Finding
Diseases with Symptoms

Introduction

- The slides contain code for data preprocessing, feature extraction, and building different machine learning models such as Logistic Regression, Decision Tree, Random Forest, and K-Nearest Neighbors. It also includes code for model evaluation and comparison of different models.
- The dataset used in this project is from the UCI Machine Learning Repository, which contains symptom and disease data for various diseases such as diabetes, hepatitis, and breast cancer.
- Overall, this slide provides a good starting point for anyone interested in building a disease prediction model using machine learning techniques.

Hypothesis

- This hypothesis assumes that there is a strong relationship between patient symptoms and the underlying diseases that cause this relationship can be captured by a machine-learning model. It further assumes that the dataset used for training the model is representative of the population and contains sufficient examples of different diseases and symptom combinations to enable the model to generalize well to new data. The hypothesis would need to be tested by training and evaluating the machine learning model on the dataset and comparing its performance to other existing methods for disease prediction.

Problem statement

The objective of the problem statement would be to train a model that can accurately predict the likely diseases based on the input symptoms, in order to assist medical professionals in the diagnosis and treatment of patients. The success of the model would be measured based on its accuracy and ability to correctly identify the most likely diseases given a set of symptoms.

Predictions

From the data set taken, we can train and test the machine learning model to predict the likely diseases given a set of patient symptoms. The data set consists of symptom descriptions and their corresponding disease labels. By training a model on this data set, we can predict the likely diseases based on a set of input symptoms.

predictions made by the model are only as good as the data set and the model's accuracy. In addition, the predictions made by the model should not be considered a substitute for professional medical advice or diagnosis. The predictions made by the model should be treated as an initial screening tool, and patients should always seek the advice of a qualified medical professional for a proper diagnosis and treatment plan.