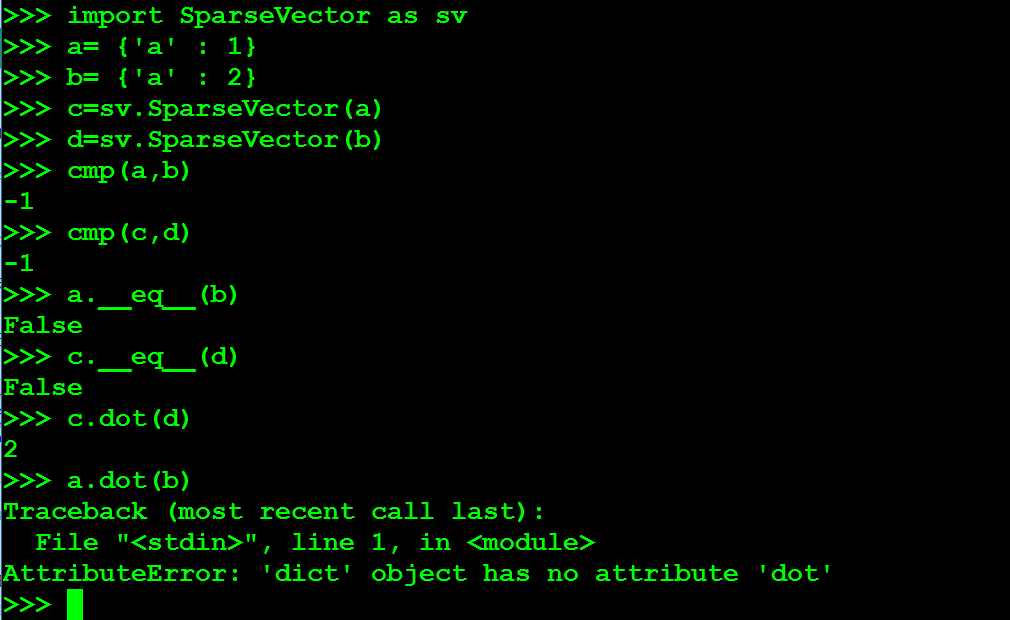
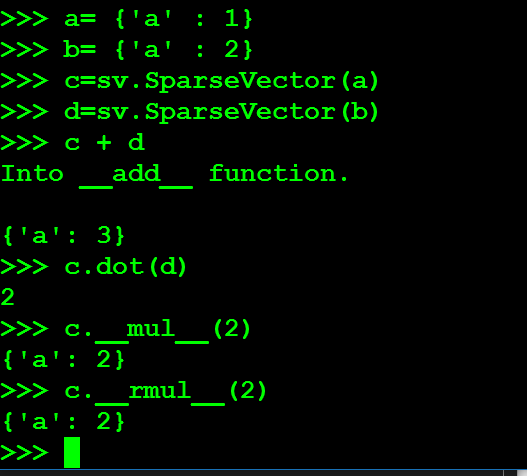
Q1

1a) The SparseVector is like a dictionary expect that it doesn’t show any non-value key.

1b) SparseVector inherit all the method in dict but dict doesn’t share the methods which are built in SparseVector

 The example is shown below:

1c)

Q2

2a)

and since

Then we have

So, the function is convex

2b) as we proved on the above

2c)

