Model Optimization and Tuning Phase Template

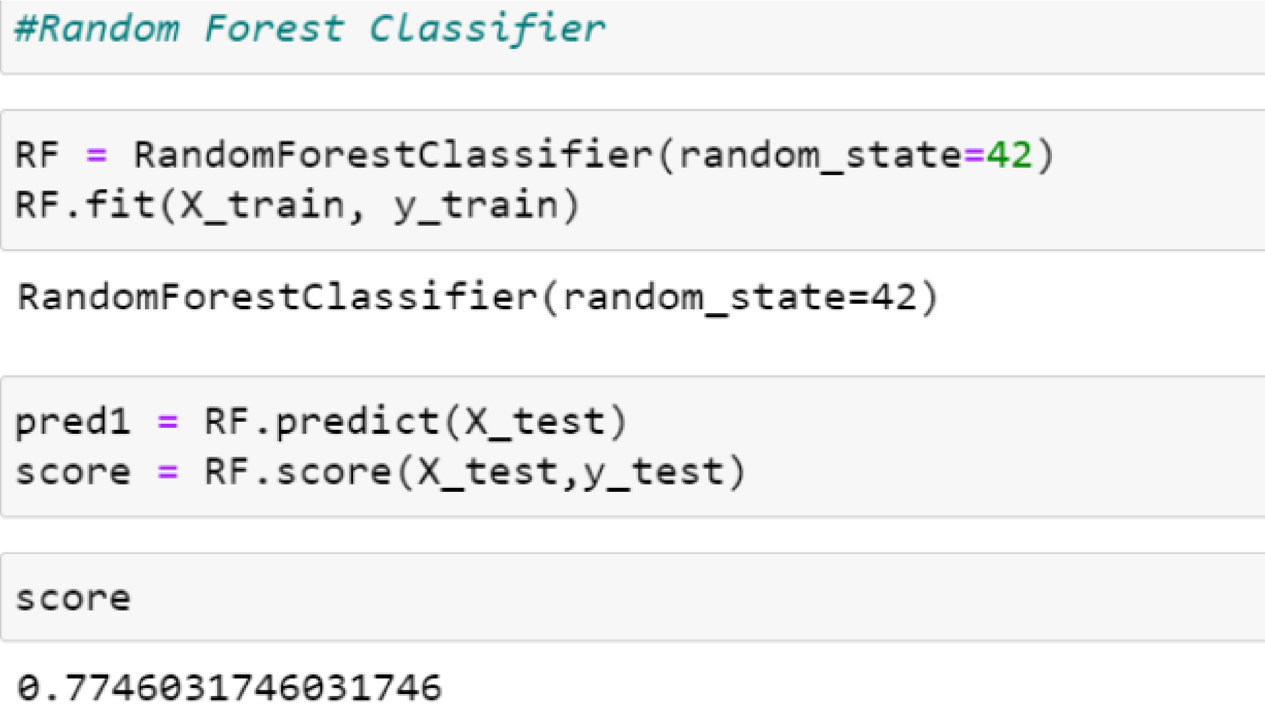
Date	15 July 2024
Team ID	740097
Project Title	Predicting baseline histological stage in HCV patients
Maximum Marks	10 Marks

