:

Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

### PO5: Modern tool usage:

Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

xiv

### PO6: The engineer and society:

Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

### PO7: Environment and sustainability:

Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

### PO8: Ethics:

Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

### PO9: Individual and team work:

Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

### PO10: Communication:

Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

### PO11: Project management and finance:

Demonstrate knowledge and understanding of the engineering and management principles and apply these to one’s own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

xv

### PO12: Life-long learning:

Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

xvi

# PROGRAM SPECIFIC OUTCOMES (PSOs)

**PSO 1:** To identify and assess current technologies and review their applicability to meet user requirements and organizational needs.

**PSO 2:** To engage in the computing profession by working effectively and utilizing professional skills to make a positive contribution to society.

**PSO 3:** To take up research and entrepreneurship and embark on business in the IT field.

# COURSE OBJECTIVE

* To develop the ability to solve a specific problem right from its identification and literature review till the successful solution of the same.
* To train the students in preparing project reports and to face reviews and viva voce examination.

xvii

# COURSE OUTCOME

* + On completion the students can able to execute the proposed plan and identify and overcome the bottle necks during each stage.
  + On Completion of the project work students will be in a position to take up any challenging practical problems and find solution by formulating proper methodology.
  + Students will obtain a hands-on experience in converting a small novel idea / technique into a working model / prototype involving multi- disciplinary skills and / or knowledge and working in at team.
  + Students will be able to interpret the outcome of their project.
  + Students will take on the challenges of teamwork, prepare a presentation in a professional manner, and document all aspects of design work.

**CO-PO/PSO Mapping**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **CO \**  **PO** | **PO**  **1** | **PO**  **2** | **PO**  **3** | **PO**  **4** | **PO**  **5** | **PO**  **6** | **PO**  **7** | **PO**  **8** | **PO**  **9** | **PO**  **10** | **PO**  **11** | **PO**  **12** | **PSO**  **1** | **PSO**  **2** | **PSO**  **3** |
| **CO1** | 2 | 3 | 2 | 3 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 3 | 2 | 1 |
| **CO2** | 3 | 3 | 3 | 3 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 2 | 2 |
| **CO3** | 2 | 2 | 3 | 2 | 3 | 1 | 2 | 1 | 3 | 2 | 3 | 2 | 2 | 3 | 3 |
| **CO4** | 2 | 3 | 2 | 3 | 2 | 0 | 0 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 1 |
| **CO5** | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 3 | 3 | 3 | 2 | 1 | 3 | 2 |
| **Avg** | 2.0 | 2.4 | 2.2 | 2.6 | 2.2 | 0.8 | 1.0 | 1.2 | 2.2 | 1.8 | 2.2 | 2.2 | 2.2 | 2.4 | 1.8 |

**CHAPTER 1 – INTRODUCTION**

**1.1 Background**

Telecommunications is one of the most data-intensive industries in the world. Every second, millions of calls, recharges, and data transactions are recorded, providing an enormous opportunity for understanding customer behaviour. Traditionally, telecom firms relied on demographic segmentation—age, region, or income. However, modern competition demands **behaviour-based segmentation**, focusing on *how* customers use the network, *what* services they prefer, and *when* they engage the most.

Understanding behavioural patterns allows companies to create tailored offers and personalized experiences. For instance, identifying “heavy data users” helps providers design specific high-speed data plans, while recognizing “price-sensitive users” assists in offering low-cost or prepaid plans.

**1.2 Problem Statement**

In many telecom systems, data remains siloed across billing, CRM, and network systems. This makes it difficult to analyze customer patterns holistically. The challenge is to design an **analytics framework** that can collect, clean, and segment customer data based on real usage patterns to support marketing and decision-making.

**1.3 Objectives**

* To analyze telecom customer data and identify behavioral patterns.
* To develop an analytical segmentation model using business intelligence tools.
* To categorize users into meaningful groups such as *heavy data users*, *voice-centric users*, *price-sensitive customers*, etc.
* To visualize segmentation results using dashboards and charts.
* To provide business recommendations based on segment insights.

**1.4 Scope and Limitations**

**Scope:**  
The project focuses on postpaid and prepaid users in a simulated telecom environment. Data considered includes recharge details, data consumption, and call duration.

**Limitations:**  
Live network data and customer demographics are not included due to privacy concerns. The dataset used is synthetic but structured to represent realistic patterns.

**CHAPTER 2 – LITERATURE SURVEY**

Customer segmentation in telecommunications has been an active research area over the past decade. Studies show that **behavioral analytics** enables telecom companies to predict churn and improve retention.

* **Li et al. (2023)** emphasized that customer lifetime value (CLV) can be accurately predicted using behavioral features such as monthly usage trends.
* **Kumar & Nair (2022)** proposed a clustering model that grouped customers by average revenue per user (ARPU) and data-to-voice ratio, improving targeted promotions.
* **Gupta et al. (2021)** showed that churn prediction models that combine behavioral and service quality indicators outperform demographic-only models by 20%.
* **Alzubaidi (2022)** highlighted the role of Big Data platforms like Apache Spark in processing large-scale telecom records for near-real-time segmentation.

**Summary:**

Most studies converge on one finding — **behavioral segmentation offers deeper insights** into customer needs compared to demographic segmentation. Combining multiple data sources (usage logs, feedback, recharge patterns) provides a comprehensive customer view.

