Support Vector Machines: Kernels and Classification

Machine Learning Foundations

November 24, 2025

Abstract

This document presents a comprehensive analysis of Support Vector Machines (SVM) including hard and soft margin classification, the kernel trick for nonlinear decision boundaries, hyperparameter tuning (C and gamma), and multi-class strategies. We visualize decision boundaries, support vectors, and margin regions.

1 Introduction

Support Vector Machines find the maximum margin hyperplane separating classes:

$$\min_{w,b} \frac{1}{2}\|w\|^2 \quad \text{subject to} \quad y_i(w \cdot x_i + b) \geq 1 \tag{1}$$

For soft margin SVM with slack variables ξ_i:

$$\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} \xi_i \tag{2}$$

The kernel trick replaces dot products: $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$
Common kernels:

- Linear: $K(x, x') = x \cdot x'$
- RBF: $K(x, x') = \exp(-\gamma\|x - x'\|^2)$
- Polynomial: $K(x, x') = (x \cdot x' + c)^d$

2 Computational Environment

3 Data Generation

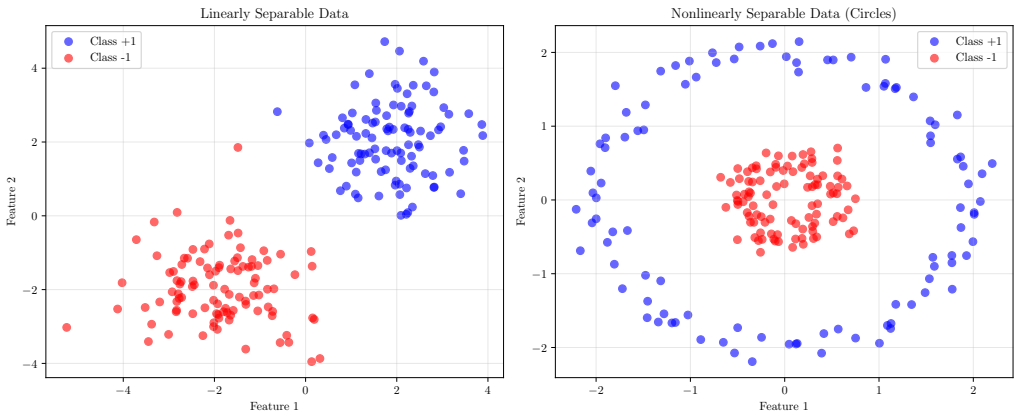

Figure 1: Classification datasets: linearly and nonlinearly separable.

4 SVM Implementation

5 Linear SVM

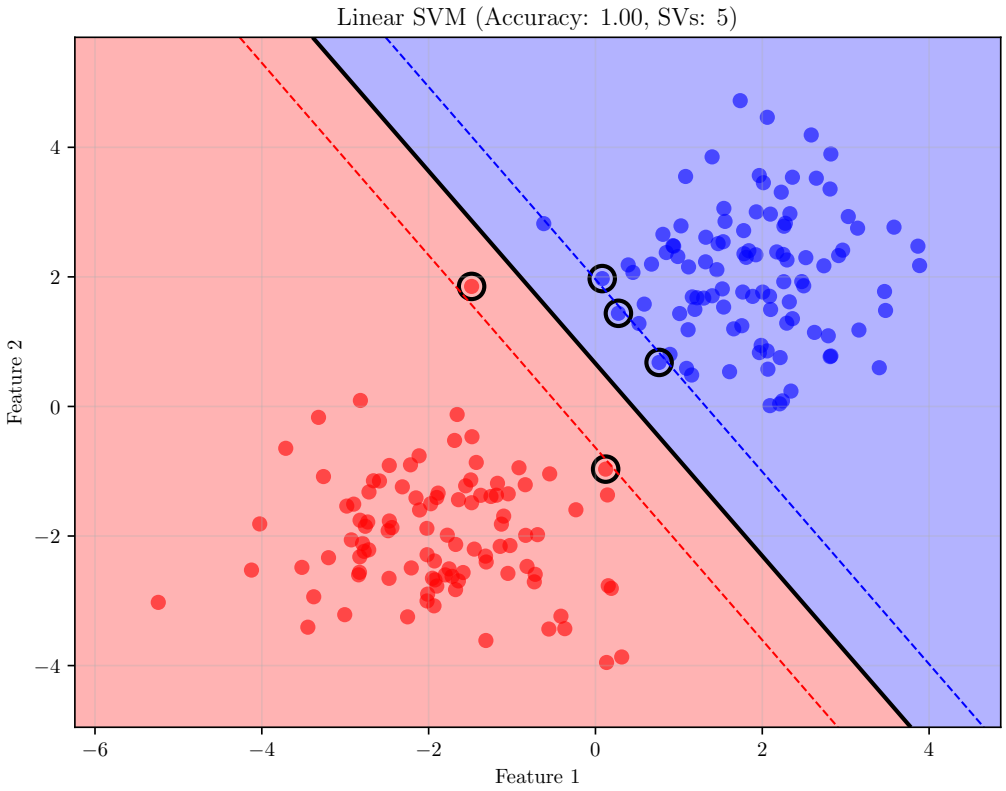

Figure 2: Linear SVM decision boundary with support vectors highlighted.