When x = 0,

So, we know that

When x = 0

So, we know that

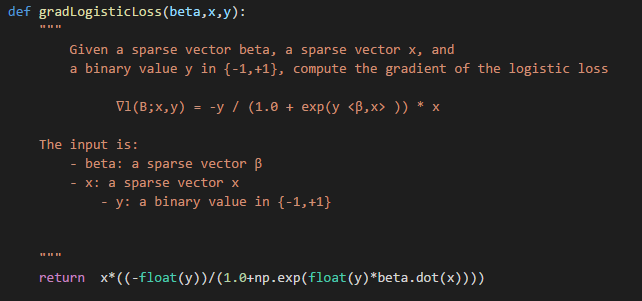
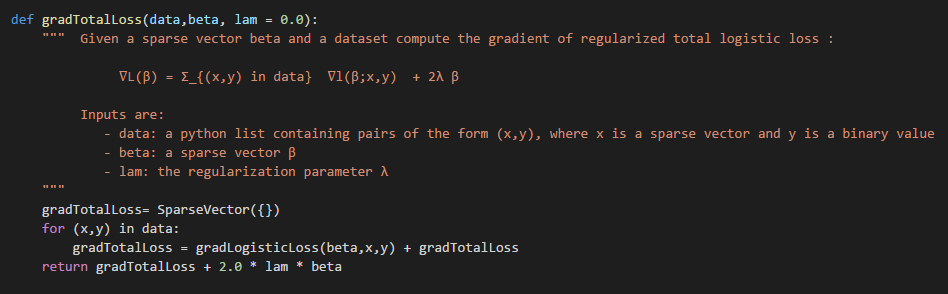
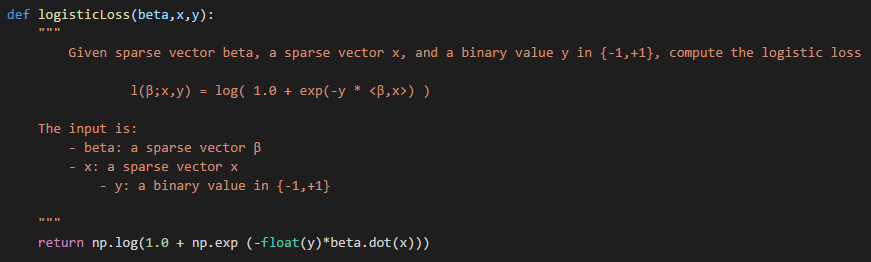
2d)

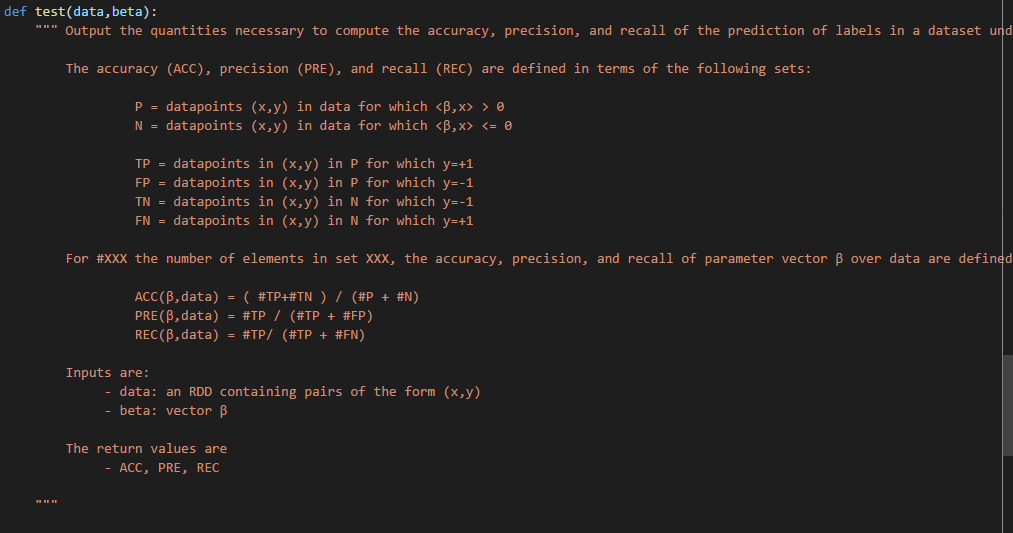
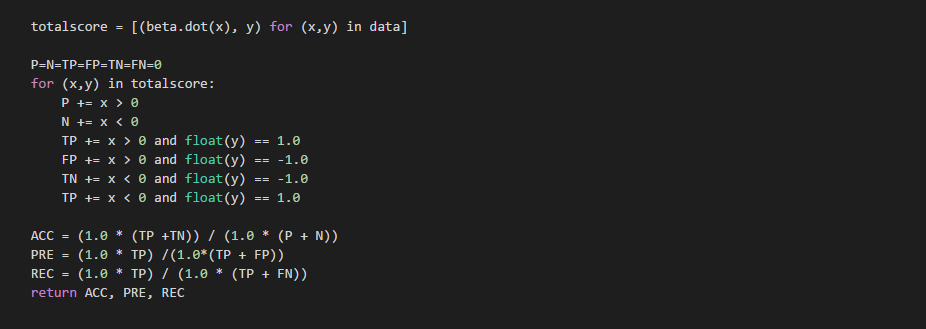
Since is convex so is convex.

