```
   region  tenure  age  marital  address  income  ed  employ  retire  gender  \
0       2      13   44        1        9    64.0   4       5     0.0       0
1       3      11   33        1        7   136.0   5       5     0.0       0
2       3      68   52        1       24   116.0   1      29     0.0       1
3       2      33   33        0       12    33.0   2       0     0.0       1
4       2      23   30        1        9    30.0   1       2     0.0       0

   reside  custcat
0       2        1
1       6        4
2       2        3
3       1        1
4       4        3
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 12 columns):
 #   Column   Non-Null Count  Dtype
---  ------   --------------  -----
 0   region   1000 non-null   int64
 1   tenure   1000 non-null   int64
 2   age      1000 non-null   int64
 3   marital  1000 non-null   int64
 4   address  1000 non-null   int64
 5   income   1000 non-null   float64
 6   ed       1000 non-null   int64
 7   employ   1000 non-null   int64
 8   retire   1000 non-null   float64
 9   gender   1000 non-null   int64
 10  reside   1000 non-null   int64
 11  custcat  1000 non-null   int64
dtypes: float64(2), int64(10)
memory usage: 93.9 KB
None
Accuracy of KNN (k=4): 0.3250
Classification Report:
              precision    recall  f1-score   support

           1       0.37      0.48      0.42        60
           2       0.15      0.13      0.14        39
           3       0.32      0.31      0.31        55
           4       0.40      0.30      0.35        46

    accuracy                           0.33       200
   macro avg       0.31      0.31      0.30       200
weighted avg       0.32      0.33      0.32       200
```

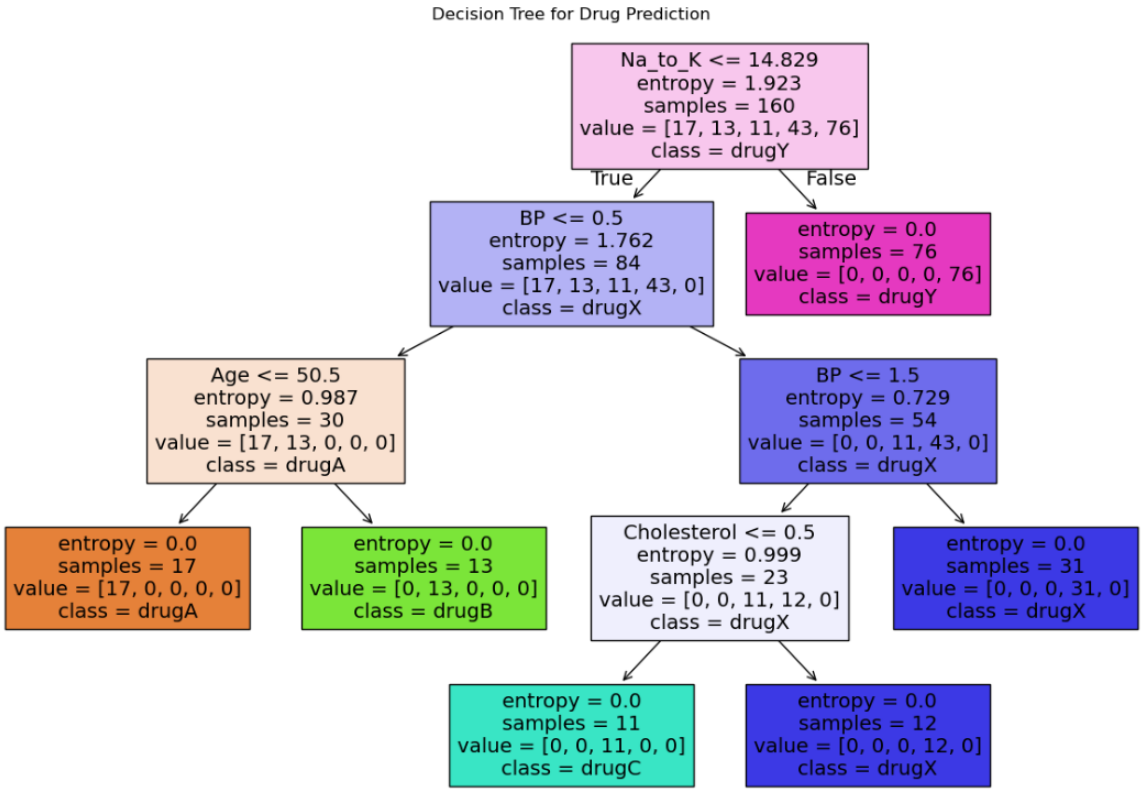
```
   Age Sex      BP Cholesterol  Na_to_K   Drug
0   23   F    HIGH        HIGH   25.355  drugY
1   47   M     LOW        HIGH   13.093  drugC
2   47   M     LOW        HIGH   10.114  drugC
3   28   F  NORMAL        HIGH    7.798  drugX
4   61   F     LOW        HIGH   18.043  drugY
