Model Optimization and Tuning Phase Template

Date	15 July 2024
Team ID	740032
Project Title	Genetic Classification of An Individual By Using Machine Learning
Maximum Marks	10 Marks

Model Optimization and Tuning Phase

Model	Tuned Hyperparameters

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):

Random forest regression

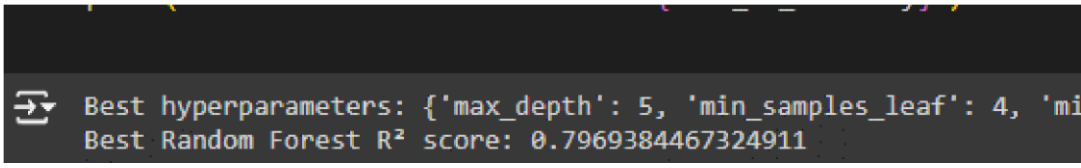
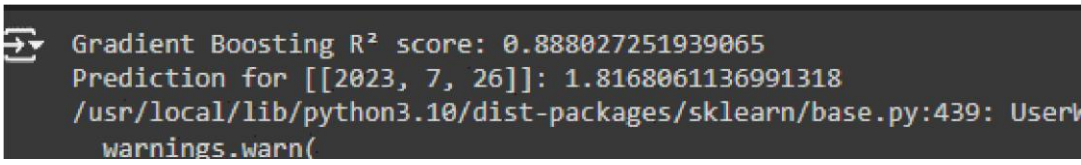
```
from sklearn.model_selection import GridSearchCV

# Define the hyperparameter grid to search over
param_grid = {
    'n_estimators': [50, 100, 200],
    'max_depth': [None, 5, 10],
    'min_samples_split': [2, 5],
    'min_samples_leaf': [1, 2, 4]
}

# Create a Random Forest Regressor model
rf_model = RandomForestRegressor(random_state=0)

# Perform grid search cross-validation
grid_search = GridSearchCV(estimator=rf_model, param_grid=param_grid, cv=5, s
grid_search.fit(x_train, y_train)

# Print the best hyperparameters
print("Best hyperparameters:", grid_search.best_params_)

# Get the best model
best_rf_model = grid_search.best_estimator_

# Make predictions on the test data using the best model
y_pred_best_rf = best_rf_model.predict(x_test)
```

Gradient tree boosting

```
# Define the hyperparameter grid to search over for Gradient Boosting
param_grid_gb = {
    'n_estimators': [50, 100, 200],
    'learning_rate': [0.01, 0.1, 0.2],
    'max_depth': [3, 5, 7]
}

# Create a Gradient Boosting Regressor model
gb_model = GradientBoostingRegressor(random_state=0)

# Perform grid search cross-validation
grid_search_gb = GridSearchCV(estimator=gb_model, param_grid=param_grid_gb, cv=5
grid_search_gb.fit(x_train, y_train)

# Print the best hyperparameters
print("Best hyperparameters for Gradient Boosting:", grid_search_gb.best_params_
```

Performance Metrics Comparison Report (2 Marks):

Model	Baseline Metric	Optimized Metri

RandomForest regression	Not provided (initial metric)	
Gradient tree boosting	Not provided (initial metric)	

Final Model Selection Justification (2 Marks):

Final Model	

Decision tree	**Reasoning** The decision tree was choosen as it given best accuracy without any tunning.