Model Optimization and Tuning Phase

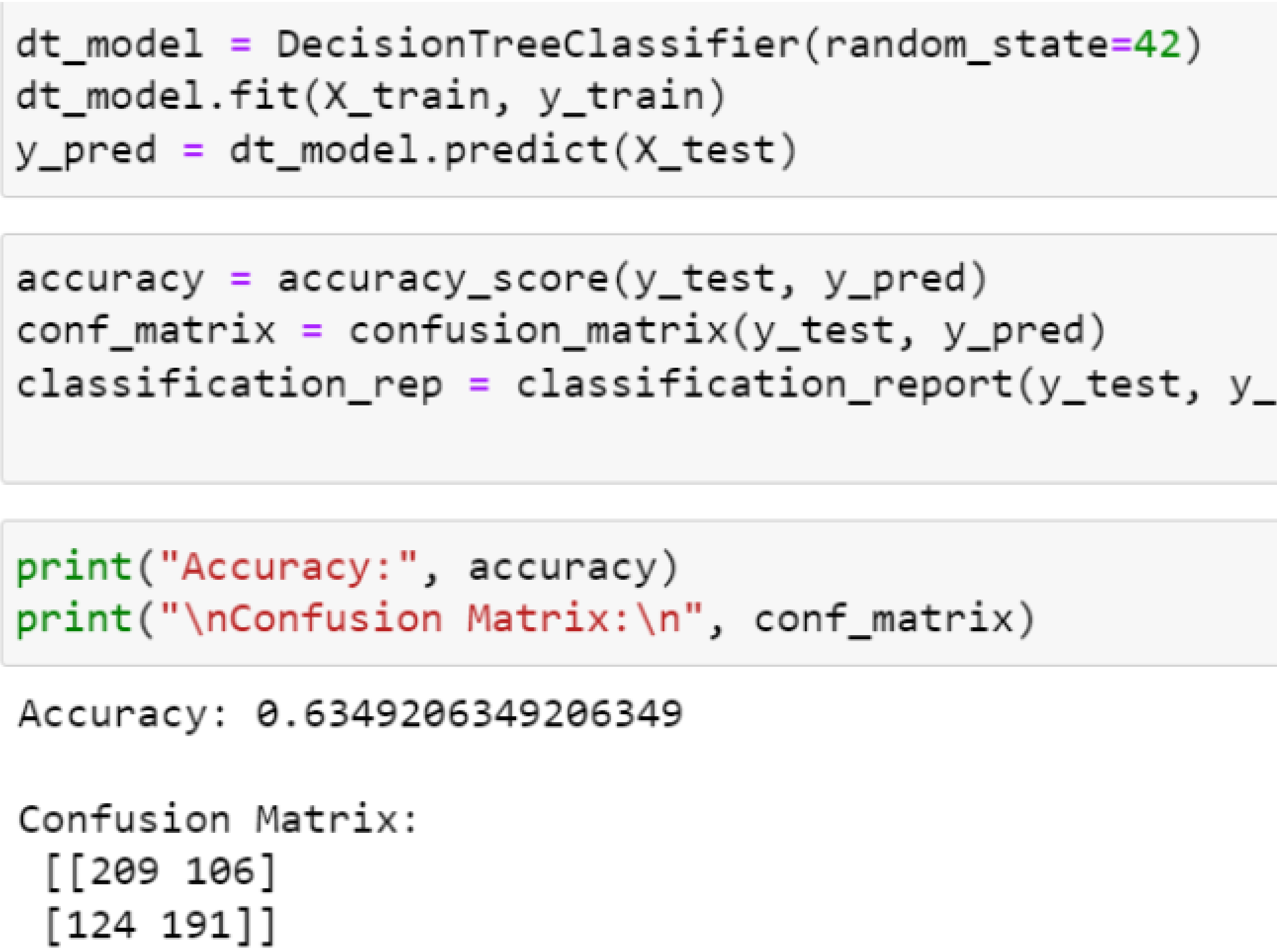
The Model Optimization and Tuning Phase involves refining neural network models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy.

Hyperparameter Tuning Documentation(8 Marks)

Model	Tuned Hyperparameters

Random forest	
	`#Random Forest Classifier`
	`RF = RandomForestClassifier(random_state=42)`
	`RF.fit(X_train, y_train)`
	`RandomForestClassifier(random_state=42)`
	`pred1 = RF.predict(X_test)`
	`score = RF.score(X_test,y_test)`
	`score`
	`0.7746031746031746`

Decision Tree



```
dt_model = DecisionTreeClassifier(random_state=42)
dt_model.fit(X_train, y_train)
y_pred = dt_model.predict(X_test)
```

