Name:- MOHAMMED SHAZI UL ISLAM

SRN:- PES2UG23CS348

SVM LAB {ML}

MOON Dataset

1. **Inferences about the Linear Kernel's performance**
   The Linear Kernel does not fit the Moons dataset well and its performance is lower compared to non-linear kernels. A linear kernel draws a straight decision boundary, which cannot effectively separate the two crescent-shaped classes of the Moons dataset. Because of this, several data points that lie along the curves are misclassified, resulting in lower accuracy, precision and recall scores. In the classification report, the accuracy for the linear kernel was around 0.87. The plot of the decision boundary also showed a straight line cutting across the crescents, confirming that a linear decision boundary is not appropriate for this dataset.

2. **Comparison between RBF and Polynomial kernel decision boundaries**
   Both RBF and Polynomial kernels handle non-linear data better than the linear kernel. However, they differ in the way they shape the decision boundary.
   The Polynomial kernel creates a curved polynomial boundary which fits the data to a certain extent but is less flexible. It achieved an accuracy of around 0.89. On the other hand, the RBF kernel creates a smooth and flexible non-linear boundary that follows the shape of the data more closely. It achieved an accuracy of around 0.97, which is the best among the three kernels. The plots showed that the RBF boundary smoothly enclosed the data, fitting the crescents naturally, whereas the polynomial boundary was more rigid. This shows that for this type of non-linear dataset, the RBF kernel is the most effective.

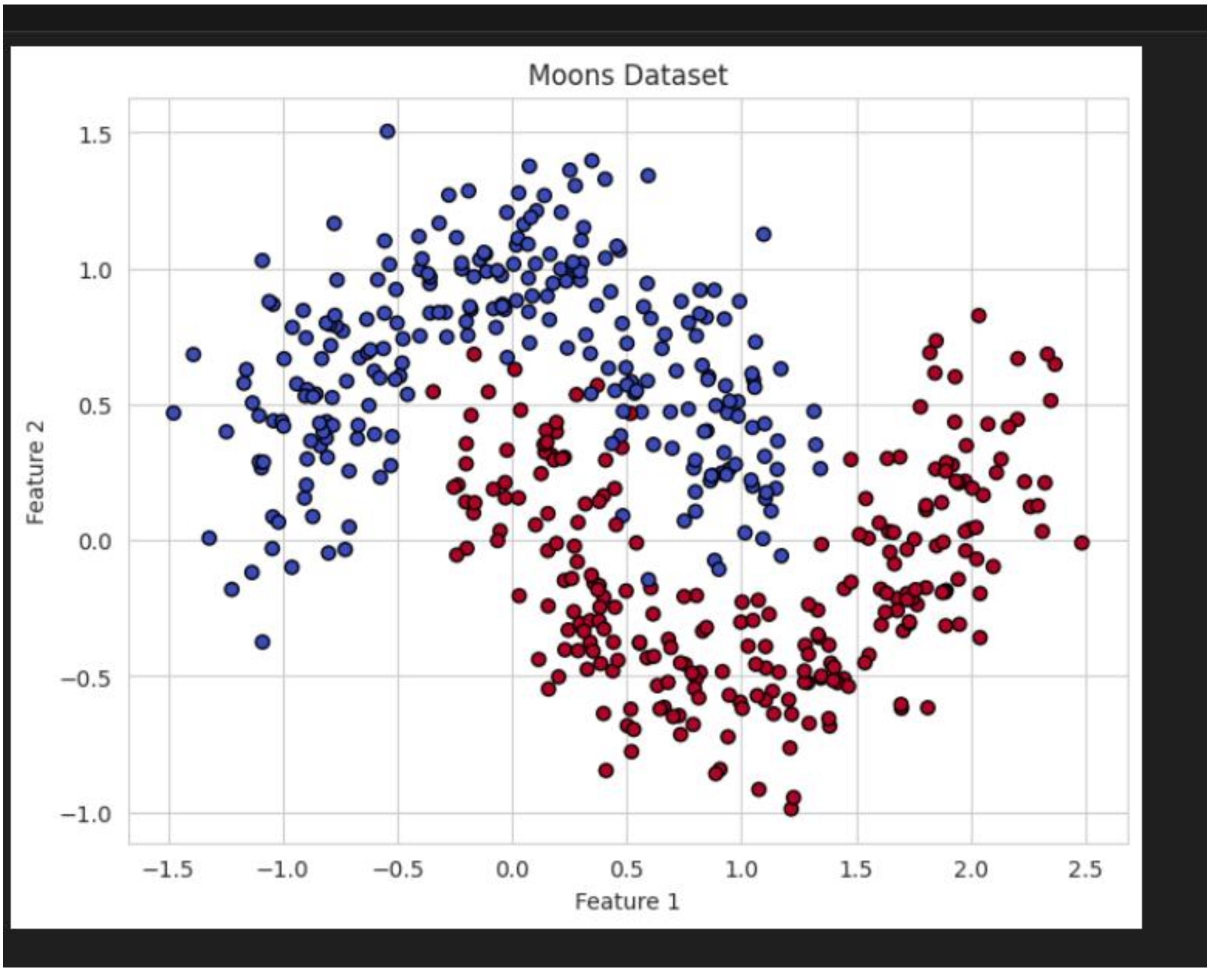
BANKNOTE Dataset

1. **Which kernel was most effective for this dataset**
   The RBF kernel was the most effective for this dataset. Although the Banknote dataset is almost linearly separable, there are slight curves and overlaps in the data. The linear kernel achieved an accuracy of around 0.88 and performed quite well, but the RBF kernel performed better with an accuracy of around 0.93 because of its flexibility in capturing slight non-linear patterns. The polynomial kernel gave the lowest performance at around 0.84, as it introduced unnecessary complexity where it was not needed. The RBF decision boundary followed the separation closely, while the linear boundary missed some overlapping points.

