

# REPORT

## MODEL COMPARISON TABLE

Model	Kernel	C	$\gamma$	Validation Accuracy
1. Linear SVC	linear	1	N/A	{Insert linear_accuracy value}
2. Default RBF	rbf	1	'scale'	{Insert rbf_accuracy value}
3. Polynomial SVC	poly	1	N/A	{Insert poly_accuracy value}
4. GridSearchCV Best	rbf	{Insert \$C\$ from best_params}	{Insert \$\gamma\$ from best_params}	{Insert final tuned accuracy value}

### **Q1 Why did the linear SVM fail, and why did the RBF kernel succeed?**

Linear SVM failed because the "moons" data is non-linearly separable. As shown in the plot, a linear boundary is a straight line, which cannot effectively separate the two interlocked classes, leading to high error.

RBF Kernel succeeded because the kernel trick implicitly mapped the 2D data into a higher-dimensional space where the classes become linearly separable. The resulting decision boundary, when projected back to 2D, is the complex, curved line that perfectly fits the 'moons' shape.

### **Q2 What did the GridSearchCV tell you? What were the best C and gamma?**

GridSearchCV systematically explored the hyperparameter space to find the optimal balance between regularization ( $C$ ) and kernel influence ( $\gamma$ ).

Best Parameters:  $C = \text{best\_params}['C']$ ,  $\gamma = \text{best\_params}['\gamma']$ .

The corresponding best cross-validated accuracy was  $\text{best\_score}: .4f$ .

### **Q3 What happens if $\gamma$ is set too high (e.g., 1000)? What happens if $C$ is set too low (e.g., 0.01)? How would this change the decision boundary?**

$\gamma$  too high (e.g., 1000): A very high  $\gamma$  means the influence of a single training example reaches only a very small, local area. The decision boundary becomes extremely irregular, jagged, and tightly follows every individual data point. This results in severe overfitting and poor generalization on new, unseen data.

$C$  too low (e.g., 0.01): A very low  $C$  forces a very large margin and prioritizes smoothness over classifying every training point correctly (high regularization). The decision boundary becomes very smooth and overly simple, potentially underfitting the data.