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 6 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   Age          200 non-null    int64
 1   Sex          200 non-null    object
 2   BP           200 non-null    object
 3   Cholesterol  200 non-null    object
 4   Na_to_K      200 non-null    float64
 5   Drug         200 non-null    object
dtypes: float64(1), int64(1), object(4)
memory usage: 9.5+ KB
None
Accuracy of Decision Tree: 1.0000

Classification Report:
              precision    recall  f1-score   support

       drugA       1.00      1.00      1.00         6
       drugB       1.00      1.00      1.00         3
       drugC       1.00      1.00      1.00         5
       drugX       1.00      1.00      1.00        11
       drugY       1.00      1.00      1.00        15

    accuracy                           1.00        40
   macro avg       1.00      1.00      1.00        40
weighted avg       1.00      1.00      1.00        40
```

```
Final predictions:
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 190ms/step
Input: [0 0], Predicted Output: 0.0515, Rounded: 0.0
Input: [0 1], Predicted Output: 0.9150, Rounded: 1.0
Input: [1 0], Predicted Output: 0.9150, Rounded: 1.0
Input: [1 1], Predicted Output: 0.0945, Rounded: 0.0
```



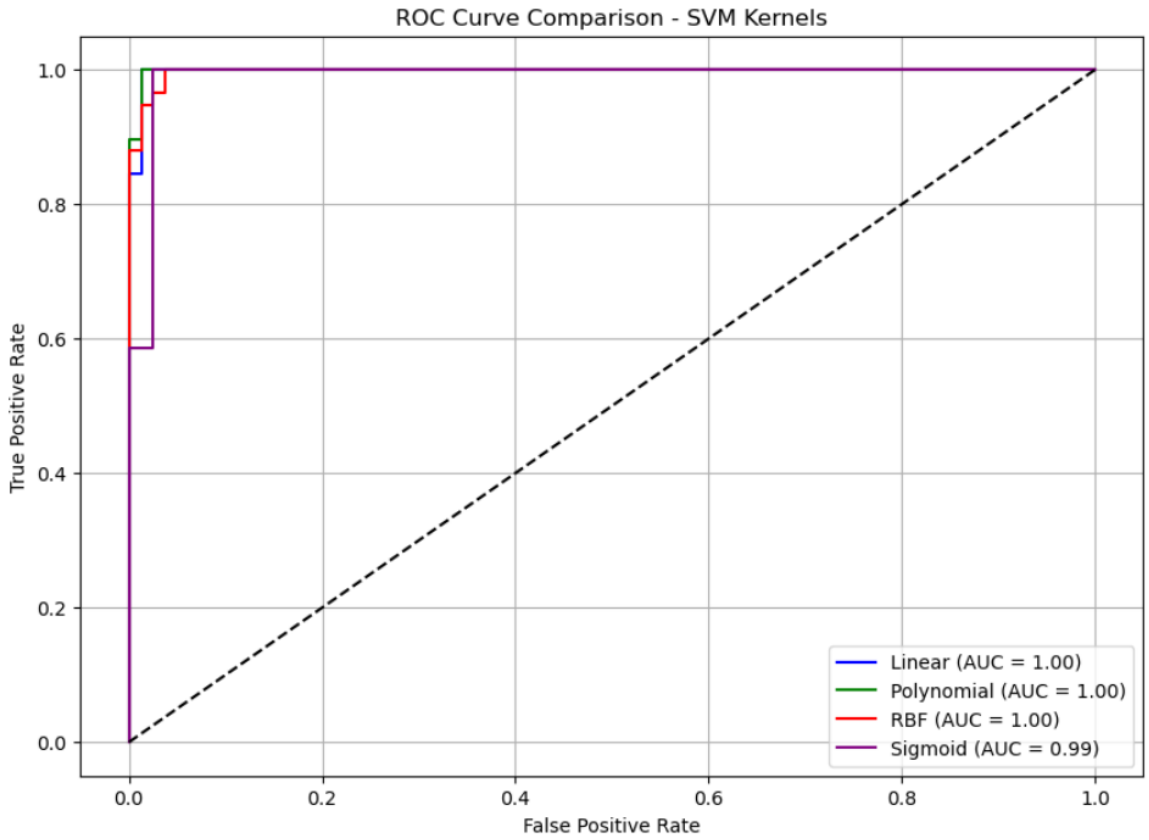
	precision	recall	f1-score	support
0	0.83	0.80	0.81	99
1	0.66	0.71	0.68	55
accuracy			0.77	154
macro avg	0.75	0.75	0.75	154
weighted avg	0.77	0.77	0.77	154

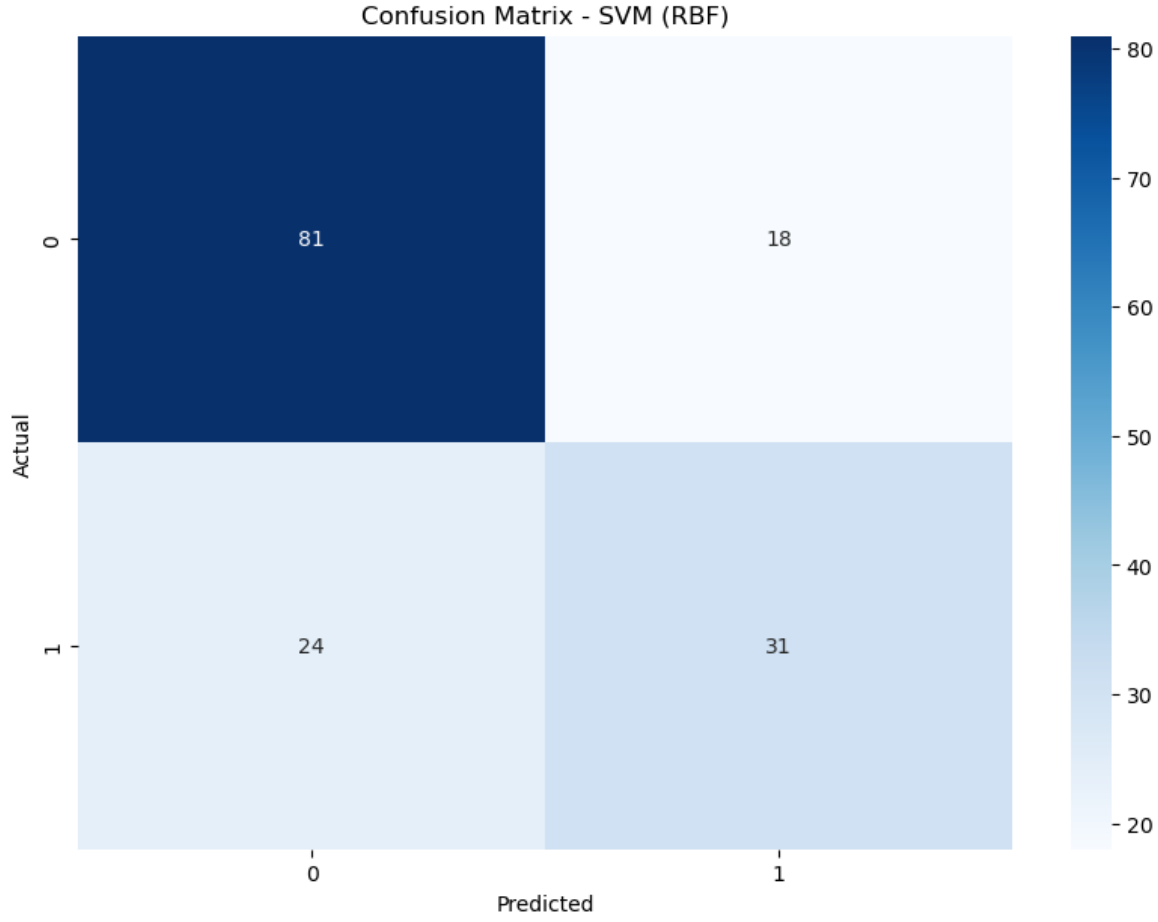
Multinomial Naive Bayes
Accuracy: 0.6623
F1 Score: 0.5185

	precision	recall	f1-score	support
0	0.73	0.75	0.74	99
1	0.53	0.51	0.52	55
accuracy			0.66	154
macro avg	0.63	0.63	0.63	154
weighted avg	0.66	0.66	0.66	154

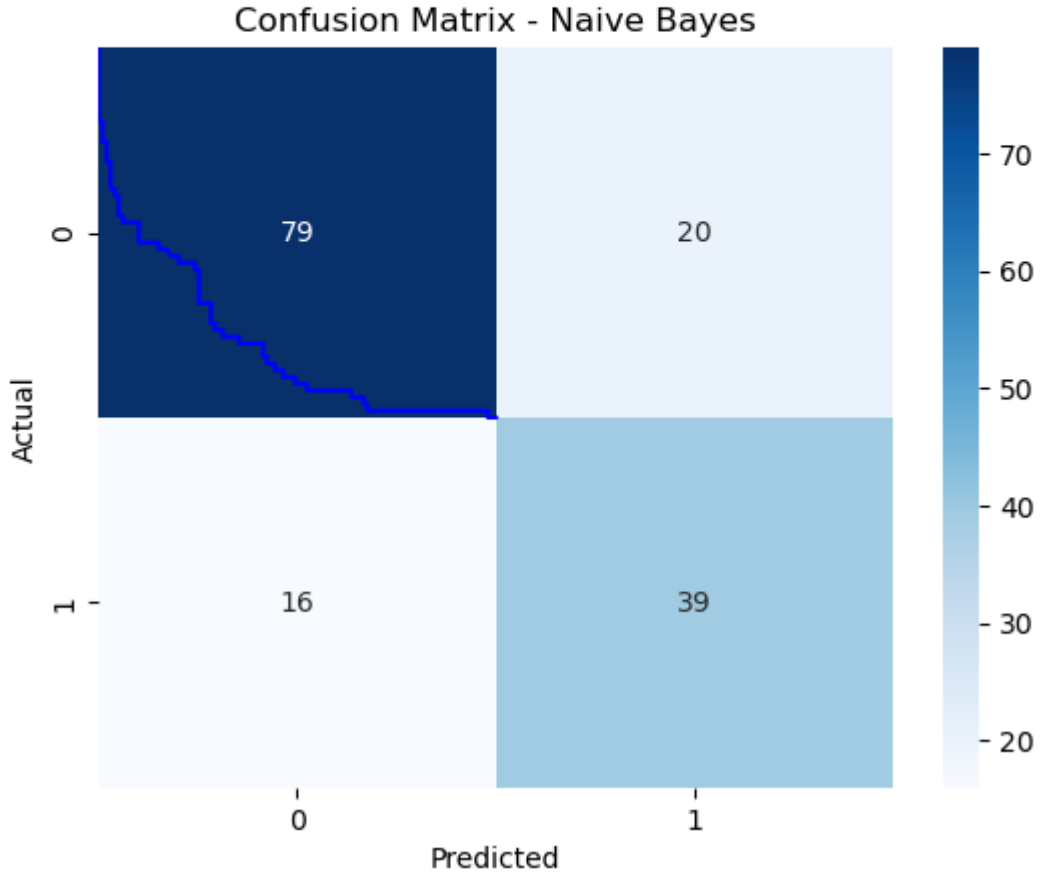
```
   Pregnancies  Glucose  BloodPressure  SkinThickness  Insulin   BMI  \
0            6      148             72             35        0  33.6
1            1       85             66             29        0  26.6
2            8      183             64              0        0  23.3
3            1       89             66             23       94  28.1