Let . Then and = which is a constant

So is convex

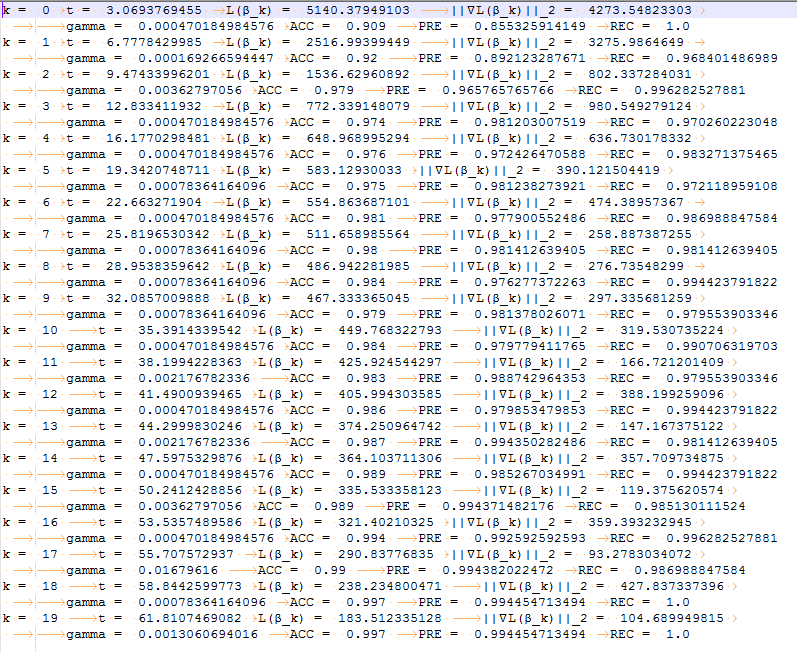
Q3

3a)

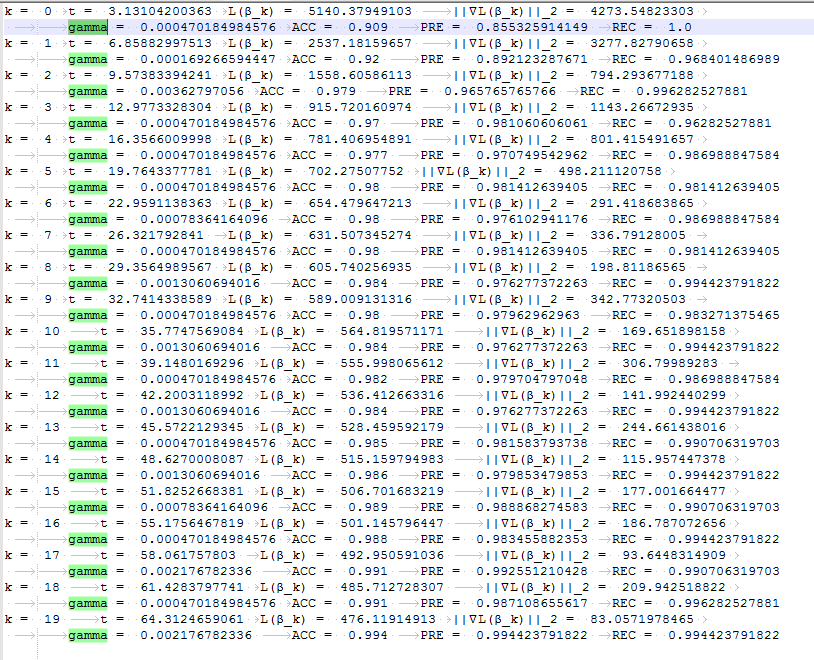
3b)

3c)

