

Identifying Key Drivers of Customer Satisfaction through Survey Data Analysis

Student Name: Preethisha.V

Register Number:613523106029

Institution: Government College of Engineering ,Dharmapuri.

Department: Electronics and Communication Engineering

Date of Submission: 24/04/2025

1. Problem Statement

Customer satisfaction plays a vital role in the success and growth of any business. Understanding what factors influence customer satisfaction helps companies enhance service quality, improve customer retention, and gain competitive advantage. This project aims to analyze customer survey data to identify the key drivers that most significantly impact satisfaction levels. The insights can help businesses prioritize improvements and allocate resources more efficiently.

2. Objectives of the Project

- To analyze customer survey data and discover patterns and correlations between features and satisfaction levels.
- To identify the most influential factors that contribute to high or low customer satisfaction.
- To create a predictive model that can estimate satisfaction based on survey inputs.
- To visualize insights in a clear and actionable format for business decision-making.

3. Scope of the Project

- **Features Analyzed:** Customer demographics, product/service feedback, service quality ratings, and overall satisfaction score.
- **Constraints:** The analysis is limited to the provided survey dataset. No external real-time data or advanced deployment will be used.
- **Output:** A final report with insights, visualizations, and a basic predictive model.

4. Data Sources

- **Dataset Source:** [Insert dataset source – e.g., Kaggle, internal survey, etc.]
- **Data Type:** Static, structured survey data.
- **Access:** Self-generated
- The dataset includes customer responses to multiple survey questions related to their experience with a product or service.

5. High-Level Methodology

- **Data Collection:** Dataset is self generated.
- **Data Cleaning:** Handling missing values, encoding categorical data, removing duplicates, and standardizing formats.
- **Exploratory Data Analysis (EDA):** Using histograms, box plots, heatmaps, and correlation matrices to uncover patterns.
- **Feature Engineering:** Creating new indicators (e.g., average satisfaction score), normalizing data, and deriving categorical variables.
- **Model Building:** Logistic Regression, Decision Trees, Random Forest for classification of satisfaction levels.
- **Model Evaluation:** Accuracy, Precision, Recall, F1-score, and ROC-AUC curves.
- **Visualization & Interpretation:** Insights presented via matplotlib, seaborn, and possibly dashboards (e.g., Streamlit if deployed).
- **Deployment:** Optional; may present results in an interactive notebook or Streamlit dashboard.

6. Tools and Technologies

- **Programming Language:** Python
- **Notebook/IDE:** Jupyter Notebook / Google Colab
- **Libraries:** pandas, numpy, seaborn, matplotlib, scikit-learn
- **Optional Deployment Tools:** Streamlit or Gradio (if visualization is made interactive)

7. Team Members and Roles

1. Preethisha.V – Data Acquisition and Preprocessing Lead

- Responsible for collecting the dataset from the specified source and ensuring data quality before analysis.
- Tasks include identifying and handling missing values, removing duplicates, converting categorical variables into numerical format (e.g., one-hot encoding or label encoding), and ensuring uniform formatting across the dataset.
- Works closely with the EDA lead to prepare clean data for initial analysis.

2. Vaishali.S – Exploratory Data Analysis (EDA) and Visualization Expert

- In charge of exploring the cleaned dataset to identify trends, relationships, and outliers.
- Uses visualization libraries such as matplotlib and seaborn to present data distributions, correlations, and feature importance visually.
- Summarizes insights from EDA that will inform feature engineering and model design.

3.Roja.S – Feature Engineering and Model Developer

- Develops new features or transforms existing ones to improve model accuracy and efficiency.
- Selects suitable machine learning algorithms (such as Logistic Regression, Decision Tree, or Random Forest) to build predictive models.
- Trains, tests, and tunes models to predict customer satisfaction based on input features.

4. Soundarya – Model Evaluation and Report Presentation Lead

- Responsible for evaluating model performance using metrics like Accuracy, Precision, Recall, F1-Score, and AUC-ROC.
- Interprets model outcomes to identify key drivers of customer satisfaction.
- Designs the final report and presentation material, including dashboards and charts that summarize the results. May also handle optional deployment (e.g., Streamlit) if decided.