Project-2

NGASSEU NDIFO Lyse Priscille

January 2023

1 Project-2:Whole Sales Customers

Context:

This project was developed as part of the DATA MINING INF3256 course in the Computer Science - Data Science Option program at the University of Yaoundé 1. We used the R programming language, including R-Shiny. The project involved a team of three participants:

- **NGASSEU NDIFO Lyse Priscille** (Email: lysendifo8@gmail.com)

- **TAMWO FEUWO Franck Valere** (Email: valeretamwo@gmail.com)

- **BAHAOUDDYN** (Email: bahaouddyn@facsciences-uy1.cm)

Objective:

Our objective was to predict the regions of the wholesale distributor's customers based on their product purchases.

Dataset:

The dataset has 440 instances and included annual spending (in monetary units) on various product categories, such as fresh products, milk products, grocery products, frozen products, detergents and paper products, and delicatessen products. The 'REGION' feature, our target had three nominal values: Lisbon, Oporto, or Other. Our dataset had no missing values, making it suitable for analysis.

Methodology:

Following data preprocessing, we utilized various machine learning algorithms to address the classification problem. Specifically, we employed decision trees, K-nearest neighbors (KNN), neural networks, and Bayesian networks for supervised classification. For unsupervised classification, K-means clustering and hierarchical clustering were applied.

My Contribution:

This project was a team effort, and I worked on the unsupervised classification part. I also helped clean the data and do Data visualizations. This project gave me important experience in both supervised and unsupervised classification, helping me understand customer behavior and preferences