

Model Development Phase

Date	11 July 2024
Team ID	SWTID1720013031
Project Title	Prediction and Analysis of Liver Patient Data Using Machine Learning
Maximum Marks	6 Marks

Model Selection Report

In the forthcoming Model Selection Report, various models will be outlined, detailing their descriptions, hyperparameters, and performance metrics, including Accuracy or F1 Score. This comprehensive report will provide insights into the chosen models and their effectiveness.

Model Selection Report:

Model	Description	Hyperparameters	Performance Metric (e.g., Accuracy, F1 Score)
Logistic Regression	Logistic Regression is a linear model that predicts the probability of a binary outcome based on input features. It uses a logistic function to model the relationship between the dependent variable and independent variables.	max_iter=1000, penalty='l1', solver='liblinear'	Accuracy - 69.7%, f1-score = 0.81

Support Vector Classifier	Support Vector Classifier, also known as Support Vector Machine (SVM), is a powerful supervised learning algorithm used for classification tasks. It finds the optimal hyperplane in a high-dimensional space that best separates classes.	<code>C=7.41993941811405,</code> <code>kernel='linear'</code>	Accuracy - 69.7%, f1-score = 0.82
Random Forest Classifier	Random Forest Classifier is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes (classification) or mean prediction (regression) of the individual trees.	<code>max_depth=10,</code> <code>min_samples_leaf=2,</code> <code>min_samples_split=10</code>	Accuracy - 69.7%, f1-score = 0.81
K Neighbors Classifier	K Nearest Neighbors Classifier is a simple and intuitive algorithm that stores all available cases and classifies new cases based on a similarity measure (e.g., distance functions like Euclidean distance).	<code>metric='manhattan',</code> <code>n_neighbors=3,</code> <code>weights='distance'</code>	Accuracy - 66.2%, f1-score = 0.77