

IOT Analytics

Project 5 SVM

1. Visualize Data

Figures 1 and 2 show the data in a 2D scatter diagram, before and after scaling the feature values in range $[0,1]$ respectively.

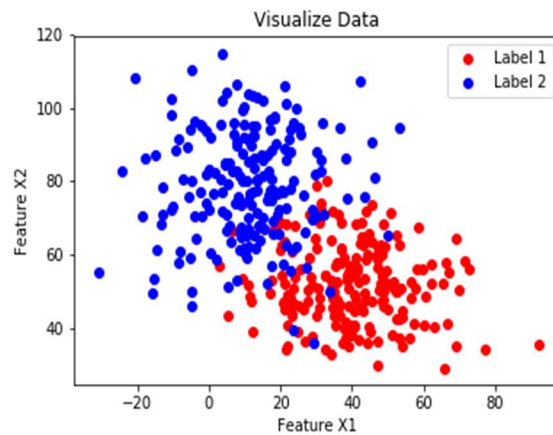


Fig. 1 Data

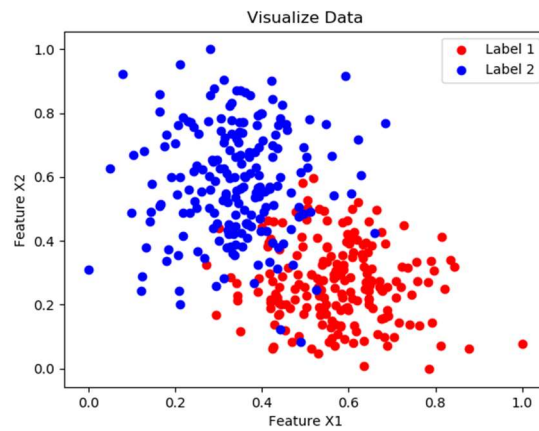


Fig. 2 Scaled Data

2. Cross-Validation and Grid Search

a. Coarse Grid Search

1. Range of C and γ values:

Values for C = 2^{-1} , 2^1 , 2^3 , 2^{17} , a total of 10 values.

Values for γ = 2^{-12} , 2^{-10} , 2^{-8} , 2^6 , a total of 10 values.

2. 3D plot:

Figure 3 shows the 3D plot to visualize the accuracy. Accuracy is plotted against the C and γ values. From the plot a better region (the darkest yellow region) was identified for fine grid search.

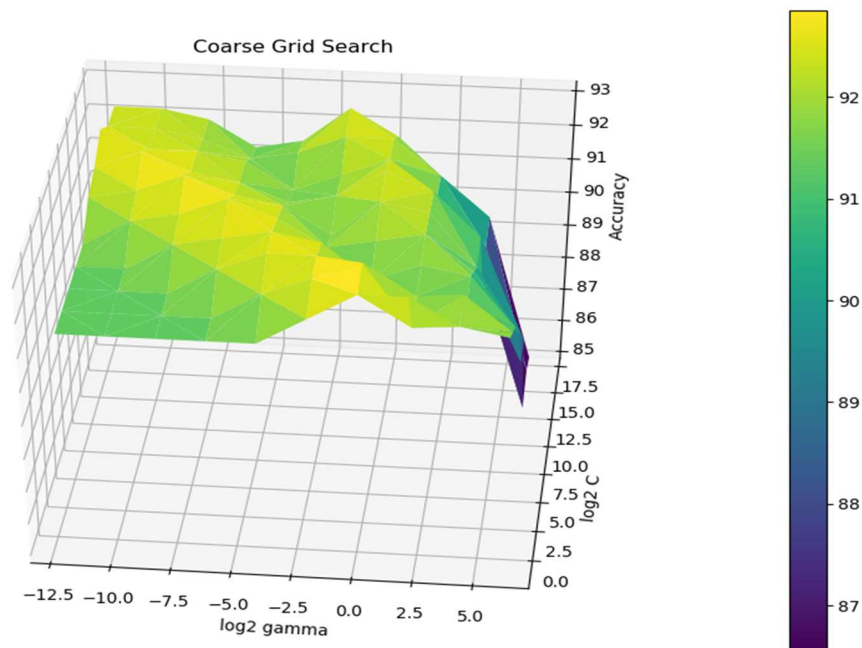


Fig. 3 Coarse Grid Search

3. Best values of C and γ

The maximum cross validation accuracy obtained: 93.112%

The best parameters:

C: 2.0

γ : 1.0

b. Fine Grid Search

1. Range of C and γ values:

Values for $C = 2^0, 2^{0.25}, 2^{0.5}, \dots, 2^{2.25}$, a total of 10 values.

Values for $\gamma = 2^{-1}, 2^{-0.75}, 2^{-0.5}, \dots, 2^{0.7}$, a total of 8 values.

2. 3D plot:

Figure 4 shows the 3D plot to visualize the accuracy. Accuracy is plotted against the C and γ values.

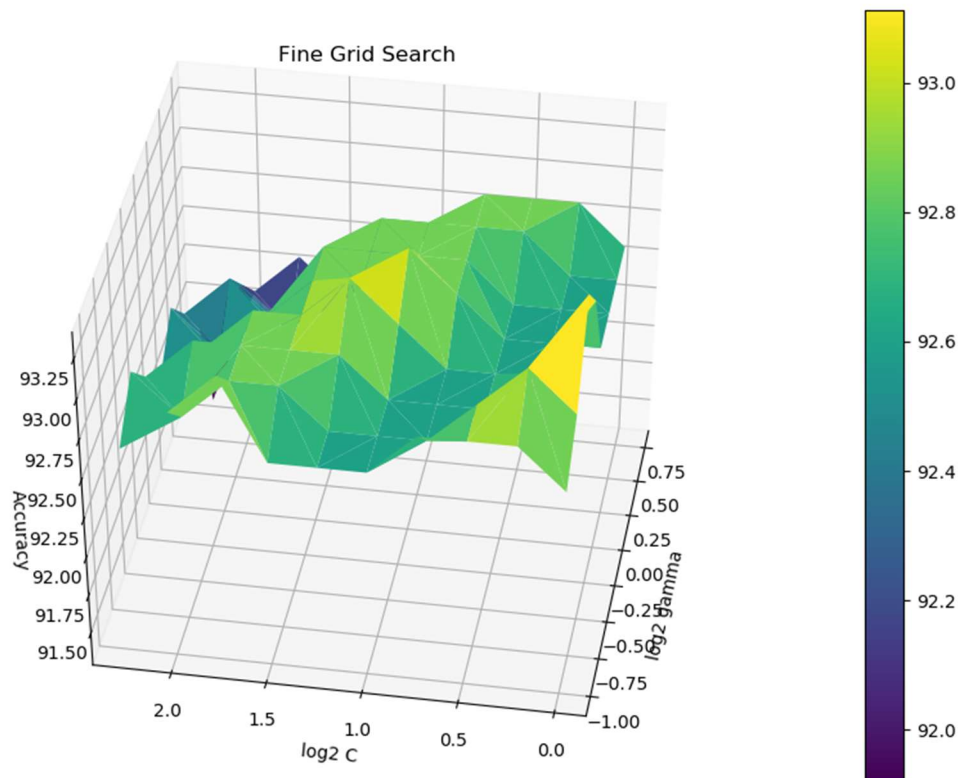


Fig. 4 Fine Grid Search

3. Best values of C and γ

The maximum cross validation accuracy obtained: 93.367%

The best parameters:

C: 1.0

γ : 0.707

3. Discussion

For the coarse grid search we obtain a cross validation accuracy of 93.112% for C and γ equal to 2 and 1 respectively. A finer grid search is conducted in the neighborhood of the best parameters obtained in coarse grid search and from visually recognizing the best region from the plot. In the fine grid search a better cross validation accuracy of 93.367% was obtained for C and γ equal to 1 and 0.707 respectively.