# UE23CS352A: MACHINE LEARNING

# Week 6: SVM

Project Title- SVM Implimentation

Name- Nisschay Khandelwal

SRN-PES2UG23CS394

Date-11/10/2025

Analysis Questions

. Moons Dataset Questions (2 questions):

1. Inferences aboutthe Linear Kernel's performance.

Based on the metrics and visualizations, the Linear Kernel shows several key characteristics:

- Lower accuracy compared to RBF and Polynomial kernels because the Moons dataset is inherently non-linear

- Straight-line decision boundary that cannot capture the curved, interlocking moon shapes effectively

- Higher misclassification rate in the overlapping regions where the two moons intersect

- Limited flexibility - Linear kernels are designed for linearly separable data, making them unsuitable for this complex geometric pattern

- The linear boundary essentially tries to draw a straight line through curved data, resulting in many points being misclassified

2. Comparison between RBF and Polynomial kernel decision boundaries.

Comparing the decision boundaries:

- RBF (Radial Basis Function) Kernel typically performs better and captures the moon shapes more naturally because:

  - Creates smooth, circular/curved boundaries that follow the data distribution

  - Better handles the non-linear, curved nature of the interlocking moons

  - More flexible in creating complex decision boundarie

- Generally provides higher accuracy and better generalization

- Polynomial Kernel may show:
  - More rigid, polynomial-shaped boundaries
  - Potentially more complex but less smooth decision boundaries
  - May be more prone to overfitting depending on the degree parameter
  - Can create curved boundaries but may not be as naturally suited to the circular moon patterns as RBF

Conclusion: The RBF kernel typically captures the shape of the Moons data more naturally due to its ability to create smooth, curved boundaries that better match the circular/curved nature of the half-moon patterns.

. Banknote Dataset Questions (2 questions):

1. Which kernel was most effective for this dataset?

- Linear excels here:
  - Real-world financial data often has linear relationships between features
  - Simpler decision boundaries are sufficient for this type of classification
  - Better generalization to new banknote samples
  - Faster training and prediction times

2. Why might the Polynomial kernel have underperformed here?

**Data Distribution Mismatch:**

- Banknote data is more **linearly separable** and doesn't require complex polynomial curves
- Polynomial kernels are designed for data with **polynomial relationships**, which may not exist in this financial dataset
- The **variance vs skewness** features likely have simpler relationships than polynomial functions

- **Overfitting Issues:**
  - Polynomial kernels can **overfit** to training data when the underlying pattern is simpler

  - Creates unnecessarily **complex decision boundaries** for linearly separable data

  - **Higher degree polynomials** may capture noise rather than genuine patterns

- **Feature Characteristics:**

  - Financial features (variance, skewness) typically have **linear or simple non-linear relationships**

  - Unlike the curved moon shapes, banknote features don't require polynomial transformations

  - **Simpler kernels** (linear/RBF) are more appropriate for this data type

- **Computational Complexity:**

  - Polynomial kernels add **unnecessary complexity** without performance benefits

  - May lead to **poor generalization** on new banknote samples

. Hard vs. Soft Margin Questions (4 questions):

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

The Soft Margin (C=0.1) produces a wider margin compared to the Hard Margin (C=100).

- Soft Margin (C=0.1):

  - Creates a wider decision boundary with more space between the classes

  - More tolerant of data points that fall within or cross the margin

  - Prioritizes generalization over perfect classification of training data

  - The margin bands are visibly wider in the visualization

- Hard Margin (C=100):

  - Creates a narrower decision boundary that fits tightly around the data

  - Less tolerant of misclassifications, trying to classify every point correctly

  - Prioritizes training accuracy over generalization

  - The margin bands appear much narrower or almost non-existent

2. Why does the soft margin model allow "mistakes"?

The Soft Margin SVM (C=0.1) allows some points inside the margin or on the wrong side because:

- **Primary Goal - Generalization:** The model prioritizes **better performance on unseen data** rather than perfect training accuracy

- **Noise Tolerance:** It recognizes that some data points might be **outliers or noise** and shouldn't dictate the entire decision boundary

- **Bias-Variance Tradeoff:** It accepts some **bias (training errors)** to reduce **variance (overfitting)**

- **Regularization Effect:** The low C value acts as **regularization**, preventing the model from becoming too complex

- **Real-world Robustness:** In practice, data often contains noise, and perfect separation may not be achievable or desirable

2. Which model is more likely to be overfitting and why?

The Hard Margin (C=100) is more likely to overfit to the training data.

