Model Optimization and Tuning Phase Template

Date	15 March 2024
Team ID	LTVIP2024TMID24955
Project Title	SMS Spam Detection - AIML
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	Tuned Hyperparameters	Optimal Values
Multinomial Naive Bayes	**Alpha** (Laplace smoothing) **Fit Prior** (Whether to learn class prior probabilities)	**Alpha**: 0.5 **Fit Prior**: True
SVC (Sigmoid Kernel)	**C** (Regularization parameter) **Gamma** (Kernel coefficient)	**C**: 0.1 **Gamma**: Scale
SVC (RBF Kernel)	**C** (Regularization parameter) **Gamma** (Kernel coefficient)	**C**: 1.0 **Gamma**: Auto

Decision Tree Classifier	**Max Depth** (Maximum depth of the tree) **Min Samples Split** (Minimum number of samples)	**Max Depth**: 10 **Min Samples Split**: 4

*****In above table TUNED HYPER PARAMATERS*****

```
grid_search_mnb = GridSearchCV(mnb, param_grid_mnb, cv=5)
grid_search_mnb.fit(X_train, y_train)
#Get the best parameters and score
print("Best parameters for MultinomialNB:", grid_search_mnb.best_params_)

grid_search_svc_rbf = GridSearchCV(svc_rbf, param_grid_svc_rbf, cv=5)
grid_search_svc_rbf.fit(X_train, y_train)
#Get the best parameters and score
print("Best parameters for SVC(rbf):", grid_search_svc_rbf.best_params_)

grid_search_svc_sigmoid = GridSearchCV(svc_sigmoid, param_grid_svc_sigmoid, cv=5)
grid_search_svc_sigmoid.fit(X_train, y_train)
#Get the best parameters and score
print("Best parameters for SVC(sigmoid):", grid_search_svc_sigmoid.best_params_)


grid_search_dt = GridSearchCV(dt, param_grid_dt, cv=5)
grid_search_dt.fit(X_train, y_train)
#Get the best parameters and score
print("Best parameters for DecisionTreeClassifier:", grid_search_dt.best_params_)
```

*****In above table OPTIMAL VALUES*****

```
Best parameters for MultinomialNB: {'alpha': 1.0}
Best parameters for SVC(rbf): {'C': 10, 'gamma': 0.1}
Best parameters for SVC(sigmoid): {'C': 10, 'gamma': 0.1}
Best parameters for DecisionTreeClassifier: {'max_depth': None, 'min_samples_split': 5}
```

Performance Metrics Comparison Report (2 Marks):

Model	Baseline Metric	Optimized Metric
Multinomial Naive Bayes	**Accuracy**: 93% **F1 Score**: 0.92	**Accuracy**: 96% **F1 Score**: 0.96
SVC (Sigmoid Kernel)	**Accuracy**: 85% **F1 Score**: 0.80	**Accuracy**: 96% **F1 Score**: 0.91
SVC (RBF Kernel)	**Accuracy: 89%** **F1 Score: 0.85**	**Accuracy: 96%** **F1 Score: 0.92**
Decision Tree Classifier	**Accuracy: 87%** **F1 Score: 0.83**	**Accuracy: 96%** **F1 Score: 0.94**

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
Multinomial Naive Bayes (MNB)	1. **Superior Performance for Text Data**: 2. **Simple and Fast**: 3. **Optimized Performance**: