SMAI (201501020)

# ASSIGNMENT 1

## Question 1

(b)

**--Ionosphere Dataset--**

**Accuracy of Voted Perceptron**

Epoch 5 -> 0.854841269841 +/- 0.0727200702474

Epoch 10 -> 0.863412698413 +/- 0.0645953741244

Epoch 15 -> 0.86626984127 +/- 0.0705421184555

Epoch 20 -> 0.860555555556 +/- 0.072545239619

Epoch 25 -> 0.854920634921 +/- 0.0801132984564

Epoch 30 -> 0.860476190476 +/- 0.0640365541867

Epoch 35 -> 0.851984126984 +/- 0.0786291274436

Epoch 40 -> 0.849206349206 +/- 0.0937986457314

Epoch 45 -> 0.846428571429 +/- 0.10326308782

Epoch 50 -> 0.849206349206 +/- 0.0937986457314

**Accuracy of Vanilla Perceptron**

Epoch 5 -> 0.823253968254 +/- 0.122492074805

Epoch 10 -> 0.823253968254 +/- 0.127712297133

Epoch 15 -> 0.828968253968 +/- 0.149333363703

Epoch 20 -> 0.823333333333 +/- 0.144212617033

Epoch 25 -> 0.820476190476 +/- 0.147052676942

Epoch 30 -> 0.812063492063 +/- 0.153140961846

Epoch 35 -> 0.811825396825 +/- 0.145531319926

Epoch 40 -> 0.820714285714 +/- 0.120957067776

Epoch 45 -> 0.826111111111 +/- 0.124176196959

Epoch 50 -> 0.846031746032 +/- 0.126132061952

**--Breast Cancer Dataset--**

**Accuracy of Voted Perceptron**

Epoch 5 -> 0.957060041408 +/- 0.0405260980198

Epoch 10 -> 0.962774327122 +/- 0.0389299144358

Epoch 15 -> 0.967101449275 +/- 0.0314069598933

Epoch 20 -> 0.969958592133 +/- 0.0298225487108

Epoch 25 -> 0.971387163561 +/- 0.0285725087562

Epoch 30 -> 0.972836438923 +/- 0.029793989367

Epoch 35 -> 0.974265010352 +/- 0.0279606146178

Epoch 40 -> 0.974265010352 +/- 0.0279606146178

Epoch 45 -> 0.974265010352 +/- 0.0279606146178

Epoch 50 -> 0.972836438923 +/- 0.029793989367

**Accuracy of Vanilla Perceptron**

Epoch 5 -> 0.945672877847 +/- 0.0456216993819

Epoch 10 -> 0.961407867495 +/- 0.0313647581688

Epoch 15 -> 0.961407867495 +/- 0.0338675856995

Epoch 20 -> 0.958530020704 +/- 0.0392257539625

Epoch 25 -> 0.948530020704 +/- 0.0463421195295

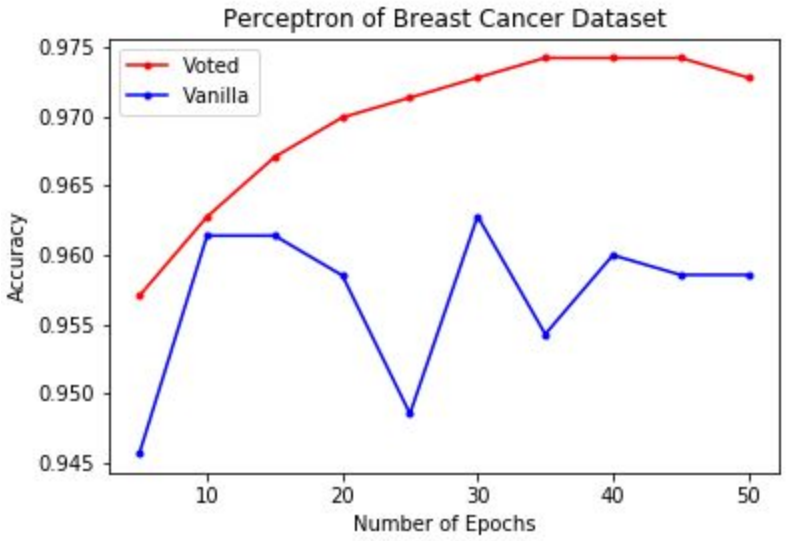
Epoch 30 -> 0.96281573499 +/- 0.0227963003353