4            0      137             40             35      168  43.1

   DiabetesPedigreeFunction  Age  Outcome
0                     0.627   50        1
1                     0.351   31        0
2                     0.672   32        1
3                     0.167   21        0
4                     2.288   33        1
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 768 entries, 0 to 767
Data columns (total 9 columns):
 #   Column                    Non-Null Count  Dtype
---  ------                    --------------  -----
 0   Pregnancies               768 non-null    int64
 1   Glucose                   768 non-null    int64
 2   BloodPressure             768 non-null    int64
 3   SkinThickness             768 non-null    int64
 4   Insulin                   768 non-null    int64
 5   BMI                       768 non-null    float64
 6   DiabetesPedigreeFunction  768 non-null    float64
 7   Age                       768 non-null    int64
 8   Outcome                   768 non-null    int64
dtypes: float64(2), int64(7)
memory usage: 54.1 KB
None
Gaussian Naive Bayes
Accuracy: 0.7662
F1 Score: 0.6842
--- Linear Kernel ---
Accuracy:              0.9708
Recall:                0.9483
Precision:             0.9821
Jaccard Score:         0.9322
Error Rate:            0.0292
Confusion Matrix:
[[78  1]
 [ 3 55]]
--- Polynomial Kernel ---
Accuracy:              0.9124
Recall:                0.7931
Precision:             1.0000
Jaccard Score:         0.7931
Error Rate:            0.0876
Confusion Matrix:
[[79  0]
 [12 46]]
--- RBF Kernel ---
Accuracy:              0.9635
Recall:                0.9483
Precision:             0.9649
Jaccard Score:         0.9167
Error Rate:            0.0365
Confusion Matrix:
[[77  2]
 [ 3 55]]
--- Sigmoid Kernel ---
Accuracy:              0.9635
Recall:                0.9483
Precision:             0.9649
Jaccard Score:         0.9167
Error Rate:            0.0365
Confusion Matrix:
[[77  2]
 [ 3 55]]
```


ROC Curve Comparison - SVM Kernels
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
Linear (AUC = 1.00)
Polynomial (AUC = 1.00)
