Final Proposal

Team: Nachiketh Reddy – Graduate Student, Aniket Surve – Graduate Student
Utkarsha Deshkar – Graduate Student

Project Proposal: Supermarket Sales Analysis Using Machine Learning

Dataset Link: https://www.kaggle.com/code/reguita/supermarket-sales-k-means-clustering/input

Introduction:

Our project focuses on analyzing supermarket sales data to better understand customer behavior. We'll use machine learning techniques to group customers based on their buying patterns and predict which group they belong to.

Data Description:

We have a dataset from Kaggle that contains information about supermarket sales, like the date, branch, customer details, product categories, and sales amounts. This data will help us learn about customers and improve marketing and sales strategies.

Project Goals:

Customer Segmentation with K-Means Clustering:

We'll start by using K-Means clustering to group customers based on how they shop. This will help us find common shopping patterns among customers.

Predicting Customer Groups:

After creating these customer groups, we'll turn this into a problem where we predict which group a customer belongs to.
We'll use different machine learning models like Decision Trees, Random Forest, Logistic Regression, Support Vector Machines, and Neural Networks to do this.

Comparing Models:

We'll see how well each model can predict customer groups.
We'll compare the models to find the one that works best for our task.
We'll also figure out which factors are most important for making these predictions.

Making Sense of the Results:

We'll use charts and graphs to help us understand the results.
By doing this, we can make smarter marketing and sales decisions based on what we've learned.

Project Timeline:

Week 5: Data Preparation and Exploration
Week 6: Clustering Customers
Week 7: Predicting Customer Groups
Week 8: Comparing Models
Week 9: Visualizing Results
Week 10: Creating Reports and Presentations

Conclusion:

This project will help us understand our customers better, making it easier to create marketing and sales strategies that work. By comparing different models, we'll find the best way to predict customer groups, and this will help us make smarter business decisions.