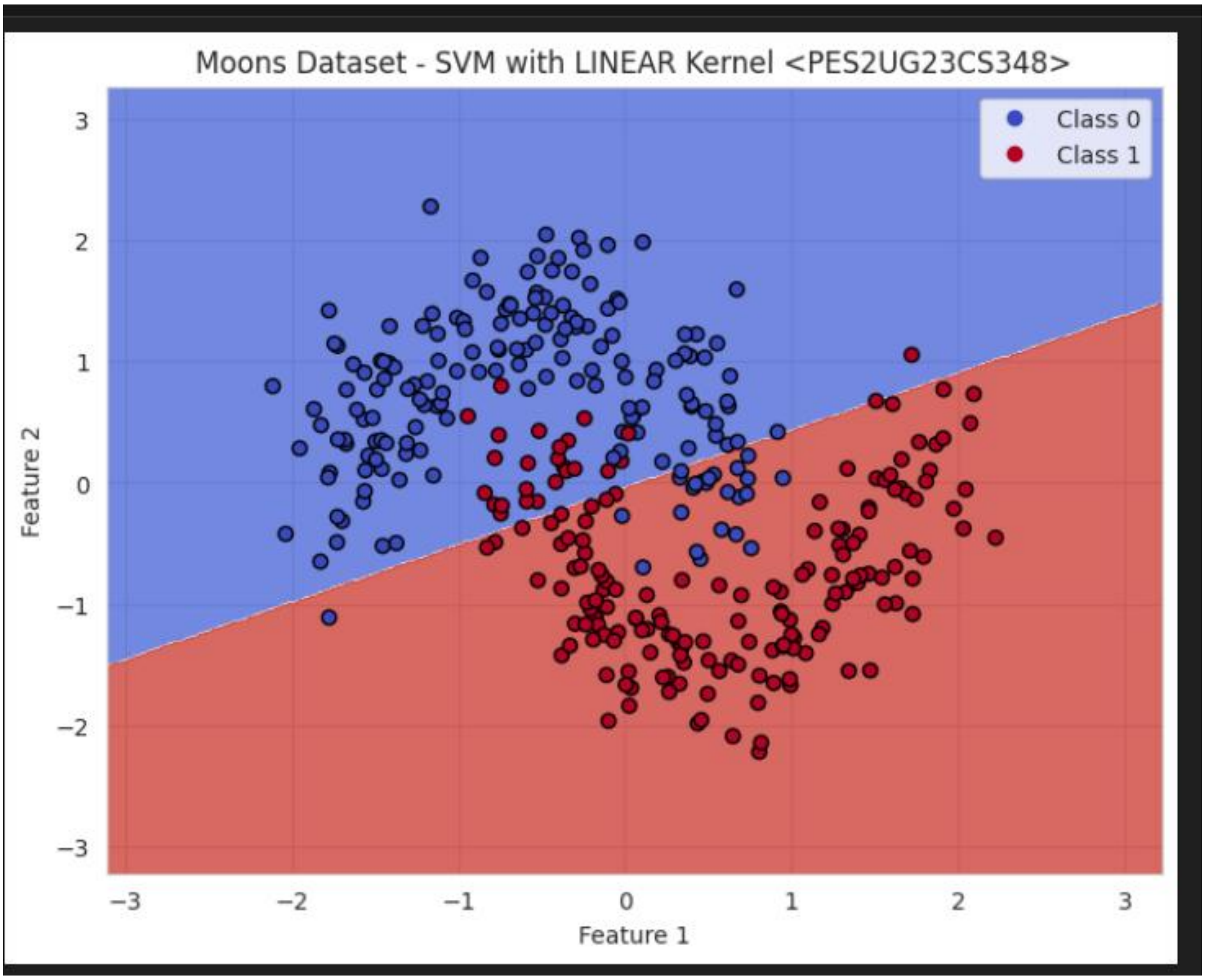
2. **Why the Polynomial kernel underperformed**
   The polynomial kernel underperformed because the Banknote dataset does not have a distinct polynomial structure. It is nearly linear, so trying to fit a polynomial curve results in overfitting. Instead of focusing on the main separation, the polynomial kernel fits to noise and outliers in the data, which leads to a decrease in test accuracy and F1-score. This is why the polynomial kernel performed worse than both the linear and RBF kernels.

