Cover page for answers.pdf CSE512 Fall 2019 - Machine Learning - Homework 4

Your Name: Shubham S Bagi

Solar ID: 112672171

NetID : sbagi

email address: shubham.bagi@stonybrook.edu

Names of people whom you discussed the homework with: Gurusangama

Homework-4

Shubham . Bagi
112672171

(1). SVM Equation

$$y = w^T x + b$$

SVM dual equation and lagrengian equation

$$w = \sum \alpha y z$$

$$y = \sum_{i=1}^{n} \alpha_i y_i x + b$$

The pred which we have takes value ~~from~~ 1 and -1

$$pred = \begin{cases} 1 & \text{if } y \geq 0 \\ -1 & \text{if } y \leq 0 \end{cases}$$

LOOCV error is given by

$$\frac{1}{n} \sum_{i=1}^{n} \delta(y, pred).$$

There are 2 cases:-

Case 1 ⇒ When you do not pick any support vector for LOOCV, $\alpha = 0$, hence no effect on error

Case 2 ⇒. When any support vector is chosen, $\alpha > 0$ and lets assume that they are misclassified, then maximum error of LOOCV is given by.

$$LOOCV = \frac{\text{Number of support vector}}{n}$$

$$= \frac{m}{n} \quad \Rightarrow \text{This is maximum error}$$

(2). Now, lets consider the general kernel.

The bound in previous section will hold.

The kernel :- $K(x, y) = \phi(x)^T \phi(y)$.

Hence the SVM equation will be

$$y = w^T \phi(x) + b$$

also $y = \sum_{i=1}^{n} \alpha_i y^i K(x^i, x) + b$.

Now the prediction function of general SVM is.

$$f(\phi(x) : w, b) = \begin{cases} 1 & \text{iff } y > 0 \\ -1 & \text{otherwise} \end{cases}$$

LOOCV error = $\frac{1}{n} \sum_{i=1}^{n} \delta(y^i, f(\phi(x); w^{-i} b^{-i}))$

Again there are 2 cases to consider.

(1). When you do not pick any support vector for LOOCV, it lies outside the margin and will be classified correctly. Hence $\alpha = 0$. and will not affect the eqn.

(2). When you remove a support vector data point. The point lies on margin, and may be classified wrong. Hence $\alpha > 0$ & will affect eqn.

So, when we consider case (2), all 'm' support vectors are classified wrong.

Hence this will lead to LOOCV error of $\frac{m}{n}$ //.

2. – Implementation of SVMs

2.1 . We have

n = number of samples

d = number of features

x – input(d,n)

y – output(n,1)

H = (transpose(x)*x).*(y*transpose(y))

f = -1 * ones(n,1)

A = []

b = []

Aeq = transpose(y)

beq = 0

lb = zeros(n,1)

ub = C * ones(n,1)

