ML-Semester-5

Lab-10 **Date:** 12/10/2025

Name: Mohammed Musharraf	**SRN:** PES2UG23CS915	**Section:** F

Analysis Questions:

Moons Dataset Questions

1. Inferences about the Linear Kernel's performance.

The linear kernel achieves only 87% accuracy because it can only create a straight decision boundary. Since the Moons dataset has interlocking crescent shapes (non-linear structure), the straight line misclassifies many points in the overlapping region.

2. Comparison between RBF and Polynomial kernel decision boundaries.

The RBF kernel (97% accuracy) captures the moon shapes more naturally than the Polynomial kernel (89% accuracy). RBF creates smooth, curved boundaries that wrap around the crescent patterns (non-linear structure), the straight line misclassifies many points in the overlapping region.

Banknote Dataset Questions

1. Which kernel was most effective for this dataset?

The RBF kernel is the most effective with 93% accuracy. It creates smooth, curved boundaries that capture the slight non-linearity in the Banknote data without overfitting.

2. Why might the Polynomial kernel have underperformed here?

The polynomial kernel achieved only 84% accuracy because the banknote dataset is relatively linearly separable with a diagonal pattern. The polynomial's complex curved boundaries add unnecessary complexity that doesn't match the simpler data structure. This is different from the Moons dataset where the polynomial's curves were more useful for capturing the crescent shapes.

Hard vs. Soft Margin Questions

1. Which margin (soft or hard) is wider?

The soft margin {C=0.1} produces a wider margin with a broader transition zone between classes.

2. Why does the soft margin allow “mistakes”?

It prioritizes generalization over perfect training accuracy. By tolerating some misclassifications, it creates a more stable boundary that’s less sensitive to outliers and performs better on new data.

3. Which model is more likely to be overfitting?

The hard margin {C=100} is more prone to overfitting because it tries to classify every training point correctly, even outliers, creating a narrower boundary that fits training data too closely.

4. Which model would you trust more for new data?

The soft margin {C=0.1} is more trustworthy because it learns general patterns rather than memorizing noise. In real-world scenarios with noisy data, starting with a low C value is preferred for better generalization.

Screenshots:

Training Results

Moons Dataset

```
SVM with LINEAR Kernel <PES2UG23CS915>
              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

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

1. Classification Report for SVM with LINEAR Kernel with SRN
2. Classification Report for SVM with RBF Kernel with SRN

```
SVM with RBF Kernel <PES2UG23CS915>
              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 <PES2UG23CS915>
...
weighted avg       0.89      0.89      0.89       150