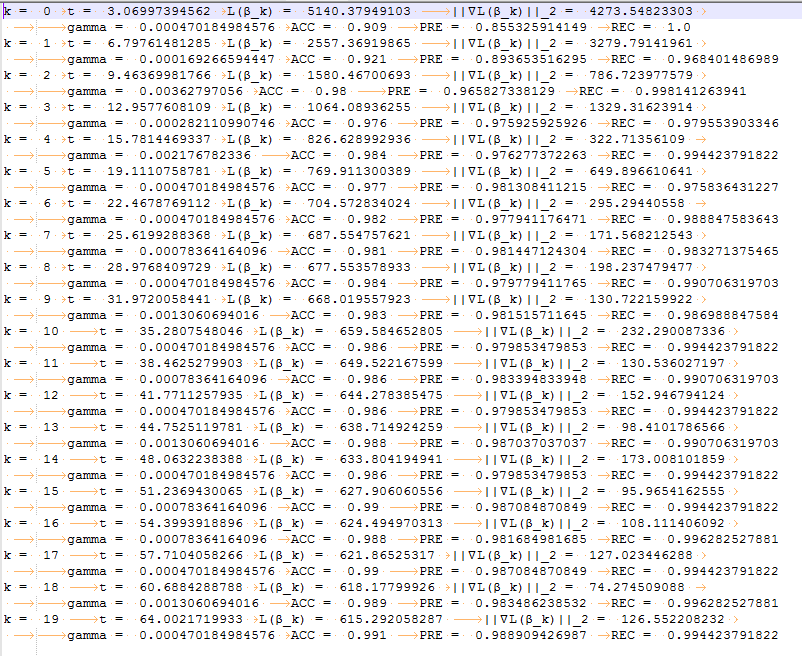
1. lambda = 0



1. lambda = 5



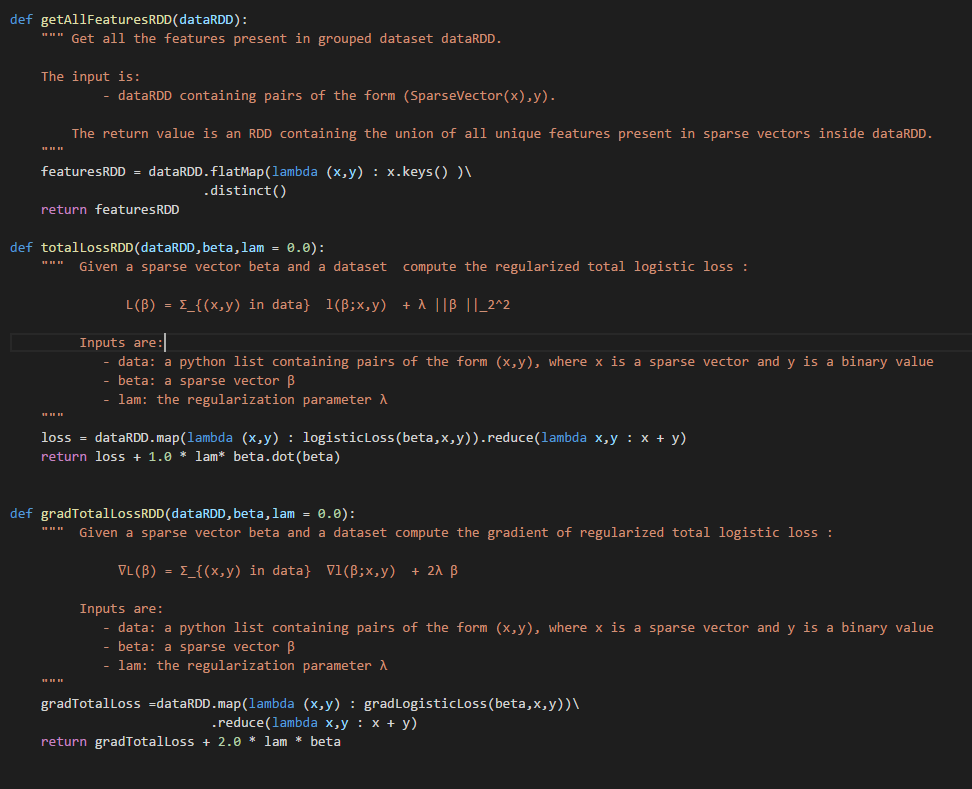