2.4

For C = 0.1

Objective Function Value - 24.7648

Number of Support Vectors - 213

Accuracy Score - 0.5204

Confusion Matrix

184	176
0	7

2.5

For C = 10

Objective Function Value =112.1461

Number of Support Vectors - 129

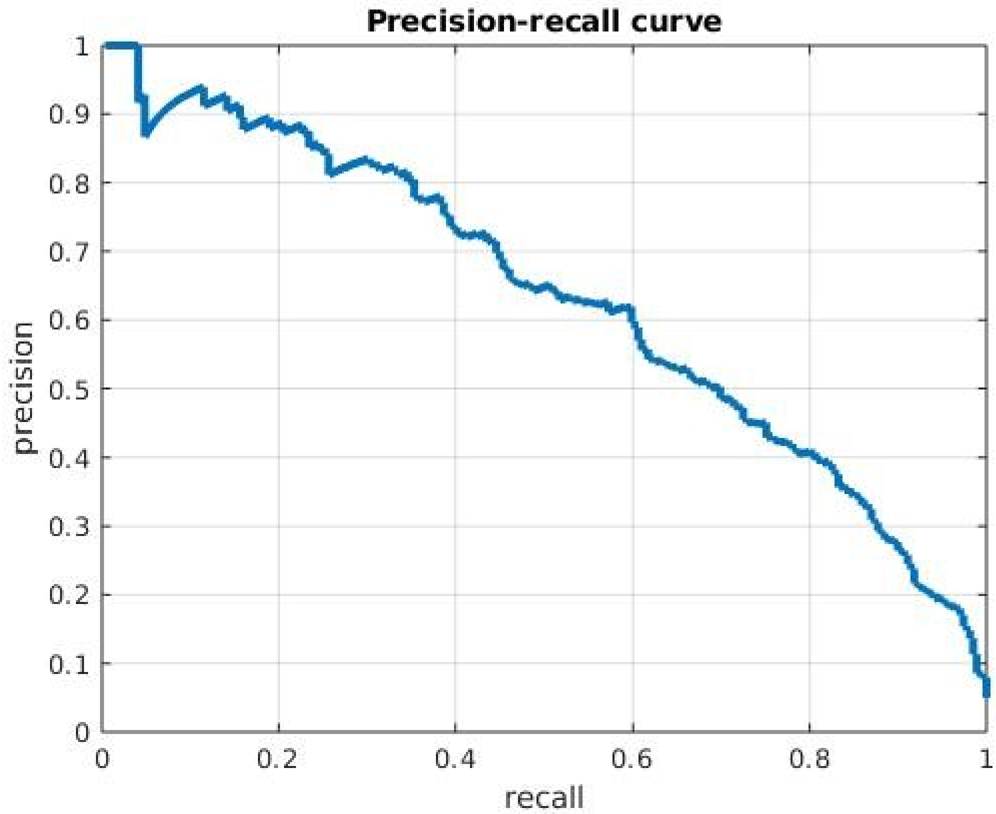
Accuracy Score - 0.9782

Confusion Matrix

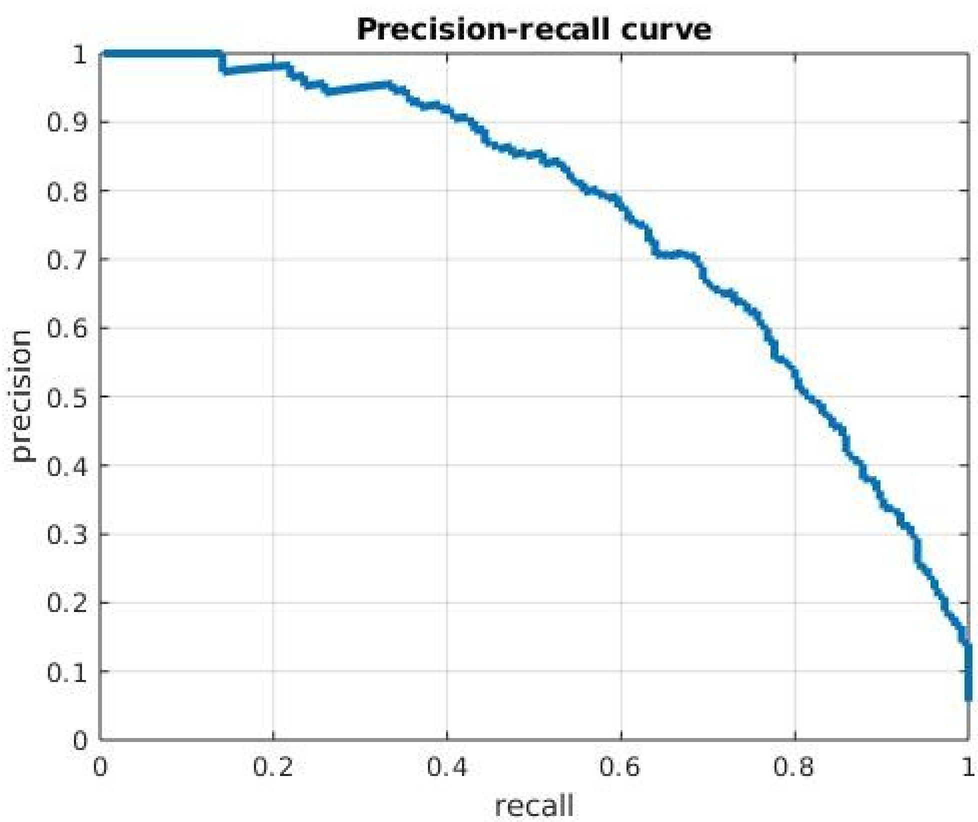
180	4
4	179

3. – SVM for object detection