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

3. Classification Report for SVM with POLY Kernel with SRN

Banknote Dataset

4. Classification Report for SVM with LINEAR Kernel
5. Classification Report for SVM with RBF Kernel

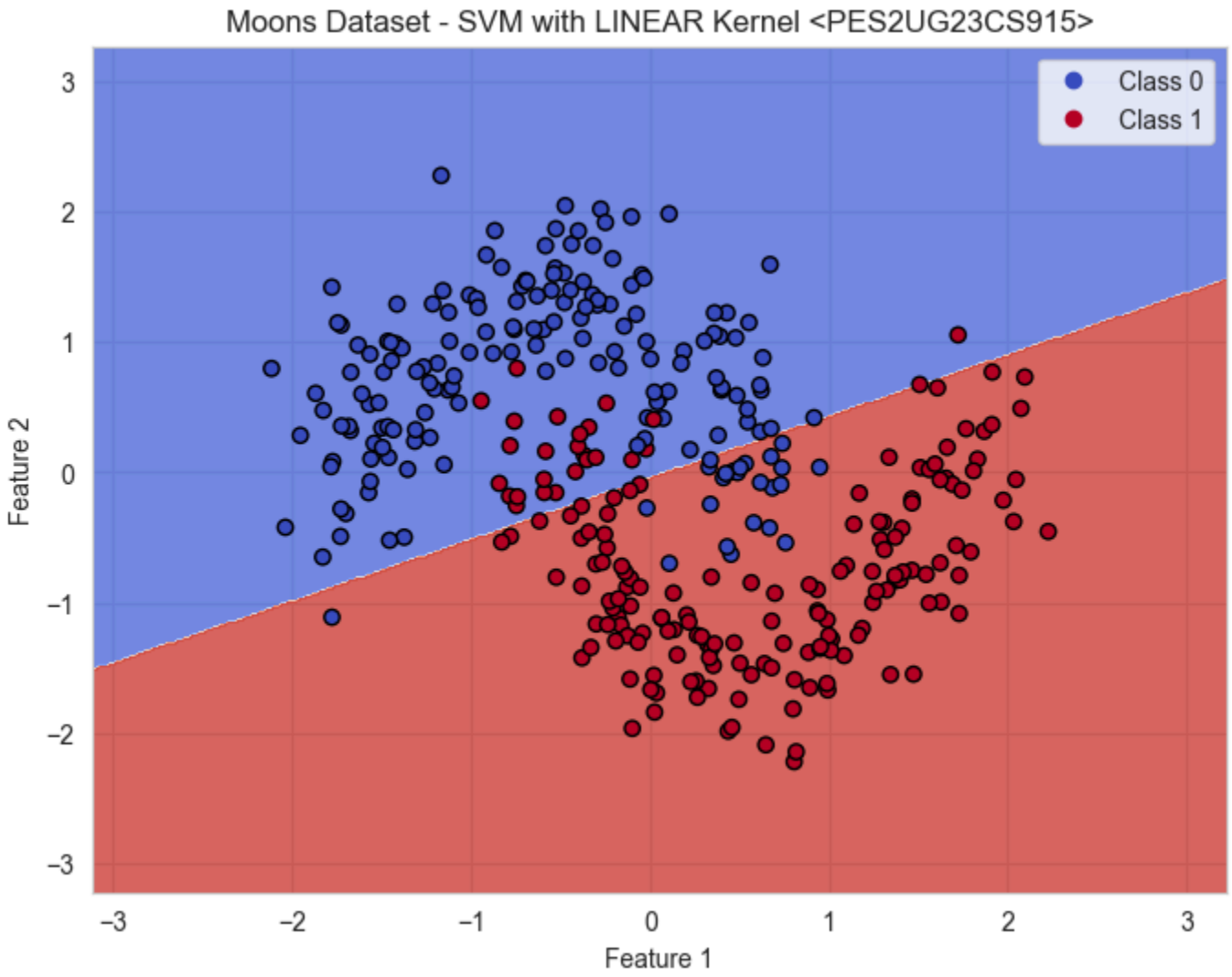
```
SVM with LINEAR Kernel <PES2UG23CS915>
              precision    recall  f1-score   support

      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

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

