BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

HYDERABAD CAMPUS

Second Semester 2020-21

BITS F464 - Machine Learning

Assignment - 2A

Group Members

Danish Mohammad 2018A7PS0103H

Mridul Kumar Rai 2018AAPS0359H

Ketan Goyal 2018A8PS0900H

Logistic Regression

Problem Statement

1) In this assignment, you will implement logistic regression from scratch for binary classification. To train the model, you will use Gradient Descent and Stochastic Gradient Descent with appropriate learning rates. Plot the loss and accuracy for your model every 50 iterations to visualize the training better. You are expected to create 10 independent random 70:30 splits on the given data, train the model and report the average loss and accuracy over all those 10 splits.
2) Try to vectorize your code as much as possible to make your computations faster and efficient. Do not hard code any parts of the implementation unless it is absolutely necessary.

What needs to be documented

1) A very brief description of your model and its implementation.

We will be calculating the accuracy by adjusting the weights and bias for both Gradient Descent and Stochastic Gradient Descent separately. We predict the class using the weight vector and bias. The weight and bias errors are calculated by comparing the predicted with the actual class. The weights and bias are updated after each iteration. The cost function gives us the loss. We classify points as true positives, true negatives, false positives and false negatives which help us in calculating f score and precision. The data is divided into 10 independent sets. The GD and SGD train and test metrics are calculated at three different learning rates on the 10 datasets.

1) Weights and biases are randomly initialised
2) In each iteration, the predicted class is compared against the actual class to determine the loss.
3) Loss (or cost) is calculated using cross-entropy loss
4) Based on the loss, the weights and biases are updated.
5) Once training is complete, after some number of iterations, we use the trained model to find true positives, true negatives, false positives and false negatives.

6) These values help us calculate accuracy, F-score, Precision, Recall

Accuracy = (TP+TN)/(TP+TN+FP+FN)
Precision = TP/(TP+FP)
Recall = TP/(TP+FN)
F-score = (2 ∗Precision ∗Recall)/(Precision+Recall)

2) The most important feature in the dataset.

Gradient Descent: The most important attribute is found to be the attribute 1 as the weight value for attribute 1 is the highest amongst all attributes

Stochastic Gradient Descent: The most important attribute is found to be the attribute as the weight value for attribute 1 is the highest amongst all attributes

