Gautam Menon **PES2UG23CS196** **Section C**

ML LAB Week - 10

(10/09/2025)

Analysis Questions :

Moon Dataset:

Q1: Inferences about the Linear Kernel's performance.
Ans : The Linear Kernel is not able to correctly classify the data , the moons dataset is non linearly seperable. The model achieves an overall accuracy of 87%. While not terrible, it indicates that a noticeable portion of the data points are being misclassified.

Q2: Comparison between RBF and Polynomial kernel decision boundaries.
Ans : The RBF kernel captures the shape of the data more naturally. Its ability to create a complex, non-linear boundary allows it to perfectly model the crescent shapes of the Moons dataset. The Polynomial kernel, while better than a linear one, is too constrained and cannot adapt as effectively to this specific data distribution.

Bank Note Dataset:

Q1:Which kernel was most effective for this dataset?
Ans : The RBF Kernel is most effective.It has an accuracy of 0.93
and its flexible, non-linear decision boundary does the best job of separating the complex, intertwined data clusters.

Q2:Why might the Polynomial kernel have underperformed here?
Ans:The Polynomial kernel performs worse here because it's best for data with a simple, predictable curve (like the Moons dataset). The boundary in the Banknote data is too irregular and complex. A polynomial function is too rigid to capture this shape, whereas the more adaptable RBF kernel is a better fit.

Hard VS Soft Margin Questions :

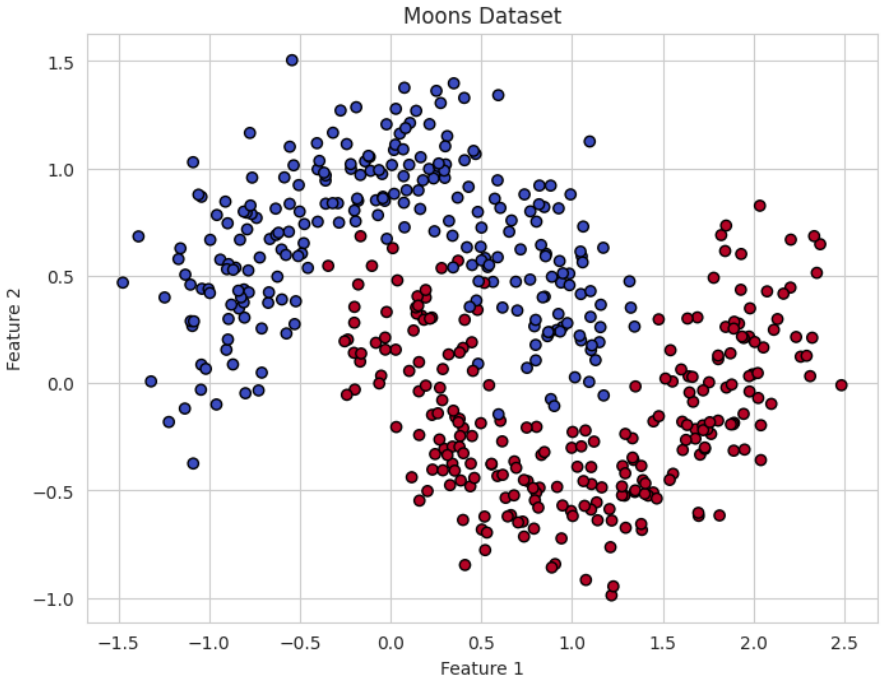
Q1:Which margin (soft or hard) is wider?
Ans: The soft margin model creates a wider margin.
The SVM has to find the optimal margin and its priority is to make this margin as wide as possible to create a generalized boundary , in soft margin a few errors could be made in classification in hard margin its prohibited to have points that end up on the wrong side.

Q2:Why does the soft margin model allow "mistakes"?
Ans:The soft margin model allows mistakes (violations) because its primary objective is generalization, not perfect classification of the training data.
Real-world data is often noisy. By allowing a few outliers or misclassifications, the model avoids ruining its decision boundary to fit this noise. This is part of the bias-variance trade-off. The model accepts a small amount of bias (making some errors on the training set) to significantly reduce its variance (being less sensitive to the specific data it was trained on).