SVM with RBF Kernel <PES2UG23CS915>
              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

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

SVM with POLY Kernel <PES2UG23CS915>
...
weighted avg       0.85      0.84      0.84       412

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

6. Classification Report for SVM with POLY Kernel

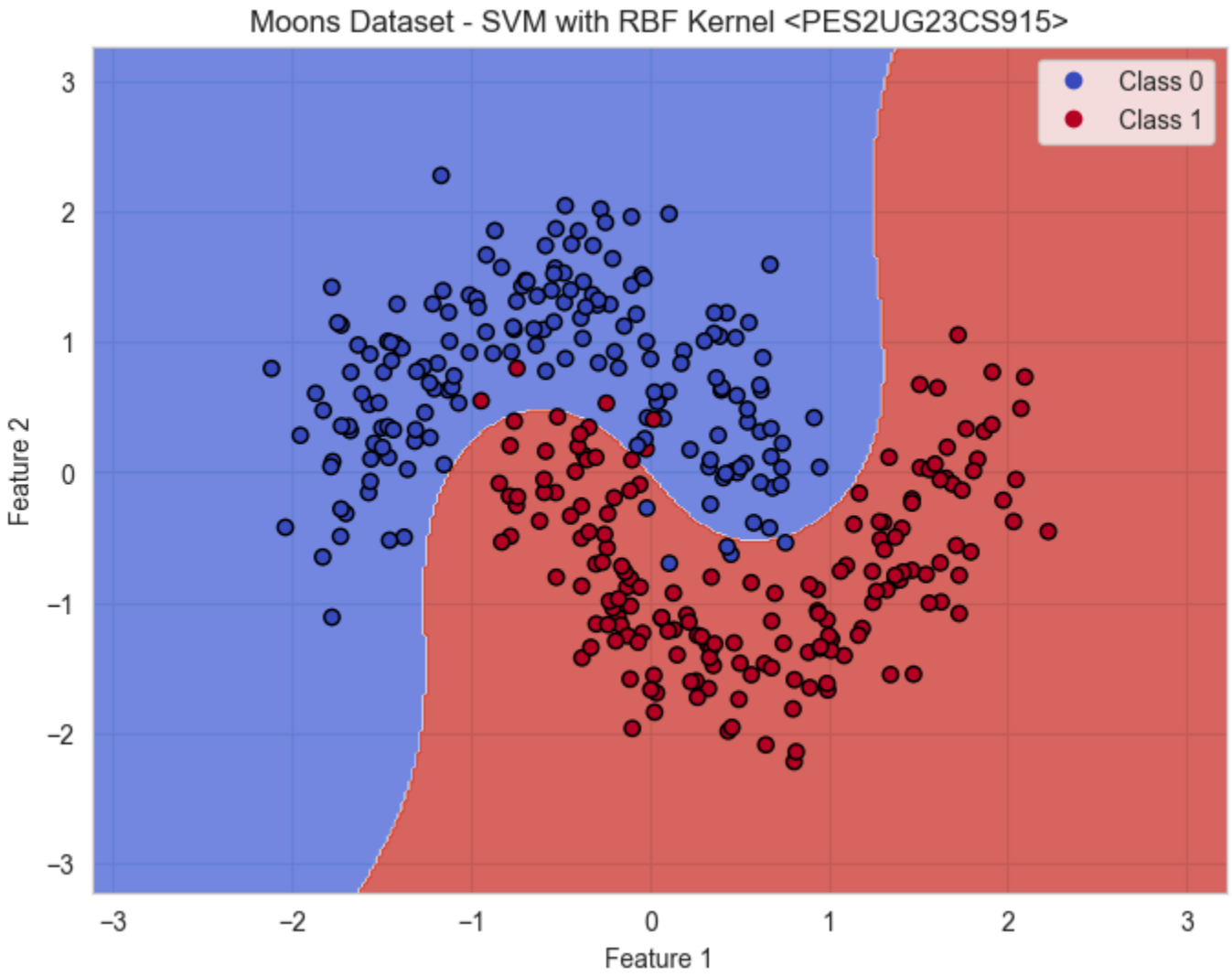
Decision Boundary Visualizations

Moons Dataset

