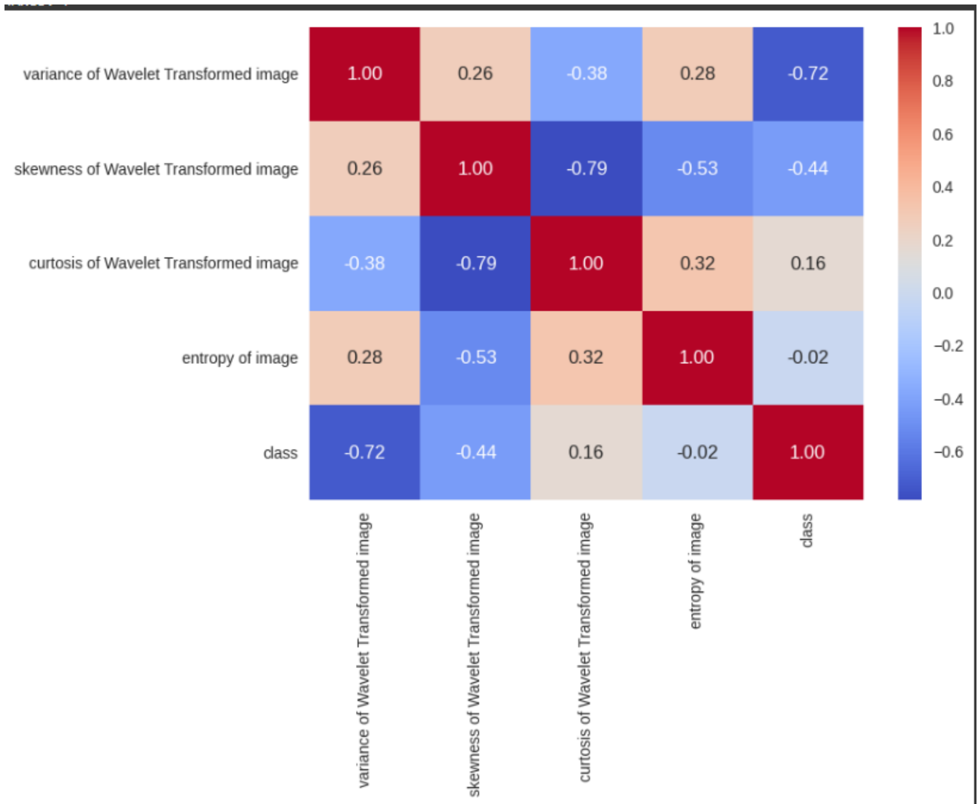
LAB Assignment 11
CSL 2050

By :- Akshat Jain B21CS005

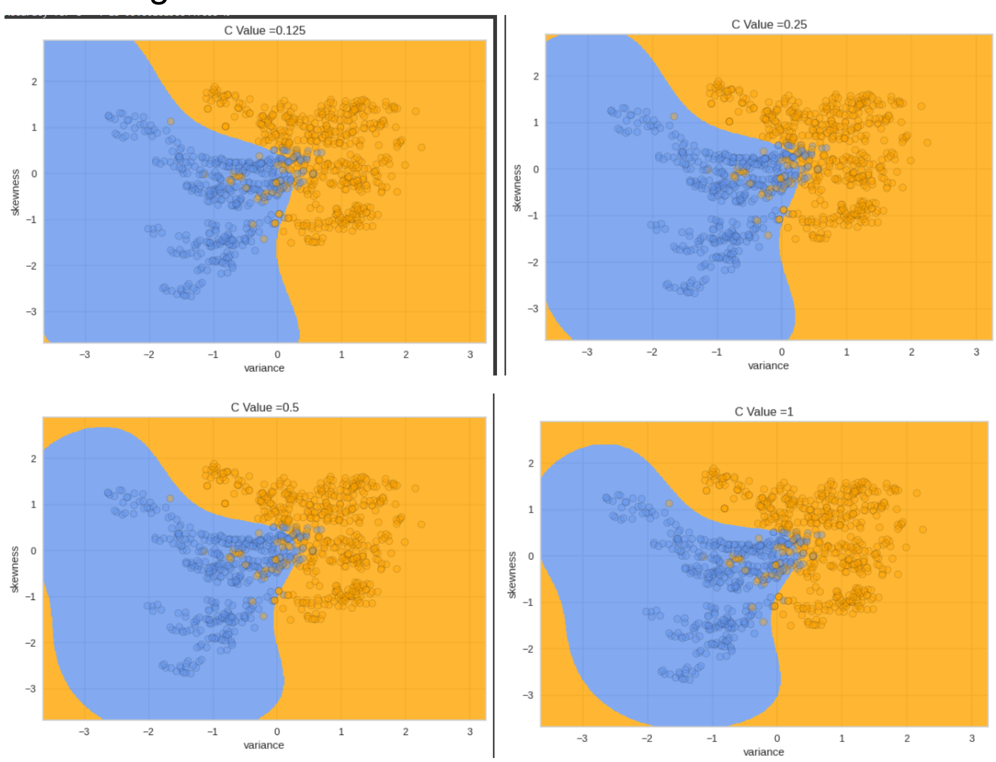
1. In this question, we must build an SVM model to identify fake notes based on several features extracted from bank notes' specimen pictures.
   1.1. We load the dataset as a pandas data frame. The names are taken from the dataset description given at the UCL machine learning repository. We perform data preprocessing, checking for NaN and Null values and then applying a StandardScaler to the data. We set our target variable to be 'class'. We then perform a 70:20:10 split for train-test-validation datasets.
   1.2. We then select our C-values as [0.125,0.25,0.5,1,2,4] and get the following results:-

