# UE23CS352A: Machine Learning
# Week 10 Lab Report – Support Vector Machine (SVM)

Name: Neha Rajkumar Patil
SRN: PES2UG23CS379
Date: 10-10-2025

## Analysis Questions:

**Moons Dataset**

**1. Inferences about the Linear Kernel's performance:**
The Linear kernel performs poorly on the Moons dataset because the data is non-linearly separable (the two classes form interlocking moon shapes). It produces a straight-line decision boundary, misclassifying many points and showing low accuracy and high bias.

**2. Comparison between RBF and Polynomial kernel decision boundaries:**
Both RBF and Polynomial kernels handle non-linearity, but:

- The RBF kernel captures the curved structure of the moons more naturally, forming smooth circular boundaries around clusters.
- The Polynomial kernel may fit well but tends to create more complex, wavy, or overfitted boundaries depending on the degree used. Thus, RBF gives a cleaner and more generalizable separation for this dataset.

**Banknote Dataset**

**1. Which kernel was most effective for this dataset?**
The Linear kernel was most effective.
The Banknote data is largely linearly separable (features like variance and skewness already differentiate classes well), so complex kernels like RBF or Polynomial do not add much benefit and may slightly reduce performance.

**2. Why might the Polynomial kernel have underperformed here?**
Because the dataset does not require complex non-linear transformations, the Polynomial kernel introduces unnecessary complexity, leading to overfitting and poorer generalization on test data.

## Hard vs. Soft Margin

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

The soft margin is wider.
It allows the decision boundary to be more flexible, tolerating small misclassifications to achieve a larger overall separation.

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

The soft margin SVM allows some points to be misclassified to maximize the overall margin and improve generalization.
This helps prevent overfitting when data is noisy or not perfectly separable.

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

The hard margin model is more likely to overfit because it forces perfect classification of training data, fitting even to noise or outliers, which reduces performance on unseen data.

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

The soft margin model, because it generalizes better, handles noise gracefully, and avoids overfitting — making it more reliable for real-world, unseen data.