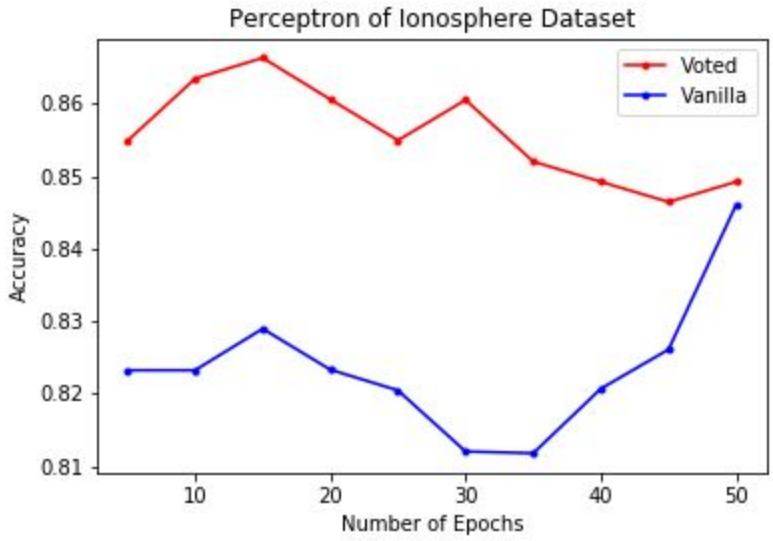
Epoch 35 -> 0.954285714286 +/- 0.0438922614164

Epoch 40 -> 0.96 +/- 0.0419912527334

Epoch 45 -> 0.958571428571 +/- 0.0392792202425

Epoch 50 -> 0.958571428571 +/- 0.0468606270482




d)

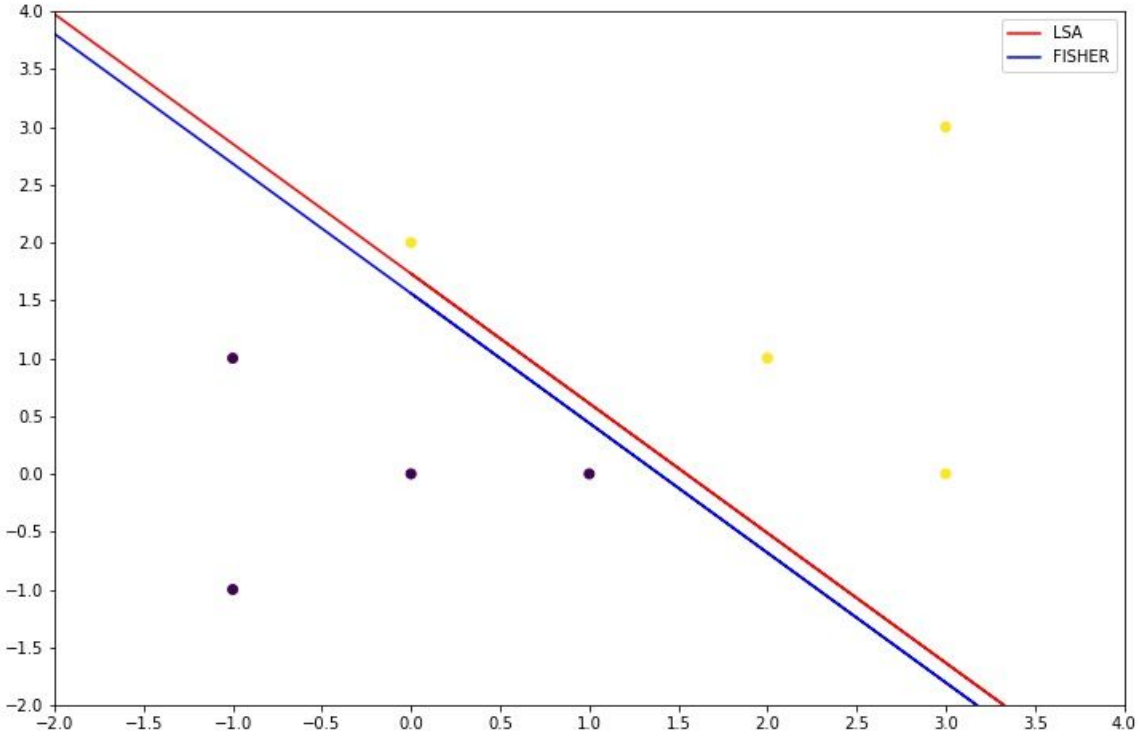
In terms of accuracy, Performance of Voted Perceptron is better than Vanilla Perceptron, and it is easy to generalize in voted perceptron than vanilla perceptron, as we increase number of epochs, deviation in vanilla is more, that makes accuracy of voted Perceptron more generalized.

