EXPERIMENT - 8

AIM: Support Vector Machines (SVMs) and the Kernel Trick

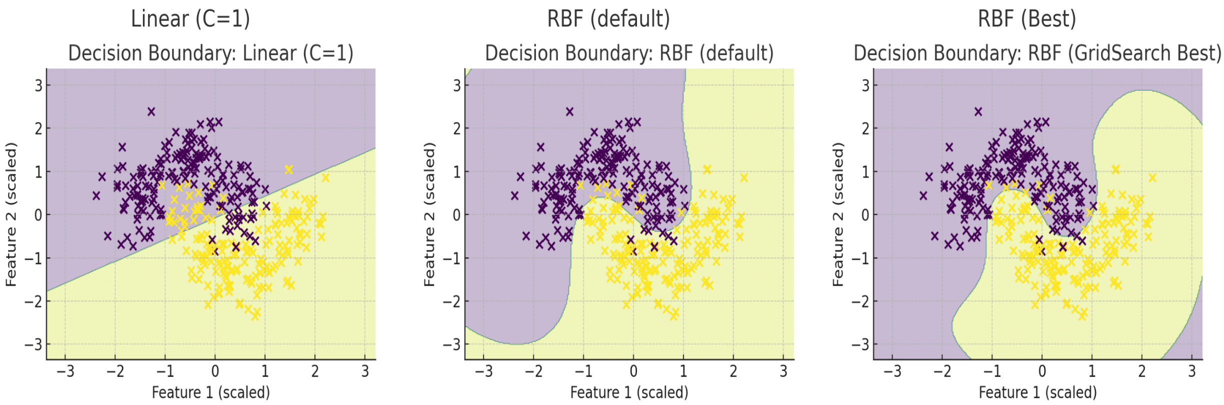
1. Learning Objectives

- Articulate SVM core concepts: maximal margin hyperplane, support vectors, and the soft-margin (C).
- Understand and implement the kernel trick for non-linear classification.
- Implement & compare Linear, Polynomial, and RBF SVMs.
- Tune hyperparameters (C, gamma, degree) using GridSearchCV.
- Visualize 2D decision boundaries.
- Rigorously evaluate on a hold-out test set.

Model Comparison (Validation Accuracy)

Model	Validation Accuracy
SVC (kernel='linear', C=1)	0.8400
SVC (kernel='rbf', default)	0.9467
SVC (kernel='poly', degree=3)	0.8400
GridSearchCV Best Model	0.9533

Decision Boundary Plots (Linear, default RBF, GridSearch best RBF)

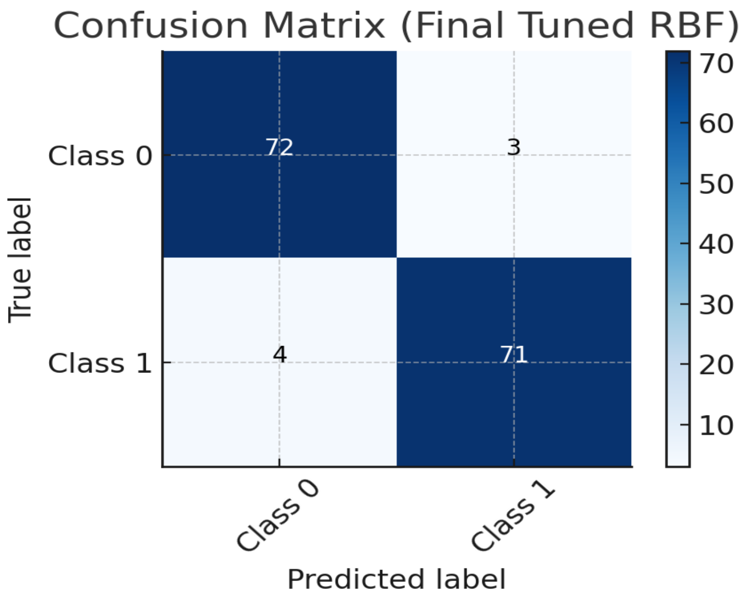


Final Performance (Hold-out Validation Set)

Classification Report (Final Tuned Model)

precision recall f1-score support 0 0.95 0.96 0.95 75 1 0.96 0.95 0.95 75 accuracy 0.95 150 macro avg 0.95 0.95 0.95 150 weighted avg 0.95 0.95 0.95 150

Confusion Matrix (Final Tuned Model)



Conclusions & Answers

Why did Linear SVM fail while RBF succeeded?
The moons dataset is non-linearly separable; a single straight hyperplane cannot split the classes. RBF implicitly maps data into a higher-dimensional space where a linear separator exists, producing curved, flexible boundaries as seen in the plots.

GridSearchCV Findings (Best C and gamma)
Best Params: C = 1, gamma = 1. Best Cross-Validated Accuracy: 0.9543.

Effect of gamma too high (e.g., 1000)
The decision boundary becomes extremely wiggly and overfits to individual points, harming generalization.

Effect of C too low (e.g., 0.01)
The model allows many violations to maximize margin, underfitting the data and yielding a smoother but inaccurate boundary.