Machine Learning Lab

Name - G S S Surya Prakash	**SRN** - PES2UG23CS192	**Class** - 5 'C'	**Topic** - Support Vector Machines

Lab - 10

Analysis Questions

1. **Inferences about the Linear Kernel's Performance**
   The Linear Kernel creates a straight decision boundary.
   It performs reasonably well but fails to fully separate overlapping data points. This indicates that the dataset is not linearly separable and requires a non-linear kernel for better accuracy.

2. **Comparison Between RBF and Polynomial Kernel Decision Boundaries**
   - The **RBF Kernel** forms a smooth, adaptive boundary that fits the complex non-linear patterns effectively.
   - The **Polynomial Kernel** creates a curved but less flexible boundary, sometimes overfitting or underfitting parts of the data.
   - Overall, the **RBF Kernel** achieves better separation and generalization.

3. **Which Kernel Was Most Effective?**
   RBF Kernel was the most effective for this dataset. It captured complex feature relationships better than the linear and polynomial kernels.

4. **Why Might the Polynomial Kernel Have Underperformed?**
   The Banknote dataset has complex, real-world feature interactions not easily expressed by polynomial relationships. A polynomial kernel (with fixed degree) cannot flexibly adapt to such

variability.Hence, it underperforms compared to the more adaptable RBF Kernel.

5. **Which Margin Is Wider?**
   The Soft Margin (C = 0.1) is wider because it allows more flexibility and tolerates misclassifications.
6. **Why Does the Soft Margin Model Allow "Mistakes"?**
   The Soft Margin SVM prioritizes a larger margin and better generalization, even if a few points are misclassified. It balances margin width and classification accuracy to avoid overfitting.
7. **Which Model Is More Likely to Overfit and Why?**
   The Hard Margin (C = 100) model is more likely to overfit, since it tries to perfectly classify all training points, making it sensitive to noise.
8. **Which Model Would You Trust More for New Data and Why?**
   Soft Margin SVM is more reliable for unseen data. It generalizes better and handles noisy or imperfect datasets effectively.

Analysis ScreenShots

1.

```
SVM with LINEAR Kernel <PES2UG23CS192>
              precision    recall  f1-score   support

           0     0.8481    0.8933    0.8701        75
           1     0.8873    0.8400    0.8630        75

    accuracy                         0.8667       150
   macro avg     0.8677    0.8667    0.8666       150
weighted avg     0.8677    0.8667    0.8666       150

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

2.

```
SVM with RBF Kernel <PES2UG23CS192>
              precision    recall  f1-score   support

           0     0.9615    1.0000    0.9804        75
           1     1.0000    0.9600    0.9796        75

    accuracy                         0.9800       150
   macro avg     0.9808    0.9800    0.9800       150
weighted avg     0.9808    0.9800    0.9800       150

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

3.

```
SVM with POLY Kernel <PES2UG23CS192>
              precision    recall  f1-score   support

           0     0.9296    0.8800    0.9041        75
           1     0.8861    0.9333    0.9091        75

    accuracy                         0.9067       150
   macro avg     0.9078    0.9067    0.9066       150
weighted avg     0.9078    0.9067    0.9066       150

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

```
SVM with LINEAR Kernel <PES2UG23CS192>
              precision    recall  f1-score   support

      Forged       0.90      0.88      0.89       229
     Genuine       0.86      0.88      0.87       183

    accuracy                           0.88       412
   macro avg       0.88      0.88      0.88       412
weighted avg       0.88      0.88      0.88       412

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

4.