Reasons for Hard Margin Overfitting:

- High Sensitivity to Outliers: Tries to classify every single training point correctly, including potential outliers

- Complex Decision Boundaries: Creates overly complex boundaries to accommodate all training points

- Poor Generalization: May perform well on training data but poorly on new, unseen data

- Memorization vs Learning: Tends to memorize training patterns rather than learn generalizable patterns

Soft Margin Advantages:

- Better Generalization: More likely to perform well on new data

- Noise Resistance: Less affected by outliers and noisy data points

- Simpler Model: Creates simpler, more robust decision boundaries

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

For **new, unseen data points**, I would **trust the Soft Margin (C=0.1) model more**.

**Reasons for Trusting Soft Margin:**

- **Better Generalization:** Designed to perform well on unseen data rather than just training data

- **Noise Robustness:** Less likely to be misled by outliers in the training set

- **Stable Predictions:** More consistent performance across different datasets

- **Realistic Assumptions:** Acknowledges that perfect separation may not always be possible

. Moons Dataset (3 screenshots):

1. Classification Report for SVM with LINEAR Kernel with SRN

```
SVM with LINEAR Kernel PES2UG23CS394
              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

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

2. Classification Report for SVM with RBF Kernel with SRN

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

SVM with RBF Kernel PES2UG23CS394
              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

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

3. Classification Report for SVM with POLY Kernel with SRN

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

SVM with POLY Kernel PES2UG23CS394
              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

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

· Banknote Dataset (3 screenshots):

4. Classification Report for SVM with LINEAR Kernel

```
SVM with LINEAR Kernel PES2UG23CS394
              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

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

5. Classification Report for SVM with RBF Kernel

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

SVM with RBF Kernel PES2UG23CS394
              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

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

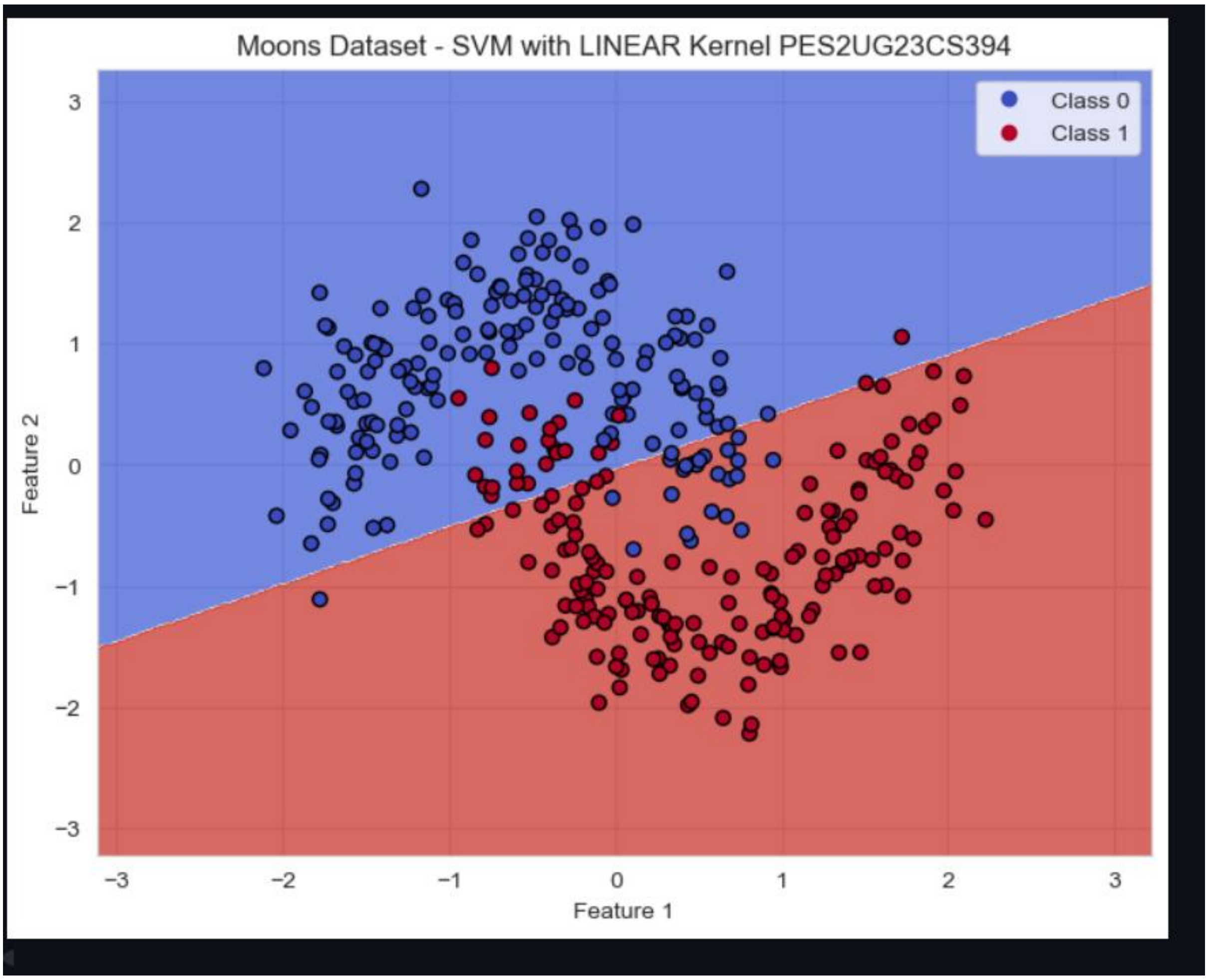
6. Classification Report for SVM with POLY Kernel

