

MMDT0091

Mentor - Ma Nuwai Thet

In this lab, I had practical experience of using SVM kernels in classification. I understand the SVM kernel as a classification tool between two groups of data to find decision boundaries.

Answer the following questions based on your analysis:

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

I have tested three different kernels.

- Linear Kernel
- Polynomial Kernel
- RBF Kernel

Here's are the different predictions results from three different kernels

```
print("Linear Kernel Predictions:", svcL_pipeline.predict(X_test[:5]))  
print("Polynomial Kernel Predictions:", svcPoly_pipeline.predict(X_test[:5]))  
print("RBF Kernel Predictions:", svcRBF_pipeline.predict(X_test[:5]))
```

```
Linear Kernel Predictions: [0. 0. 0. 0. 1.]  
Polynomial Kernel Predictions: [1. 0. 1. 0. 0.]  
RBF Kernel Predictions: [1. 0. 0. 0. 0.]
```

Linear Kernel Predictions: [0. 0. 0. 0. 1.]

Polynomial Kernel Predictions: [1. 0. 1. 0. 0.]

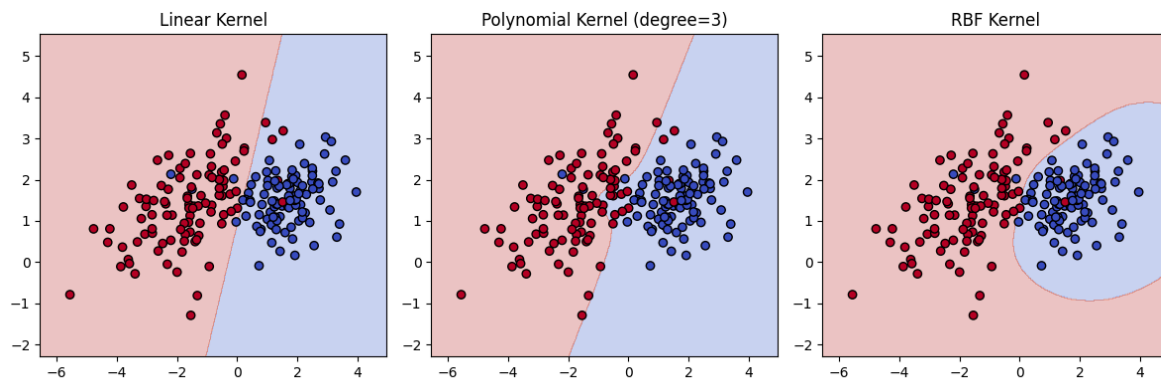
RBF Kernel Predictions: [1. 0. 0. 0. 0.]

In our fraud detection results, as it is a straight line, only the 5th prediction shows fraud.

In polynomials, a few mixed points are marked as fraud in the plot and also in the results first and third are highlighted as fraud too.

In RBF, we can see that no fraud zone is bigger than fraud zones and only the first one is a fraud. Apart from that from 2 to 5 there are areas situated outside as not fraud.

Also along with this visualizations shows how each models are different



In the linear kernel, we can see a straight line between classes (red and blue dots) . Here, we can see fraud from the other class (not fraud).

In the Polynomial kernel, we can see a curved line not a straight line. As we use degree 3 here. We can also say that some classes are mixed here.

In the RBF kernel, I think this one is flexible as we can see the irregular shoes here not like poly kernel and linear kernel. It seems adjustable and adaptable. We can also interpret that we have a small group of classes surrounded by another class because of the circle shape boundary.

By interpreting from the results and plots, this is how we can see the difference between kernels.

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

We can use the linear kernel when we want to separate our data by a straight boundary and when we have more features and simple. However, if our datasets seem to have non-linear patterns or if we want to capture those, the Polynomial model is better as it is not too simple and not too complex. But we need to be careful of the use of degrees. For RBF, we can use it in complex modeling when we can't decide how the boundary

line will be. If we also don't know the data structure we can also use RBF. One useful point of RBF is that it is really flexible.