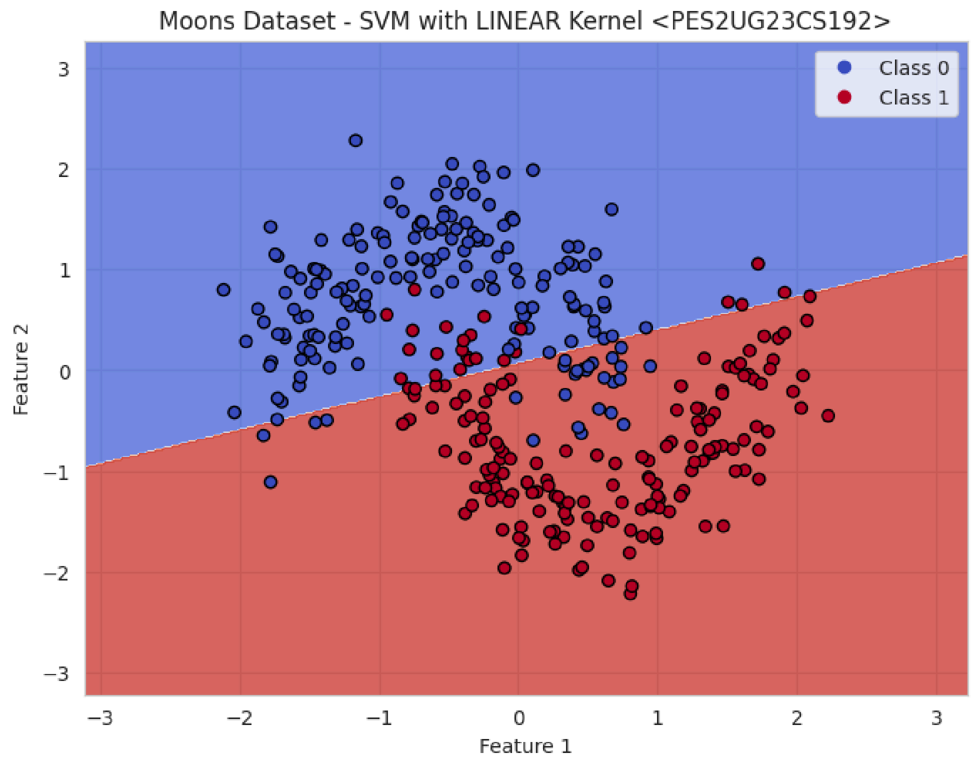
```
SVM with RBF Kernel <PES2UG23CS192>
              precision    recall  f1-score   support

      Forged       0.96      0.91      0.94       229
     Genuine       0.90      0.96      0.93       183

    accuracy                           0.93       412
   macro avg       0.93      0.93      0.93       412
weighted avg       0.93      0.93      0.93       412

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

5.

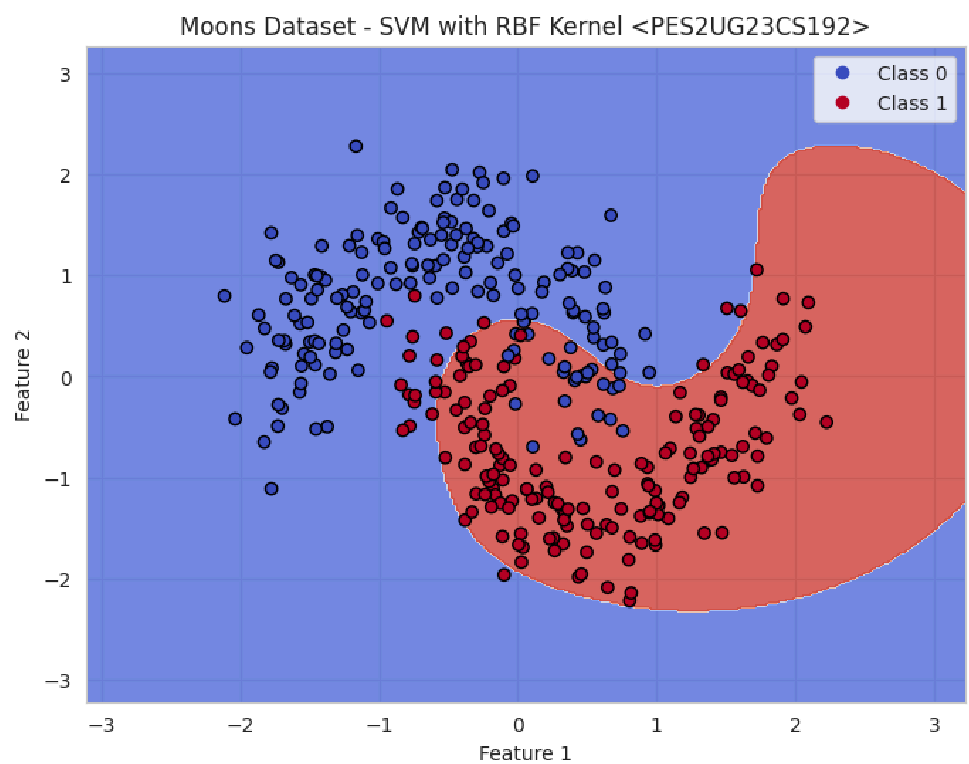
```
SVM with POLY Kernel <PES2UG23CS192>
              precision    recall  f1-score   support

      Forged       0.96      0.81      0.88       229
     Genuine       0.80      0.96      0.88       183

    accuracy                           0.88       412
   macro avg       0.88      0.89      0.88       412
weighted avg       0.89      0.88      0.88       412

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

6.