In terms of speed, Vanilla Perceptron works better because If there are 1000 updates made during perceptron learning, the voted perceptron requires that you store 1000 weight vectors,

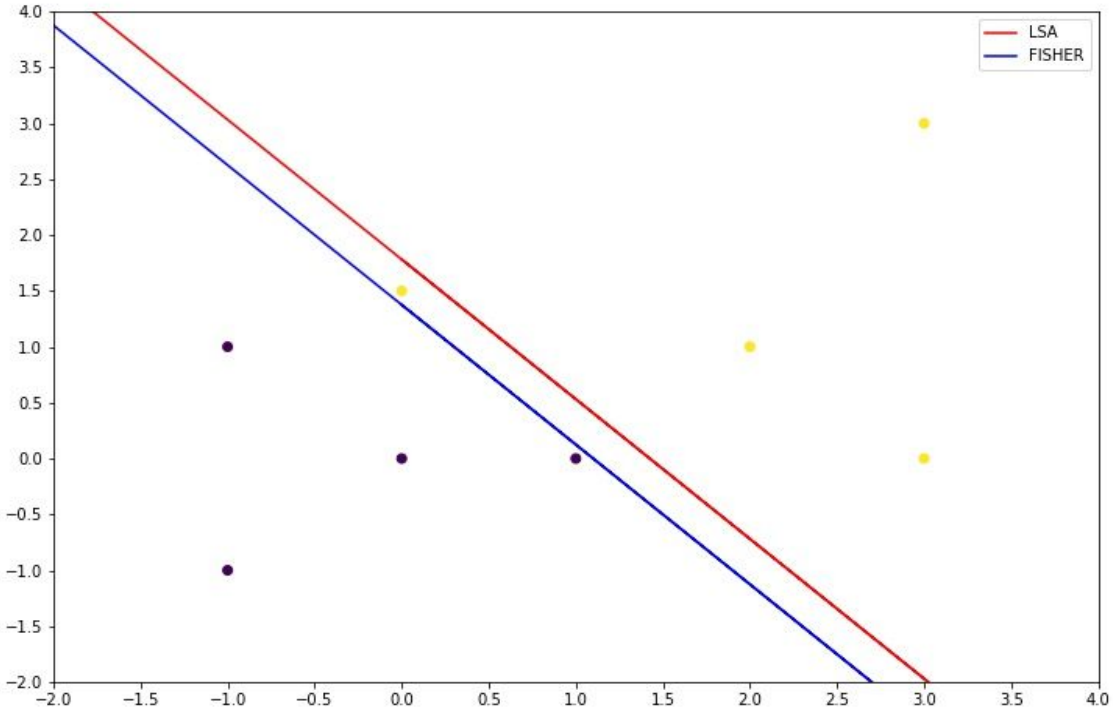
together with their counts. This requires an absurd amount of storage, and makes prediction 1000 times slower than the vanilla perceptron

# Question 2

3)



4)




5) Although slope of fisher and LSM are same in both, they are not classifying the points by same equation. That is happening due to the variation of threshold in case of least square method. In fisher, we find the threshold by projecting the points on 1 dimensional space on w find by fisher LDA approach and by passing through the middle of 2 closest points in different classes, Thus fisher takes cares of the expreme point by moving parallel to separate closest points.

