



LINUXIFY
TEAM

DECEMBER 2023

PROJECT PROPOSAL

Customer Segmentation for E-commerce Platform

PRESENTED TO:
Mr.kassahun

PRESENTED BY:
Linuxify IT team

TABLE OF CONTENTS

Introduction

3

Problem of the Existing System

4

Objective of the Data Analysis

5

Data Source

6

Data Preprocessing

7

Machine Learning Model Selection

8

Model Development Steps

9

The Model Result and Findings

11

Conclusion

12

INTRODUCTION

The e-commerce industry has witnessed tremendous growth, presenting vast opportunities for businesses to connect with customers and drive sales. However, the challenge lies in effectively analyzing the massive amount of data generated by e-commerce platforms to understand customer behavior and preferences. This project proposal aims to address this challenge through customer segmentation, enabling our organization to tailor marketing strategies and enhance customer engagement for improved business performance.

Customer segmentation in an e-commerce platform is the process of dividing a company's customer base into distinct groups or segments based on specific characteristics, behaviors, or demographics.

customer segmentation is to identify groups of customers who share similar traits and behaviors.

E-commerce platforms typically use various criteria to segment their customers. These criteria can include demographic factors (e.g., age, gender, location), psychographic factors (e.g., interests, values, lifestyle), purchase history, browsing behavior, engagement level, and more. The specific criteria chosen depend on the platform's objectives and the type of products or services offered.

LINUXIFY IT TEAM

Birhan Aschalew.....UGR/9050/13

Ayan Abas.....UGR/0999/13

Yonas Sisay.....UGR/3279/13

Kena Tekalign.....UGR/6147/13



PROBLEM OF THE EXISTING SYSTEM

Our organization currently faces two significant challenges



Firstly

the large volume of data generated by our e-commerce platform makes it difficult to extract meaningful insights without appropriate analytical models



Secondly

the absence of such models hinders our ability to accurately target customers and develop personalized marketing strategies based on their unique characteristics and preferences.



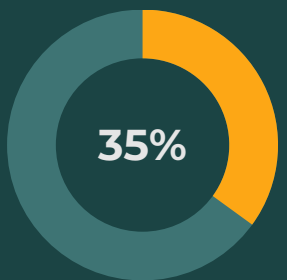
OBJECTIVE OF THE DATA ANALYSIS

The primary objective of this project is to leverage data analysis techniques to achieve comprehensive customer segmentation. By analyzing the available data, we aim to achieve the following:



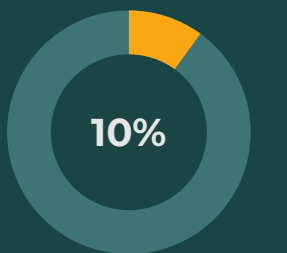
Goal #1

Identify distinct customer segments based on their characteristics, behavior, and preferences.



Goal #2

Gain a deeper understanding of the needs and preferences of each customer segment.



Goal #3

Develop targeted marketing strategies and personalized experiences tailored to each segment.



Goal #4

Improve customer engagement, conversion rates, and overall business performance.



DATA SOURCE

The data for this analysis will be sourced from our e-commerce platform's customer database

This dataset includes various types of information, such as transactional data, customer demographics, browsing behavior, purchase history, and other relevant customer metrics. The data will be collected and stored in a structured format, ensuring data quality and maintaining the confidentiality of customer information.

1. **Purchase History:** Analyzing customers' past purchase behavior provides valuable insights into their preferences, buying patterns, and product interests.
2. **Website Analytics:** Website analytics tools, such as Google Analytics, can provide data on customer browsing behavior, page views, click-through rates, and conversion rates.
3. **Customer Surveys:** Conducting surveys allows you to gather direct feedback from customers about their preferences, satisfaction levels, and needs.
4. **CRM Data:** If our e-commerce platform has a customer relationship management (CRM) system, it can be a valuable source of customer data. CRM data includes customer contact information, communication history, interactions, and customer support tickets, providing insights into customer interactions and preferences.



DATA PREPROCESSING

Data preprocessing is an essential step in customer segmentation, as it involves cleaning, transforming, and organizing the collected data to prepare it for analysis and segmentation.