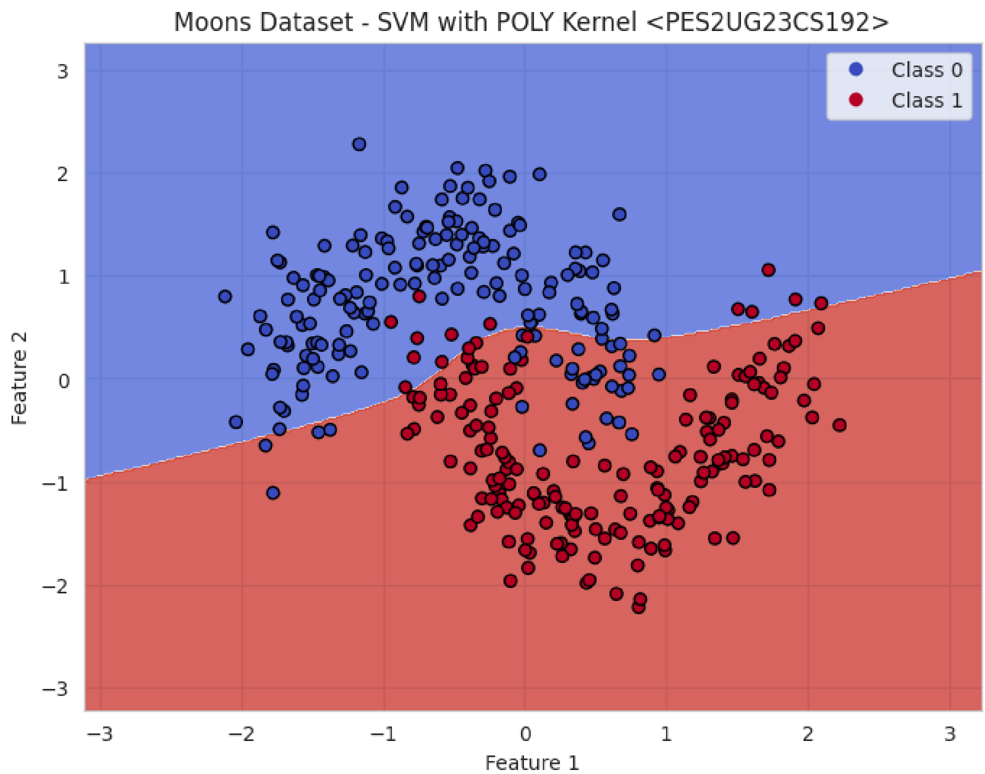
Plot ScreenShots



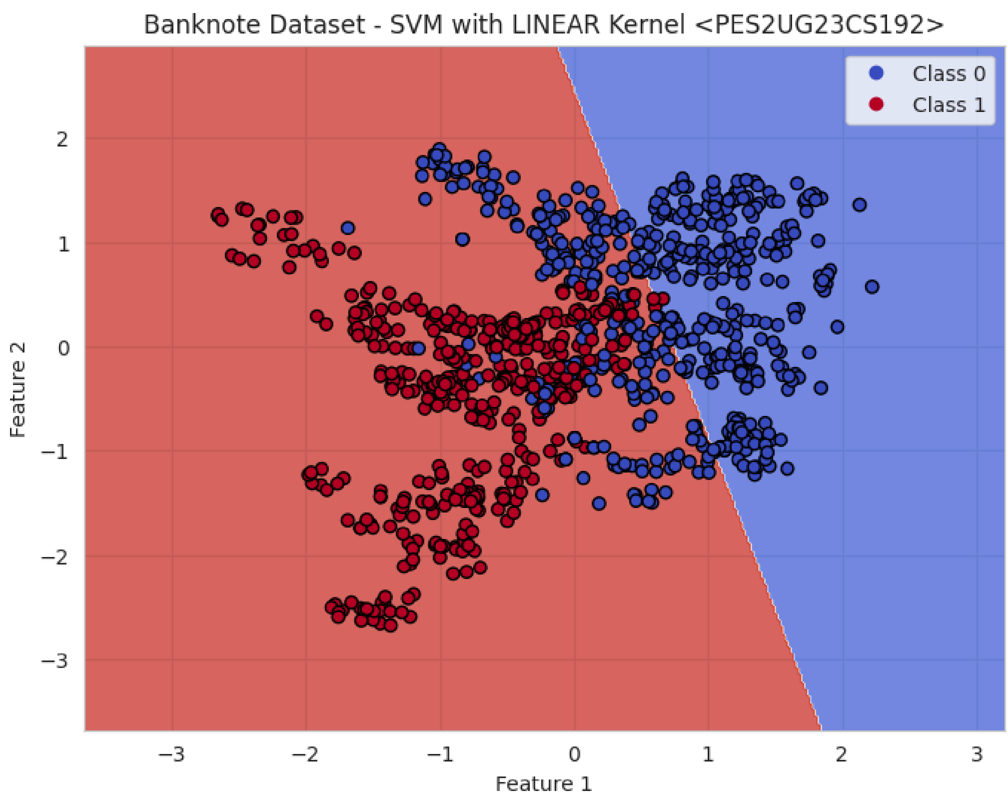
1.