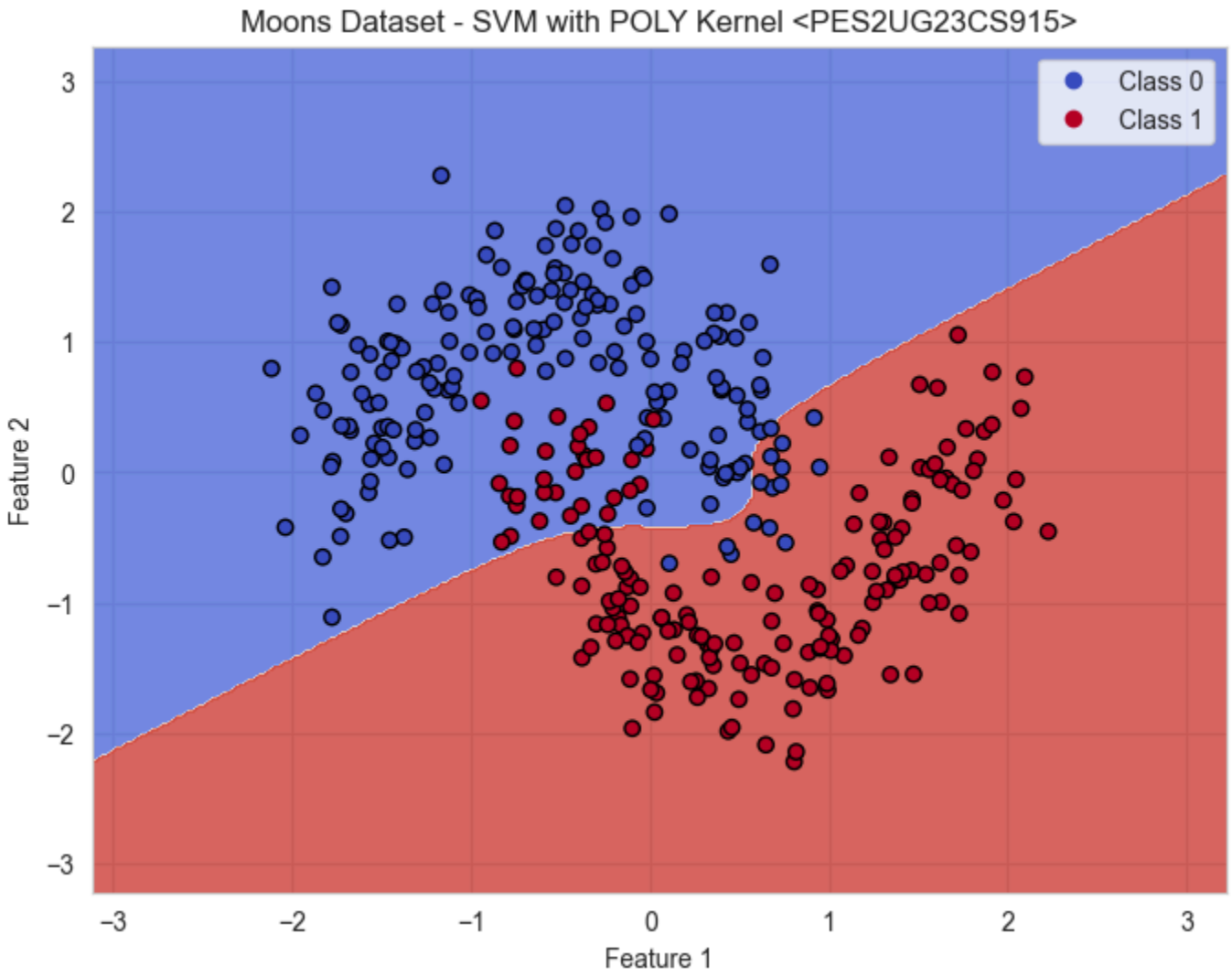
7. Moons Dataset - SVM with LINEAR Kernel



8. Moons Dataset - SVM with RBF Kernel

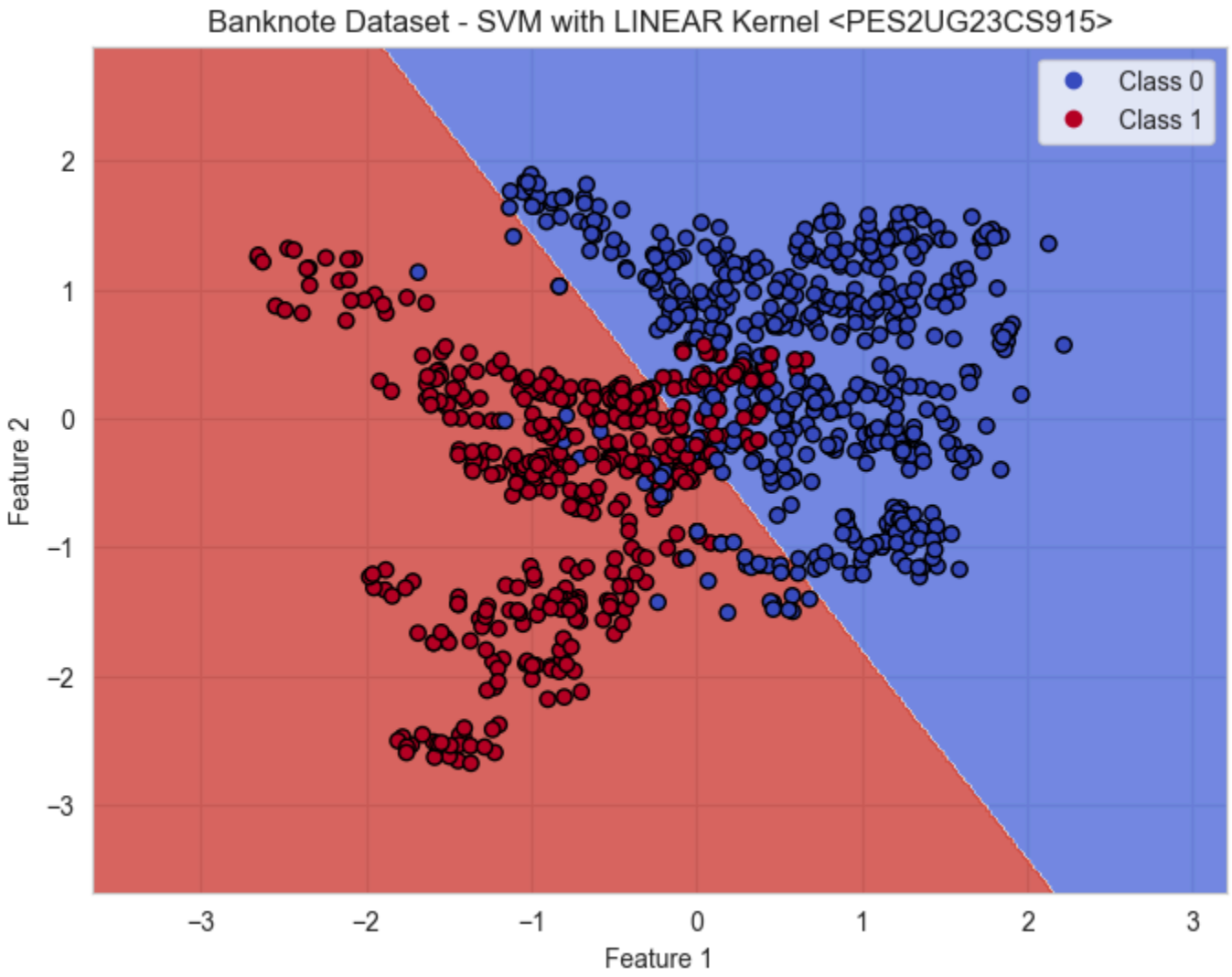


9. Moons Dataset - SVM with POLY Kernel

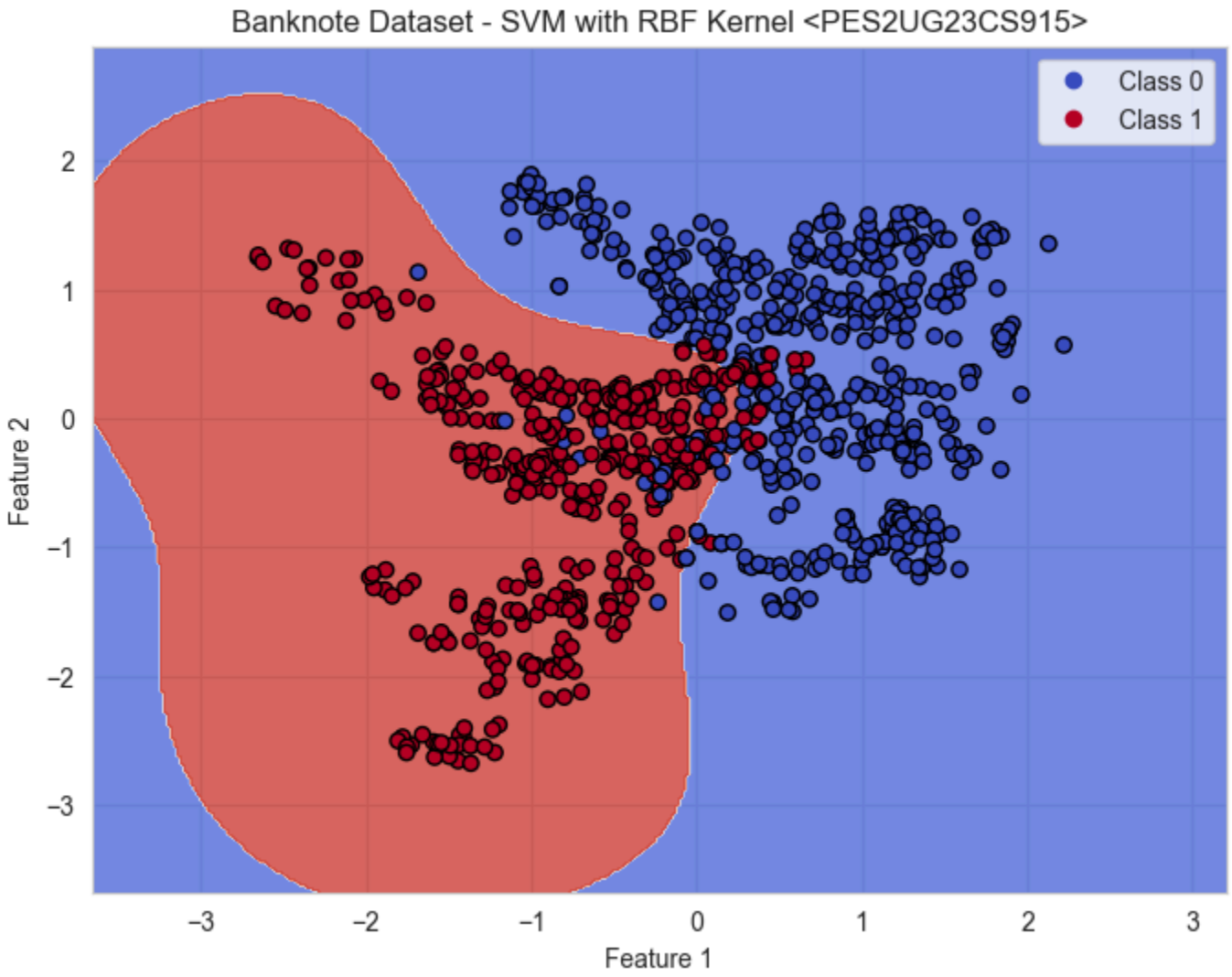