3.lambda = 10

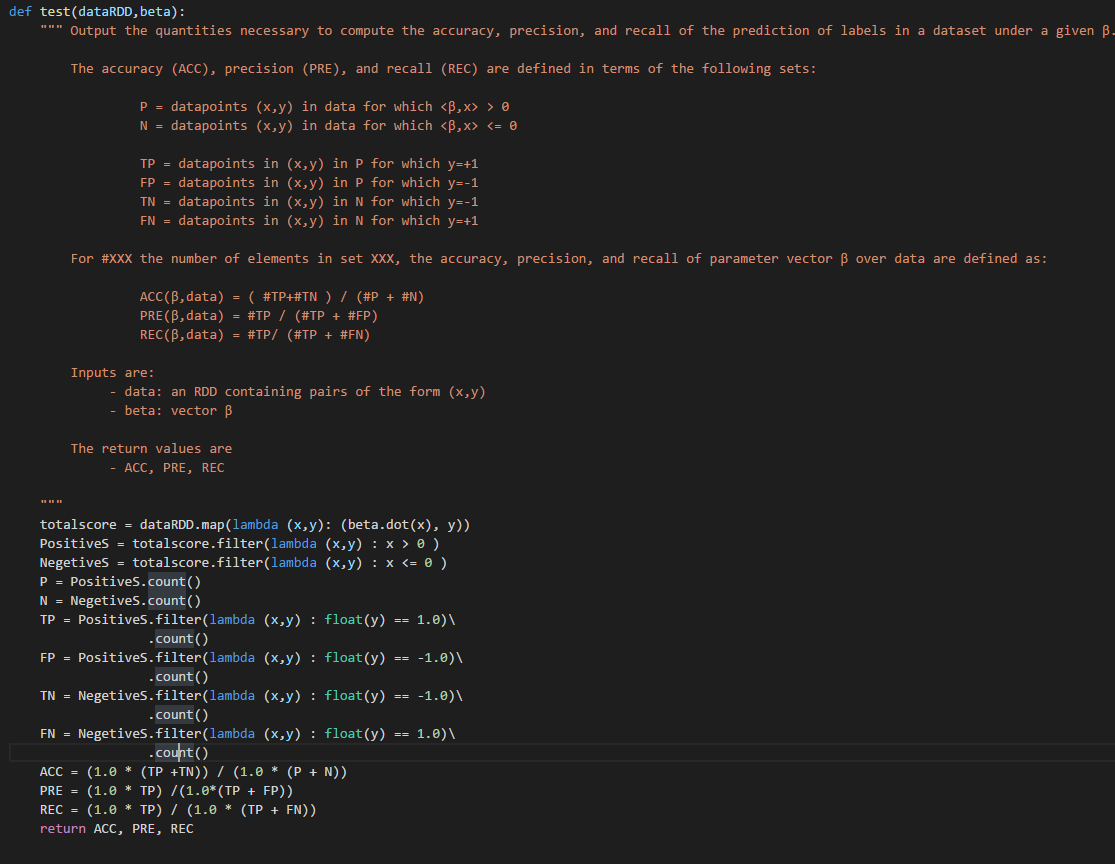
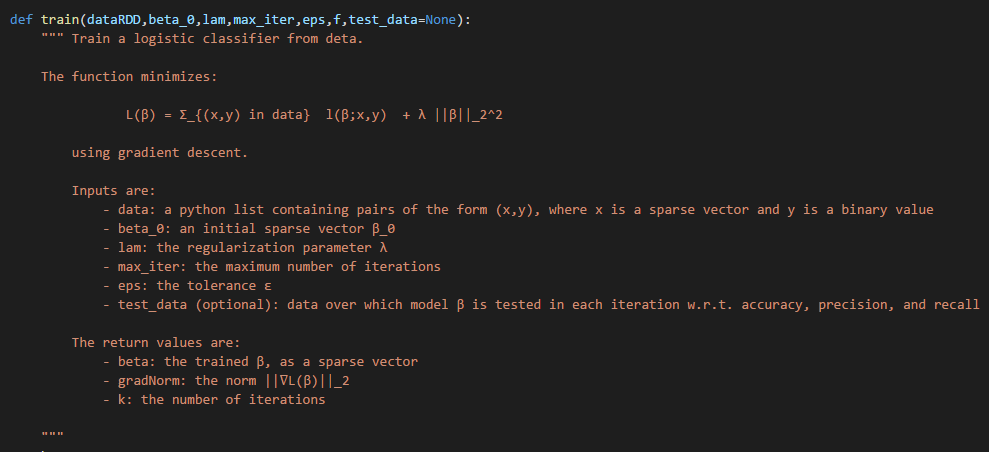