# Question 3

**Approach 1 (Perceptron) for 3588 features:**

Epoch 5

Training Accuracy->  0.998579545455

Testing Accuracy-> 0.942613636364 +/- 0.0257631455654

Epoch 10:

Training Accuracy->  1.0

Testing Accuracy -> 0.946022727273 +/- 0.0160706086633

Epoch 15:

Training Accuracy->  1.0

Testing Accuracy -> 0.945454545455 +/- 0.0170075335762

Epoch 20:

Training Accuracy->  1.0

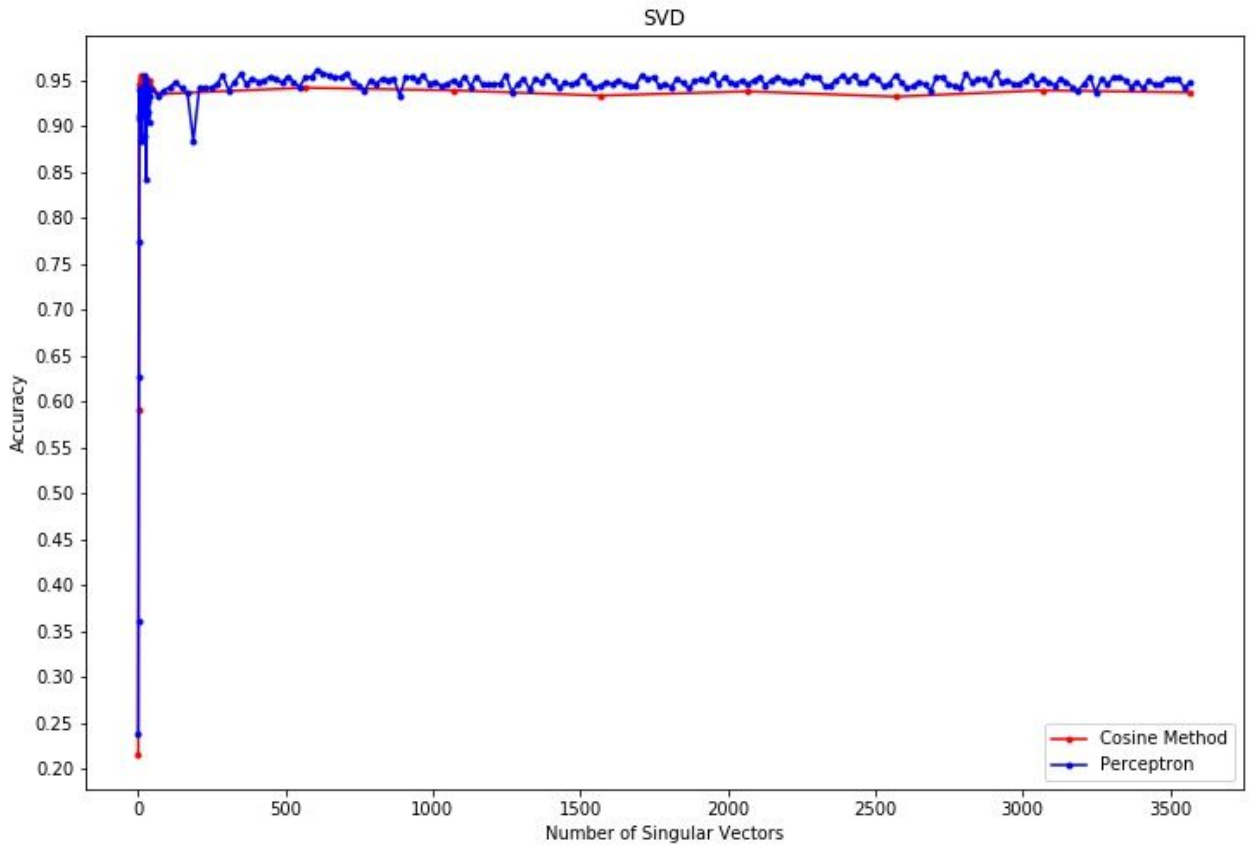
Testing Accuracy -> 0.945454545455 +/- 0.0170075335762