Banknote Dataset

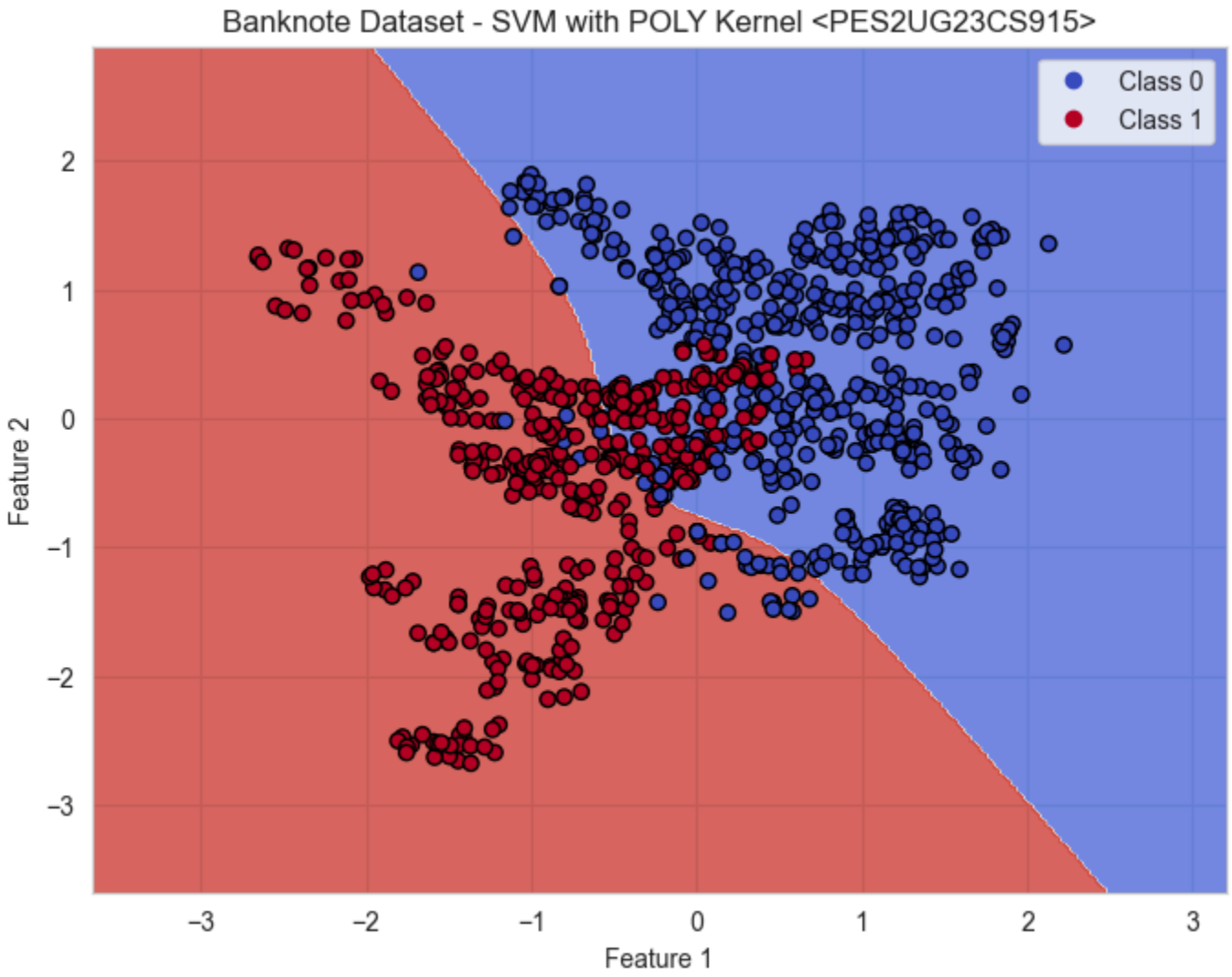
10. Banknote Dataset - SVM with LINEAR Kernel



11. Banknote Dataset - SVM with RBF Kernel

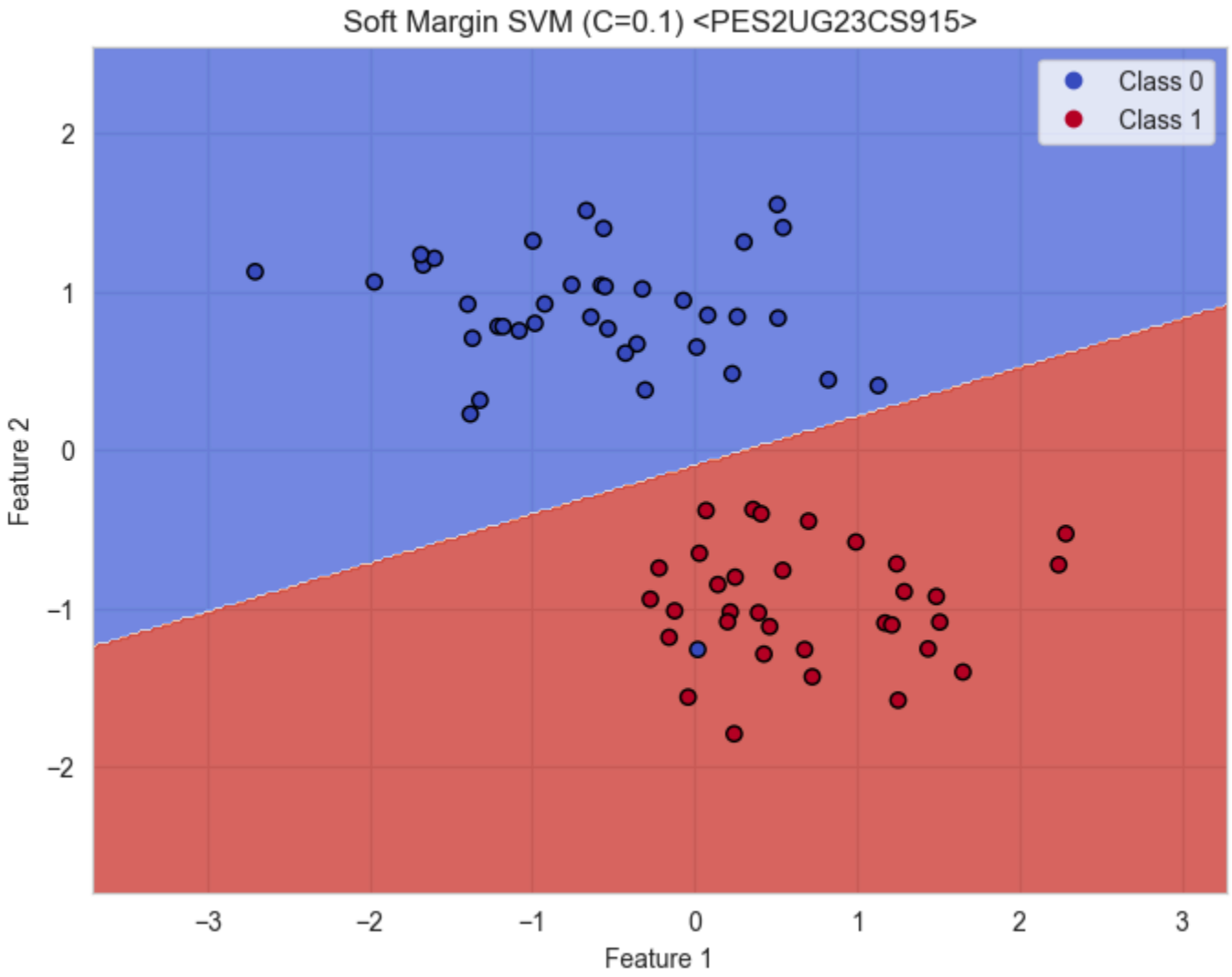


12. Banknote Dataset - SVM with POLY Kernel