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

SVM with POLY Kernel PES2UG23CS394
              precision    recall  f1-score   support

      Forged       0.82      0.91      0.87       229
     Genuine       0.87      0.75      0.81       183

    accuracy                           0.84       412
   macro avg       0.85      0.83      0.84       412
weighted avg       0.85      0.84      0.84       412

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

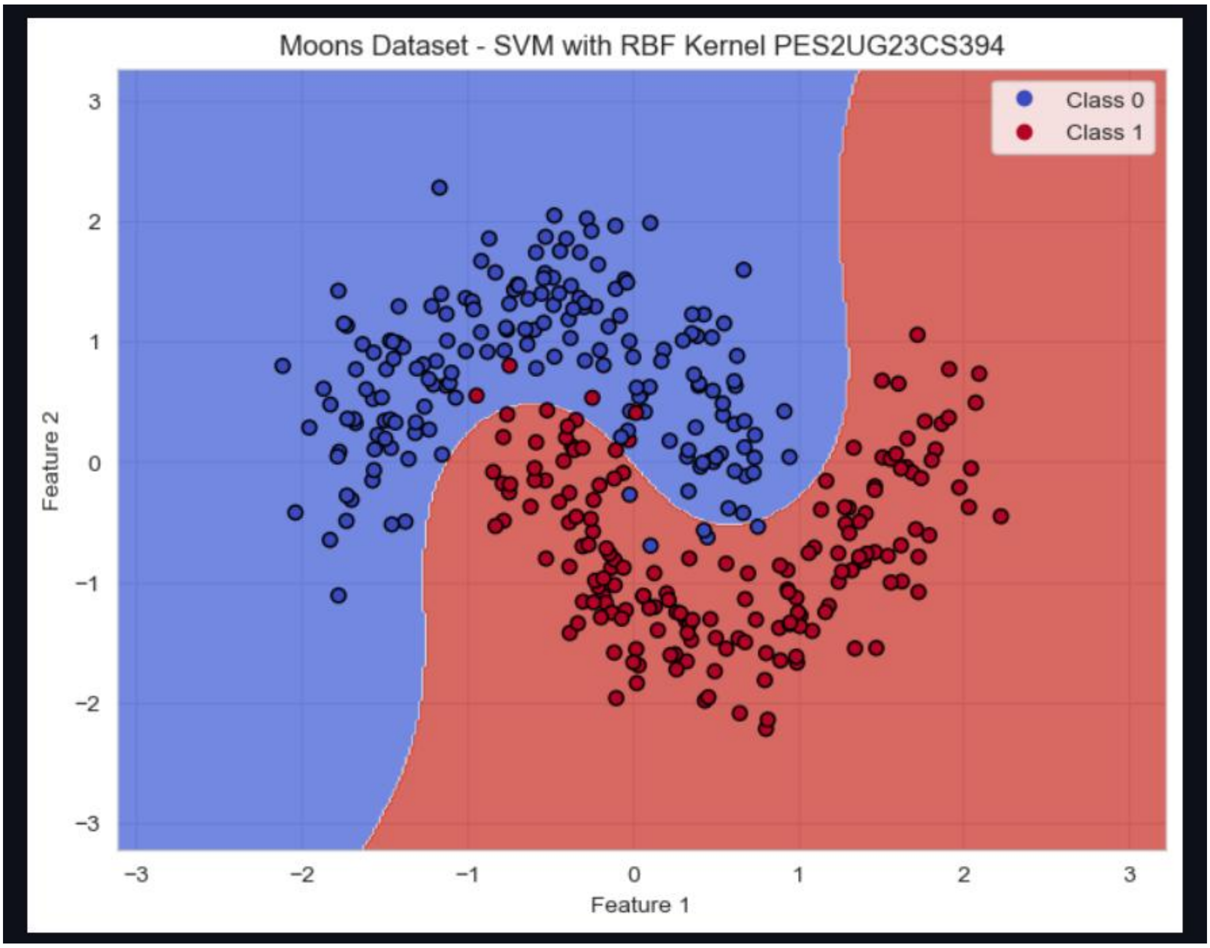
. Decision Boundary Visualizations (8 Screenshots): Capture the plot for each model's decision boundary.