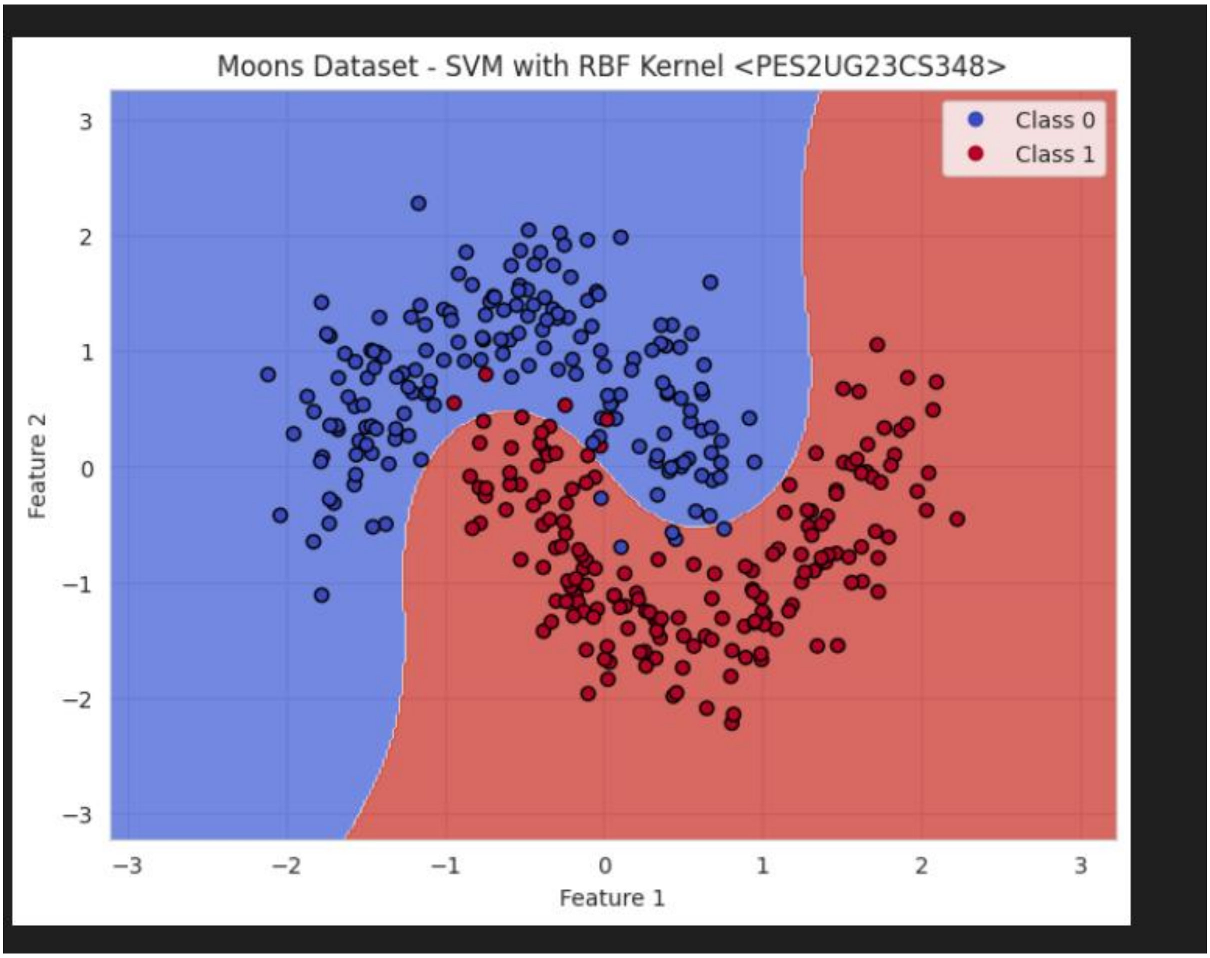
Hard vs. Soft Margin

1. **Which margin (soft or hard) is wider**
   The soft margin produced a wider margin. A smaller C value like 0.1 allows the boundary to be wider and more general.
2. **Why does the soft margin model allow mistakes**
   The soft margin model allows some mistakes to make the model more generalizable. In real-world datasets, some noise or outliers are always present. By allowing misclassifications, the model avoids fitting to these outliers too strictly and learns the overall pattern better.
3. **Which model is more likely to be overfitting and why**
   The hard margin model with a large C value is more likely to overfit. A high C forces the model to classify all training points correctly, even noisy ones, which leads to a very narrow boundary and a poor fit on new data.
4. **Which model would you trust more for new data and why**
   The soft margin model would be more reliable for new data. Its wider boundary and tolerance for a few misclassifications make it more robust and less likely to overfit. It generalizes better than the hard margin model and performs more consistently on unseen samples.