Q3 Which model is more likely to be overfitting and why?
Ans:The hard margin model is much more likely to be overfitting.
Overfitting occurs when a model learns the training data too well, including its noise and random fluctuations. A hard margin SVM, with its high penalty for errors (a large C parameter), is forced to create a complex decision boundary that correctly classifies every single training point. This process is like "memorizing" the training data instead of learning the underlying pattern.

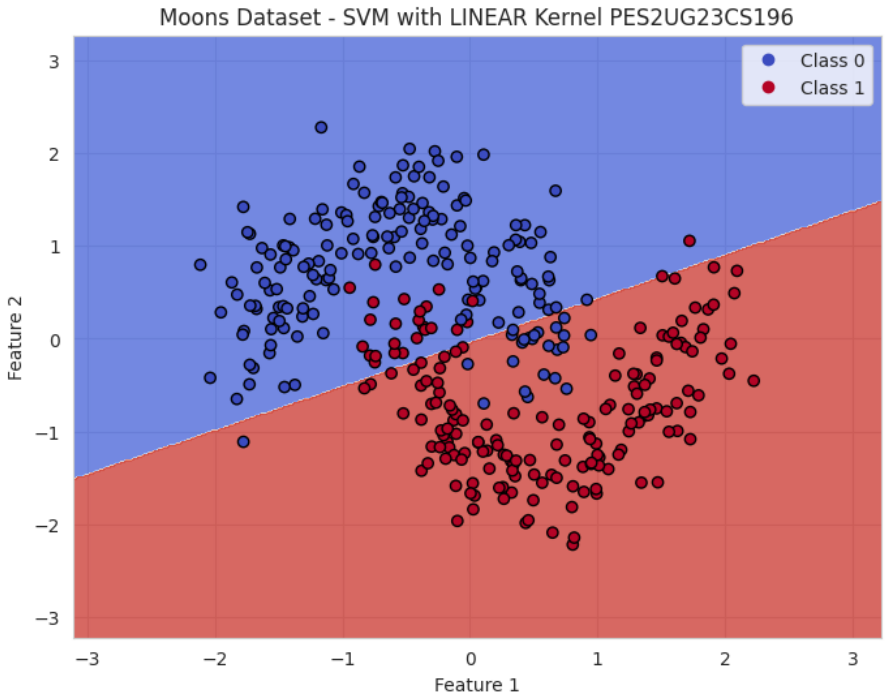
Q4:Which model would you trust more for new data and why?
I would trust the soft margin model more for classifying new data as its able to generalize reliably.