**Approach 2(cosine similarity):**

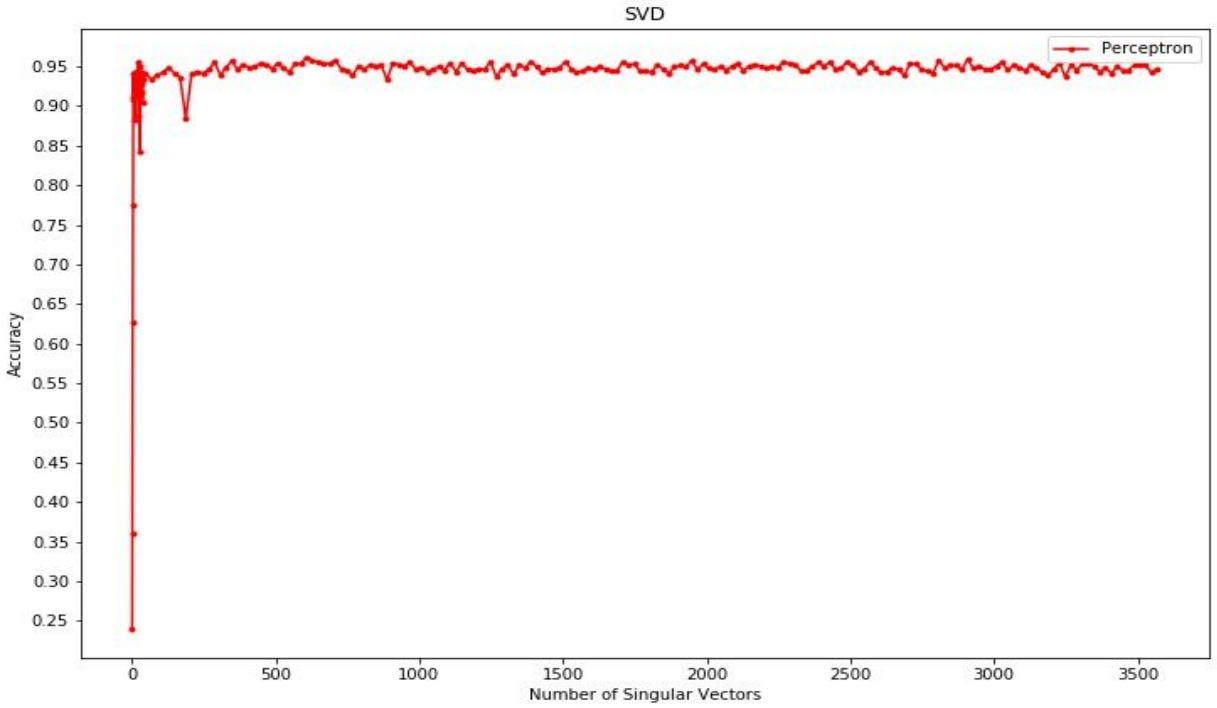
features -> accuracy

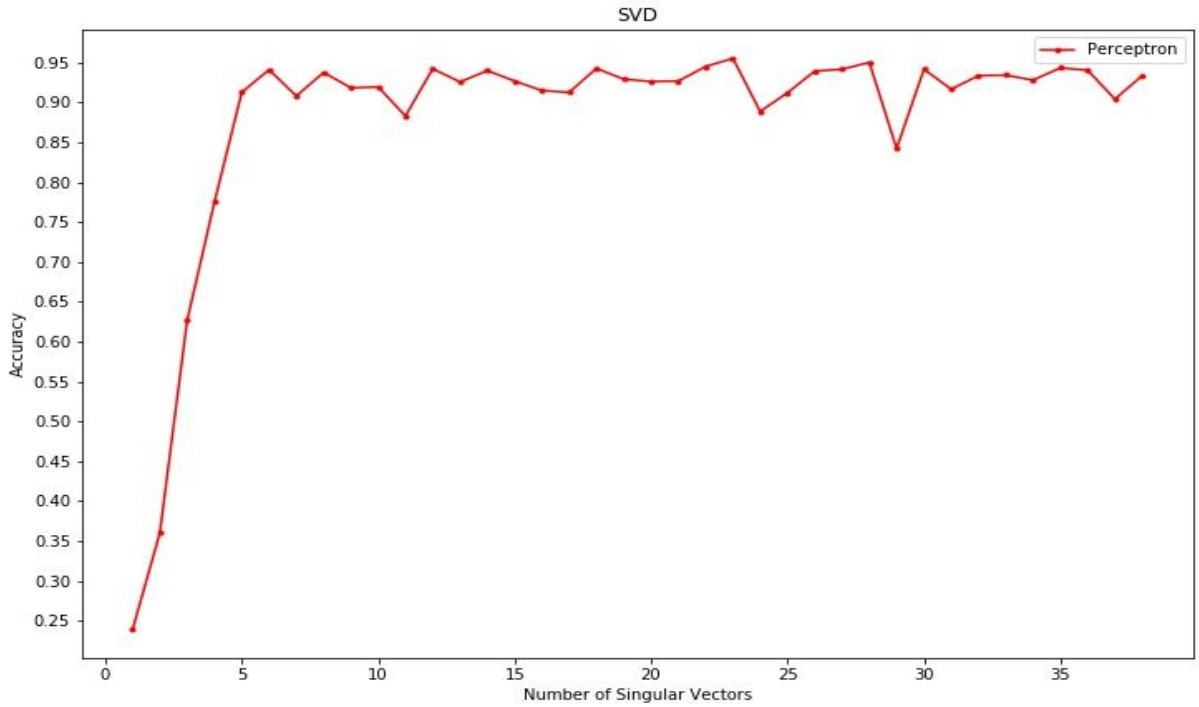
3568 -> 0.93693181818181814,
568 -> 0.94204545454545452

