UE23CS352A: MACHINE LEARNING

Week 10: SVM(Support Vector Machines) - Classifiers

Name : CHARAN M REDDY

SRN : PES2UG23CS146

Section : C

Date : 10/10/2025

---

2. Analysis Questions

1. Based on the metrics and the visualizations, what inferences about the performance of the Linear Kernel can you draw?

>> The Linear Kernel was able to separate the two classes to some extent, but not perfectly. From the visualization, we can see that the decision boundary is a straight line, while the data points are curved like two half-moons. Because of this, some points near the overlapping region were misclassified. The accuracy of around 87% shows that the model performed well but not the best for this type of non-linear data.

2. Compare the decision boundaries of the RBF and Polynomial kernels. Which one seems to capture the shape of the data more naturally?

>> The RBF kernel creates a smooth, flexible boundary that bends naturally around the points of the two classes. It can separate complicated patterns and handles overlaps well. The Polynomial kernel has a less flexible curved boundary. It captures some of the pattern but misclassifies more points . The RBF kernel captures the shape of the data more naturally than the Polynomial kernel, making it better for this dataset.

3. In this case, which kernel appears to be the most effective?

>> The RBF kernel is the most effective for this dataset. It has the highest accuracy (93%) and the best balance of precision and recall .

4. The Polynomial kernel shows lower performance here compared to the Moons dataset. What might be the reason for this?

>> The Polynomial kernel performs worse on the Banknote dataset because the data is more irregular and not easily captured by a simple polynomial curve. The banknote features have more complex patterns, so the polynomial boundary cannot fit all points well.

5. Compare the two plots. Which model, the "Soft Margin" (C=0.1) or the "Hard Margin" (C=100), produces a wider margin?

>> The "Soft Margin" SVM model (with C=0.1) produces a wider margin. The "Hard Margin" model's boundary is much closer to the data points, resulting in a narrower margin.

6. Look closely at the "Soft Margin" (C=0.1) plot. You'll notice some points are either inside the margin or on the wrong side of the decision boundary. Why does the SVM allow these "mistakes"? What is the primary goal of this model?

>> The "Soft Margin" model allows some points to be misclassified because its main objective is to find a decision boundary that is stable and generalizes well. By allowing a few mistakes on outlier or noisy points, the model can create a wider and accurate margin that better captures the overall trend of the data. It essentially trades a little bit of training accuracy for better predictive performance .

7. Which of these two models do you think is more likely to be overfitting to the training data? Explain your reasoning.

>> The "Hard Margin” SVM model (C=100) is much more likely to be overfitting to the training data. Overfitting occurs when a model learns the training data too well, including its noise and outliers. The hard margin model has laid its decision boundary specifically to classify every single data point correctly. This extreme sensitivity to training data is a sign of overfitting.

8. Imagine you receive a new, unseen data point. Which model do you trust more to classify it correctly? Why? In a real-world scenario where data is often noisy, which value of C (low or high) would you generally prefer to start with?

>> "Soft Margin" (C=0.1) model . Because it is less likely to be overfitted. Its decision boundary was determined by the general distribution of the data, not by specific outliers. In a real-world scenario where data is almost always noisy, it is generally better to start with a lower value of C. A lower C creates a soft margin, providing a more clear and generalized model .

---

3. Screenshots

1) Classification Report for SVM with LINEAR Kernel (Moons Dataset)

```
SVM with LINEAR Kernel <PES2UG23CS146>
              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 (Moons Dataset)

```
SVM with RBF Kernel <PES2UG23CS146>
              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 (Moons Dataset)

```
SVM with POLY Kernel <PES2UG23CS146>
              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

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

4) Classification Report for SVM with LINEAR Kernel (Banknote Dataset)

```
SVM with LINEAR Kernel <PES2UG23CS146>
              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