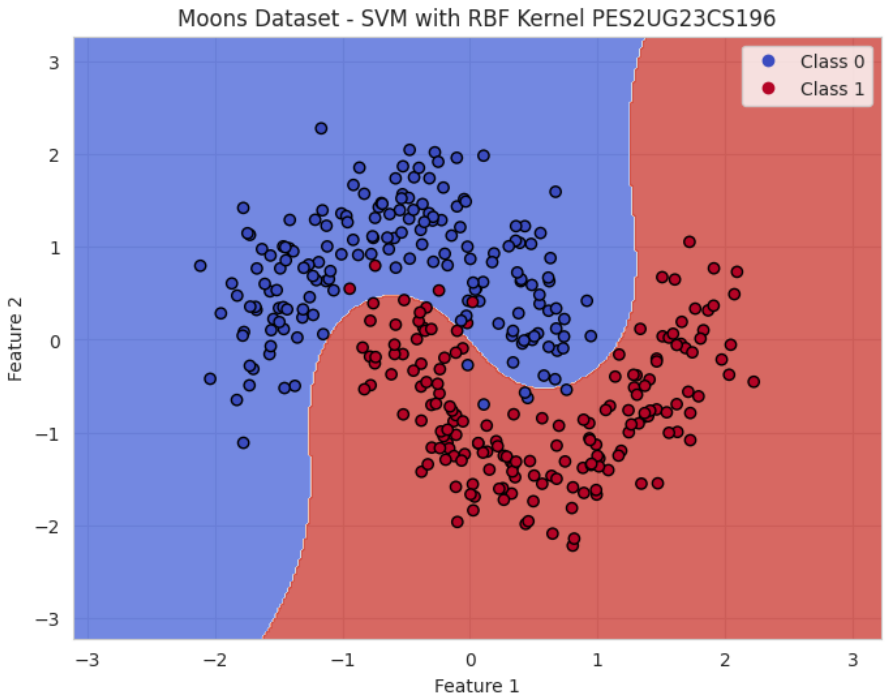
Moons Dataset

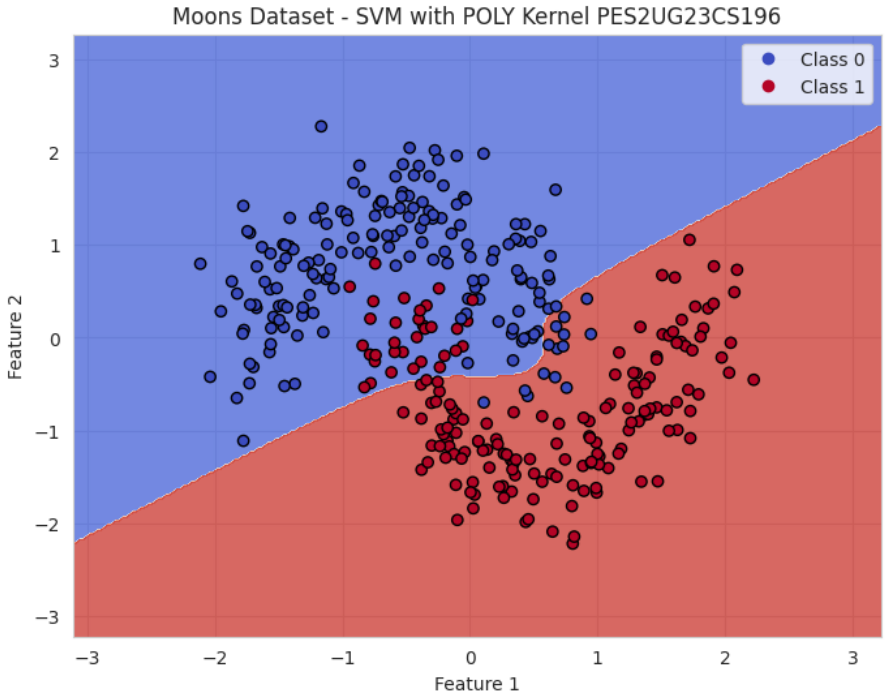



```
SVM with LINEAR Kernel PES2UG23CS196
              precision    recall  f1-score   support

           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 PES2UG23CS196
              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 PES2UG23CS196
              precision    recall  f1-score   support