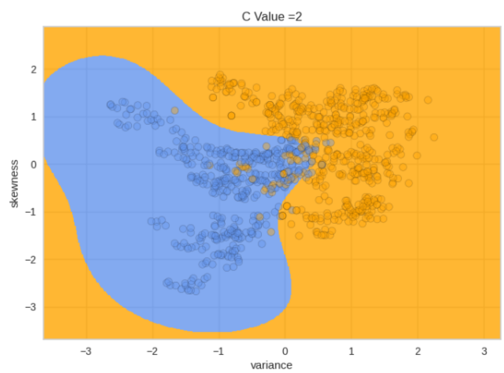
```
Accuracy for C = 0.125 is 99.27797833935018 %
Accuracy for C = 0.25 is 99.63898916967509 %
Accuracy for C = 0.5 is 99.63898916967509 %
Accuracy for C = 1 is 100.0 %
Accuracy for C = 2 is 100.0 %
Accuracy for C = 4 is 100.0 %
```

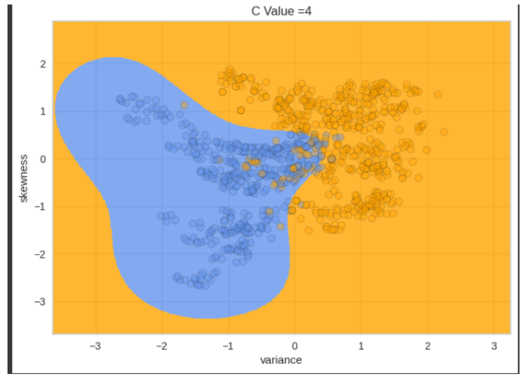
   For plotting the decision boundary, we first calculate the Correlation matrix with the target and find that the 'Variance' and 'Skewness' columns have the best absolute value for correlation with class.



We then get the decision boundaries for these 2 fetaures





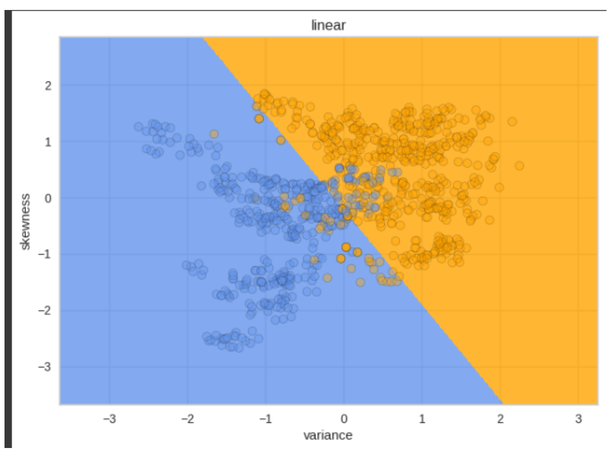


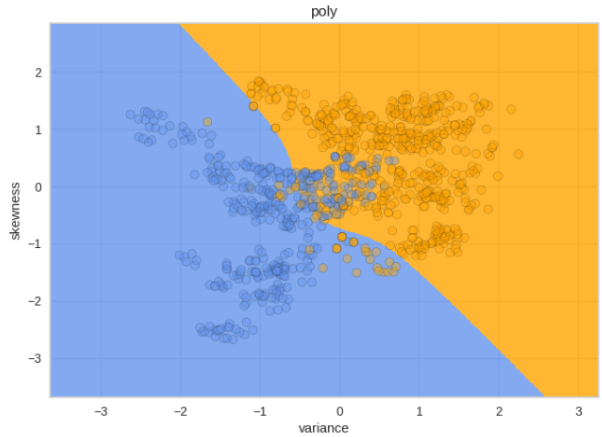
We also do the same with PCA by taking n_components = 2

