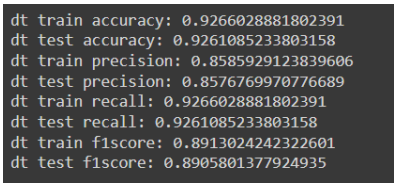
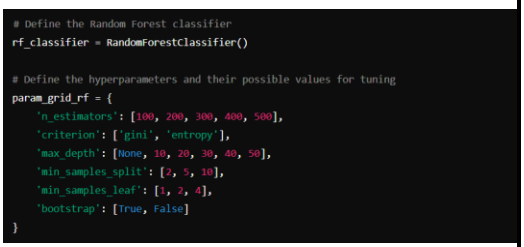
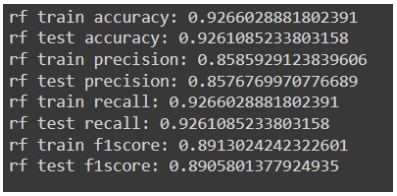
Model Optimization and Tuning Phase Report

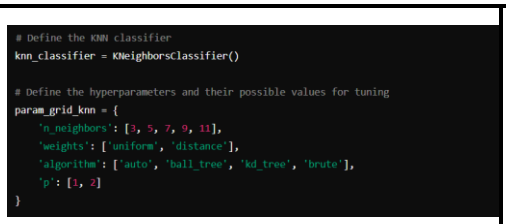
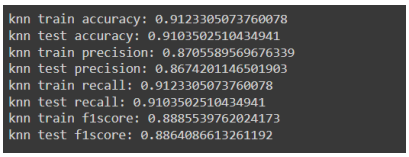
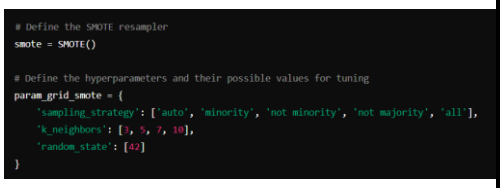
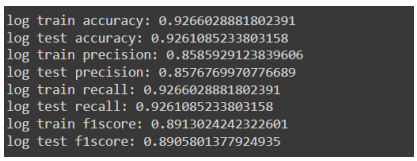
Date	09 July 2024
Team ID	739801
Project Title	Sepsis Survival Minimal Clinical Records
Maximum Marks	10 Marks

Model Optimization and Tuning Phase

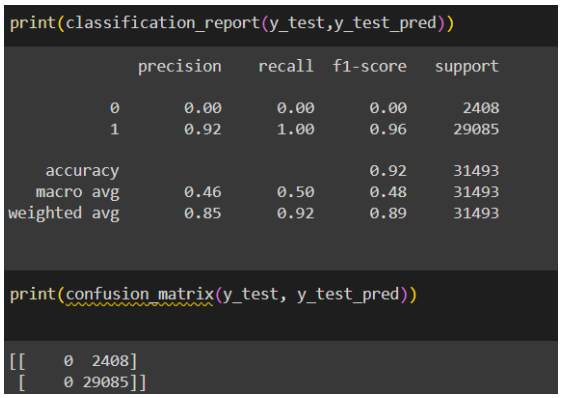
The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Hyperparameter Tuning Documentation (6 Marks):

Model	Tuned Hyperparameters	Optimal Values
Decision Tree		
Random Forest		

KNN		
Logistic Regresson		

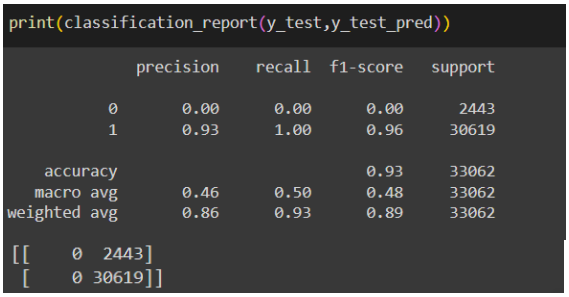
Performance Metrics Comparison Report (2 Marks):

Model	Optimized Metric
Decision Tree	

Random Forest	`print(classification_report(y_test,y_test_pred))`
	
	precision recall f1-score support
	
	0 0.00 0.00 0.00 2408
	1 0.92 1.00 0.96 29085
	
	accuracy 0.92 31493
	macro avg 0.46 0.50 0.48 31493
	weighted avg 0.85 0.92 0.89 31493
	
	`print(confusion_matrix(y_test, y_test_pred))`
	
	[[0 2408]
	[0 29085]]

KNN	`print(classification_report(y_test,y_test_pred))`
	
	precision recall f1-score support
	
	0 0.12 0.72 0.20 2408
	1 0.96 0.54 0.69 29085
	
	accuracy 0.56 31493
	macro avg 0.54 0.63 0.45 31493
	weighted avg 0.89 0.56 0.66 31493
	
	`print(confusion_matrix(y_test, y_test_pred))`
	
	[[1724 684]
	[13243 15842]]

SMOTE	`print(classification_report(y_test,y_pred1))`
	
	precision recall f1-score support
	
	0 0.11 0.73 0.20 2408
	1 0.96 0.53 0.69 29085
	
	accuracy 0.55 31493
	macro avg 0.54 0.63 0.44 31493
	weighted avg 0.89 0.55 0.65 31493
	
	`print(confusion_matrix(y_test, y_test_pred))`
	
	[[1748 660]
	[13534 15551]]

Logistic Regression	

```
print(classification_report(y_test,y_test_pred))
```

```
              precision    recall  f1-score   support

           0       0.00      0.00      0.00      2443
           1       0.93      1.00      0.96     30619

    accuracy                           0.93     33062
   macro avg       0.46      0.50      0.48     33062
weighted avg       0.86      0.93      0.89     33062

[[    0  2443]
 [    0 30619]]
```

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
Random Forest	The decision to select the Random Forest model after applying SMOTE was driven by the need to address class imbalance, leverage the robustness and performance of an ensemble method, gain insights into feature importance, and ensure reliable and correct predictions. This combination aims to enhance the overall predictive capability of the sepsis survival prediction system, contributing to better clinical outcomes.