

## 1. What's the difference between the different SVM kernels?

Support Vector Machines (SVM) use kernels to change how data is separated. The kernel decides the shape of the decision boundary.

- Linear kernel: draws a straight line (or hyperplane) between classes. It works well when data is almost linearly separable. Output is a distance from the line, which is then turned into a class label.
- Polynomial kernel: allows curved boundaries by adding polynomial terms (squares, cubes, etc.). It is useful when there are interactions between features. The output reflects these polynomial relationships.
- RBF (Radial Basis Function) kernel: creates very flexible, nonlinear boundaries using local similarity between points. It is powerful when the pattern is complex and not linearly separable. The output is based on smooth curves that adapt to the data.

When to use:

- Use Linear when features are many or the relationship is simple.
- Use Polynomial when you expect quadratic/cubic effects.
- Use RBF as a default choice for unknown or complex boundaries, but it usually needs tuning.

## 2. When would you use each one, and how do their outputs differ?

All kernels output class labels, but the decision boundary changes: Linear = straight, Polynomial = fixed curved, RBF = flexible nonlinear. This affects precision, recall, and F1 depending on the data.

Model	Accuracy	Precision (fraud)	Recall (fraud)	F1 (fraud)	ROC-AUC	PR-AUC
Linear SVM	0.970	0.334	0.883	0.485	0.952	0.845
Poly SVM (d=3)	0.990	0.661	0.796	0.722	0.921	0.796
RBF SVM	0.986	0.538	0.774	0.635	0.972	0.689

Linear → best recall, but low precision (many false alarms).

Polynomial → best balance (highest F1).

RBF → best ROC-AUC, but weaker PR-AUC without tuning.