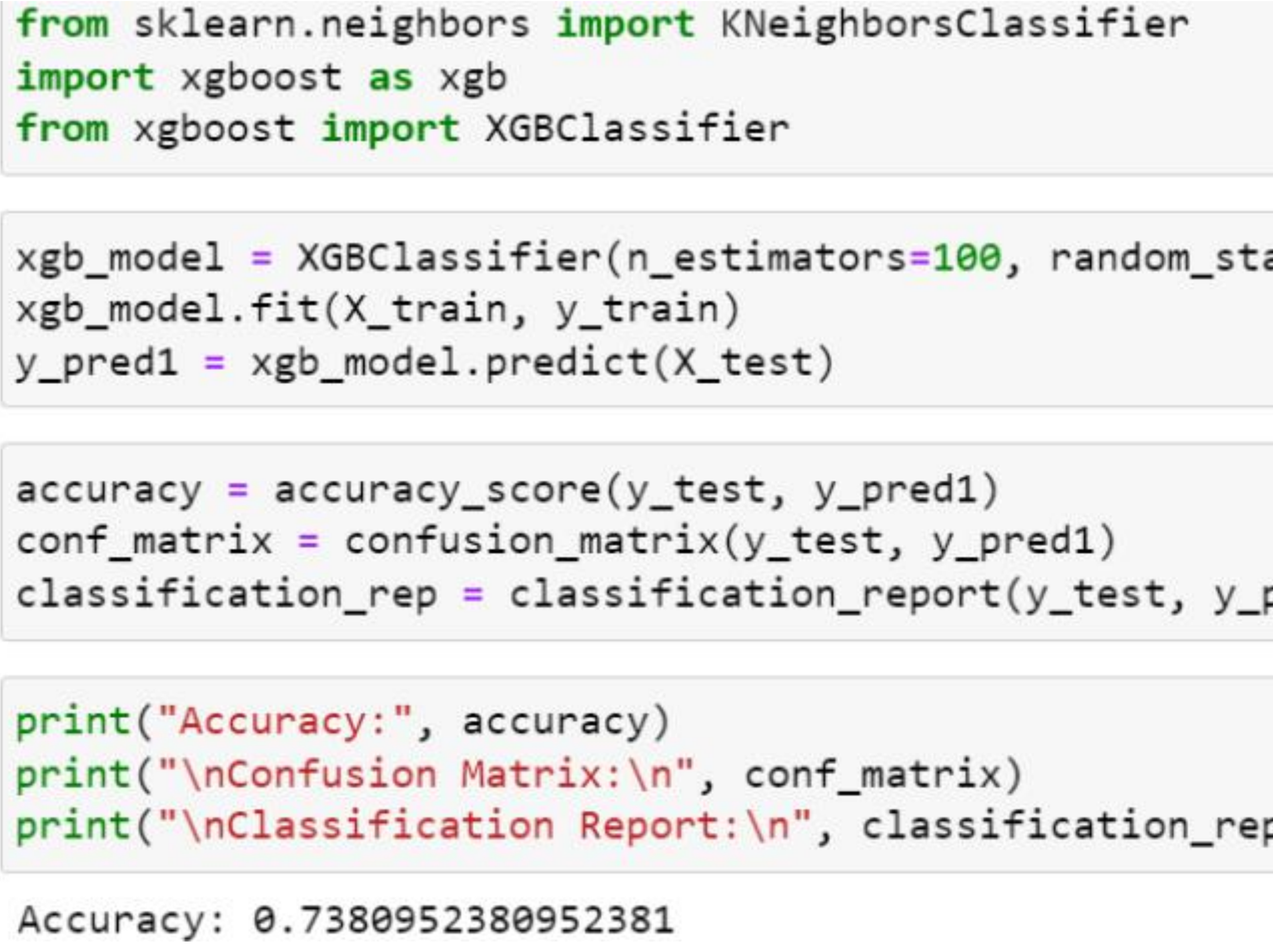
```
accuracy = accuracy_score(y_test, y_pred)
conf_matrix = confusion_matrix(y_test, y_pred)
classification_rep = classification_report(y_test, y_
```

```
print("Accuracy:", accuracy)
print("\nConfusion Matrix:\n", conf_matrix)
```

```
Accuracy: 0.6349206349206349

Confusion Matrix:
 [[209 106]
 [124 191]]
```

XGboost



```
from sklearn.neighbors import KNeighborsClassifier
import xgboost as xgb
from xgboost import XGBClassifier
```

```
xgb_model = XGBClassifier(n_estimators=100, random_sta
xgb_model.fit(X_train, y_train)
y_pred1 = xgb_model.predict(X_test)
```

```
accuracy = accuracy_score(y_test, y_pred1)
conf_matrix = confusion_matrix(y_test, y_pred1)
classification_rep = classification_report(y_test, y_p
```

```
print("Accuracy:", accuracy)
print("\nConfusion Matrix:\n", conf_matrix)
print("\nClassification Report:\n", classification_rep
```

```
Accuracy: 0.7380952380952381
```

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
Decision Tree	• while decision trees have advantages, they also face challenges like overfitting. Combining them with ensemble methods or using advanced algorithms (like XGBoost) can enhance their performance. If you'd like more details or have specific questions, feel free to ask! Decision trees help predict disease severity early.