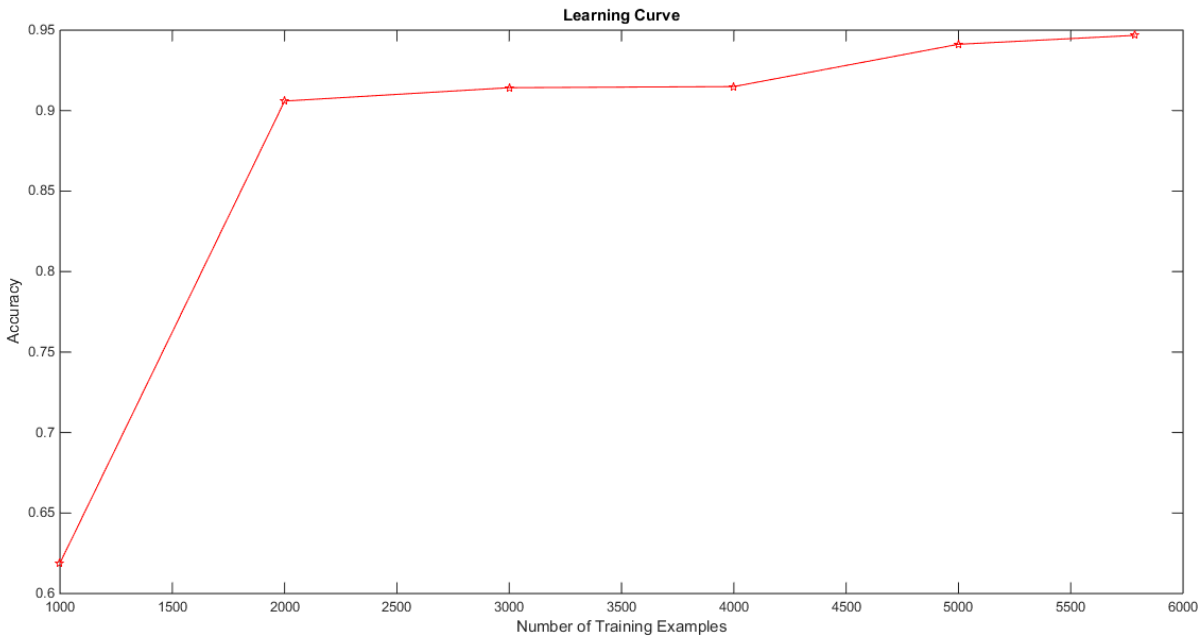
COL 774 : Assignment 3

Rahul Patidar - 2012CS10244 27th April, 2015

---

1. Newsgroup Classification
    a. Average Accuracy = 0.9586
    b. Average Accuracy on Random Prediction = 0.1124
    c. It won't affect the Naïve Bayes learner. But error will be higher with more cross-posted data.
    d. As the number of training examples increases accuracy also got increases. But the rate of increases is less after certain number of iterations. Accuracy get above 90% after 1000 iterations it shows that it