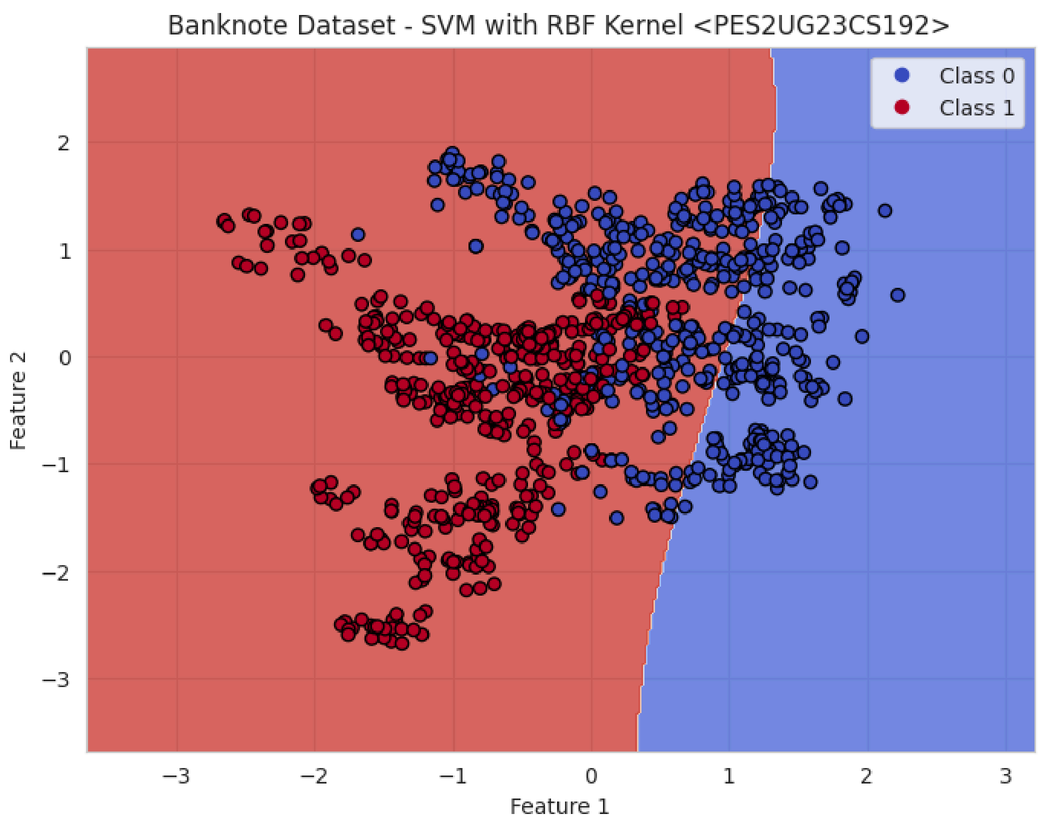
2.


