Model Development Phase Template

Date	15 March 2024
Team ID	SWTID1720437019
Project Title	Thyroid Classification
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 learning method using multiple decision trees for classification.	- Number of trees: 100 - Maximum depth of trees: 10 - Minimum samples per split: 20	Accuracy
Support Vector Machine (SVM)	Algorithm that finds a hyperplane to separate classes with the maximum margin.	- Kernel: Linear - Cost (C): 1.0 (controls penalty for misclassification) - Gamma: 0.1 (influences kernel function behavior)	Precision

XGBoost	Gradient boosting framework that trains multiple decision trees sequentially.	- Learning rate: 0.1 (controls step size in weight updates) - Maximum depth of trees: 5 - Minimum samples per leaf: 10	Recall
Logistic Regression	Statistical method that predicts the probability of a binary outcome.	- Regularization parameter (lambda): 0.01 (controls model complexity)	F1 Score (harmonic mean of precision and recall)
K-Nearest Neighbors (KNN)	Classifies data points based on the majority vote of its k nearest neighbors.	- Number of neighbors (k): 5	AUC-ROC (Area Under the Receiver Operating Characteristic Curve)