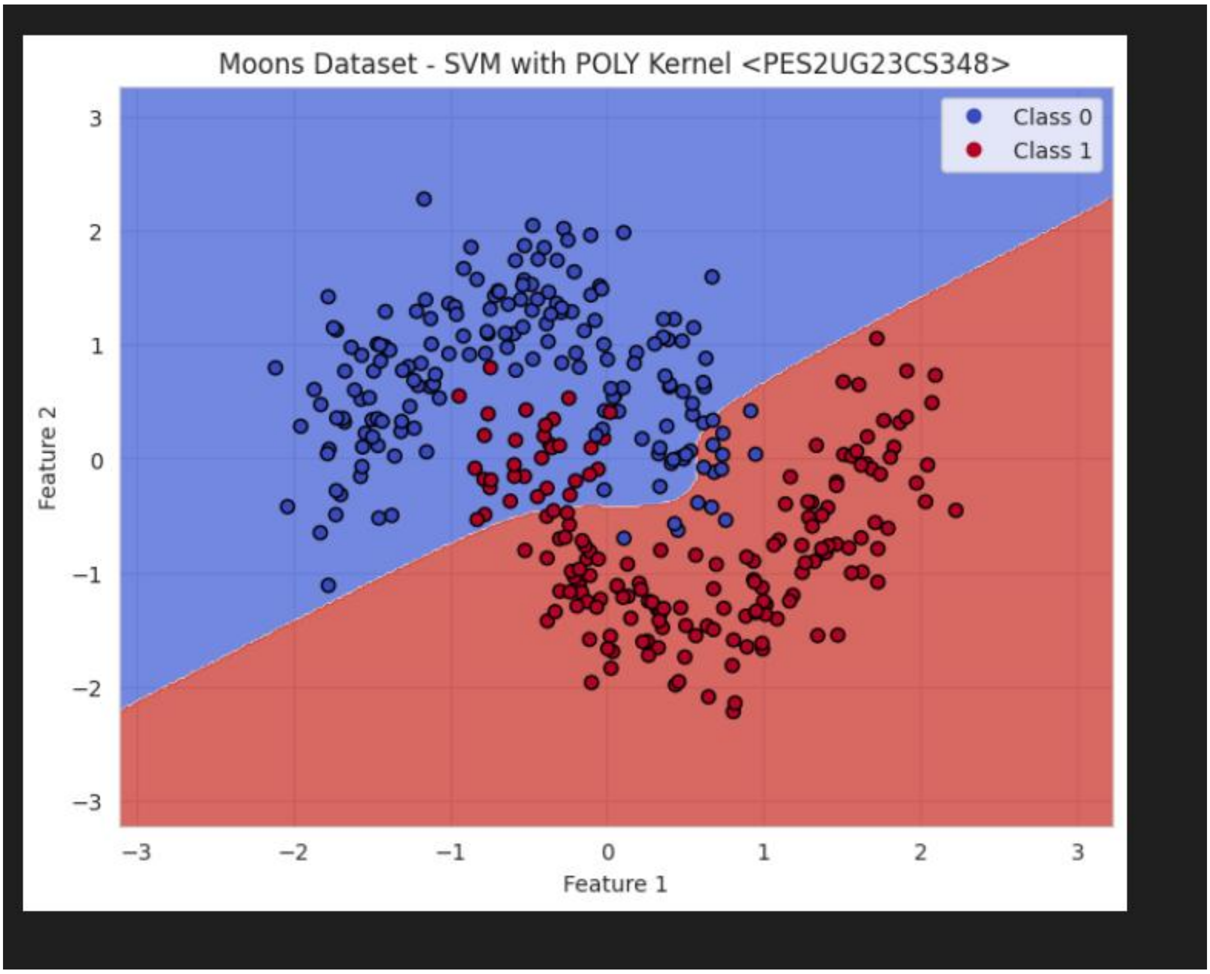


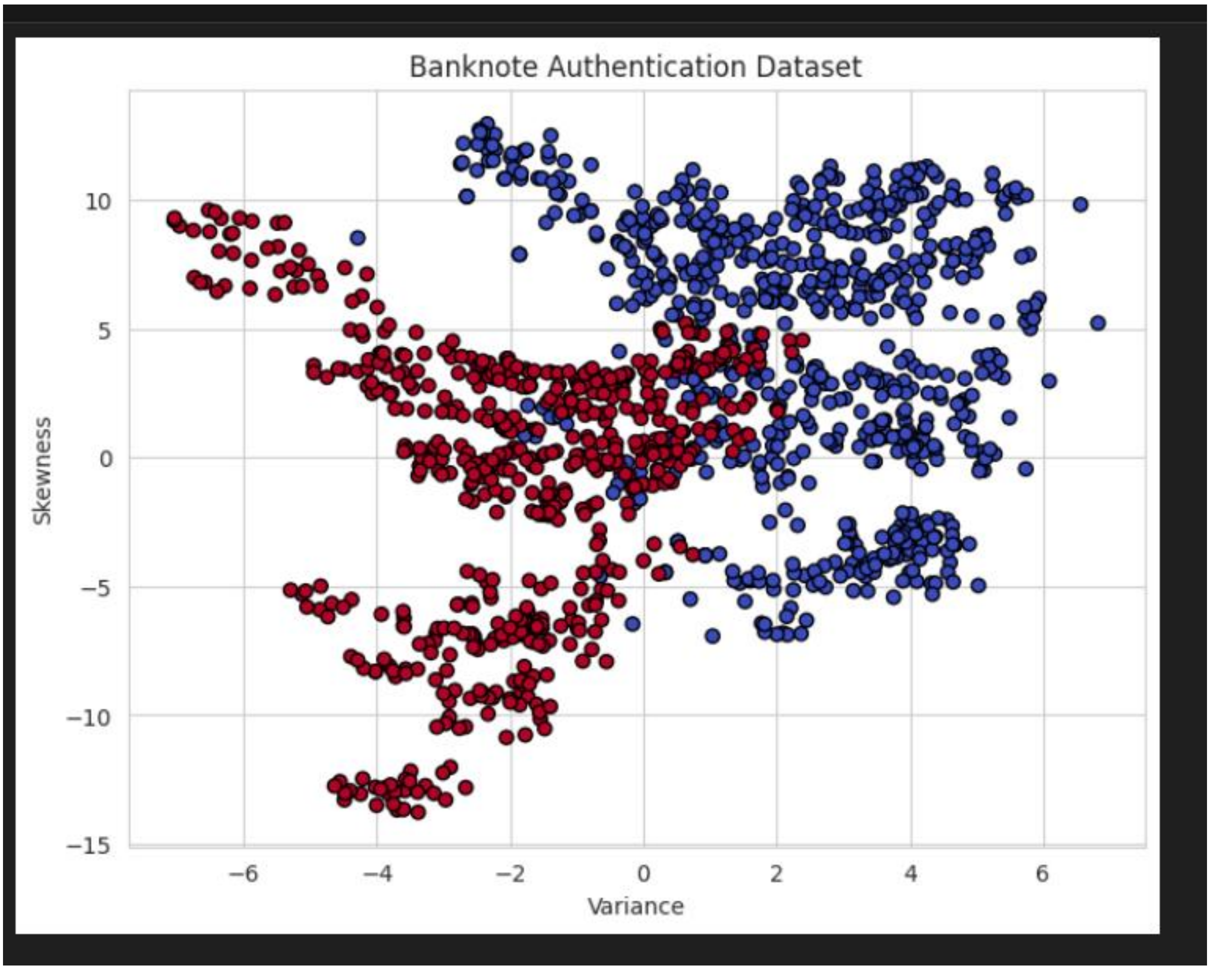
```
[7]

... SVM with LINEAR Kernel PES2UG23CS348
                  precision    recall  f1-score   support

               0       0.85      0.89      0.87        75
               1       0.89      0.84      0.86        75

        accuracy                           0.87       150
       macro avg       0.87      0.87      0.87       150
    weighted avg       0.87      0.87      0.87       150

    ----------------------------------------

    SVM with RBF Kernel PES2UG23CS348
                  precision    recall  f1-score   support

               0       0.95      1.00      0.97        75
               1       1.00      0.95      0.97        75

        accuracy                           0.97       150
       macro avg       0.97      0.97      0.97       150
    weighted avg       0.97      0.97      0.97       150

    ----------------------------------------

    SVM with POLY Kernel PES2UG23CS348
    ...
    weighted avg       0.89      0.89      0.89       150

    ----------------------------------------
```