Moons Dataset - SVM with POLY Kernel <PES2UG23CS192>
Class 0
Class 1
Feature 1
Feature 2
−3
−2
−1
0
1
2
3

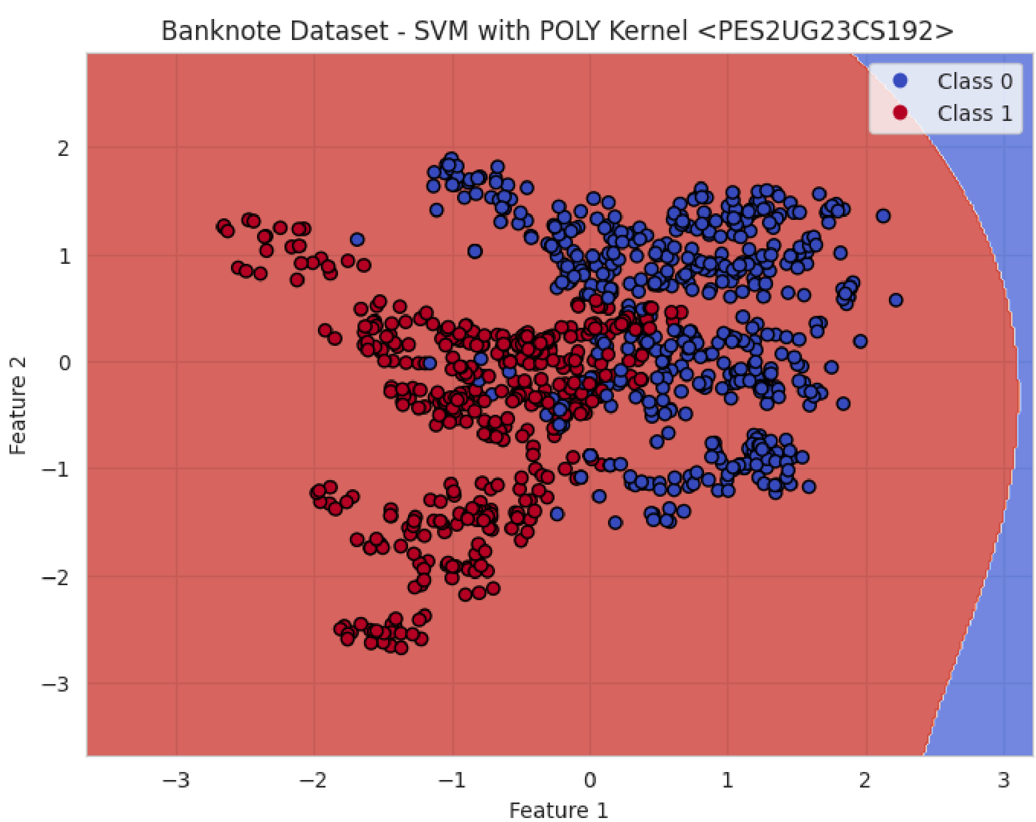
3.


Banknote Dataset - SVM with LINEAR Kernel <PES2UG23CS192>
Class 0
Class 1
Feature 1
Feature 2
−3
−2
−1
0
1
2
3

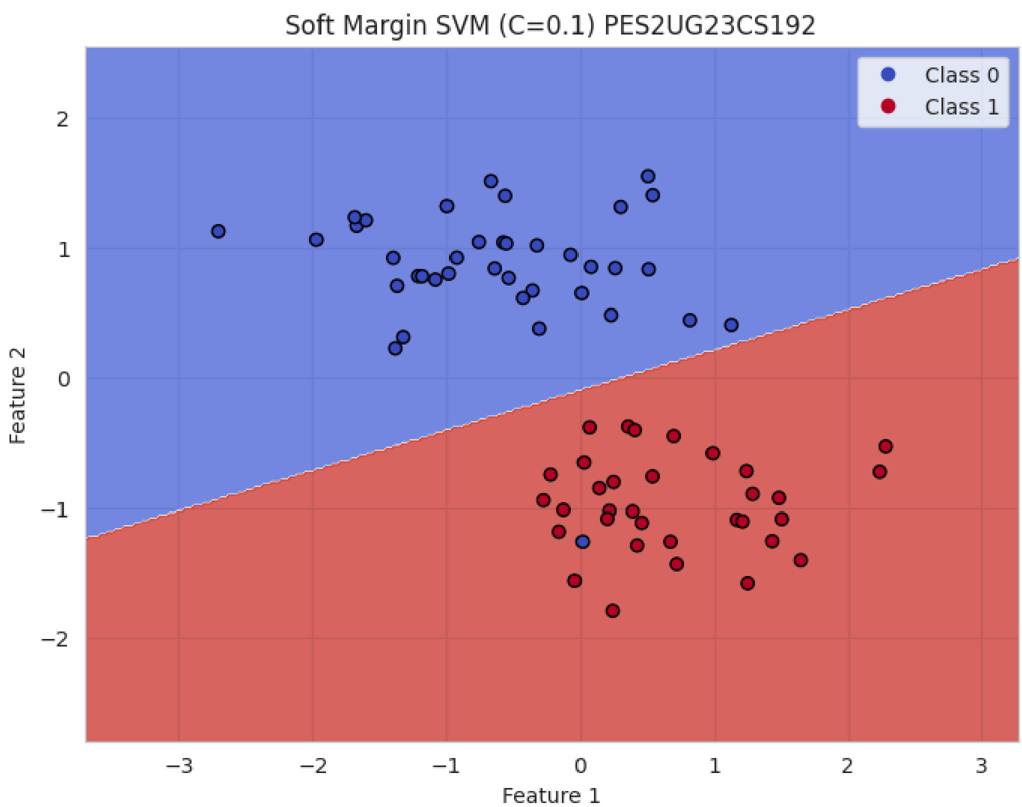
4.