**Question 4:**

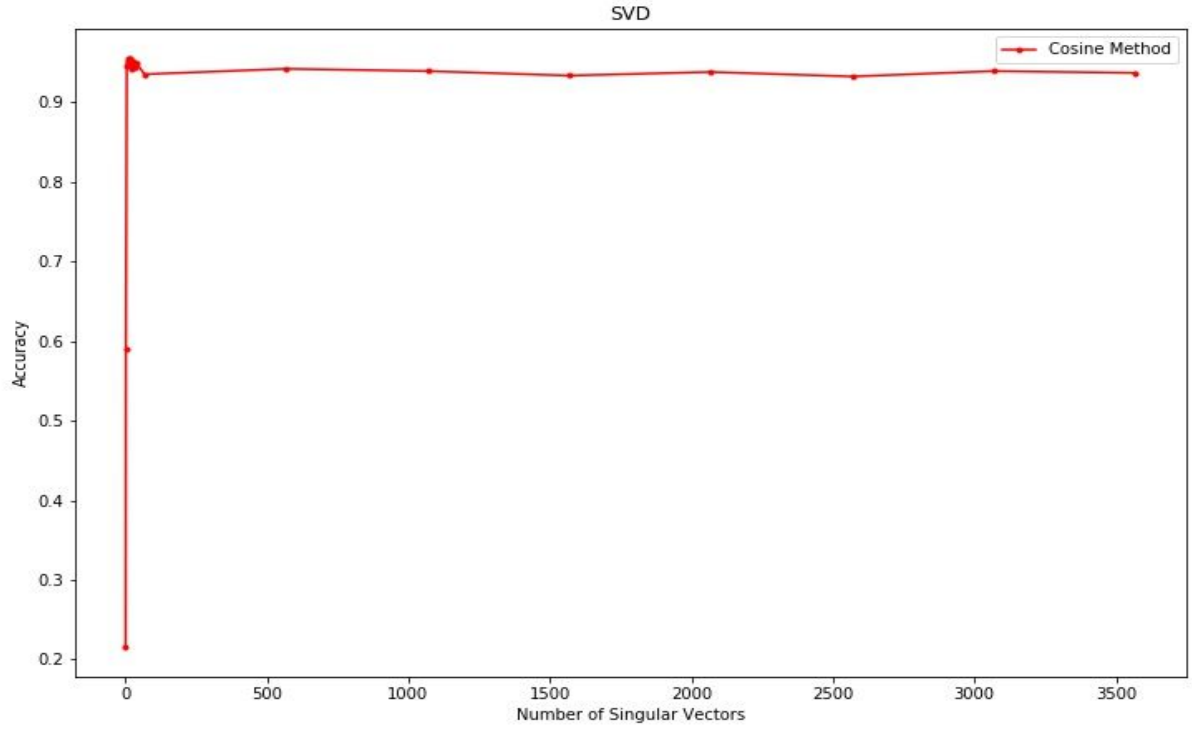


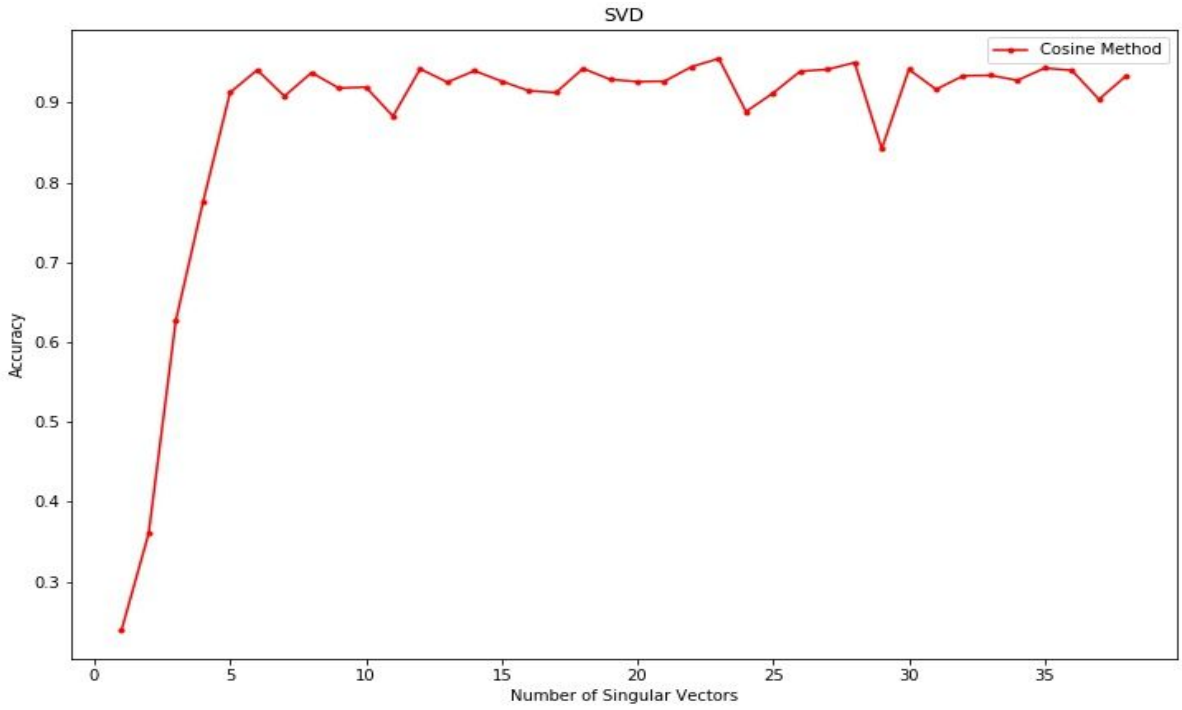

Approach 1(Perceptron) is performing better than cosine similarity approach by little margin. Performance in terms of speed is also better in perceptron than cosine. Initially there were around 17000 words that were reduced to 3500 words by removing those words which are not occuring in lot of documents say(0.5%). It helps in computing SVD.




For Perceptron Optimal will be around 5-10 features ( highest is around 7 for this experiment)




For Cosine Similarity: Optimal will be around between 6-10.