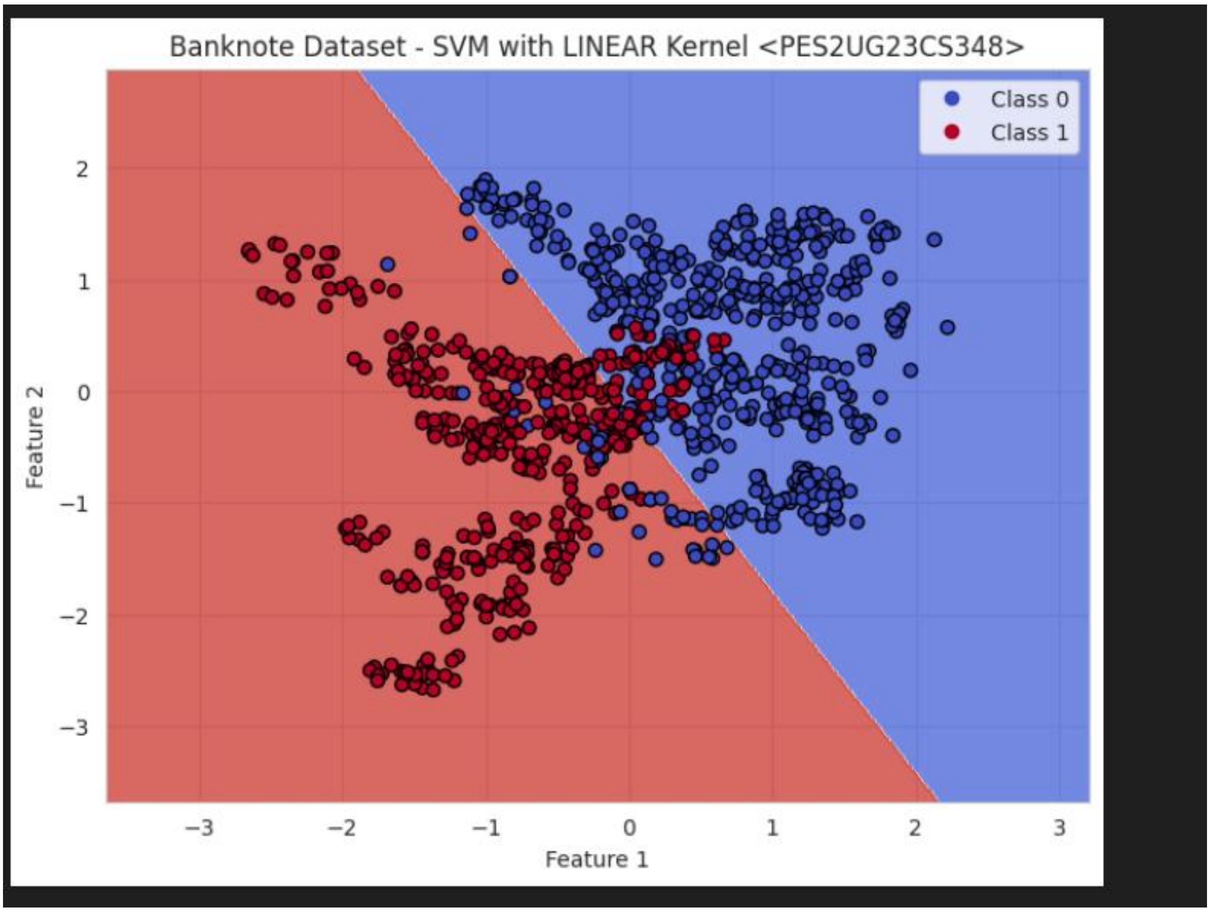

Moons Dataset - SVM with LINEAR Kernel <PES2UG23CS348>
Class 0
Class 1
Feature 2
Feature 1
3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2
3