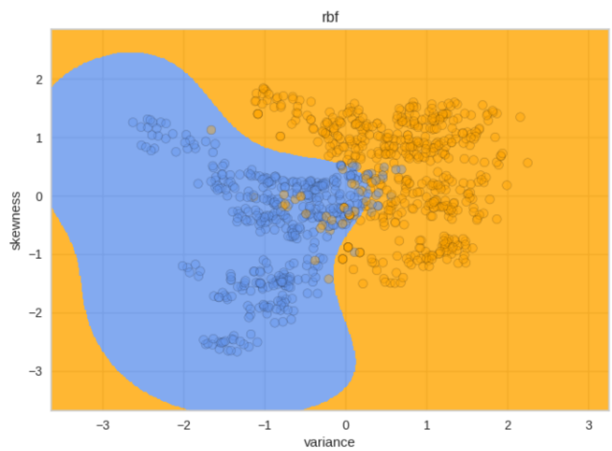
1.3. We then train the SVM classifier on the original data and see these results for all the kernel functions.

```
Accuracy for function = linear is 99.63898916967509 %
Accuracy for function = poly is 100.0 %
Accuracy for function = rbf is 100.0 %
```

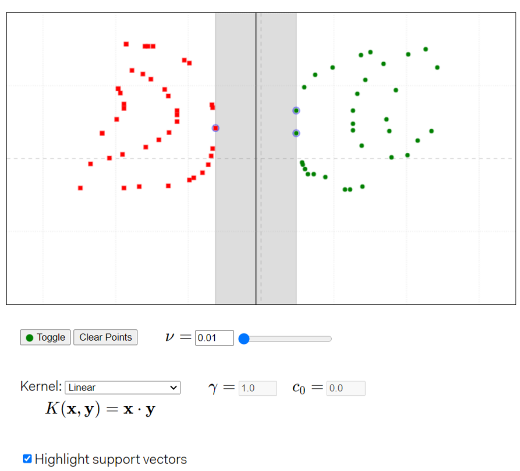
We see that the Accuracy for the ‘poly’ and the ‘rbf’ functions is coming to be 100%.
Here are the decision boundaries from the same, with two features being selected for training.