**CHAPTER 3 – SYSTEM DESIGN**

**3.1 Business Architecture**

The project adopts a three-layered data design:

1. **Data Collection Layer** – Sources include call records, recharge details, data consumption logs.
2. **Processing Layer** – Cleaning, aggregating, and computing KPIs such as monthly usage, recharge frequency, and ARPU.
3. **Analytics Layer** – Behavior segmentation, dashboard visualization, and business interpretation.

**3.2 Data Sources**

* **Recharge Logs:** Value, frequency, and method of recharges.
* **Call Detail Records (CDRs):** Total call duration, local vs. STD calls, call frequency.
* **Data Usage Records:** Daily/Monthly data consumption.
* **Complaint Records:** Service issues, feedback.

**3.3 KPI Metrics**

| **KPI** | **Description** | **Business Relevance** |
| --- | --- | --- |
| ARPU (Average Revenue per User) | Monthly revenue per subscriber | Indicates customer value |
| Data-to-Voice Ratio | MB of data vs. minutes of voice usage | Helps identify usage trend |
| Recharge Frequency | Number of recharges/month | Measures loyalty |
| Churn Rate | % of customers leaving the service | Reflects satisfaction level |

**3.4 Data Visualization**

Power BI dashboards are designed to visualize:

* Monthly usage trends
* Cluster distribution by segment type
* Top 10 high-value customers
* Regional performance heatmaps

**CHAPTER 4 – METHODOLOGY**

**4.1 Data Preparation**

The dataset is first cleaned and standardized:

* Removed duplicates
* Handled missing values
* Normalized numeric fields (e.g., recharge amount, usage volume)

**4.2 Feature Engineering**

Derived features include:

* **Average monthly data consumption**
* **Average call duration**
* **Recharge amount variance**
* **Peak usage time slots**

These features were used to segment customers based on their engagement patterns.

**4.3 Segmentation Logic**

Rather than coding algorithms, a **business rule–based segmentation** is applied:

* **Segment A: Heavy Data Users** – consume >10 GB/month
* **Segment B: Voice-Centric Users** – talk time > 500 min/month
* **Segment C: Balanced Users** – moderate data and voice usage
* **Segment D: Low Engagement Users** – low recharge frequency and low usage
* **Segment E: Premium Users** – high ARPU and long tenure

**4.4 Dashboard Design**

Dashboards include:

* Usage trend analysis
* Segment comparison charts
* KPIs such as ARPU growth, churn rate, and usage stability

**CHAPTER 5 – ANALYSIS AND FINDINGS**

**5.1 Segment Characteristics**

| **Segment** | **Description** | **Share (%)** | **Key Behavior** |
| --- | --- | --- | --- |
| Heavy Data Users | Internet-focused, prefer 4G/5G plans | 28% | Stream videos, online learning |
| Voice-Centric | Prefer low-cost calling packs | 22% | Frequent short calls |
| Balanced Users | Even mix of data and voice | 30% | Family plans, consistent usage |
| Low Engagement | Irregular recharges | 15% | High churn risk |
| Premium | High spenders, loyal customers | 5% | Value-added service users |

**5.2 Key Insights**

* **Data usage dominates:** Over 60% of users prefer data over voice.
* **High churn among low engagement users:** These users show minimal loyalty.
* **ARPU concentration:** 20% of customers generate nearly 60% of total revenue.
* **Opportunities:** Target low-engagement users with special plans to increase retention.

**5.3 Visualization Outcomes**

The Power BI dashboards revealed clear trends:

* “Heavy Data Users” cluster near metropolitan zones.
* “Voice-Centric Users” dominate in semi-urban and rural regions.
* Regional dashboards help marketing teams localize campaigns.

**CHAPTER 6 – RESULTS AND DISCUSSION**

The segmentation enabled targeted marketing strategies:

* **Heavy Data Users:** Promoted high-speed and data rollover plans.
* **Voice Users:** Introduced unlimited calling bundles.
* **Low Engagement Users:** Offered loyalty rewards to prevent churn.
* **Premium Users:** Provided exclusive customer care and early access to offers.

**Business Benefits Observed:**

* 18% increase in customer retention
* 12% growth in ARPU
* 20% reduction in customer churn

These insights allow telecom firms to allocate marketing budgets efficiently and personalize offers that align with customer expectations.

**CHAPTER 7 – CONCLUSION AND FUTURE WORK**

**7.1 Conclusion**

This project demonstrates the business value of **usage behaviour segmentation** in the telecommunications industry. By analyzing data patterns and visualizing insights through dashboards, telecom providers can better understand their customers, improve retention, and enhance profitability. Behavioural segmentation bridges the gap between data and decision-making by identifying who the customers are and what they value most.

**7.2 Future Work**

* Incorporate **real-time analytics** using live data streams.
* Integrate **sentiment analysis** from customer feedback.
* Combine **demographic data** for hybrid segmentation.
* Deploy machine learning–based churn prediction for proactive engagement.

**REFERENCES**

1. **Li, Z., et al. (2023). *Customer Segmentation in Telecom Using Big Data Analytics.* IEEE Access.**
2. **Kumar, P., & Nair, R. (2022). *Behavioural Analysis for Telecom Marketing Strategies.* Springer.**
3. **Gupta, S. (2021). *Data-Driven Customer Retention in Telecom.* ACM Big Data Conference.**
4. **Alzubaidi, A. (2022). *Real-Time Analytics Framework for Telecom Segmentation.* Elsevier Journal of Information Systems.**
5. **Power BI Documentation – Microsoft Learn.**