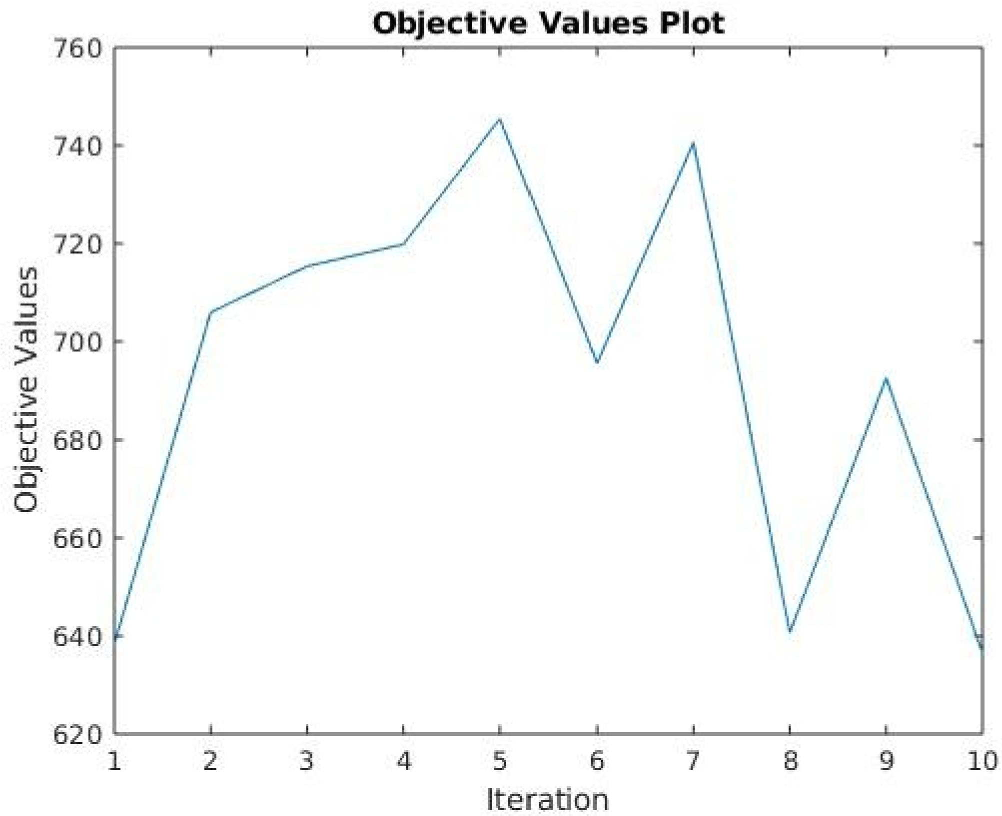
3.4.1. POC Curve- **AP value** – 0.6362



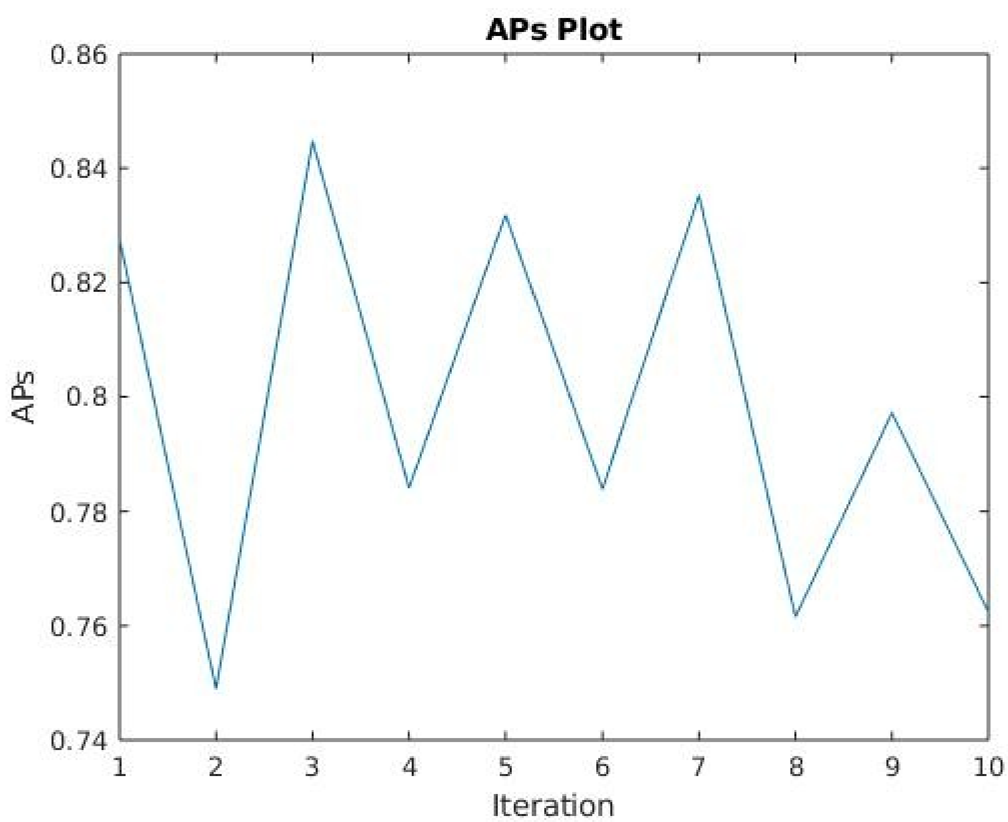
3.4.3
Graph:
POC Curve



Objective Function Value



Average Precision Curve



Objective Values:
640.2453
704.2202
710.0602
719.4210
745.6221
691.3503
738.4119
647.0789
692.7320
636.2466

Average Precision Values:
0.8258
0.7524
0.8509
0.7632
0.8376
0.7850
0.8459
0.7646
0.7998
0.7762