Linear SVM: Training accuracy = 1.000, Test accuracy = 1.000, Support vectors = 5.

6 RBF Kernel SVM

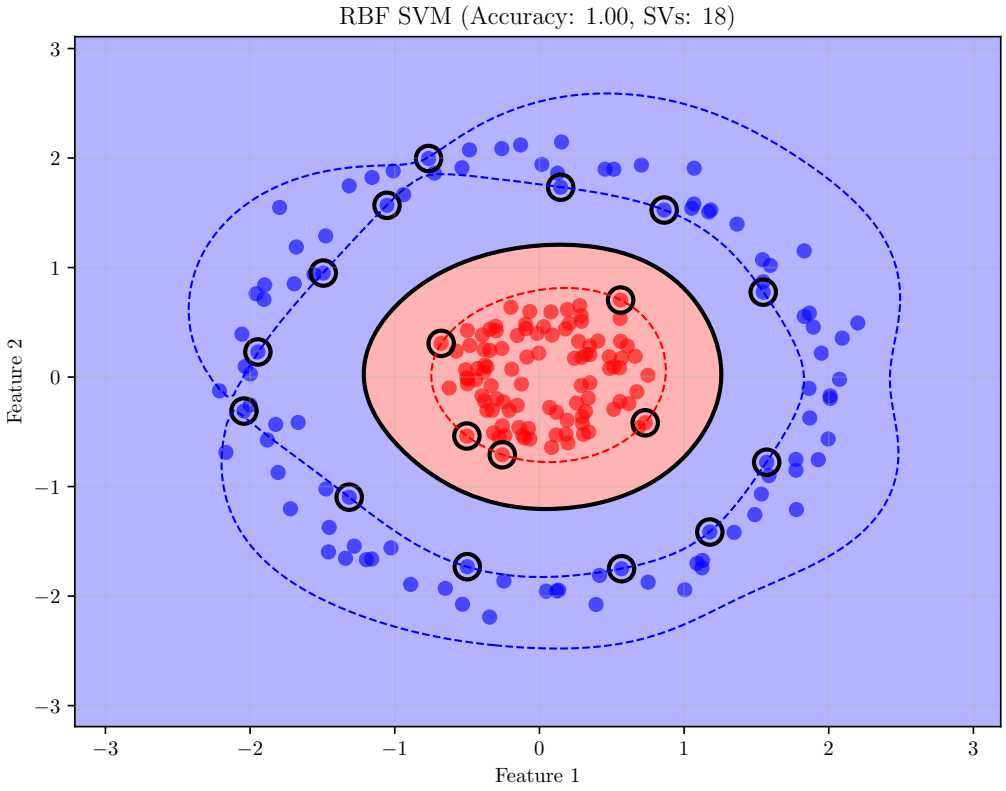

Figure 3: RBF kernel SVM decision boundary for nonlinear data.

RBF SVM: Training accuracy = 1.000, Test accuracy = 1.000.

