Model Development Phase Template

Date	15 March 2024
Team ID	739680
Project Title	Estimating presence or absence of Smoking through bio signals
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)
Random Forest	Ensemble of decision-trees : robust, handles complex relation-ships, reduces overfitting, and provides feature importance for smoking prediction	-	Accuracy score = 64%
KNN	Simple tree-structure : interpretable, captures non-linear relationships, suitable for initializing sights into smoke detection patterns	-	Accuracy Score = 69%

Linear Regression	Linear regression: predicts the value of unknown data by using another related and known data value. It mathematically models the unknown or dependent variable and the known or independent variable as a linear equation.	-	Accuracy Score = 76.4%
Gradient Boosting	Gradient boosting with trees; optimizes predictive performance, handles complex relationships, and is suitable for accurate loan approval predictions.	-	Accuracy Score = 76.4%