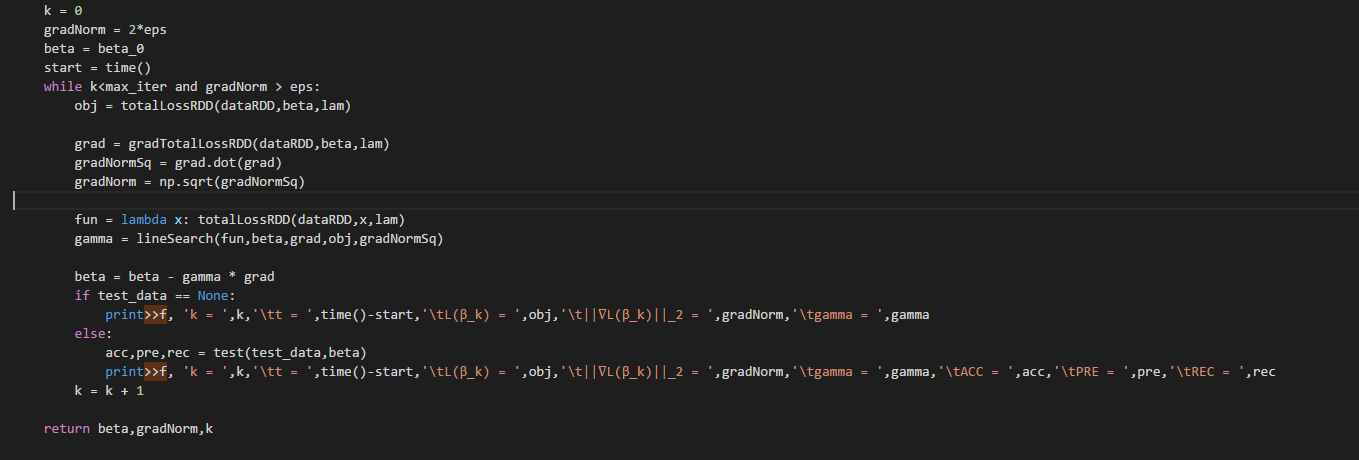
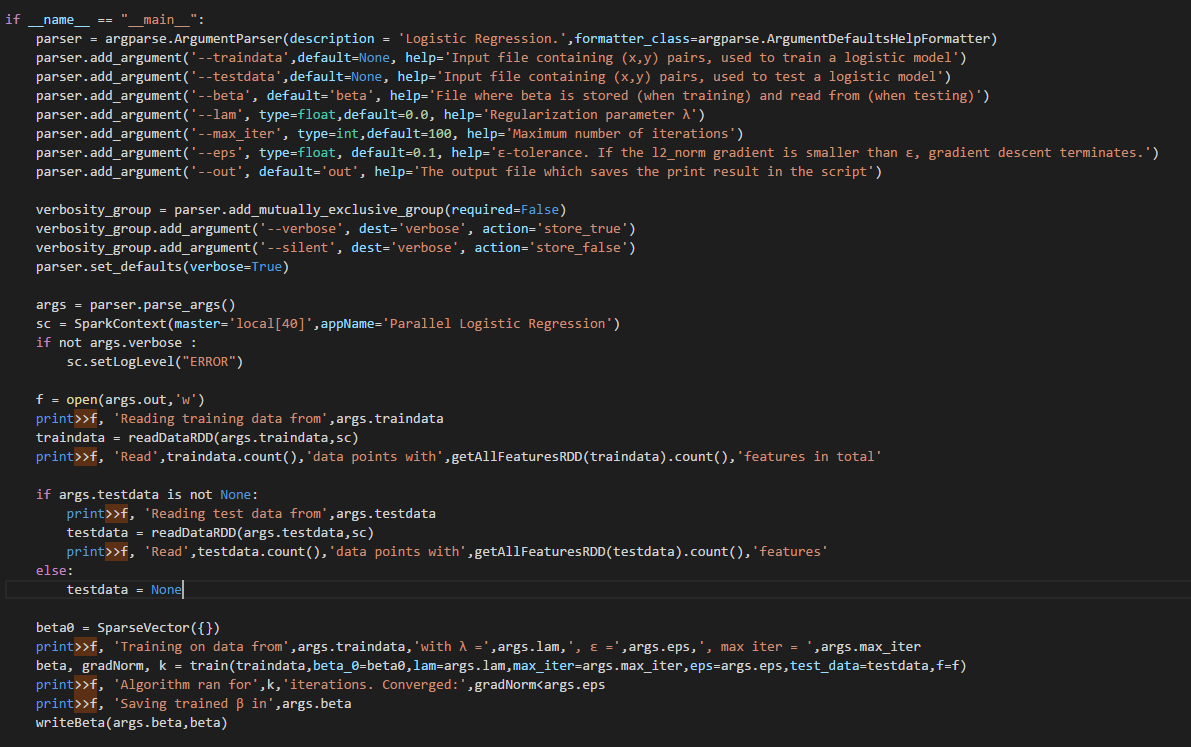
3d) Precision is  and Recall is