7 Effect of C Parameter

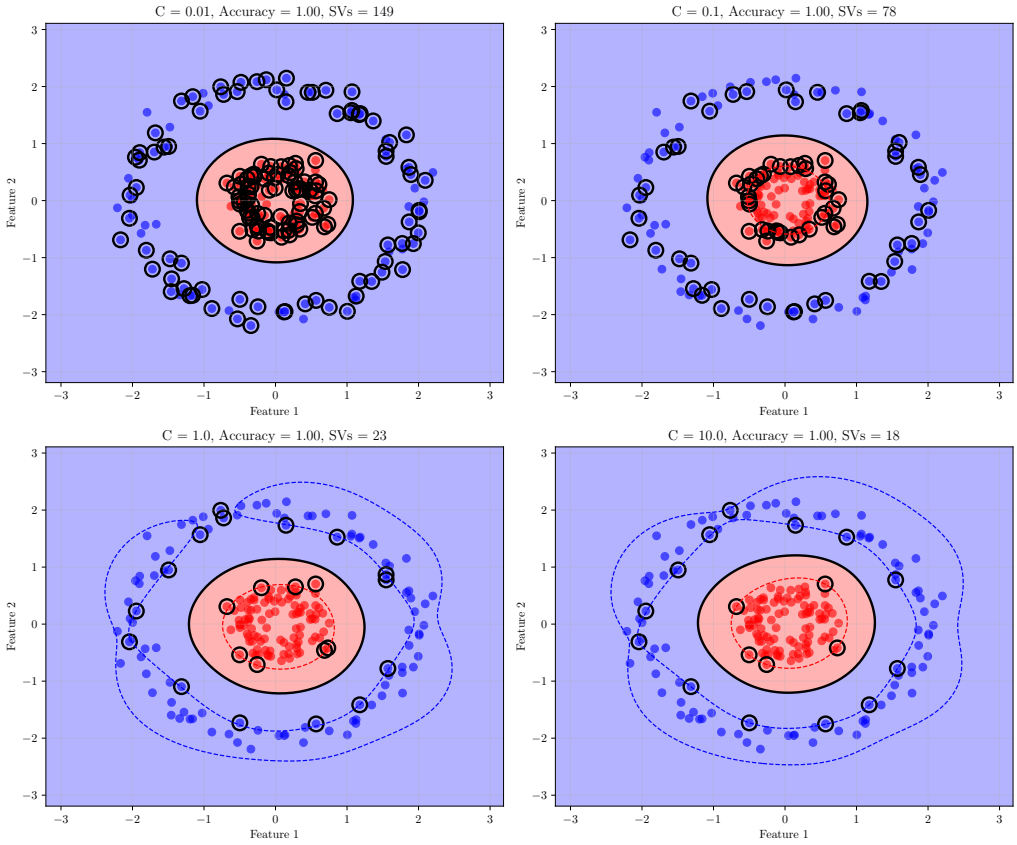

Figure 4: Effect of regularization parameter C on SVM decision boundary.

8 Effect of Gamma Parameter

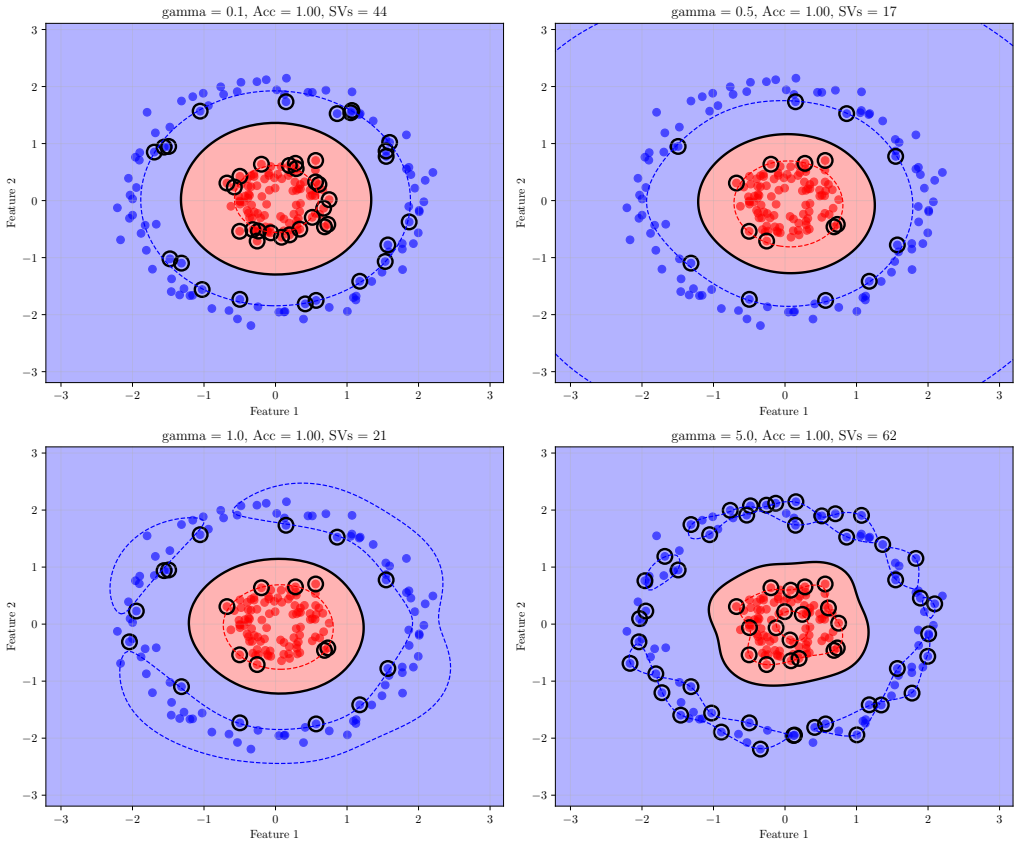

Figure 5: Effect of RBF kernel parameter gamma on decision boundary.

