LAB 6 MACHINE LEARNING REPORT(B20MT012)

METHOD

In this assignment we were supposed to classify the target variable ‘credit’ by training the model as SVM .
I took the three features -”Experience”,”Income”,”CCaverage” and stored them in x .
I took credit as target and stored it in y
Then I plotted a 3-d scatter plot of my features.
Then I splitted the dataset into testing and training(80:20).
Then using the standard scalar scaled the features variable(x_train and x_test)
Made the model with different values of C and for each C printed the model score and confusion matrix .

RESULTS

C=0.0001

CONFUSION MATRIX

```
[[704   0]
 [296   0]]
```

CLASSIFICATION REPORT

```
classification report              precision    recall  f1-score
support

           0       0.70      1.00      0.83       704
           1       0.00      0.00      0.00       296

    accuracy                           0.70      1000
   macro avg       0.35      0.50      0.41      1000
weighted avg       0.50      0.70      0.58      1000
```

C=0.001

CONFUSION MATRIX

```
[[704   0]
 [296   0]]
```

CLASSIFICATION REPORT

```
classification report              precision    recall  f1-score
support

           0       0.70      1.00      0.83       704
           1       0.00      0.00      0.00       296

    accuracy                           0.70      1000
   macro avg       0.35      0.50      0.41      1000
weighted avg       0.50      0.70      0.58      1000
```

C=0.01

CONFUSION MATRIX

```
[[704   0]
 [296   0]]
```

CLASSIFICATION REPORT

```
classification report              precision    recall  f1-score
support

           0       0.70      1.00      0.83       704
           1       0.00      0.00      0.00       296

    accuracy                           0.70      1000
   macro avg       0.35      0.50      0.41      1000
weighted avg       0.50      0.70      0.58      1000
```

C=0.1

CONFUSION MATRIX

```
[[704   0]
 [296   0]]
```

CLASSIFICATION REPORT

```
classification report              precision    recall  f1-score
support

           0       0.70      1.00      0.83       704
           1       0.00      0.00      0.00       296

    accuracy                           0.70      1000
   macro avg       0.35      0.50      0.41      1000
weighted avg       0.50      0.70      0.58      1000
```

C=1

CONFUSION MATRIX

```
[[704   0]
 [296   0]]
```

CLASSIFICATION REPORT

```
classification report              precision    recall  f1-score
support

           0       0.70      1.00      0.83       704
           1       0.00      0.00      0.00       296

    accuracy                           0.70      1000
   macro avg       0.35      0.50      0.41      1000
weighted avg       0.50      0.70      0.58      1000
```

C=10

CONFUSION MATRIX

```
[[704   0]
 [296   0]]
```

CLASSIFICATION REPORT

```
classification report              precision    recall  f1-score
support

           0       0.70      1.00      0.83       704
           1       0.00      0.00      0.00       296

    accuracy                           0.70      1000
   macro avg       0.35      0.50      0.41      1000
weighted avg       0.50      0.70      0.58      1000
```

C=100

CONFUSION MATRIX

```
[[578 126]
 [249  47]]
```

CLASSIFICATION REPORT

```
classification report              precision    recall  f1-score
support

           0       0.70      0.82      0.76       704
           1       0.27      0.16      0.20       296
```

```
    accuracy                           0.62      1000
   macro avg       0.49      0.49      0.48      1000
weighted avg       0.57      0.62      0.59      1000
```

C=1000

CONFUSION MATRIX

```
[[491 213]
 [198  98]]
```

CLASSIFICATION REPORT

```
 classification report              precision    recall  f1-score
support

           0       0.71      0.70      0.70       704
           1       0.32      0.33      0.32       296

    accuracy                           0.59      1000
   macro avg       0.51      0.51      0.51      1000
weighted avg       0.59      0.59      0.59      1000
```

GRID BEST PARAMETER CAME OUT TO BE {'C': 0.0001}

OBSERVATIONS

From the given dataset when we observed the 3d plot between the features we can clearly see that the given data is not linearly separable hence we are not able to make a good SVM model to classify the data and get a good accuracy.

We are also getting warning of unable to converge due to this

The model was unable to classify the data.It is observed that the model classifies all the data as 0 except for C=100 and C=1000

As the value of C increased it did not make any effect from 0.0001 to 10 but as we made the value of C=100 and C=1000 the model predicted some of the 1's correctly but in these cases the precision was quite low.

It was able to predict some of the 1's correctly because as we increase the value of C the margin becomes narrow and it makes constraints hard to ignore due to which some of the prediction was done right in these cases.

https://colab.research.google.com/drive/1sxaZ5Bj5Z6uWR7YIaSZvQ8j_gEZ3T7ks?usp=sharing