

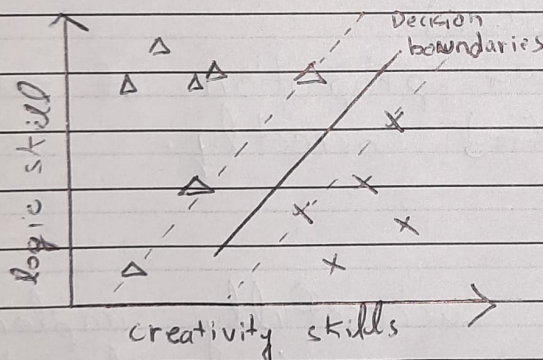


ML Exp No. 5

VUFI920014
Tejas Shinde
BE comp-A/A

Aim: To understand and implement the Support vector machine.

Theory: In the SVM algorithm, we plot each data item as point in n-dimensional space with the value of each feature being the value of particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates two classes very well.



The dataset "customer_purchases" has 3 columns Age, salary and Purchased. Here we have to predict if the given record have purchased or not.

First we used linear kernel i.e data points separated using single line-using SVC library. As it didn't fit the data very well. So we trained using Gaussian kernel.

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2}$$

It fits the data with 93% of accuracy. further ~~discuss~~ discussing on classification report.

64 data are predicted to be non-purchased and its label is actually non-purchased (True positive)

4 data are predicted to be non-purchased and its label is actually purchased. (False positive)

3 data are predicted to be purchased and its label is actually non-purchased. (False negative)

29 data are predicted to be purchased and its label is actually purchased. (True negative)

The proportion of positive prediction is actually 0.96 correct (Precision)

Recall trends to 0.94 proportion of actual positive was identified correctly.

Conclusion: Hence we successfully implemented support vector machine algorithm