Moons Dataset (3 plots):

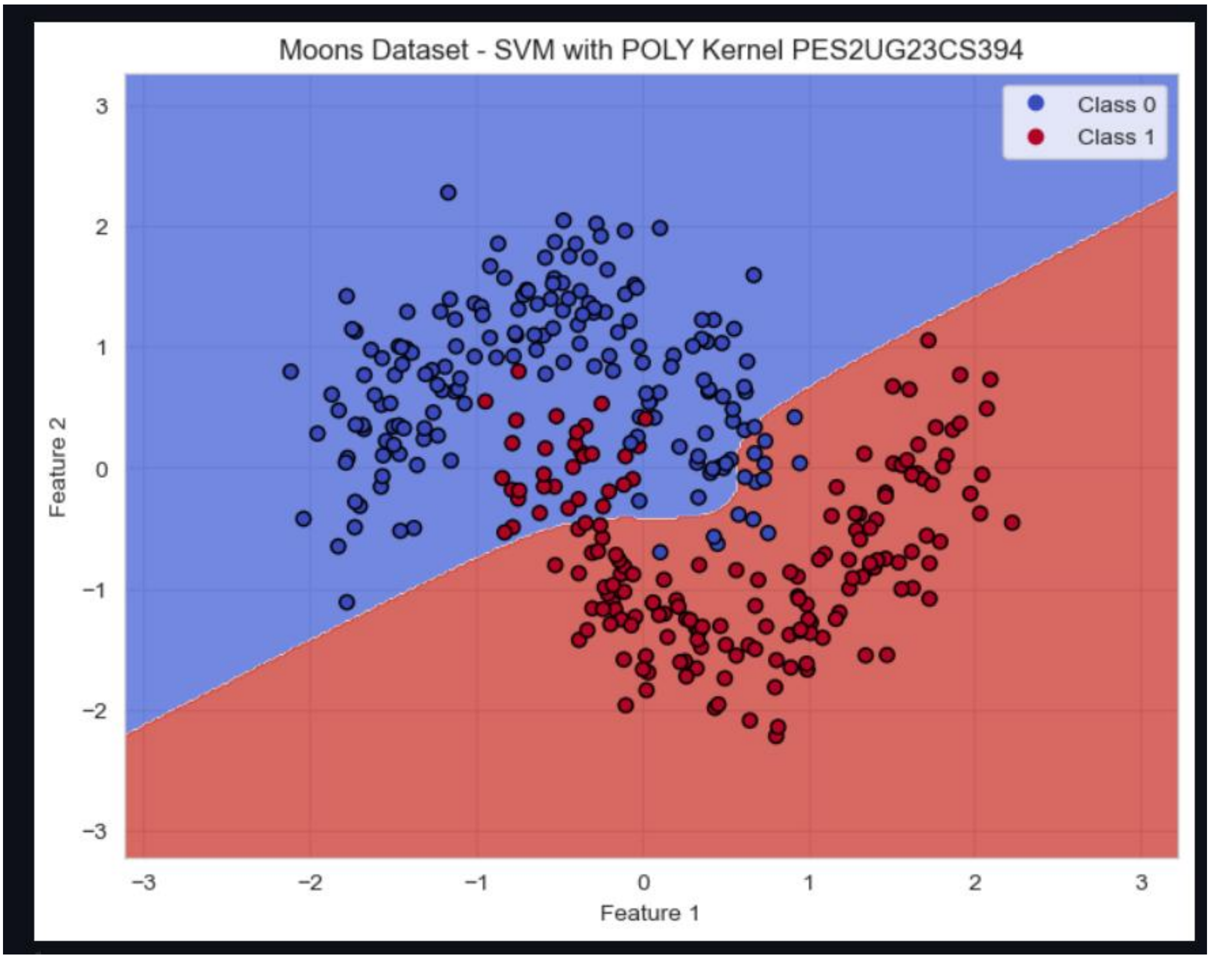
7. Moons Dataset - SVM with LINEAR Kernel