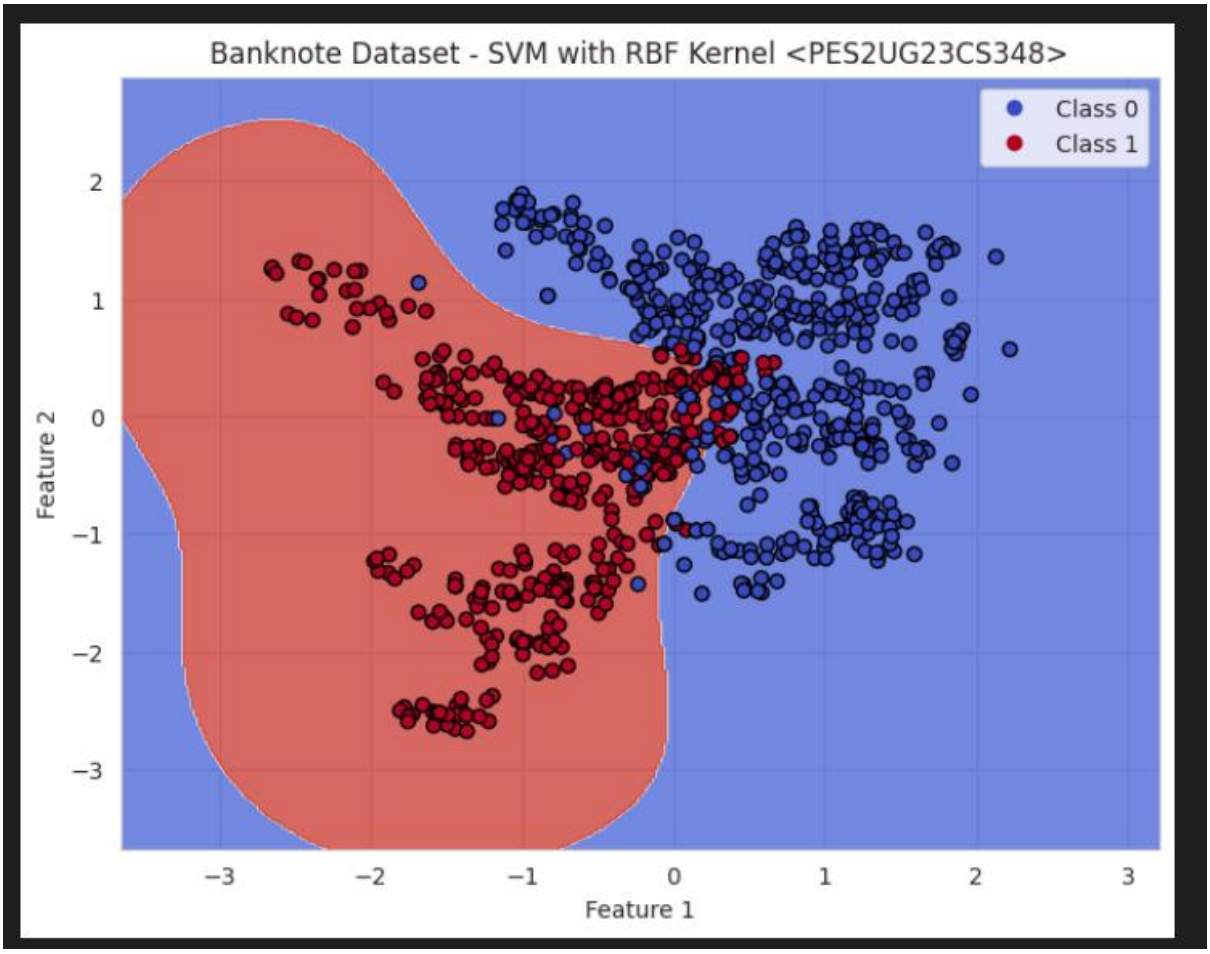


Moons Dataset - SVM with RBF Kernel <PES2UG23CS348>
Class 0
Class 1
Feature 2
Feature 1
3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2
3