           0       0.85      0.95      0.89        75
           1       0.94      0.83      0.88        75

    accuracy                           0.89       150
   macro avg       0.89      0.89      0.89       150
weighted avg       0.89      0.89      0.89       150

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

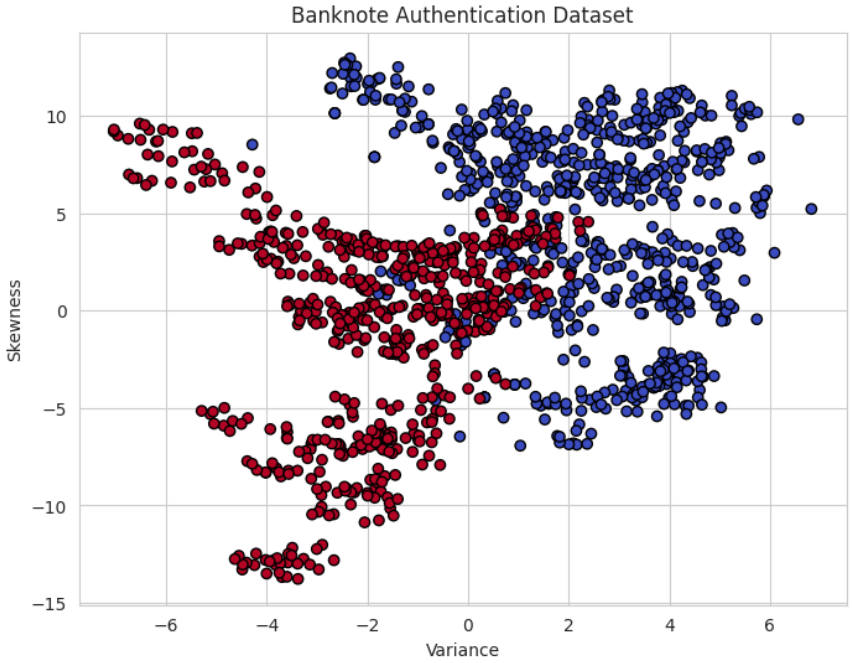
