Model Development Phase Template

Date	5 October 2024
Team ID	LTVIP2024TMID24838
Project Title	Detection of Phishing Websites from URLs
Maximum Marks	6 Marks

Model Selection Report

In the forthcoming Model Selection Report, various models will be outlined, detailing their descriptions, hyper parameters, 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	Random Forest is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of their predictions (classification) or mean prediction (regression). It is particularly useful for large datasets and when dealing with multiple variables	`'n_estimators'`
		`'max_depth'`
		`'min_samples_split'`
		`'min_samples_leaf'`
		`'criterion'`
			Accuracy score: 96

KNN	KNN operates on the principle of instance-based learning, meaning it makes predictions based on the entire dataset. This allows it	`'n_neighbors'`
		`'weights'`
		`'metric'`
			Accuracy score: 61

	to utilize all available rows to inform its predictions, making it particularly effective when there is a rich dataset with diverse instances.		

Logistic Regression	Using logistic regression on a large dataset allows for robust analysis and improved predictive capabilities. The larger sample size enhances the model's ability to learn meaningful patterns and generalize well to new data.	`'C'`
		`'penalty'`
		`'solver'`
			Accuracy score: 92

Decision Tree	Decision Trees are a popular supervised learning algorithm used for both classification and regression tasks. They work by splitting the data into subsets based on feature values, creating a tree-like model of decisions. When dealing with large datasets and multiple variables, Decision Trees offer several advantages	`'max_depth'`
		`'min_samples_split'`
		`'min_samples_leaf'`
		`'criterion'`
			Accuracy score: 95