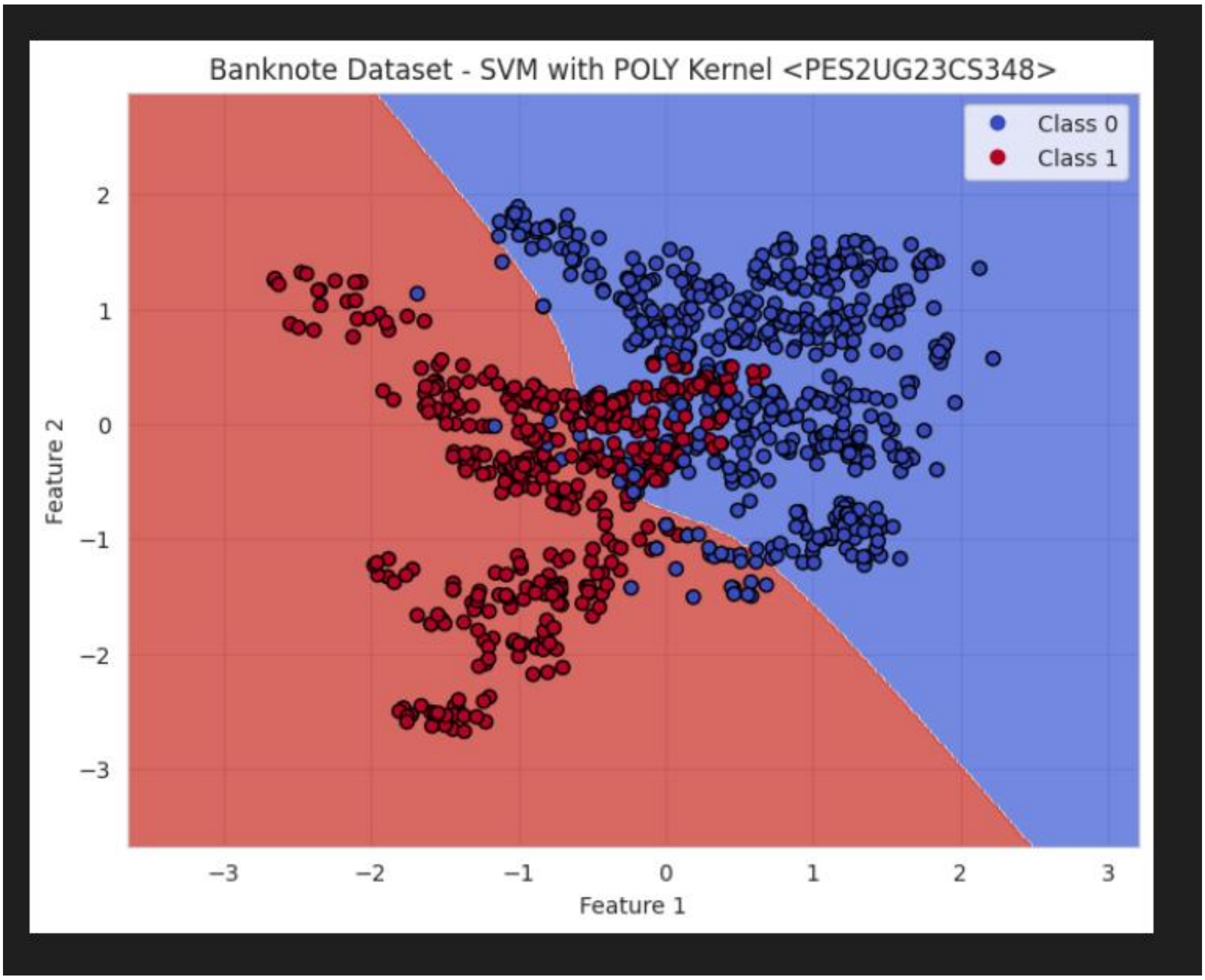

Moons Dataset - SVM with POLY Kernel <PES2UG23CS348>
Class 0
Class 1
Feature 1
Feature 2
−3
−2
−1
0
1
2
3

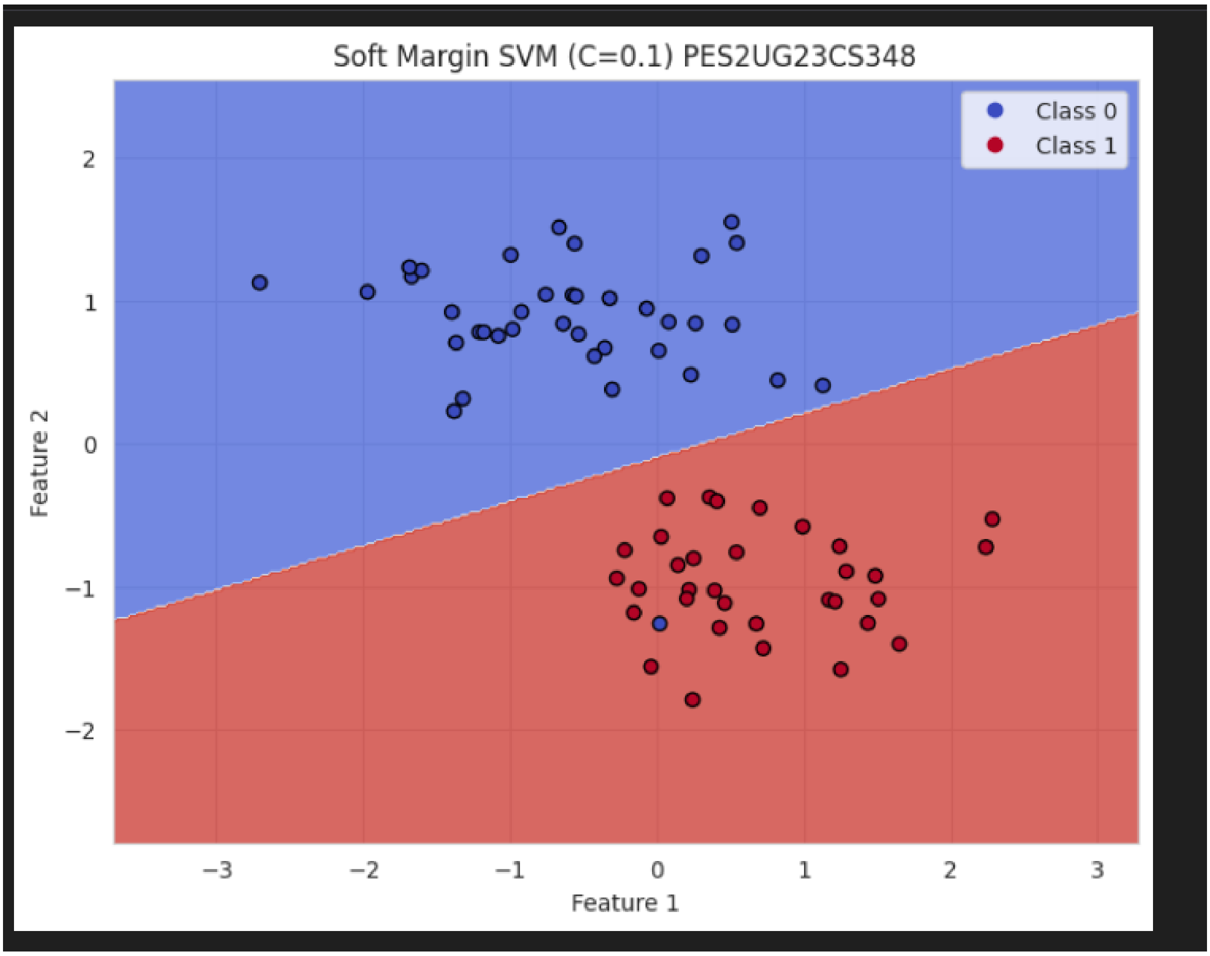

Banknote Authentication Dataset
Skewness
10
5
0
−5
−10
−15
−6
−4
−2
0
2
4
6
Variance

