**Data Processing Steps Diagram**

The data processing steps diagram shows the different steps involved in processing the heart disease patient dataset for Heart Disease Prediction project. The diagram starts with loading the dataset, which is followed by data cleaning, data preprocessing, data splitting, feature scaling, feature selection, model training, model evaluation, model tuning, and finally, the final model.

In the data cleaning step, the dataset is checked for any missing or inconsistent data, and any such data is removed or corrected. In the data preprocessing step, the dataset is prepared for use in machine learning algorithms by converting categorical variables to numerical variables, and scaling the features to a common range. In the data splitting step, the dataset is split into training and testing sets to evaluate the performance of the machine learning algorithms.

In the feature scaling step, the features are scaled to a common range to ensure that no feature has a disproportionate impact on the results. In the feature selection step, the most relevant features are selected for use in the machine learning algorithms to improve their performance. In the model training step, the machine learning algorithms are trained on the training data to predict the presence of heart disease.

In the model evaluation step, the performance of the machine learning algorithms is evaluated on the testing data using metrics such as accuracy, precision, recall, and F1-score. In the model tuning step, the hyperparameters of the machine learning algorithms are adjusted to improve their performance. Finally, in the final model step, the best performing machine learning algorithm is selected for use in predicting the presence of heart disease in a patient.

**Diagram

Description automatically generated**

**Machine Learning Algorithms Comparison Diagram**

The machine learning algorithms comparison diagram shows a pie chart that compares the different machine learning algorithms used in Heart Disease Prediction project. The pie chart is divided into eight segments, each representing a different machine learning algorithm. The size of each segment represents the proportion of the dataset that was correctly classified by that algorithm.

The eight machine learning algorithms used in the project are logistic regression, naive Bayes, support vector machine, k-nearest neighbors, decision tree, random forest, XGBoost, and artificial neural network. The pie chart shows that the random forest algorithm achieved the highest accuracy, correctly classifying 30% of the dataset. The decision tree algorithm achieved the second-highest accuracy, correctly classifying 20% of the dataset. The remaining algorithms achieved accuracies ranging from 10% to 15%.

The machine learning algorithms comparison diagram is useful for comparing the performance of different machine learning algorithms and selecting the best algorithm for a given problem. In the case of Heart Disease Prediction project, the random forest algorithm was selected as the best performing algorithm for predicting the presence of heart disease in a patient..

**Chart, pie chart

Description automatically generated**

**System Components Diagram**

The system components diagram shows the different components of Heart Disease Prediction system and how they are related to each other. The diagram consists of four main components: the heart disease prediction application, data processing, machine learning algorithms, and model evaluation.

The heart disease prediction application is the main component of the system and is responsible for predicting the presence of heart disease in a patient. It uses the machine learning algorithms and the final model to make predictions based on input features.

The data processing component is responsible for cleaning, preprocessing, and splitting the heart disease patient dataset for use in machine learning algorithms. It also performs feature scaling and feature selection to improve the performance of the machine learning algorithms.

The machine learning algorithms component consists of eight different algorithms that are used to predict the presence of heart disease. These algorithms include logistic regression, naive Bayes, support vector machine, k-nearest neighbors, decision tree, random forest, XGBoost, and artificial neural network.

The model evaluation component is responsible for evaluating the performance of the machine learning algorithms using metrics such as accuracy, precision, recall, and F1-score. It selects the best performing algorithm for use in the heart disease prediction application.

The system components diagram is useful for understanding the different components of Heart Disease Prediction system and how they work together to predict the presence of heart disease in a patient. It provides a high-level overview of the system architecture and can be used to identify areas for improvement and optimization.

Diagram

Description automatically generated