Banknote Dataset - SVM with RBF Kernel <PES2UG23CS192>
Class 0
Class 1
Feature 2
Feature 1
−3
−2
−1
0
1
2
3

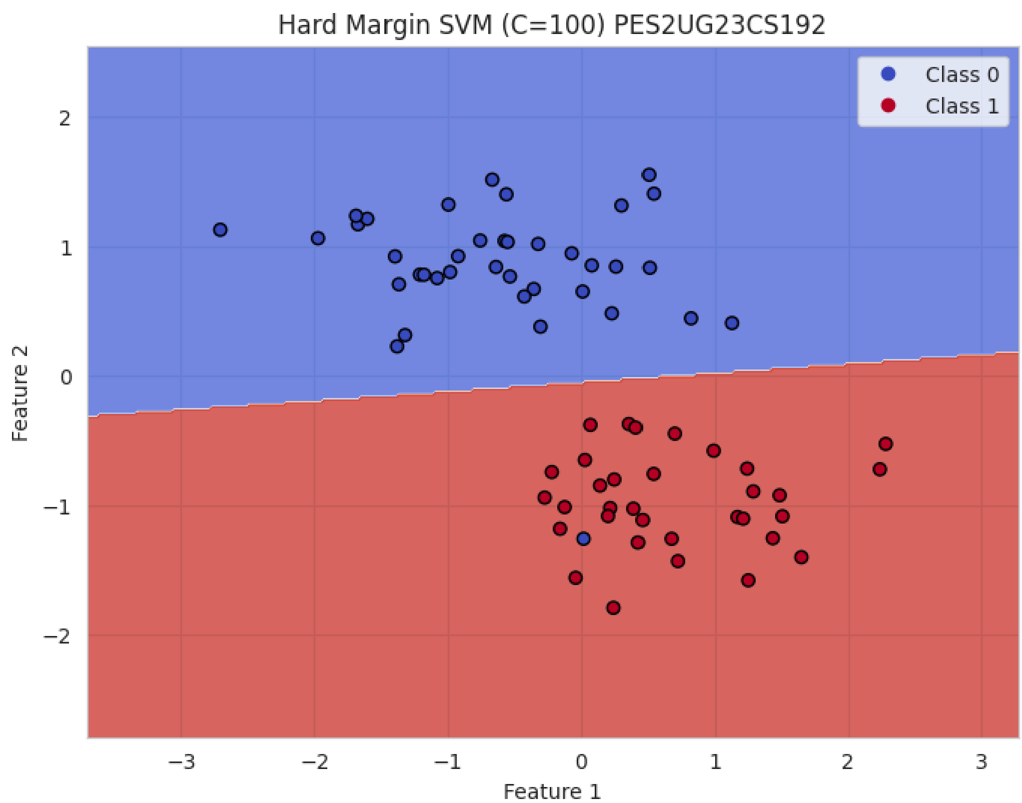
5.


Banknote Dataset - SVM with POLY Kernel <PES2UG23CS192>
Class 0
Class 1
Feature 2
Feature 1
−3
−2
−1
0
1
2
3

6.


Soft Margin SVM (C=0.1) PES2UG23CS192
Class 0
Class 1
2
1
0
−1
−2
Feature 2
−3
−2
−1
0
1
2
3
Feature 1

7.


Hard Margin SVM (C=100) PES2UG23CS192
Class 0
Class 1
2
1
0
−1
−2
Feature 2
−3
−2
−1
0
1
2
3
Feature 1

8.