RBF (AUC = 1.00)
Sigmoid (AUC = 0.99)

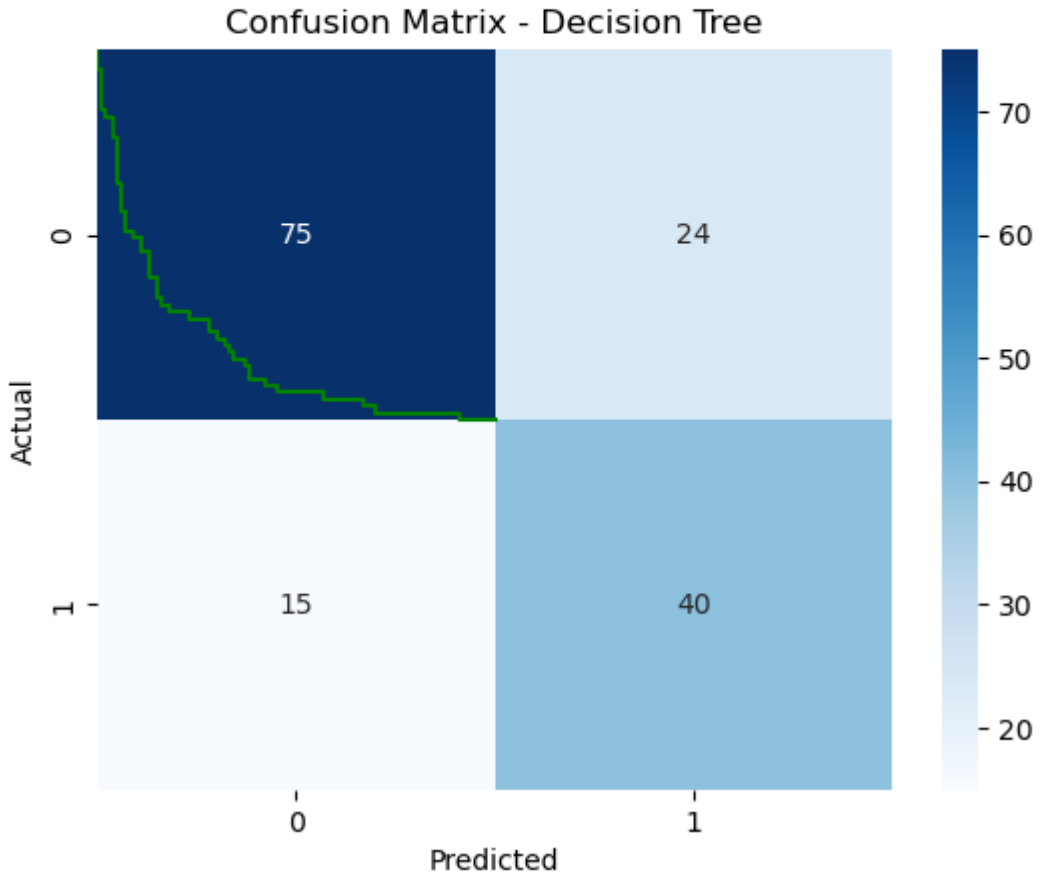

Confusion Matrix - SVM (RBF)
Actual
Predicted
81
18
24
31
0
1
80
70
60
50
40
30
20

```
--- Naive Bayes ---
Accuracy:  0.7662
Recall:    0.7091
Precision: 0.6610
F1 Score:  0.6842
Confusion Matrix:
[[79 20]
 [16 39]]
```

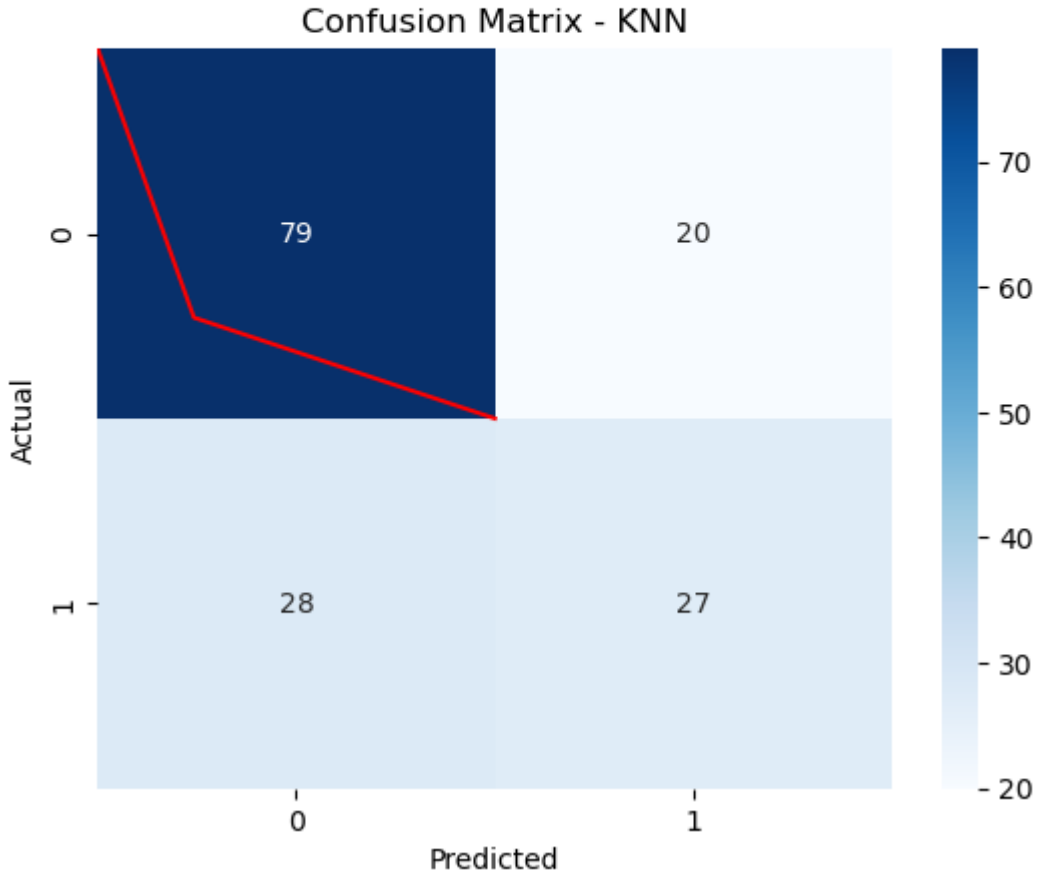


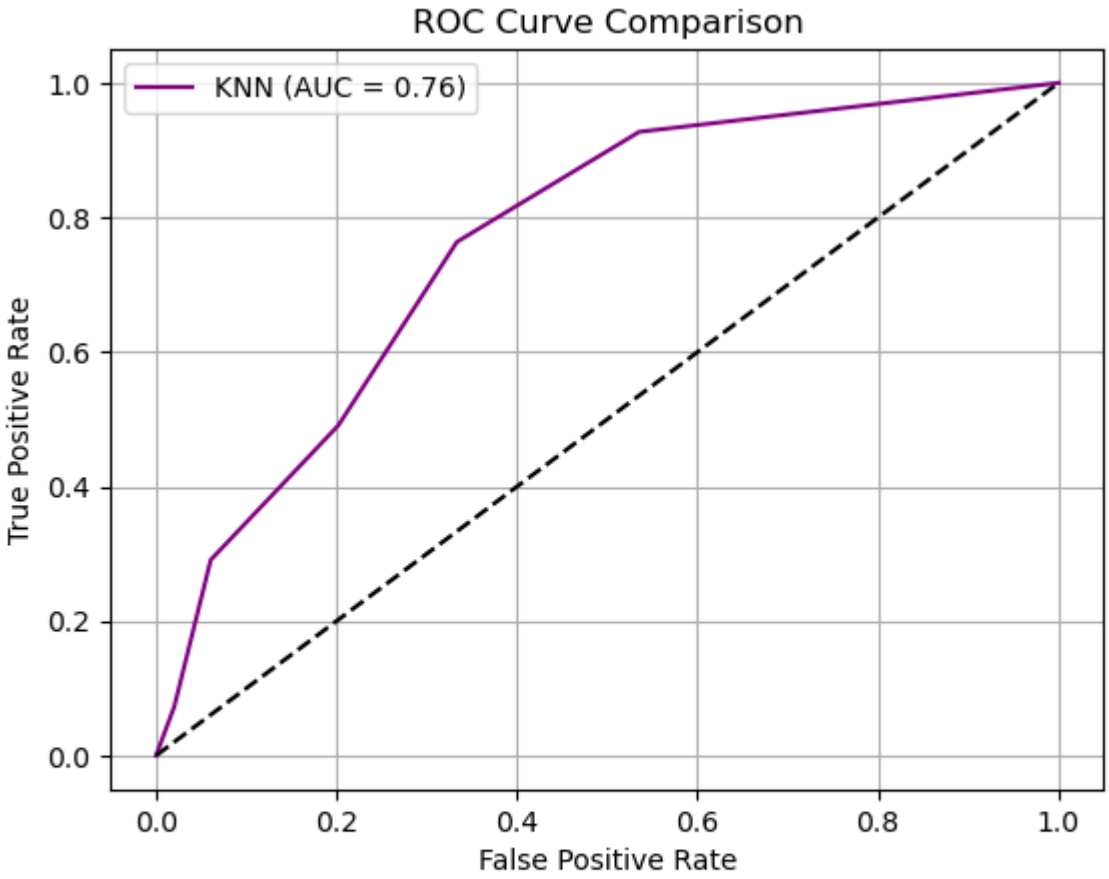

```
--- SVM (RBF) ---
Accuracy:  0.7273
Recall:    0.5636
Precision: 0.6327
F1 Score:  0.5962
Confusion Matrix:
[[81 18]
 [24 31]]
```

```
--- Decision Tree ---
Accuracy:  0.7468
Recall:    0.7273
Precision: 0.6250
F1 Score:  0.6723
Confusion Matrix:
[[75 24]
 [15 40]]
```



```
--- KNN ---
Accuracy:  0.6883
Recall:    0.4909
Precision: 0.5745
F1 Score:  0.5294
Confusion Matrix:
[[79 20]
 [28 27]]
```





	model	accuracy	recall	precision	f1_score
0	SVM (RBF)	0.727273	0.563636	0.632653	0.596154
1	Naive Bayes	0.766234	0.709091	0.661017	0.684211
2	Decision Tree	0.746753	0.727273	0.625000	0.672269
3	KNN	0.688312	0.490909	0.574468	0.529412