8. Moons Dataset - SVM with RBF Kernel

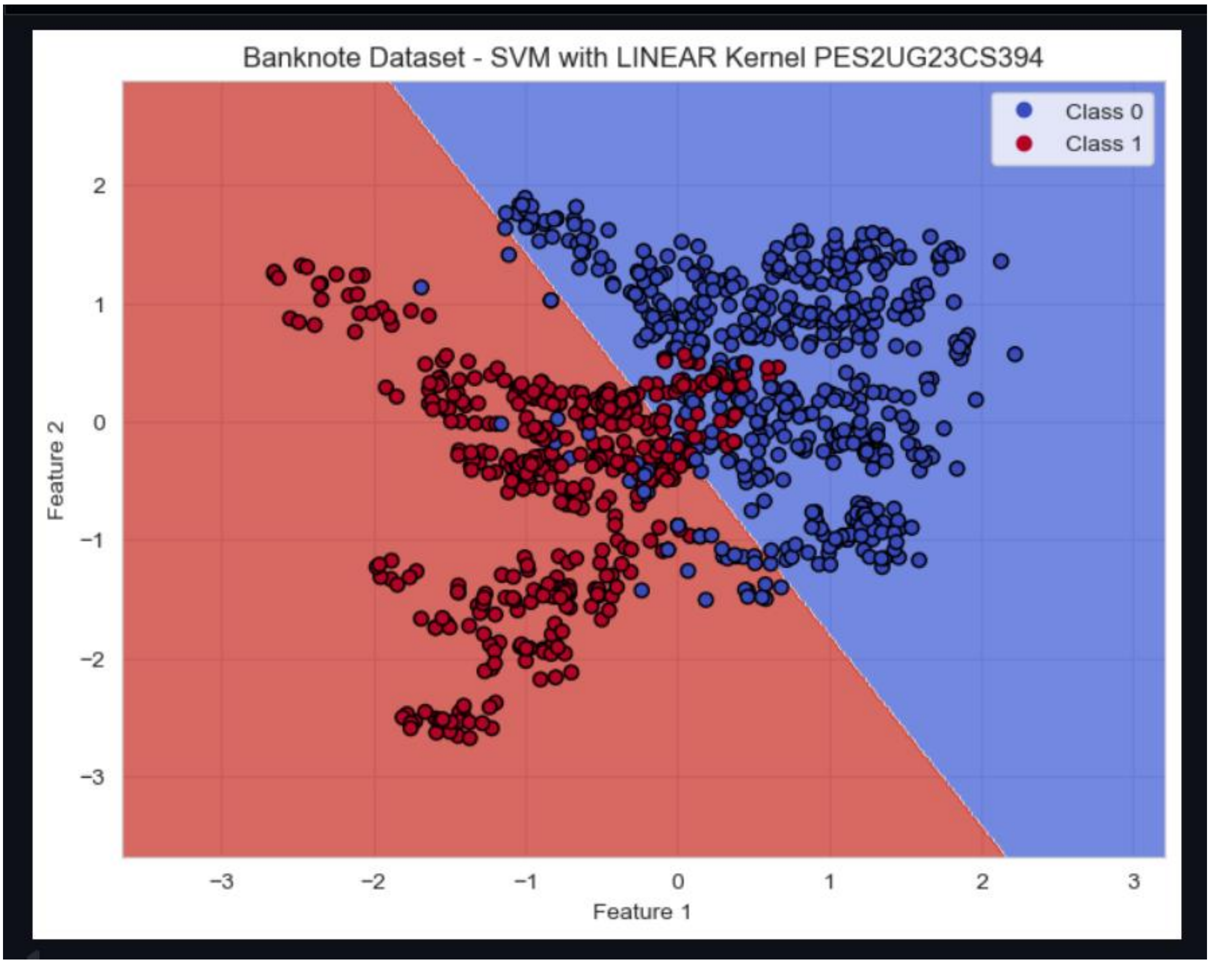



9. Moons Dataset - SVM with POLY Kernel