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

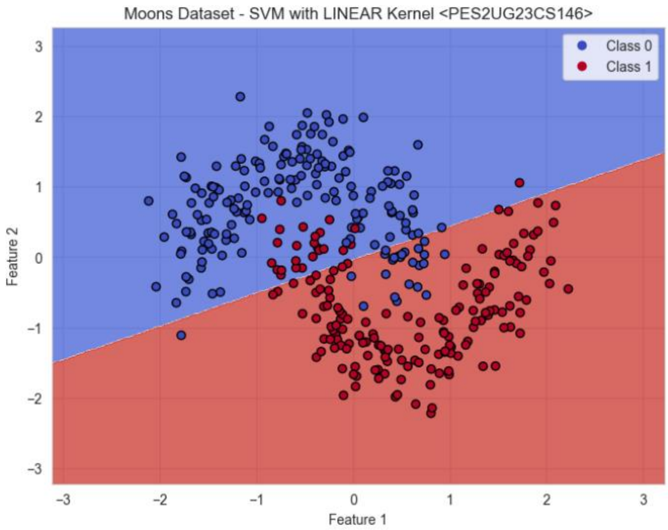
5) Classification Report for SVM with RBF Kernel (Banknote Dataset)

```
SVM with RBF Kernel <PES2UG23CS146>
              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 (Banknote Dataset)

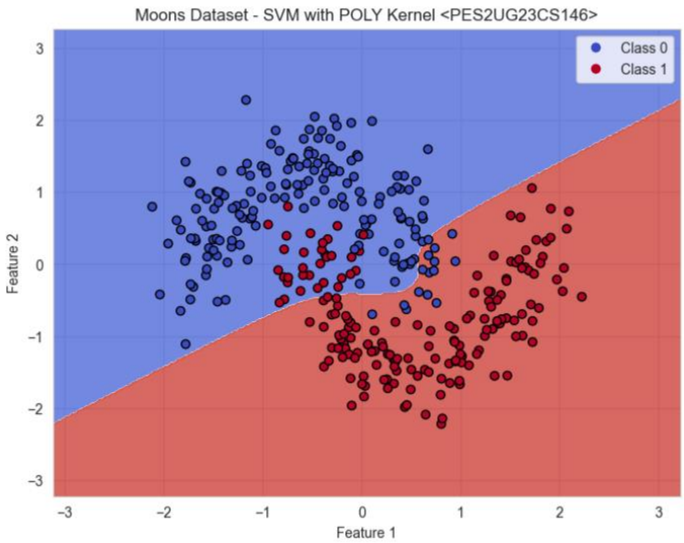
```
SVM with POLY Kernel <PES2UG23CS146>
              precision    recall  f1-score   support

      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

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

7) Decision Boundary Visualizations for SVM with LINEAR Kernel (Moons Dataset)

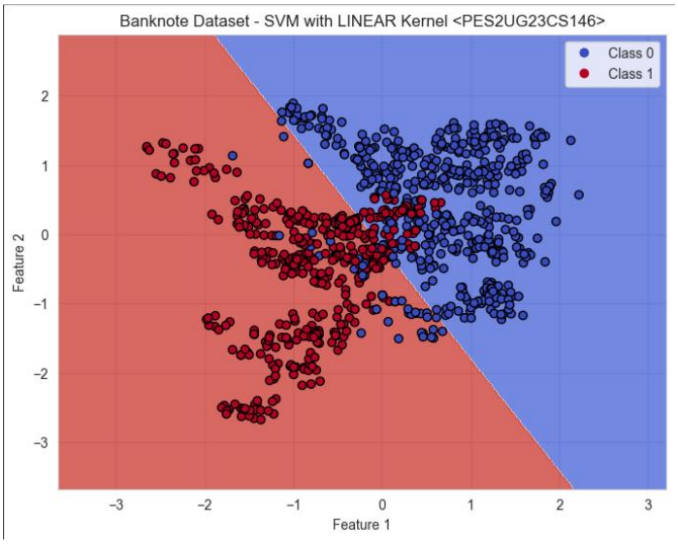


8) Decision Boundary Visualizations for SVM with RBF Kernel (Moons Dataset)

