Drug Classification Dataset Using CRISP-DM Methodology

This is a data collected to classify drugs based on Age, Sex, Blood Pressure Levels(BP), Cholesterol Levels and Na to Potassium Ration.

Summary

The Classifying Drug Dataset typically includes various features such as patient demographics (age, sex), health metrics (blood pressure, cholesterol levels), and drug types. The primary goal is to develop predictive models that classify the appropriate drug based on these attributes. The dataset is often used for machine learning tasks, allowing users to apply algorithms like decision trees and random forests to analyze patterns and improve drug prescription accuracy. Overall, it serves as a valuable resource for exploring the relationship between patient characteristics and drug effectiveness.

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

- **Overview:**
- Need to classify drugs based on patient data to optimize treatment plans.
- Addressing medication errors and improving patient outcomes.

Data Overview

- **Dataset Source:** Kaggle Drug Classification Dataset
- **Key Features:**
- Age
- Sex
- Blood Pressure (BP)
- Cholesterol levels
- Na to Potassium Ratio
- Target Variable: Drug Type

Methods

- **Approach:** Data Preprocessing: Cleaning, normalization, and splitting dataset.
- Model Selection: Choosing algorithms (e.g., Decision Trees, Random Forest, or Neural Networks).
- Evaluation Metrics: Accuracy, Precision, Recall, and F1 Score.

Results

- **Findings:**
- Performance metrics of the model.
- Comparison of different algorithms used.
- Insights gained from data trends (e.g., which factors most influence drug classification).

Conclusions

- **Summary:**
- The effectiveness of using patient data for drug classification.
- Recommendations for healthcare providers based on findings.
- Potential for future work (e.g., expanding dataset, integrating additional features).

Thank You!

Email: jacklinewaceke199@gmail.com
GitHub: @JacklineWaceke
LinkedIn: Jackline Waceke