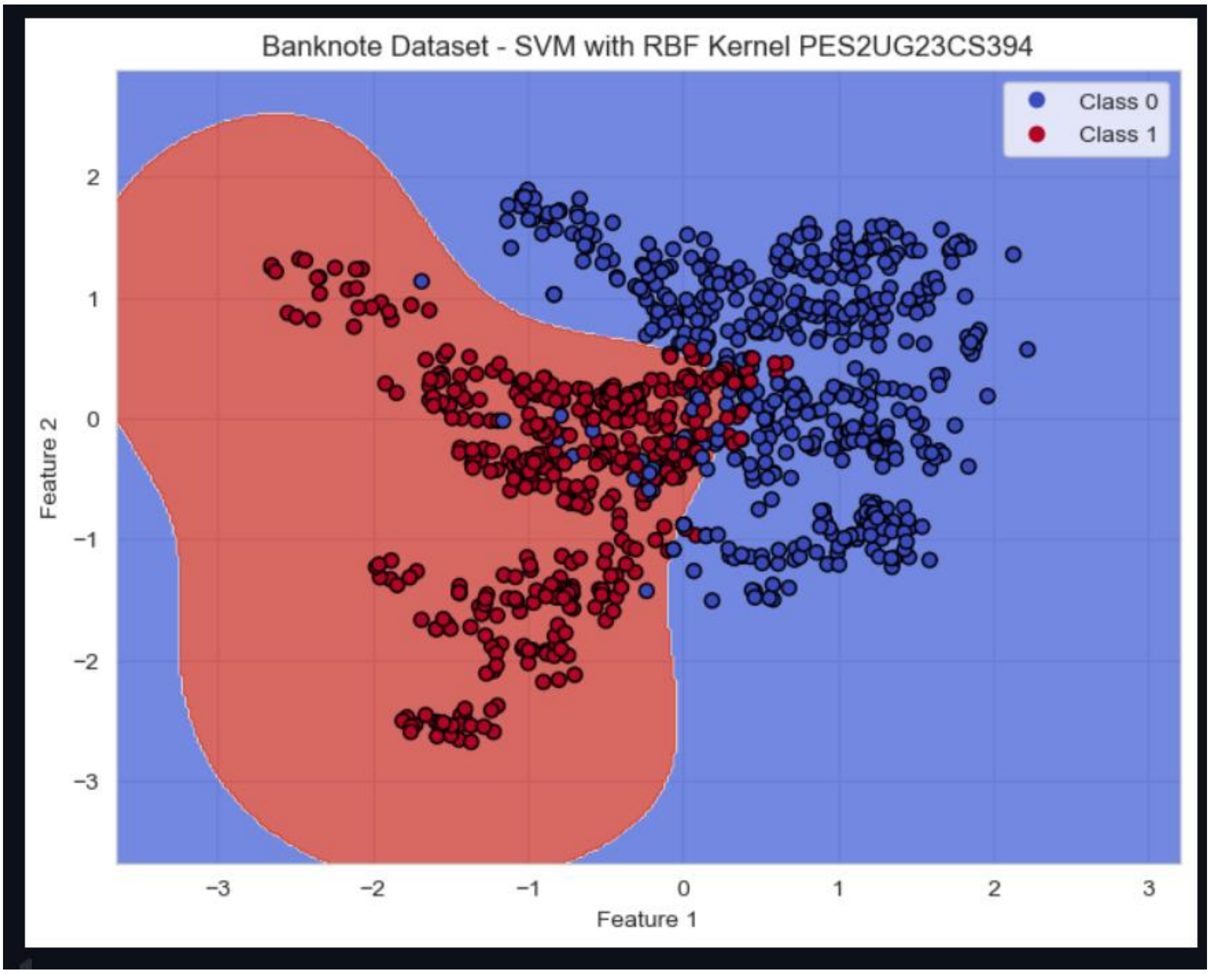



. Banknote Dataset (3 plots):

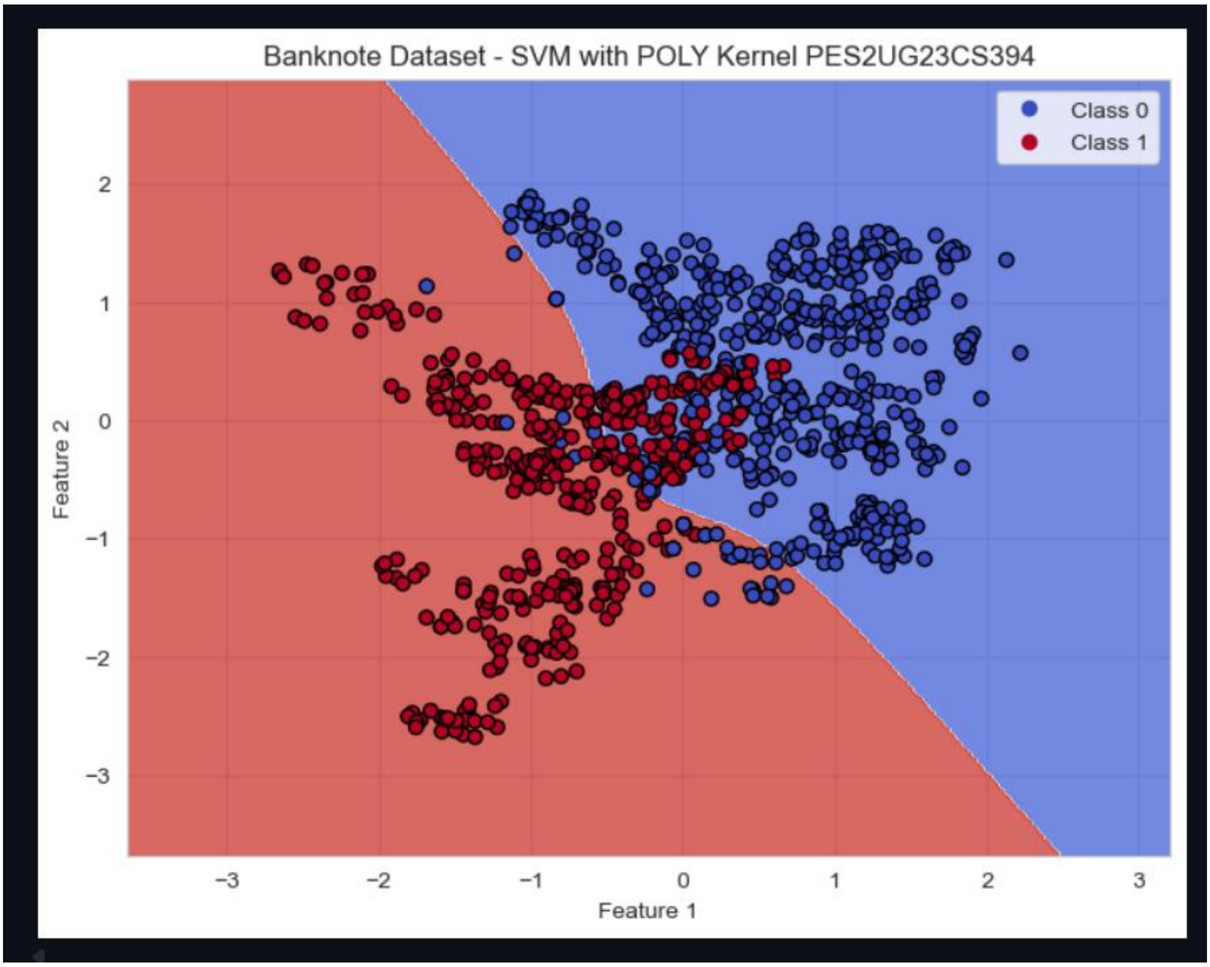
10. Banknote Dataset - SVM with LINEAR Kernel




