Model Development Phase Template

Date	12 July 2024
Team ID	SWTID1720083491
Project Title	Early Prediction of Chronic Kidney Disease 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 statistical algorithm which analyze the relationship between two data factors. It gives the probability values between 0 and 1.	C = 100 solver = 'lbfgs'	F1 Score – 88%

Decision Tree	Simple tree structure; interpretable, captures non-linear relationships.	max_depth: 4, max_features': 'sqrt', min_samples_leaf: 1, min_samples_split: 2	F1 Score – 90%
Random Forest	Ensemble of decision trees; robust, handles complex relationships, reduces overfitting, and provides feature importance for prediction.	n_estimators = 300, min_samples_split = 2 min_samples_leaf = 4 max_features = 'log2' max_depth = 8	F1 Score – 94%
SVM	SVM algorithm is to find the optimal hyperplane in an N-dimensional space that can separate the data points in different classes in the feature space.	C = 1 gamma = 0.01 kernel = 'linear'	F1 Score – 88%