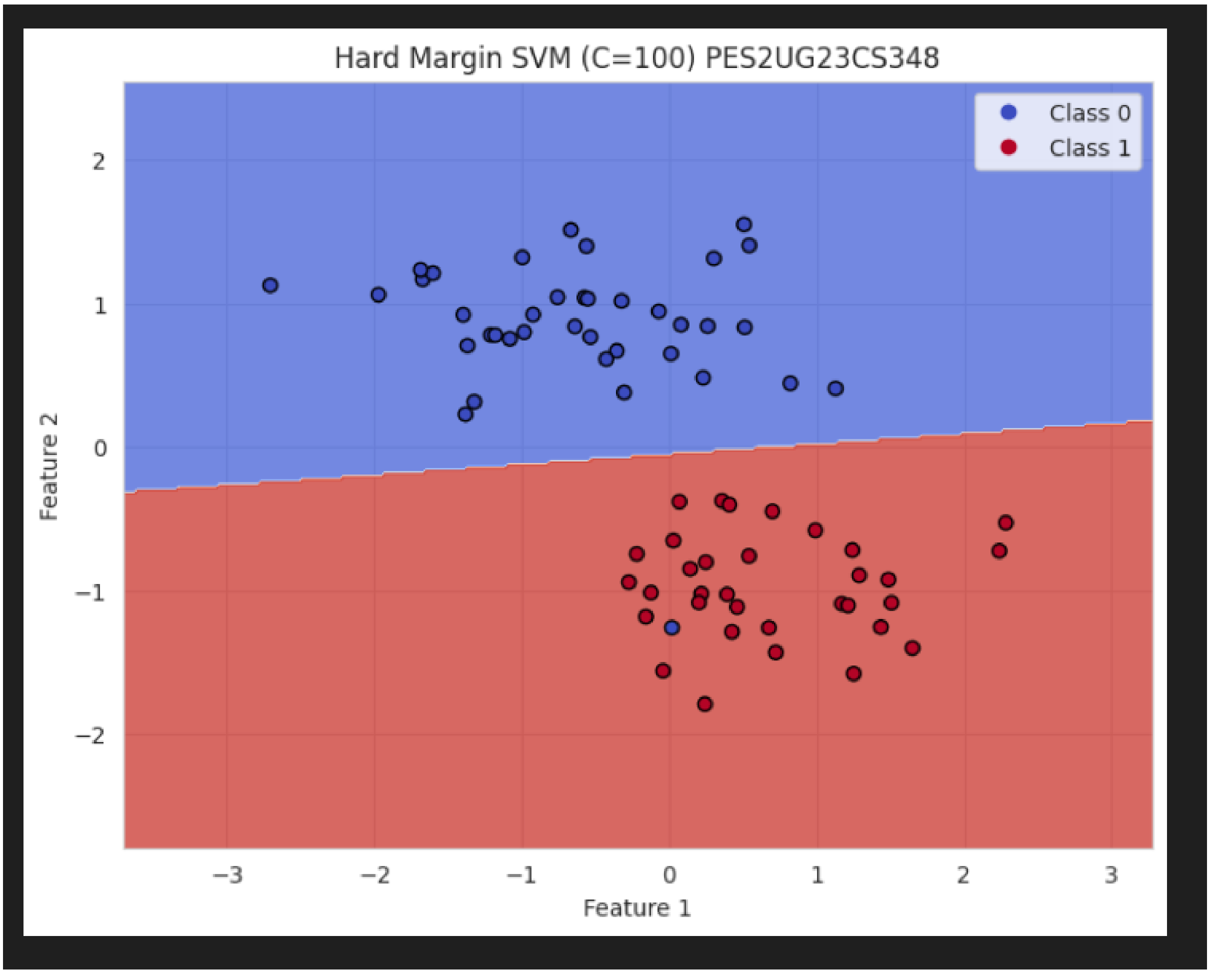

Banknote Dataset - SVM with LINEAR Kernel <PES2UG23CS348>
Class 0
Class 1
Feature 2
Feature 1
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2
3



Banknote Dataset - SVM with RBF Kernel <PES2UG23CS348>
Class 0
Class 1
2
1
0
−1
−2
−3
Feature 2
−3
−2
−1
0
1
2
3
Feature 1


Banknote Dataset - SVM with POLY Kernel <PES2UG23CS348>
Class 0
Class 1
Feature 1
Feature 2
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2
3


Soft Margin SVM (C=0.1) PES2UG23CS348
Class 0
Class 1
Feature 1
Feature 2
−3
−2
−1
0
1
2
3



Hard Margin SVM (C=100) PES2UG23CS348
Class 0
Class 1
Feature 2
Feature 1
2
1
0
−1
−2
−3
−2
−1
0
1
2
3