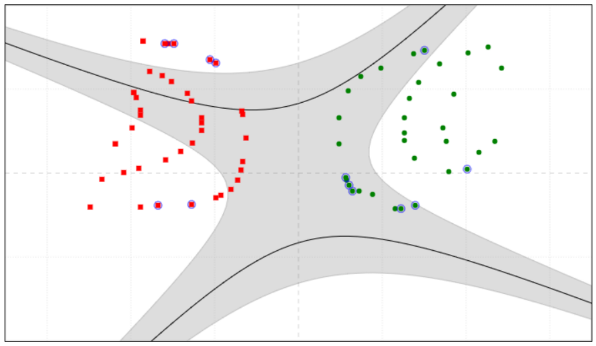






1.4. We make two datasets and get the following Observations



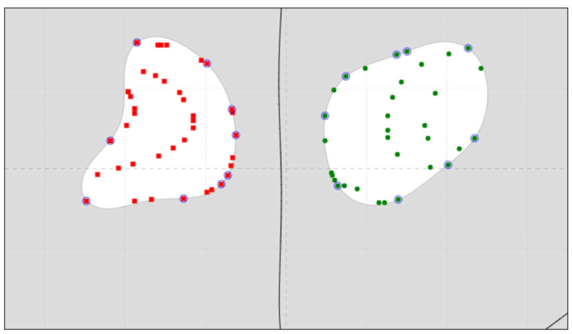


Toggle Clear Points $\nu =$ 0.06

Kernel: Quadratic $\gamma =$ 1.0 $c_0 =$ 0.0

$K(\mathbf{x}, \mathbf{y}) = (\gamma \mathbf{x} \cdot \mathbf{y} + c_0)^2$

Highlight support vectors

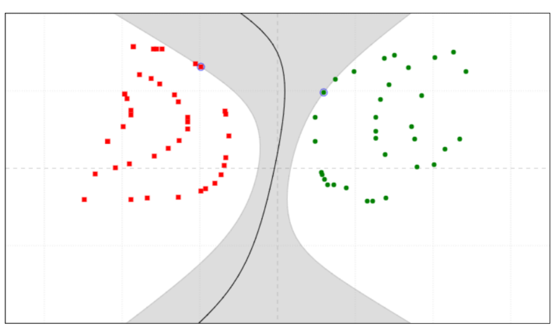


Toggle Clear Points $\nu =$ 0.01

Kernel: Radial Basis Functions $\gamma =$ 1.0 $c_0 =$ 0.0

$K(\mathbf{x}, \mathbf{y}) = \exp\left(-\gamma \|\mathbf{x} - \mathbf{y}\|^2\right)$

Highlight support vectors

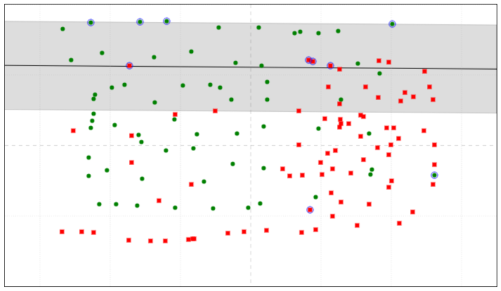


Toggle Clear Points $\nu =$ 0.01

Kernel: Sigmoid $\gamma =$ 1.0 $c_0 =$ 0.0

$K(\mathbf{x}, \mathbf{y}) = \tanh\left(\gamma \mathbf{x} \cdot \mathbf{y} + c_0\right)$

Highlight support vectors

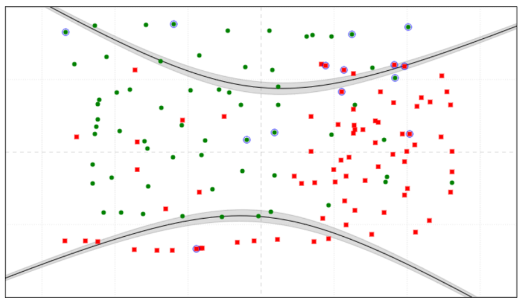


Toggle | Clear Points $\nu =$ 0.05

Kernel: Linear $\gamma =$ 1.0 $c_0 =$ 0.0

$K(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$

☑ Highlight support vectors

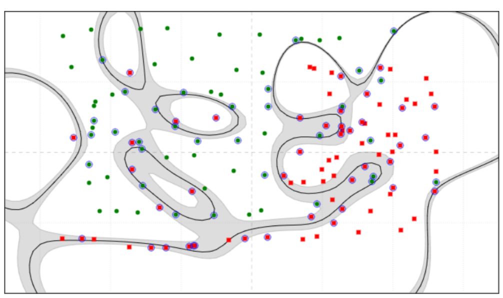


Toggle | Clear Points $\nu =$ 0.05

Kernel: Quadratic $\gamma =$ 1.0 $c_0 =$ 0.0

$K(\mathbf{x}, \mathbf{y}) = (\gamma \mathbf{x} \cdot \mathbf{y} + c_0)^2$

☑ Highlight support vectors

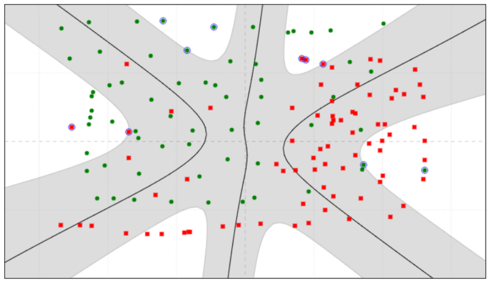


Toggle | Clear Points $\nu =$ 0.05

Kernel: Radial Basis Functions $\gamma =$ 1.0 $c_0 =$ 0.0

$K(\mathbf{x}, \mathbf{y}) = \exp\left(-\gamma \|\mathbf{x} - \mathbf{y}\|^2\right)$

☑ Highlight support vectors



Toggle Clear Points $\nu =$ 0.05

Kernel: Sigmoid $\gamma =$ 1.0 $c_0 =$ 0.0

$K(\mathbf{x}, \mathbf{y}) = \tanh(\gamma \mathbf{x} \cdot \mathbf{y} + c_0)$

Highlight support vectors