1. Data Cleaning: Removing duplicate records, handling missing values, and addressing any inconsistencies in the data to ensure data integrity

2. Data Transformation: Data transformation techniques are applied to normalize or rescale variables to ensure they have a similar scale and distribution. This helps prevent certain variables from dominating the segmentation process due to their larger ranges. Common transformation techniques include log transformation, standardization, or min-max scaling.

3. Feature Selection: Identifying the most relevant features for customer segmentation, considering factors such as demographics, purchase behavior, and engagement metrics.

4. Feature Engineering: Creating new features or transforming existing ones to capture additional insights, such as customer lifetime value or recency-frequency-monetary (RFM) metrics.

5. Normalization/Standardization: Scaling the data to ensure all features are on a similar scale, avoiding biases towards particular attributes during the analysis.

Phase 1

Data
Cleaning

Phase 2

Data
Transformation

Phase 3

Feature
Selection

Phase 4

Feature
Engineering

Phase 5

Normalization/
Standardization



MACHINE LEARNING MODEL SELECTION

MODEL SELECTION		PERCENTAGE
K-Means Clustering		20%
Hierarchical Clustering		15%
Gaussian Mixture Models (GMM)		15%
Random Forests		15%
Neural Networks		10%
Support Vector Machines		10%
Decision Trees	\$1,035,000	15%

To solve the identified business problem of customer segmentation, we propose utilizing clustering algorithms such as k-means, hierarchical clustering, or Gaussian mixture models.

These unsupervised learning algorithms are well-suited for identifying natural groupings within the customer data based on their similarity in features. By applying these models, we can segment customers into distinct groups and gain insights into their behavior and preferences.



MODEL DEVELOPMENT STEP

The steps for training, testing, and evaluating a machine learning model for customer segmentation on an e-commerce platform are as follows:

Data Preparation	Split your dataset into two parts: a training set and a testing set. The training set will be used to train the model, while the testing set will be used to evaluate its performance. The typical split ratio is around 70-80% for training and 20-30% for testing, but this can vary depending on the size of your dataset.
Model Training	Use the training set to train your machine learning model. The specific steps for training will depend on the chosen model.
Model Testing	Once the model is trained, use the testing set to evaluate its performance. Apply the trained model to the testing data and obtain segment assignments for each customer in the testing set. Ensure that the testing data is not used during the training process to provide an unbiased evaluation of the model's generalization ability.
Evaluation Metric	Calculate appropriate evaluation metrics to assess the quality of the customer segmentation.
Interpretation and Analysis	Interpret the obtained customer segments and analyze their characteristics and behaviors.



**Model
Refinement**

Based on the evaluation results and interpretation of the segments, refine the model if necessary. This may involve adjusting hyperparameters, trying different preprocessing techniques, or exploring alternative models. Iterate on the training and testing steps to improve the segmentation performance.

**Final
Evaluation**

Once you have refined the model, perform a final evaluation using the testing set to assess its performance after refinement. This evaluation provides a final measure of how well the model segments customers and informs decisions on the model's readiness for deployment.

Profit**248%****Dividend per
share****100%**

It's important to note that the steps mentioned above are iterative and may require multiple iterations to achieve satisfactory results.

Additionally, monitoring the performance of the deployed model over time and continuously updating it with new data is crucial to ensure its effectiveness in customer segmentation on your e-commerce platform



THE MODEL RESULT AND FINDINGS



IDENTIFIED CUSTOMER SEGMENTS

SEGMENT PROFILES

BEHAVIORAL PATTERNS



SEGMENT SIZE AND IMPORTANCE:

CROSS-SEGMENT COMPARISONS

MARKETING PERSONALIZATION



CUSTOMER LIFETIME VALUE



CONCLUSION

In conclusion, this project proposal aims to address the challenges faced by our organization in effectively analyzing customer data for targeted marketing strategies. By leveraging data analysis techniques and utilizing machine learning models for customer segmentation, we strive to improve customer engagement and overall business performance. We kindly request your approval to proceed with this project, and we are confident in our team's expertise and capabilities to achieve successful outcomes.