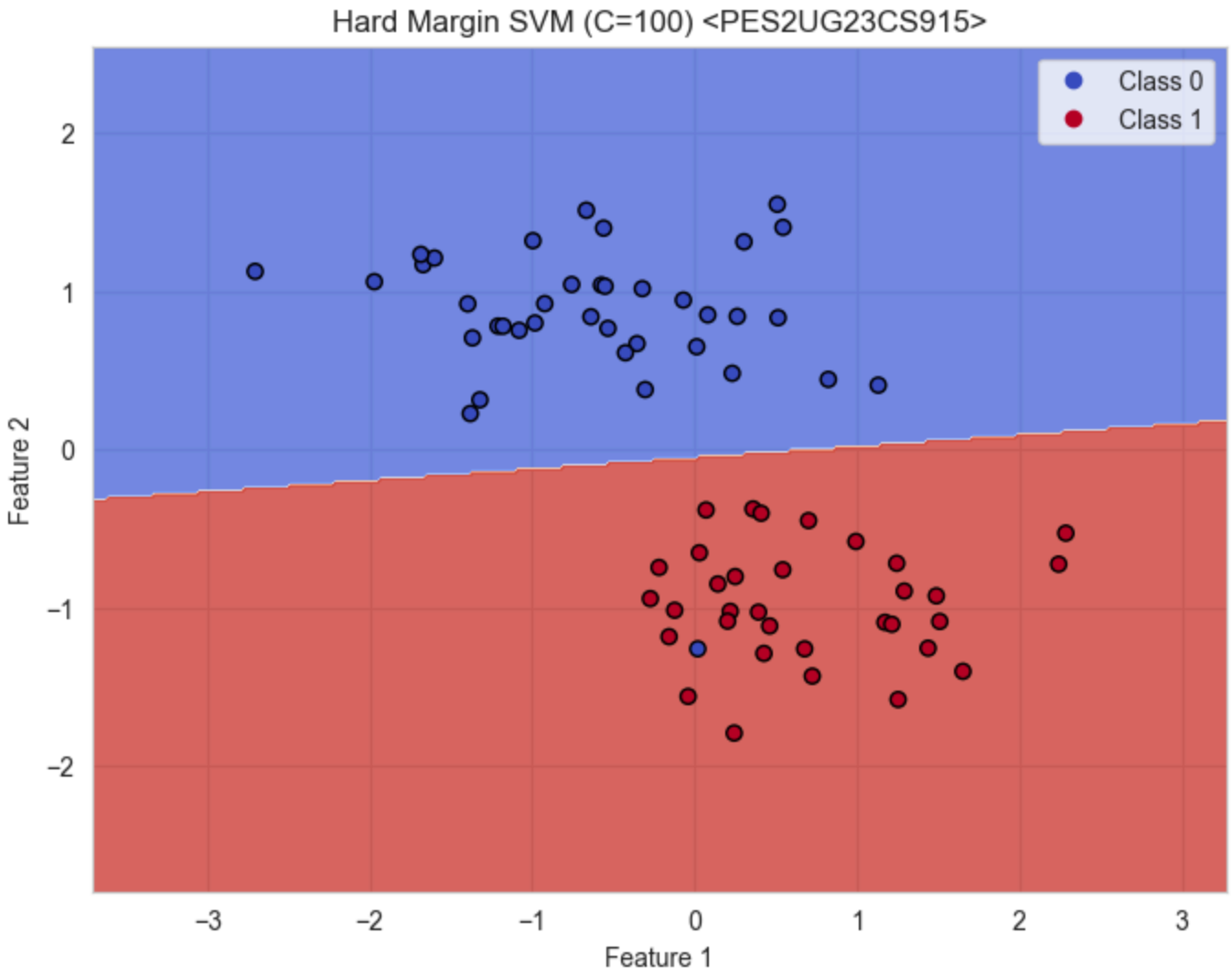


Margin Analysis:

13. Soft Margin SVM (C=0.1)



14. Hard Margin SVM (C=100)


Hard Margin SVM (C=100) <PES2UG23CS915>
Class 0
Class 1
Feature 2
Feature 1
2
1
0
−1
−2
−3
−2
−1
0
1
2
3