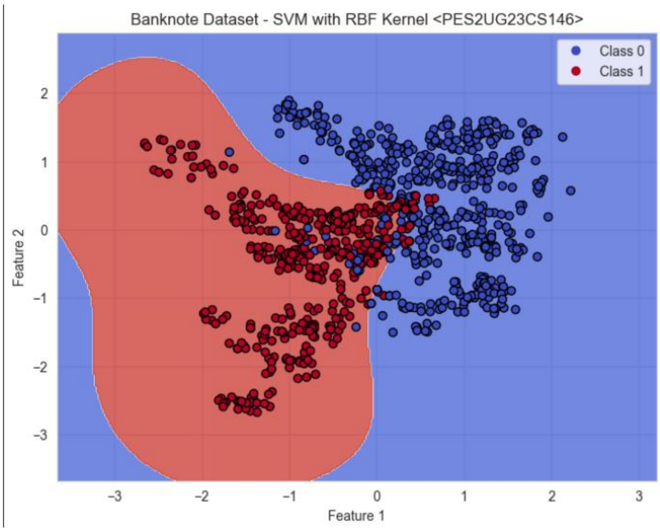
9) Decision Boundary Visualizations for SVM with POLY Kernel (Moons Dataset)



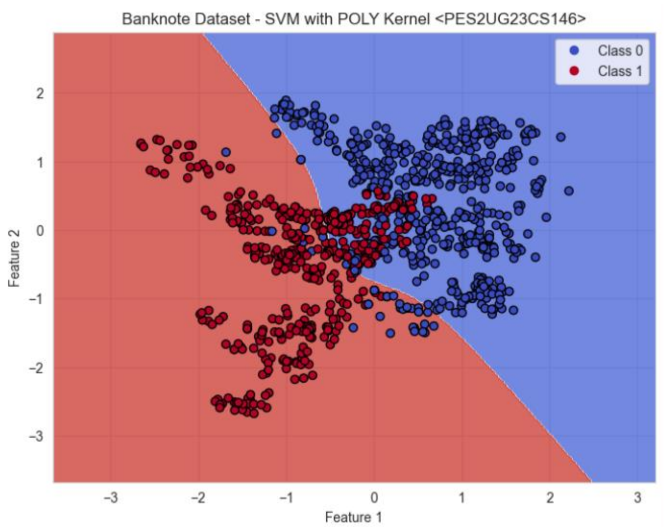
10) Decision Boundary Visualizations for SVM with LINEAR Kernel (Banknote Dataset)



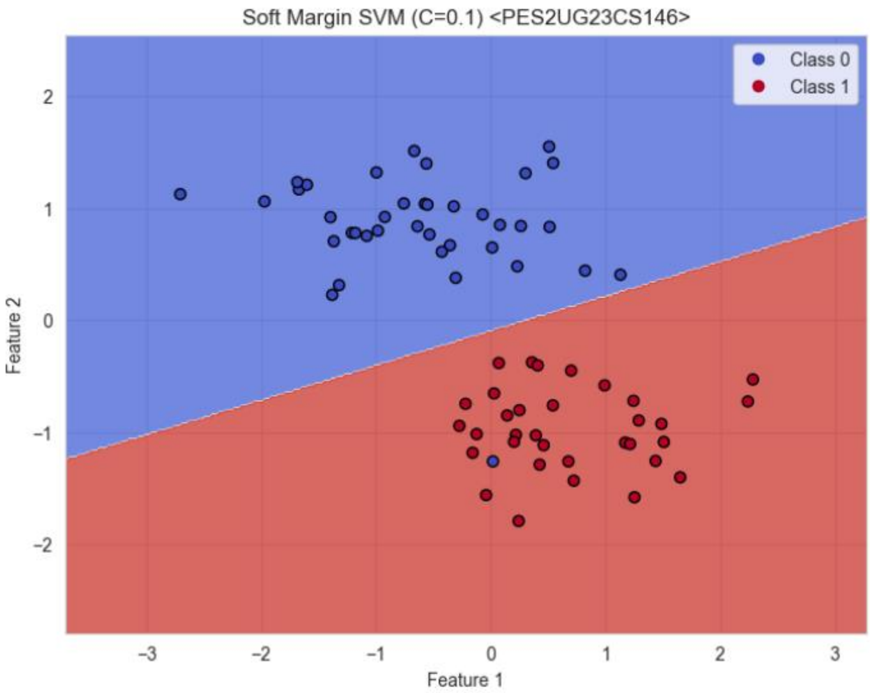
11) Decision Boundary Visualizations for SVM with RBF Kernel (Banknote Dataset)



12) Decision Boundary Visualizations for SVM with POLY Kernel (Banknote Dataset)



13) Margin Analysis for Soft Margin SVM (C=0.1)



14) Margin Analysis for Hard Margin SVM (C=100)