11. Banknote Dataset - SVM with RBF Kernel

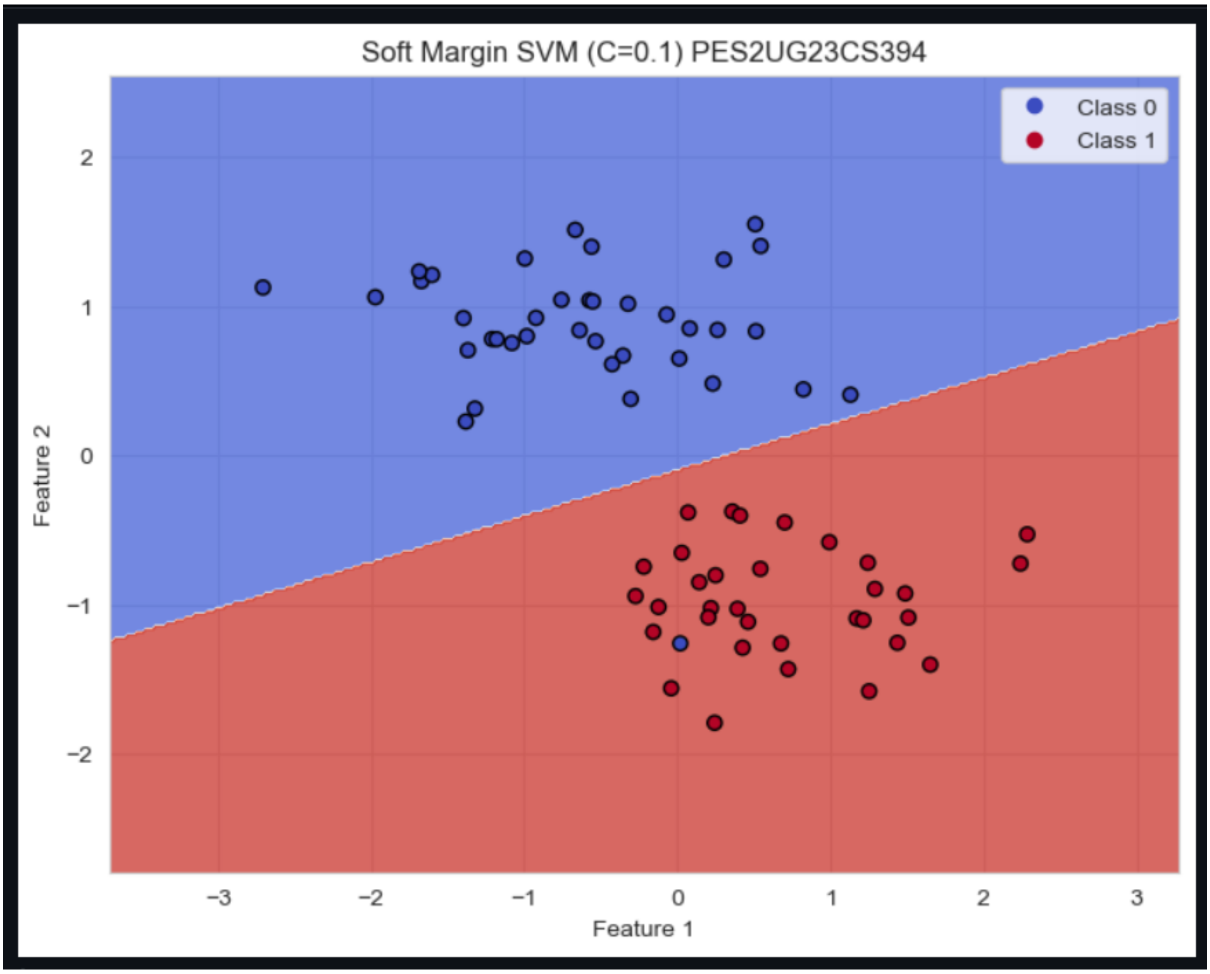



12. Banknote Dataset - SVM with POLY Kernel

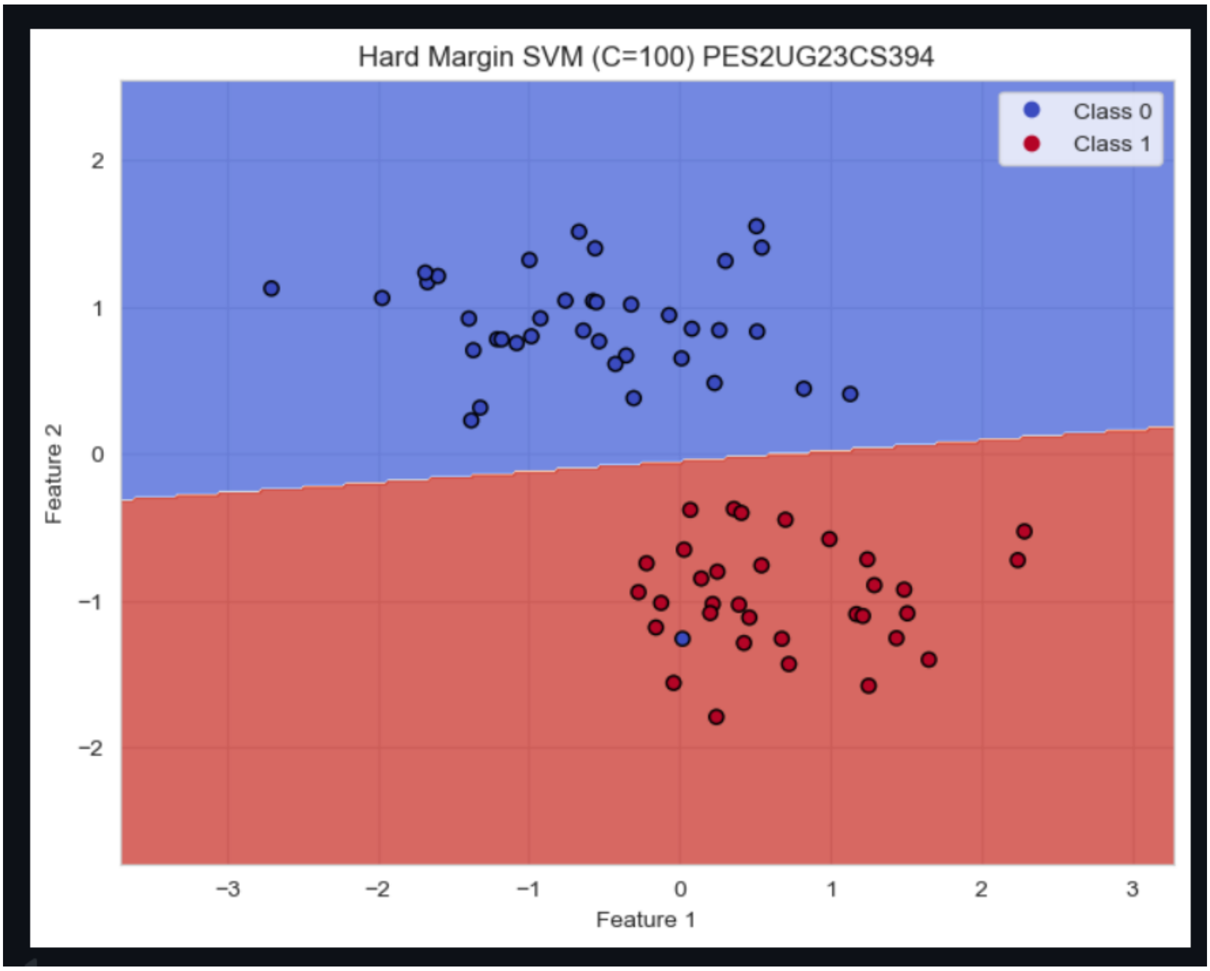



· Margin Analysis (2 plots):

13. Soft Margin SVM (C=0.1)




14. Hard Margin SVM (C=100)


Hard Margin SVM (C=100) PES2UG23CS394
Class 0
Class 1
2
1
0
−1
−2
−3
−2
−1
0
1
2
3
Feature 1
Feature 2