### Training Results:

#### Moons Dataset

**1. Classification Report for SVM with LINEAR Kernel with SRN**

SVM with LINEAR Kernel - PES2UG23CS379

| | 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 - PES2UG23CS379

| | 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 - PES2UG23CS379

| | 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

### 4. Classification Report for SVM with LINEAR Kernel

SVM with LINEAR Kernel - PES2UG23CS379

| | 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 - PES2UG23CS379

| | 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

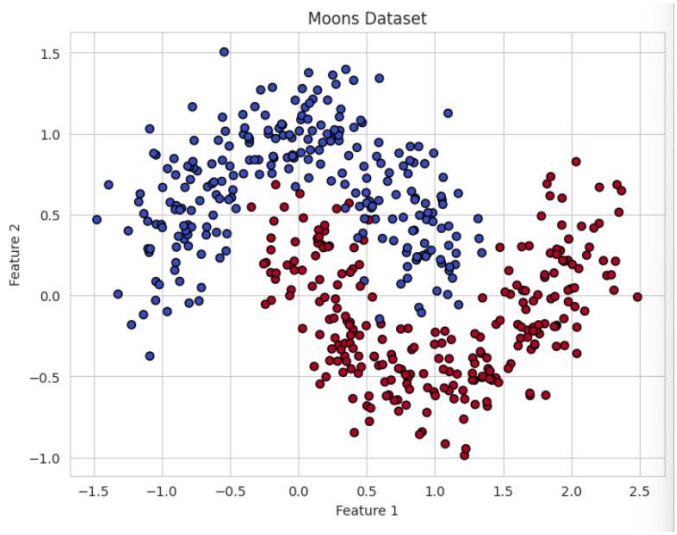
SVM with POLY Kernel - PES2UG23CS379

| | 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 |

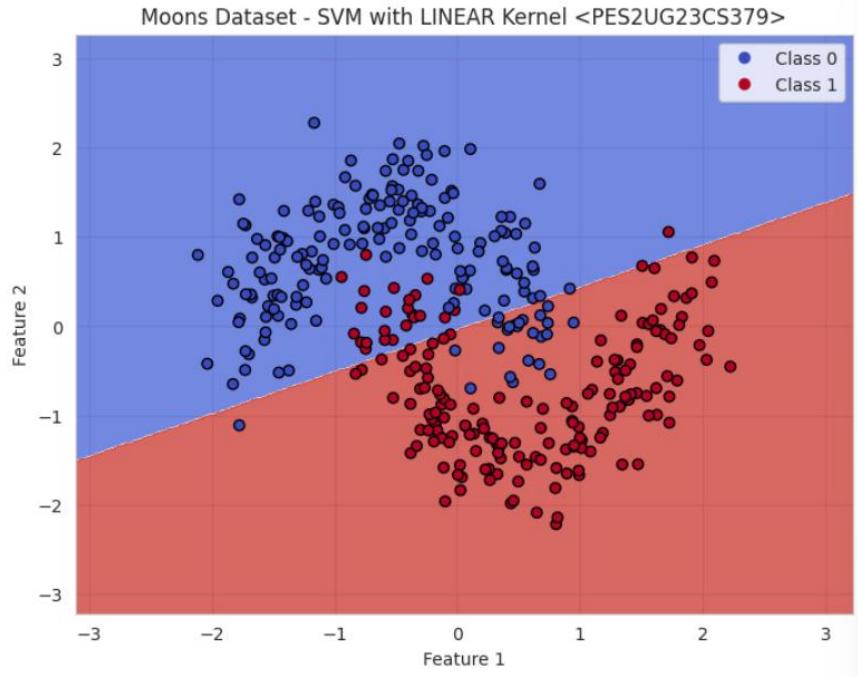
---

## Decision Boundary Visualizations

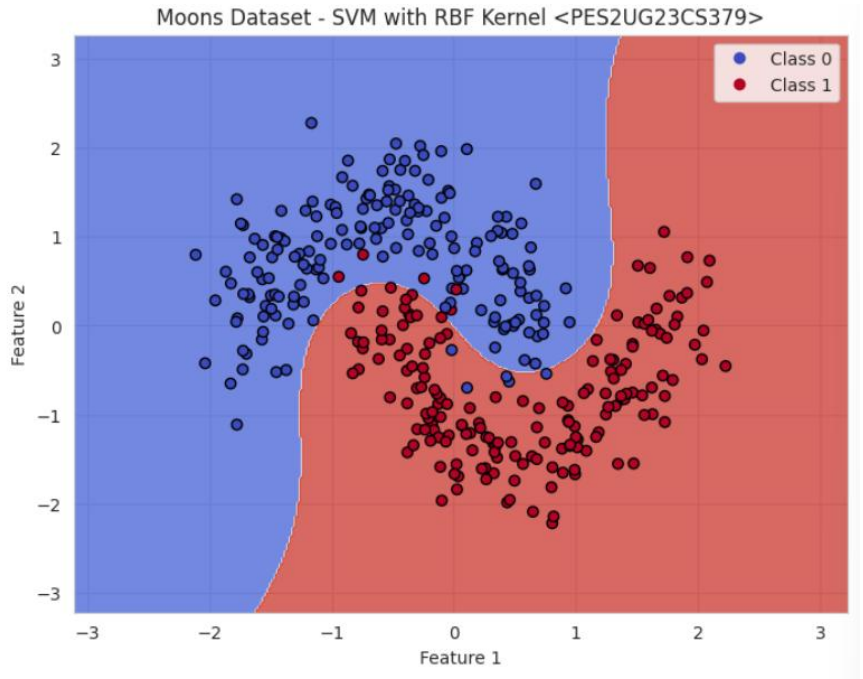
## Moons Dataset



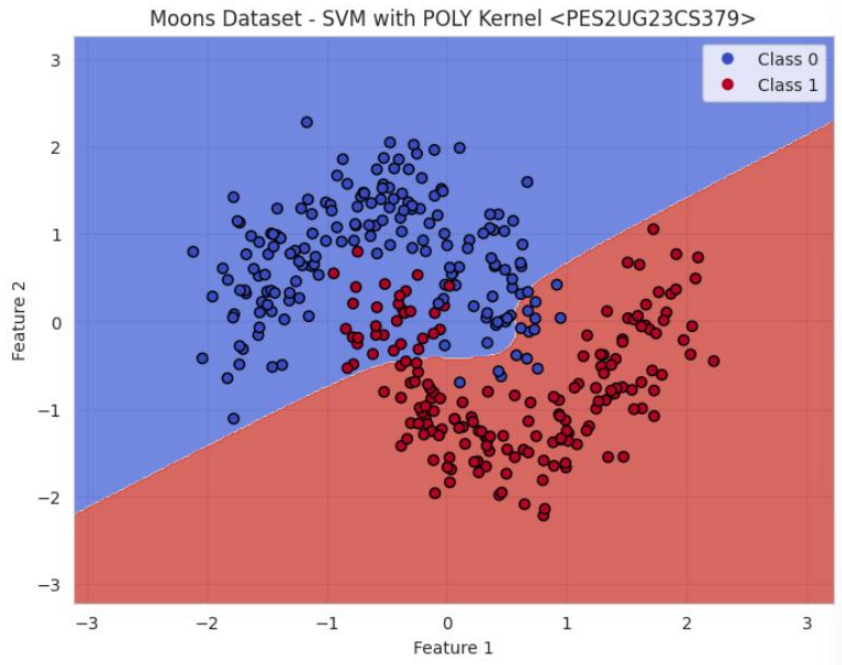