Gautam Men **Section C**



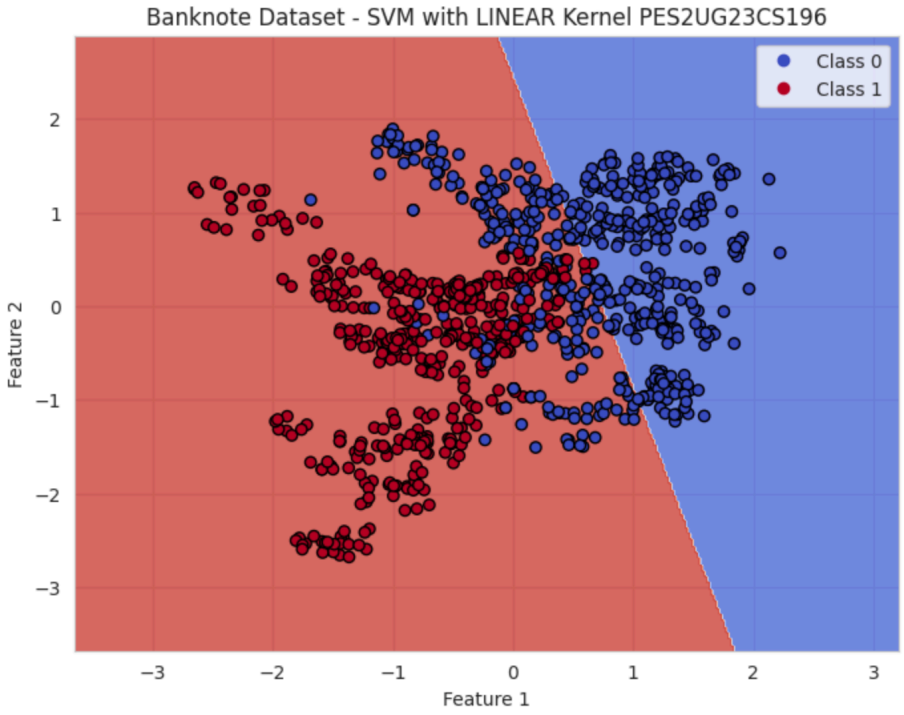


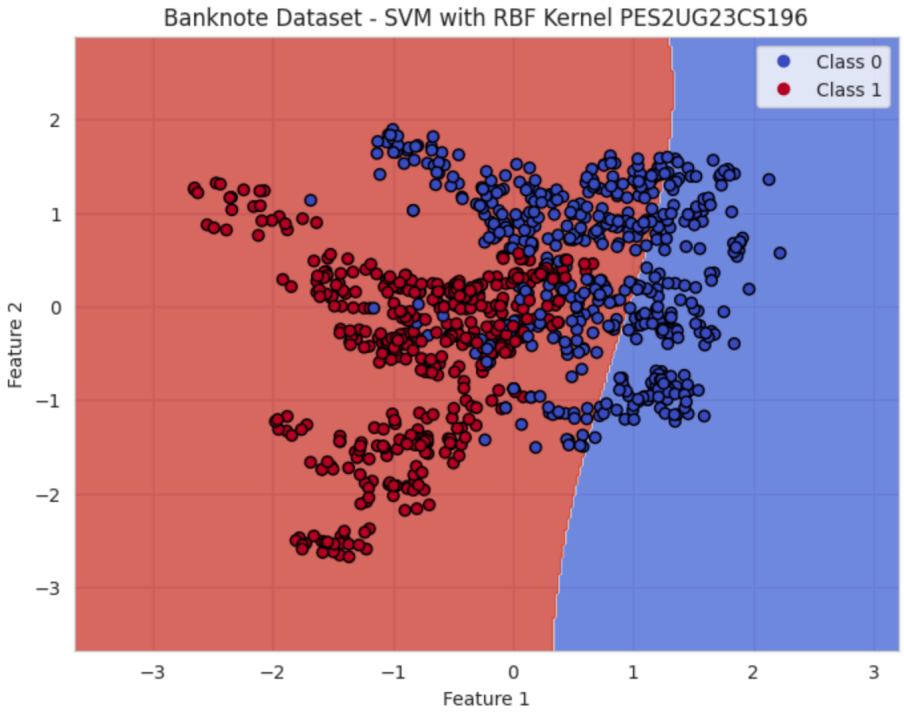


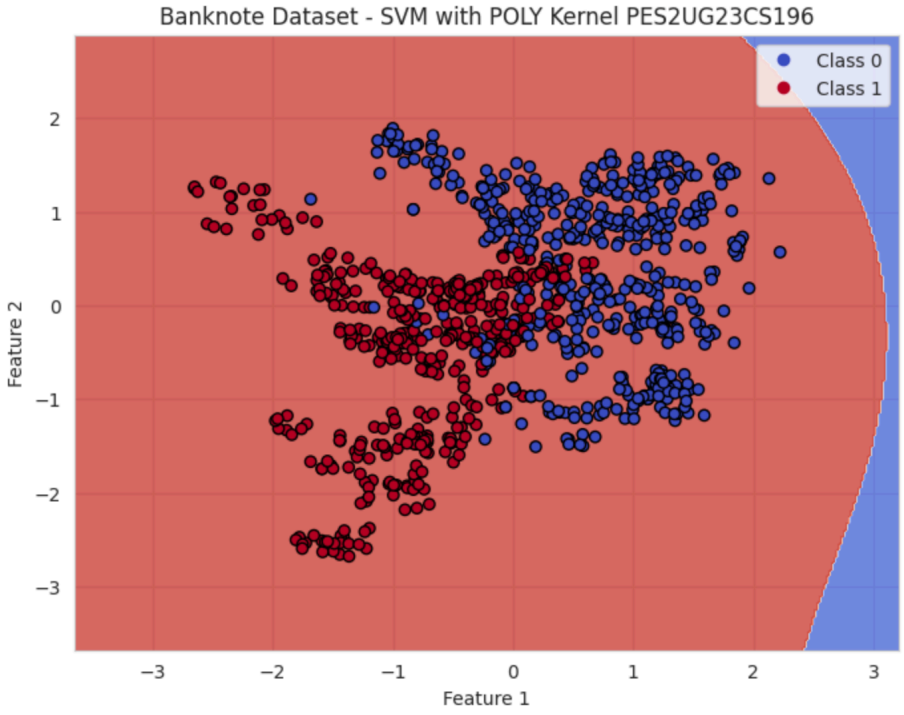
Banknote Authentication Dataset

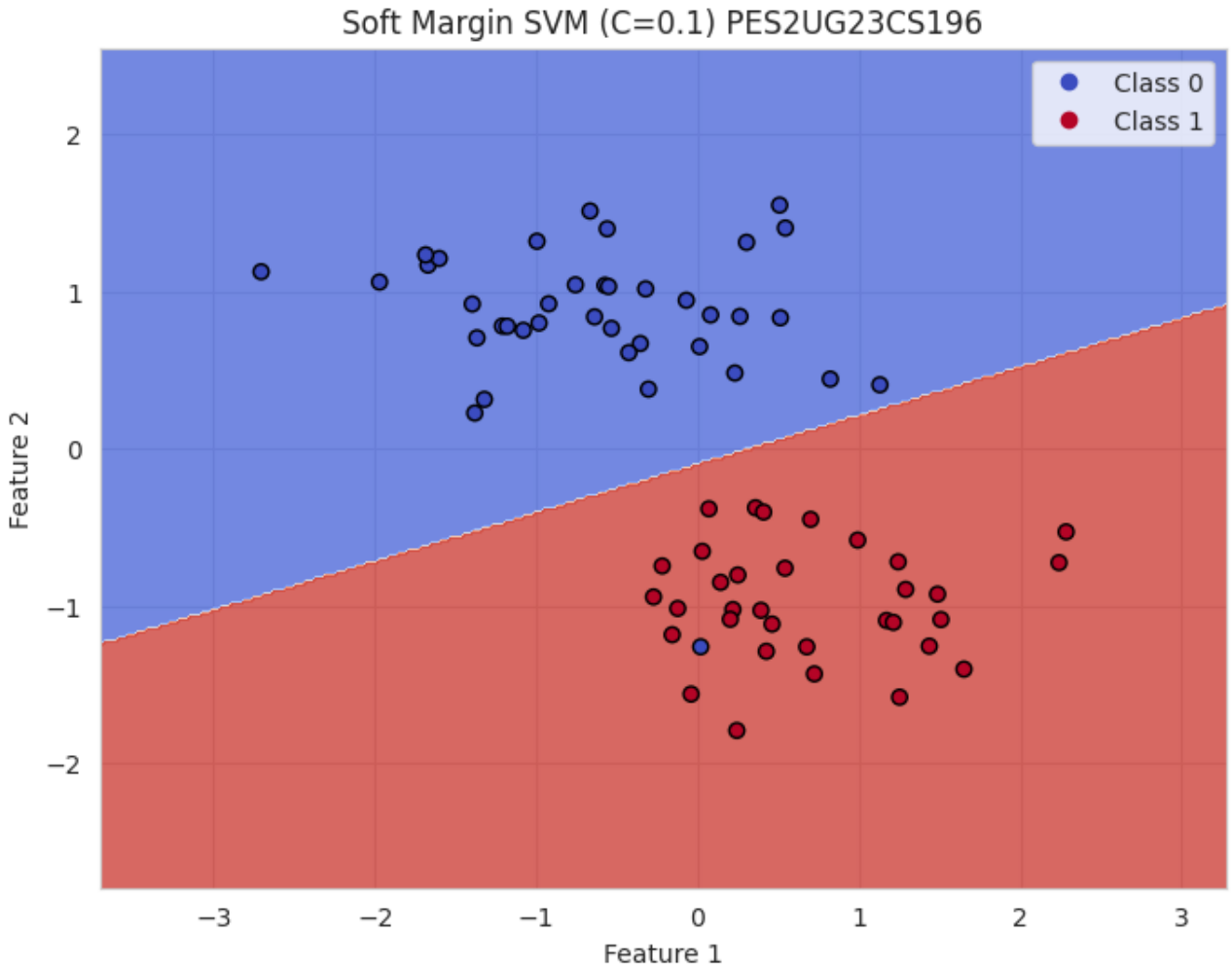


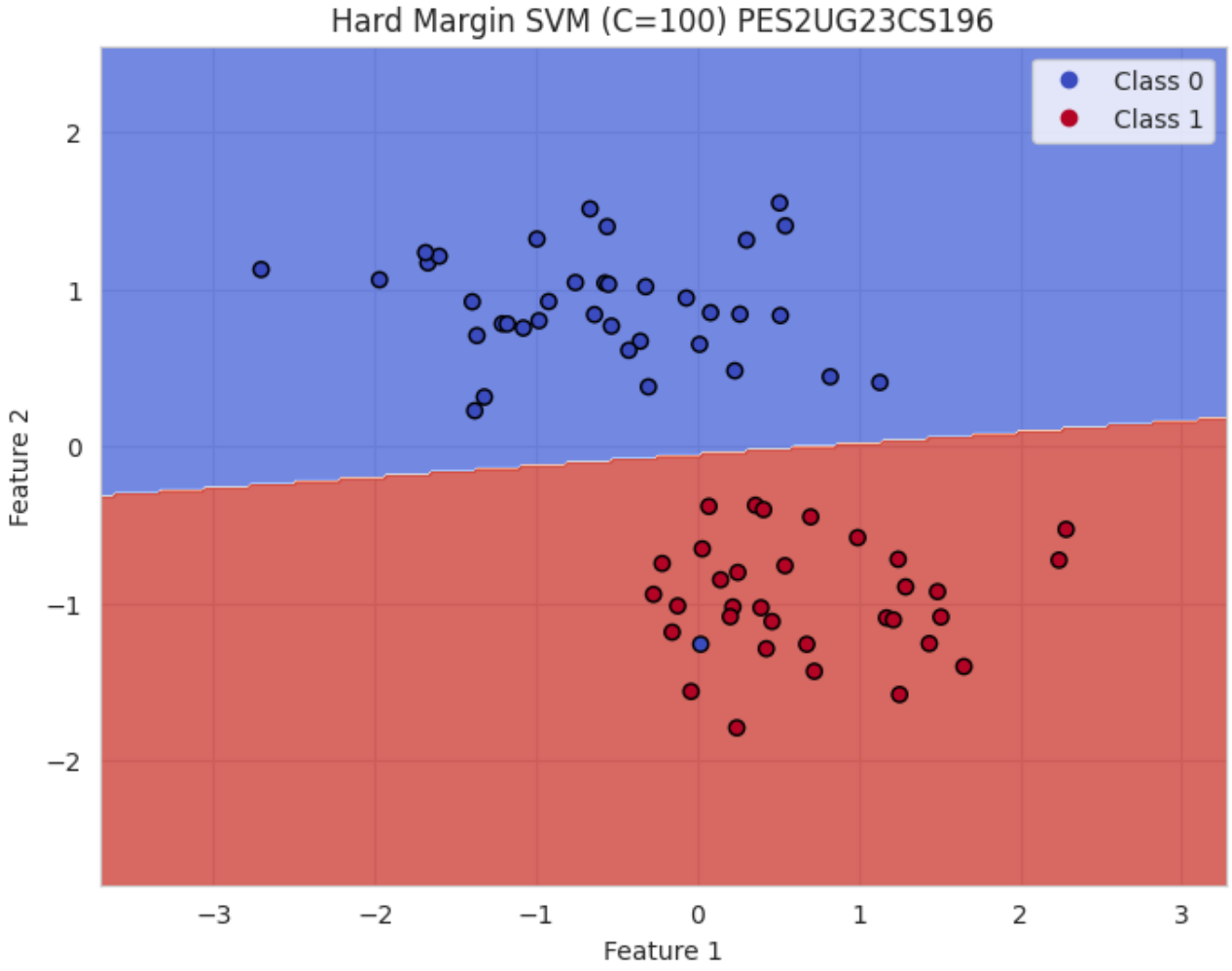

```
SVM with LINEAR Kernel PES2UG23CS196
              precision    recall  f1-score   support

      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 PES2UG23CS196
              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 PES2UG23CS196
              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

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


Banknote Dataset - SVM with LINEAR Kernel PES2UG23CS196
Class 0
Class 1
Feature 2
Feature 1
−3
−2
−1
0
1
2
3


Banknote Dataset - SVM with RBF Kernel PES2UG23CS196
Class 0
Class 1
Feature 2
Feature 1
−3
−2
−1
0
1
2
3


Banknote Dataset - SVM with POLY Kernel PES2UG23CS196
Class 0
Class 1
Feature 2
Feature 1
−3
−2
−1
0
1
2
3


Soft Margin SVM (C=0.1) PES2UG23CS196
Class 0
Class 1
Feature 2
Feature 1
2
1
0
−1
−2
−3
−2
−1
0
1
2
3


Hard Margin SVM (C=100) PES2UG23CS196
Class 0
Class 1
Feature 2
Feature 1
2
1
0
−1
−2
−3
−2
−1
0
1
2
3