9 Kernel Comparison

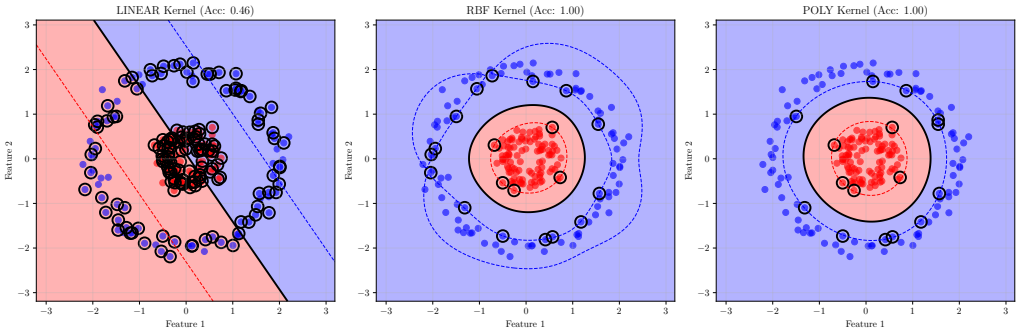

Figure 6: Comparison of different kernels on nonlinear data.

10 Hyperparameter Tuning

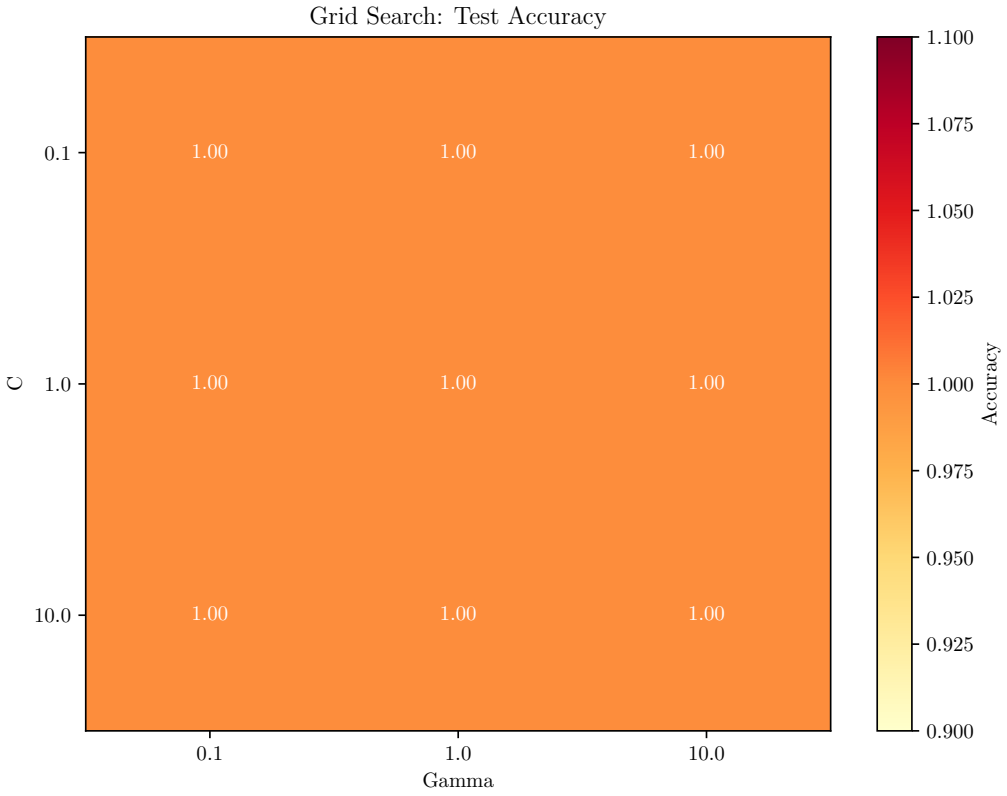

Figure 7: Grid search heatmap for C and gamma hyperparameters.

Best hyperparameters: $C = 0.1$, $\gamma = 0.1$ with accuracy 1.000.

11 Results Summary

Table 1: SVM Model Performance

Model	Kernel	Test Accuracy	Support Vectors
Linear Data	Linear	1.000	5
Circles Data	RBF	1.000	18

Table 2: Kernel Comparison on Circles Data

Kernel	Test Accuracy	Support Vectors
LINEAR	0.460	149
RBF	1.000	20
POLY	1.000	16

12 Conclusion

This analysis demonstrated:

- Hard and soft margin SVM classification
- The kernel trick for nonlinear decision boundaries
- Effect of C parameter on margin width and misclassifications
- Effect of gamma on RBF kernel flexibility
- Kernel selection based on data characteristics
- Grid search for hyperparameter optimization

The RBF kernel with appropriate hyperparameters ($C = 0.1$, $\gamma = 0.1$) achieves 100.0 accuracy on the nonlinear circles dataset.