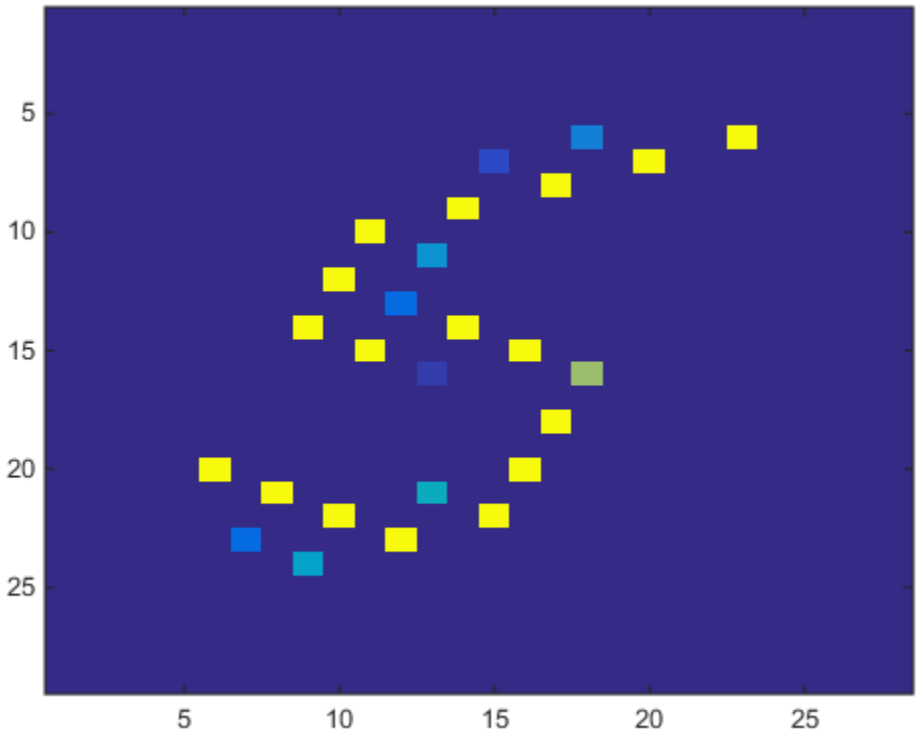
learns faster in first 1000 iterations and then learning rate decreases.

e. Where 1 : rec.autos 2 : rec.motorcycles
3 : rec.sport.baseball 4 : rec.sport.hockey
5 : talk.politics.guns 6 : talk.politics.misc
7 : talk.politics.misc 8 : talk.religion.misc

rec.autos has highest diagonal entry in confusion matrix
talk.politics.misc and **talk.politics.guns** are most confusing with each other.

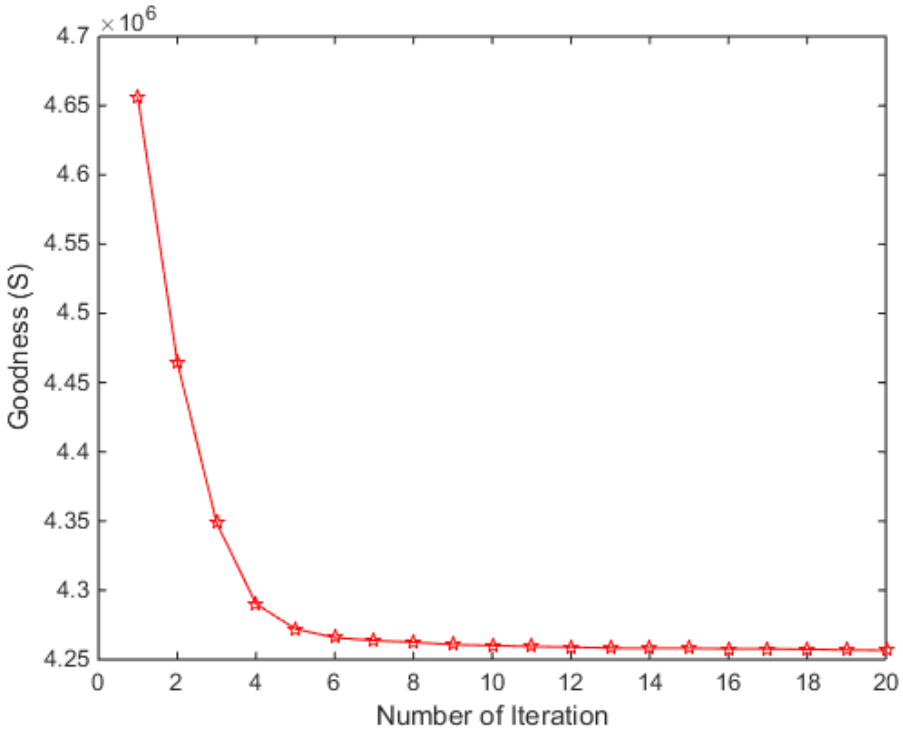
	1	2	3	4	5	6	7	8
1	216	3	0	0	0	0	4	0
2	4	205	0	0	1	0	0	0
3	1	0	210	4	1	0	1	0
4	1	2	4	196	0	0	1	0
5	0	1	0	0	174	1	3	1
6	1	1	1	1	2	169	4	3
7	0	0	0	0	11	1	141	2
8	0	1	0	0	3	2	2	67

