

Healthcare Diagnosis Decision Tree Model

Business Understanding

- To automate initial diagnosis of common illnesses based on symptoms and risk factors. Positive diagnoses are flagged and prioritized for review by a healthcare professional in order to optimize the use of doctors time and attention.
- This will have applications in the healthcare industry.
- After having close family members affected by serious diseases I am motivated to find solutions to faster diagnosis that will lead to expedited treatment that can be the difference in recovery and quality of life after an illness.

Data Understanding

- Symptom and risk factor data will be directly downloaded from several sources.

Data Preparation

- Since much of this data is categorical, encoding will be necessary to work with this data.

Modeling

- A decision tree model will be used to predict the diagnoses.
- The target variable will be a positive or negative diagnosis and corresponding priority flag.
- This use case will be solve a classification problem.

Evaluation

- The Gini Index will be the determining metric of this model.