## 7. SVM with LINEAR Kernel



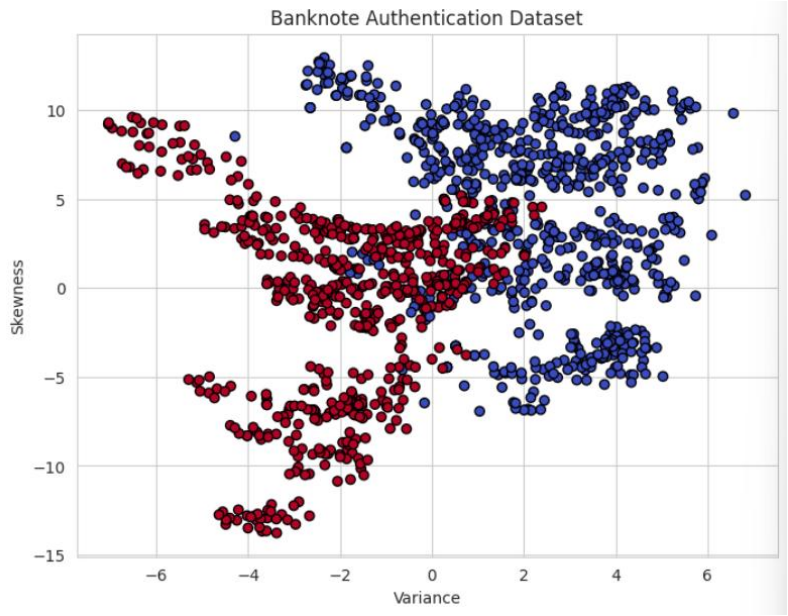
## 8. SVM with RBF Kernel



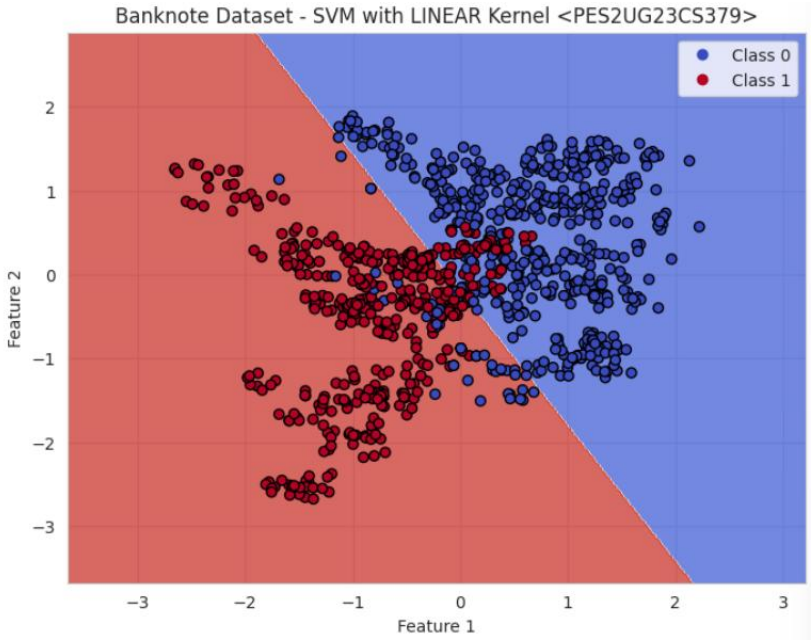

## 9. SVM with POLY Kernel



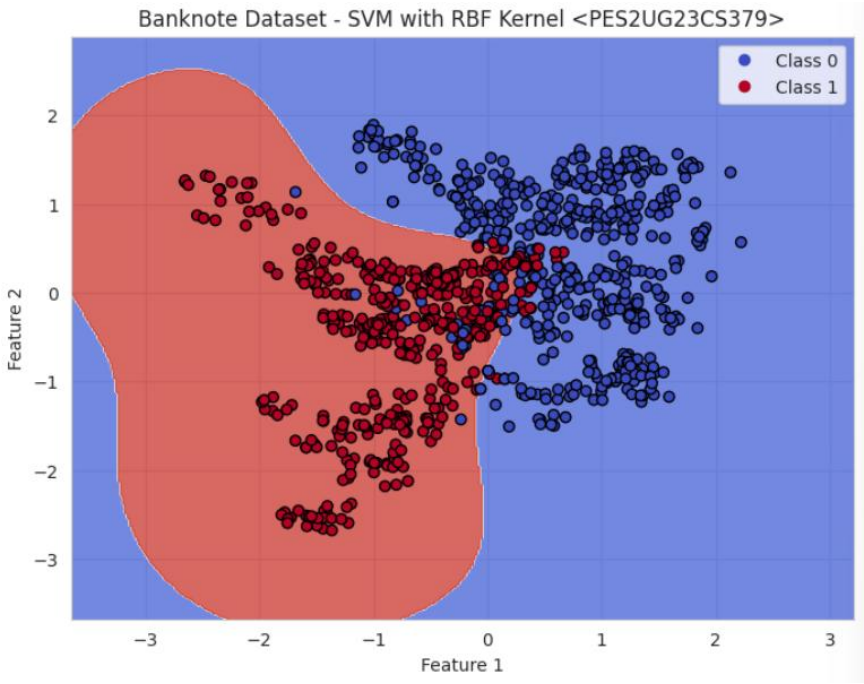
**Banknote Dataset**



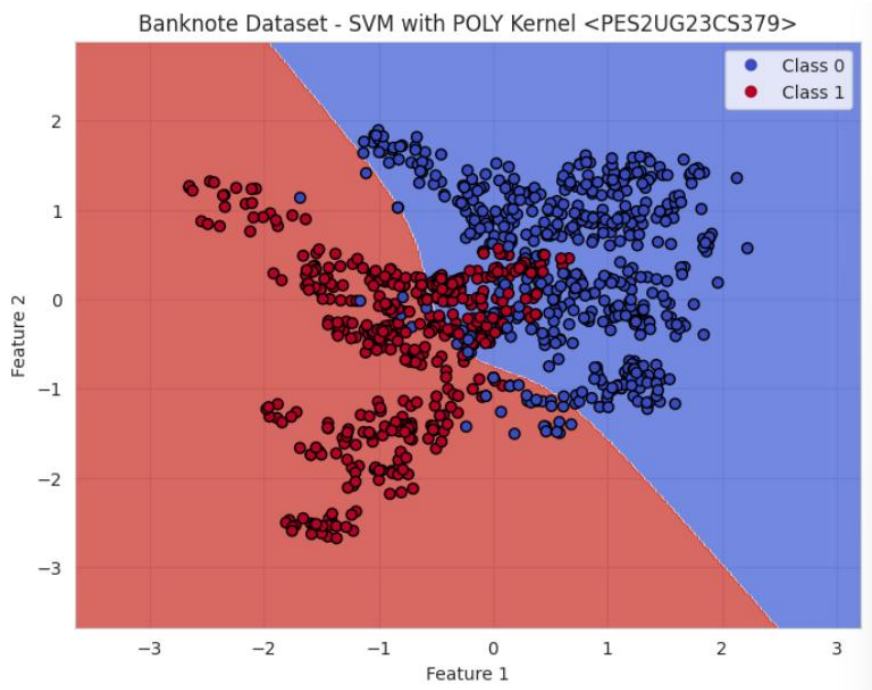

**10. SVM with LINEAR Kernel**



## 11. SVM with RBF Kernel

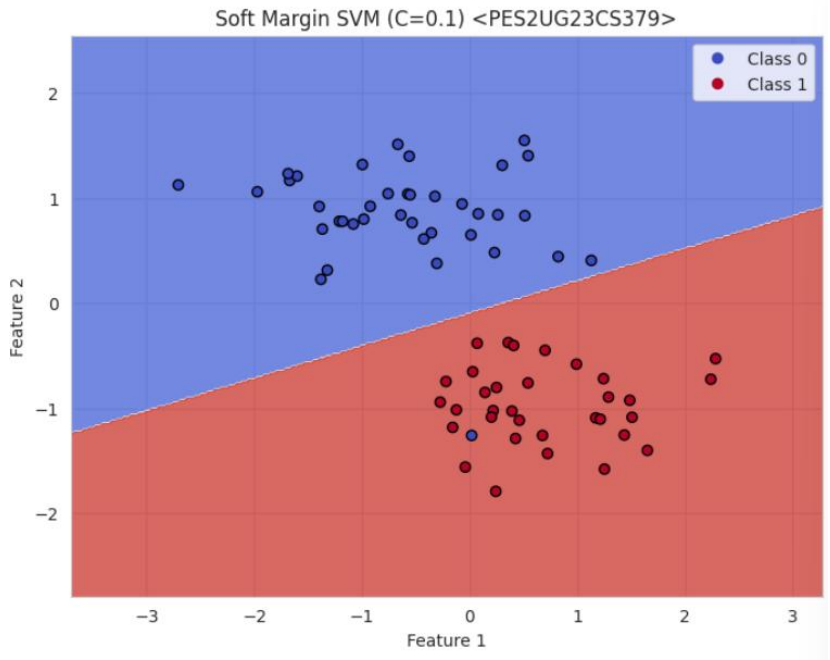



## 12. SVM with POLY Kernel



## Margin Analysis

### 13. Soft Margin SVM (C=0.1)




### 14. Hard Margin SVM (C=100)