2. K – Means for Digit Recognition
    a. Digit image of example 1

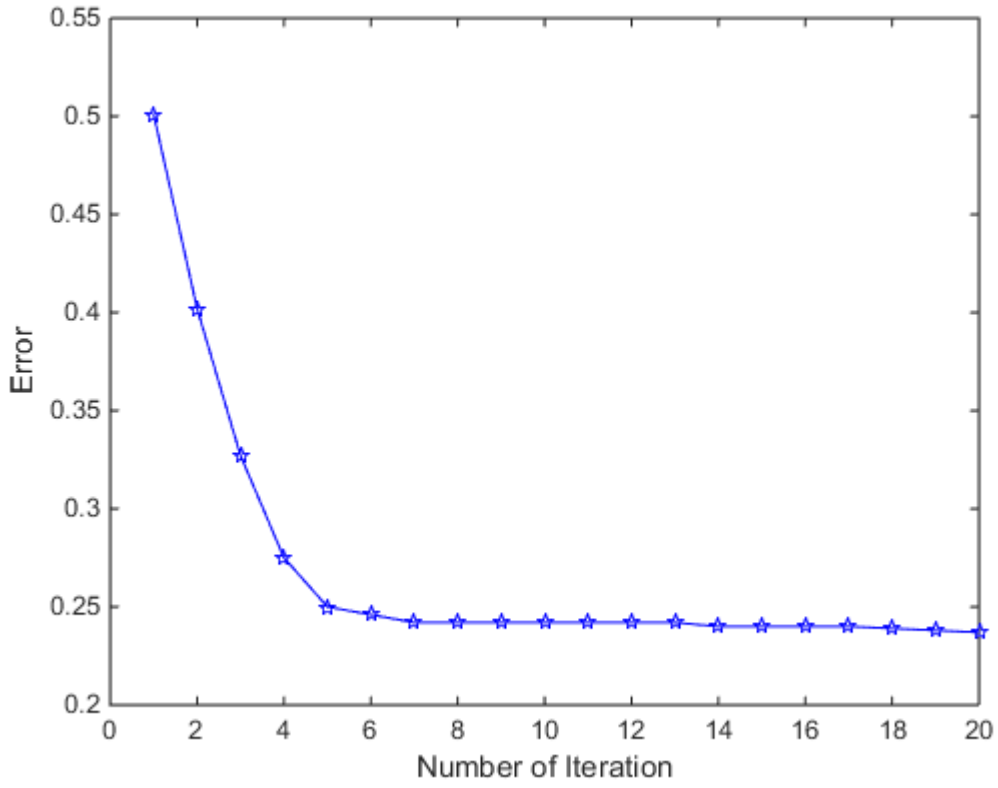


b.

c. As the number of iterations increases closeness between the points decreases. Which shows the better clustering.



d.



As the number of training example increases error term decreases but the rate of error is less after particular number of iterations. Which show the as the number of interation increases the clustering goes better and better.

3. Not Done
4. Principle Component Analysis
    a. Average Face

    b. 50 Principle Components are stored in “PC_50.txt” .
    c. Top five eigenfaces

(1)

(2)

(3)

(4)

(5)

d. Face image of training example 7

Original Image

Projected Image

too most extent projected image can easily predict the original image and both of them are more or less same

e. Real Image and Projected Image of Hands

Projected image predict the original hand image to very less extent. As it is not trained with similar training examples