

## Model Development Phase Template

Date	30 September 2024
Team ID	LTVIP2024TMID24973
Project Title	Detection of Phishing Websites from URLs Using Machine learning
Maximum Marks	6 Marks

### Model Selection Report

In the forthcoming Model Selection Report, various models will be outlined, detailing their descriptions 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	Performance Metric (e.g., Accuracy, F1 Score)
Logistic regression	Logistic regression is a statistical method used for binary classification problems, where the outcome variable is categorical with two possible outcomes (e.g., success/failure, yes/no). It models the relationship between one or more independent variables and the probability outcome occurring.	91%

KNN	K-Nearest Neighbors (KNN) is a simple yet powerful algorithm used in machine learning for classification and regression tasks. It operates on the principle of proximity, where the output for a given input is determined by the majority class (in classification) or the average (in regression) of its k nearest neighbors in the feature space.	94%
Random Forest	Random Forest is an ensemble learning method primarily used for classification and regression tasks. It builds multiple decision trees during training and merges their outputs to improve accuracy and control overfitting.	97%