High precision means in the normal mushroom label there’s majority normal mushrooms and poor poisoned mushroom. So, in this case we should use high precision.

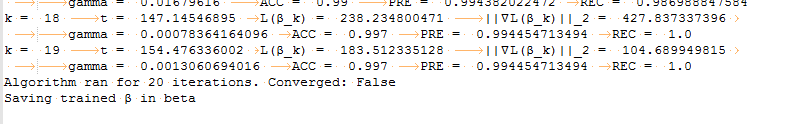
Q4

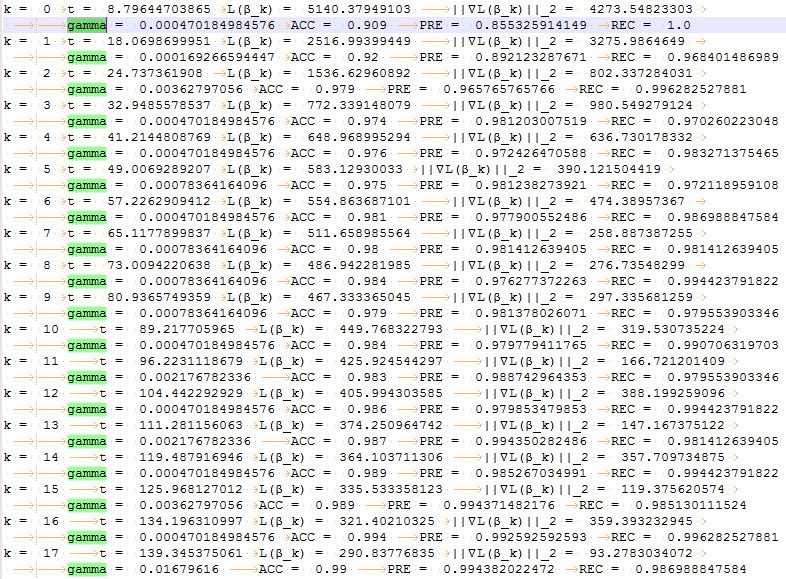
4a)



