CMPE 480 Project 4 - Decision Tree
Muhammet Ali Topcu - 2020400147

Information About the Dataset

The dataset is about a set of patients, all of whom suffered from the same illness. During their course of treatment, each patient responded to one of 5 medications, Drug A, Drug B, Drug C, Drug X or Drug Y. Aim of the project is to build a model to predict which drug might be appropriate for a future patient with the same illness. The features of this dataset are Age, Sex, Blood Pressure, Cholesterol, and the ratio of the sodium to potassium of the patients, and the target is the drug that each patient responded to.
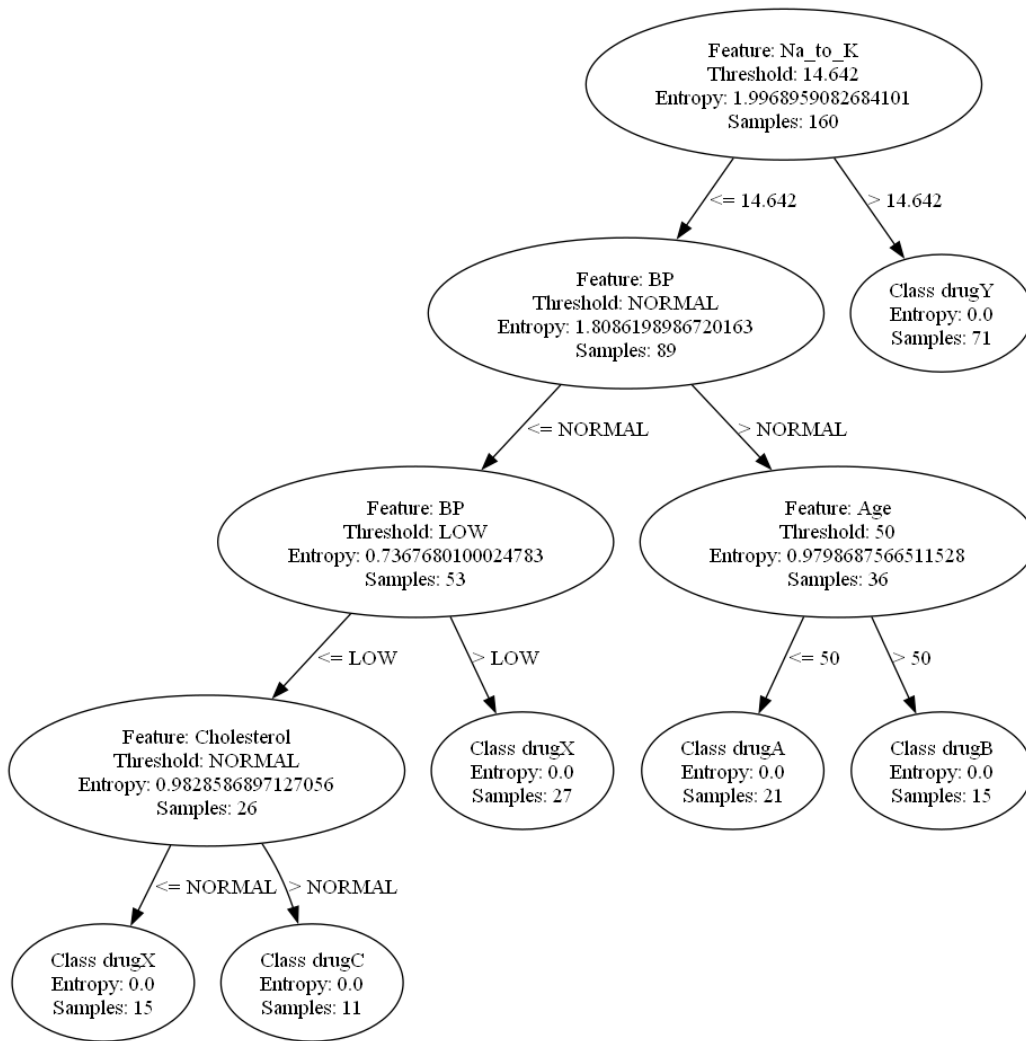
Use of 5-fold Cross Validation

In the implementation, 5-fold cross validation is applied. The dataset contains 200 records. The code generates 5 decision tree with using different parts of the dataset as training data and validation data. At the end, it chooses the decision tree with the lowest validation error.

Error plot for Validation vs Validation Set Index

Final Decision Tree

Source Code

Submitted together with this PDF.