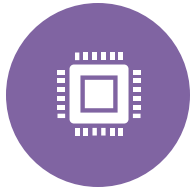
Drug Classification – ML Project Summary

TEAM: DATA CRAFTERS

SHUBHAM MISHRA (TEAM LEAD),

SAKSHAM BHARDWAJ, SURAJ TIWARI,

ANNU KUMAR



B.TECH CSE – GALGOTIAS UNIVERSITY

Problem Statement

- Predict the best drug for a patient based on health features using Machine Learning.

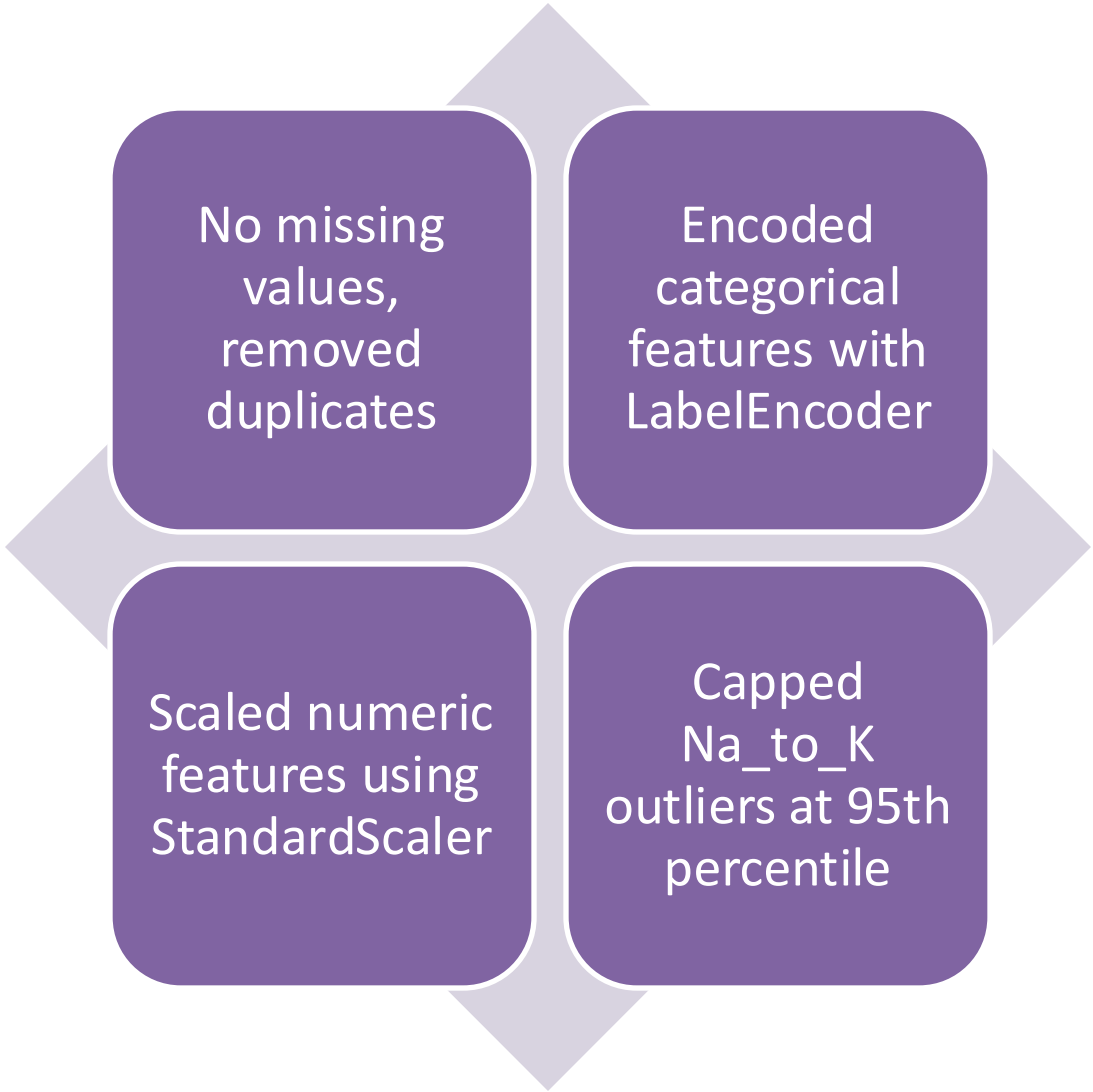
Dataset Overview

200 patient records (drug200.csv)

Features: Age, Sex, BP, Cholesterol, Na_to_K

Target: Drug (DrugA, DrugB, DrugC, DrugX, DrugY)

Data Preprocessing



Modeling

MODEL: RANDOMFORESTCLASSIFIER

TRAIN-TEST SPLIT: 70-30 (STRATIFIED)

EVALUATION: CONFUSION MATRIX AND CLASSIFICATION REPORT

ACCURACY: ~100% ON TEST SET

Key Findings

Na_to_K is the most predictive feature

DrugY associated with high Na_to_K and high BP

DrugC often prescribed for low BP patients

Good classification balance across drug types

Conclusion

- ✅ Cleaned and processed ML-ready dataset
- ✅ Achieved high accuracy and clarity
- ✅ Project meets evaluation rubric
- ✅ Ready for deployment in healthcare prediction systems