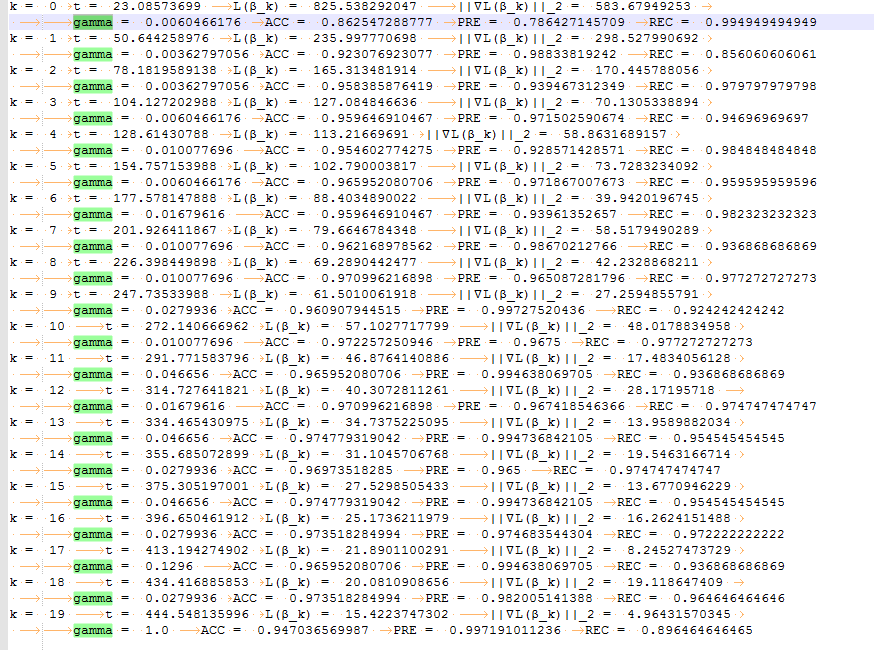


4b)

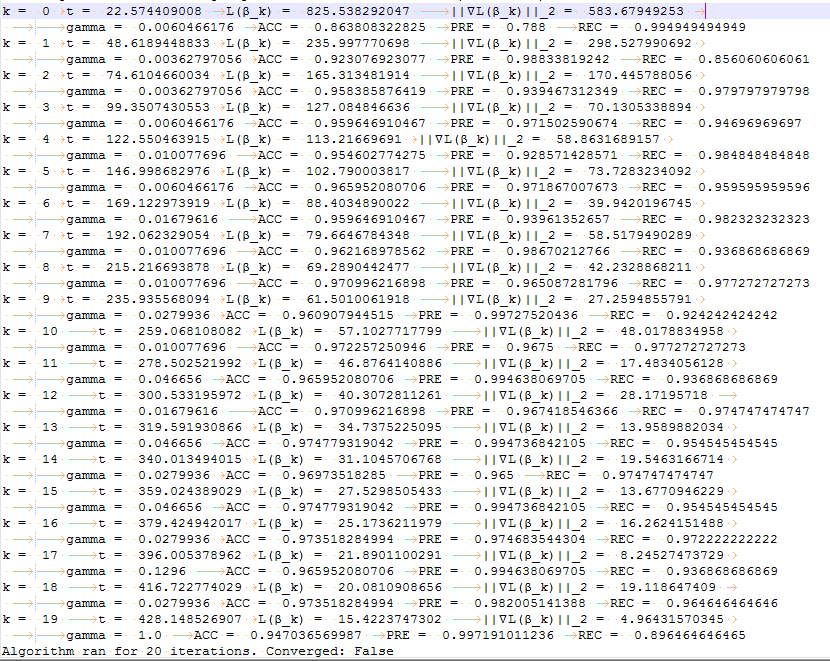
 Using Parallel

 logisticRegressison use 63s however ParallelLogisticRegression use 154s which is slower. Thus, using parallel in small data set doesn’t improve the speed

4c)

LogisticRegression

ParallelLogisticRegression



logisticRegressison use 444s however ParallelLogisticRegression use 428 which is faster. Thus, using parallel in large data set can improve the speed

Q5

5a)



Gradient norm 2

Accurate



Precision



Recall

5b)

c) lambda = 0 has the greatest result

10 Most Negative

**' basebal' -2.56823700000000**

**' game' -2.10310900000000**

**' player' -1.94976900000000**

**' team' -1.91179500000000**

**' yanke' -1.60964700000000**

**' win' -1.45783200000000**

**' philli' -1.38217800000000**

**' plai' -1.38204400000000**

**' stat' -1.30981000000000**

**' pitch' -1.24824500000000**

10 Most Positive

**' inform' 1.25335100000000**

**' doctor' 1.18740100000000**

**' effect' 1.14800100000000**

**' treatment' 1.03699300000000**

**' diseas' 1.03653500000000**

**' medic' 0.917249000000000**

**' treat' 0.884590000000000**

**' problem' 0.878204000000000**

**' health' 0.808931000000000**

**' peopl' 0.778835000000000**