

Model Optimization and Tuning Phase Template

Date	29 September 2024
Team ID	LTVIP2024TMID24973
Project Title	Detection of Phishing Websites from URLs Using Machine learning
Maximum Marks	10 Marks

Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Hyperparameter Tuning Documentation (6 Marks):

Model	Finalized Model Accuracy				
Model 1: Logistic regression	<pre>[] #Splitting the data into train and test from sklearn.model_selection import train_test_split x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,random_state=0)</pre> <pre>[] from sklearn.linear_model import LogisticRegression lr=LogisticRegression() lr.fit(x_train,y_train)</pre> <div><div>LogisticRegression</div><div>LogisticRegression()</div></div> <pre>[] y_pred1=lr.predict(x_test) from sklearn.metrics import accuracy_score log_reg=accuracy_score(y_test,y_pred1) log_reg</pre> <pre>0.9172320217096337</pre> <pre>import pickle pickle.dump(lr,open('Phishing websites.pkl','wb'))</pre> <table><tr><td>0</td><td>Logistic Regression</td><td>0.917232</td><td>0.917232</td></tr></table>	0	Logistic Regression	0.917232	0.917232
0	Logistic Regression	0.917232	0.917232		

Note:

Due to the specific requirements of this project and the dataset's structure, the other models were not implemented in this phase, as Logistic Regression demonstrated superior performance for text classification.

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
Logistic Regression	Logistic regression is justified for phishing website detection from URLs because it provides an interpretable, scalable, and efficient solution to a binary classification problem. Its probabilistic output, ease of handling various types of features, and ability to manage imbalanced data make it a strong candidate for this cybersecurity application.