3) The final train and test metrics (loss, accuracy, recall, precision and fscore) achieved by your model with GD and SGD.

```
******************10 random 70:30 train test splits with learning rate 0.001******************
Model weights(Gradient Descent):            [-1.15099771 -0.58983848 -0.62123401 -0.15286678]
Model weights(Stochastic Gradient Descent):  [-1.16751858 -0.58098724 -0.63843841 -0.1349048 ]
 No.          Stat Name                 Gradient Descent        Stochastic Gradient Descent
  1            Accuracy                0.912621359223301            0.9223300970873787
                 Loss                  105.85167975833875            99.80119820647515
  2            Accuracy                0.9296116504854369           0.9368932038834952
                 Loss                  95.84016758222486             94.37806877880848
  3            Accuracy                0.9368932038834952           0.9368932038834952
                 Loss                  91.67941838497808             89.49778076233073
  4            Accuracy                0.9393203883495146           0.9368932038834952
                 Loss                  88.57639591910213             87.95511150831624
  5            Accuracy                0.9393203883495146           0.9320388349514563
                 Loss                   88.2529770373236             88.36276328882136
  6            Accuracy                0.9393203883495146           0.941747572815534
                 Loss                  87.73983369757175             89.5597429797786
  7            Accuracy                0.9393203883495146           0.9368932038834952
                 Loss                  87.98172275063814             89.23600560528098
  8            Accuracy                0.9393203883495146           0.9393203883495146
                 Loss                  87.74959056275424             87.68454603005736
  9            Accuracy                0.941747572815534            0.9538834951456311
                 Loss                  87.53045393803983             87.42143100439796
 10            Accuracy                0.941747572815534            0.9393203883495146
                 Loss                  86.67717340291662             86.09580377264706
 All         Avg Accuracy                   0.935922                     0.937621
 All           Avg Loss                     90.787941                    89.999245
```

```
******************10 random 70:30 train test splits with learning rate 0.01******************
Model weights(Gradient Descent):             [-2.11812586 -1.17670507 -1.36818947 -0.16477877]
Model weights(Stochastic Gradient Descent):  [-2.20424345 -1.19872891 -1.45566998 -0.10017886]
 No.          Stat Name                 Gradient Descent        Stochastic Gradient Descent
  1            Accuracy                0.9781553398058253            0.9878640776699029
                 Loss                  38.67451541685887             32.507876356706774
  2            Accuracy                0.9878640776699029            0.9805825242718447
                 Loss                  30.761994088431763            33.770599099403306
  3            Accuracy                0.9878640776699029            0.9805825242718447
                 Loss                  30.385496698425968            34.477831970935924
  4            Accuracy                0.9878640776699029            0.9733009708737864
                 Loss                  30.158415550471283            37.02445226661268
  5            Accuracy                0.9878640776699029            0.9878640776699029
                 Loss                  29.92496627457698             28.044883057340495
  6            Accuracy                0.9878640776699029            0.9878640776699029
                 Loss                  29.834083858656424            33.194958285387244
  7            Accuracy                0.9878640776699029            0.9805825242718447
                 Loss                  30.196803603758323            29.953711573922025
  8            Accuracy                0.9878640776699029            0.9854368932038835
                 Loss                  29.95504240096974             29.202384669556167
  9            Accuracy                0.9878640776699029            0.9757281553398058
                 Loss                  30.010092374501667            37.48456924104304
 10            Accuracy                0.9878640776699029            0.9830097087378641
                 Loss                  30.274579236006844            28.117326122572038
 All         Avg Accuracy                    0.986893                      0.982282
 All           Avg Loss                      31.017599                     32.377859
```

```
******************10 random 70:30 train test splits with learning rate 0.1*******************
Model weights(Gradient Descent):             [-3.90184941 -2.06698732 -2.5540207  -0.08621851]
Model weights(Stochastic Gradient Descent):  [-4.23018508e+00 -2.12585863e+00 -2.99690504e+00  3.25044314e-03]
 No.          Stat Name                 Gradient Descent        Stochastic Gradient Descent
  1            Accuracy                0.9878640776699029            0.9563106796116505
                 Loss                  16.18865842209692             50.20916102857387
  2            Accuracy                0.9878640776699029            0.9781553398058253
                 Loss                  15.587521472276558            26.73742233042657
  3            Accuracy                0.9878640776699029            0.9757281553398058
                 Loss                  15.59005615300998             24.500348125690618
  4            Accuracy                0.9878640776699029            0.9878640776699029
                 Loss                  15.568440136919317            12.576364071841613
  5            Accuracy                0.9878640776699029            0.9490291262135923
                 Loss                  15.561210409809227            59.155850160473484
  6            Accuracy                0.9878640776699029            0.9854368932038835
                 Loss                  15.57570636356912             16.163749094428944
  7            Accuracy                0.9878640776699029            0.9757281553398058
                 Loss                  15.523095716661201            20.150024659096264
  8            Accuracy                0.9878640776699029            0.9805825242718447
                 Loss                  15.578512143922362            22.504402881760956
  9            Accuracy                0.9878640776699029            0.970873786407767
                 Loss                  15.593627226221829            28.708323752080982
 10            Accuracy                0.9878640776699029            0.9053398058252428
                 Loss                  15.619810755261373            131.65290939184922
 All         Avg Accuracy                    0.987864                      0.966505
 All           Avg Loss                      15.638664                     39.235856
```

```
******************10 random 70:30 train test splits with learning rate 0.001******************
Model weights(Gradient Descent):             [-1.15099771 -0.58983848 -0.62123401 -0.15286678]
Model weights(Stochastic Gradient Descent):  [-1.16751858 -0.58098724 -0.63843841 -0.1349048 ]
 No.          Stat Name                 Gradient Descent        Stochastic Gradient Descent
  1            Accuracy                0.912621359223301             0.9223300970873787
                 Loss                  105.85167975833875            99.80119820647515
  2            Accuracy                0.9296116504854369            0.9368932038834952
                 Loss                  95.84016758222486             94.37806877880848
  3            Accuracy                0.9368932038834952            0.9368932038834952
                 Loss                  91.67941838497808             89.49778076233073
  4            Accuracy                0.9393203883495146            0.9368932038834952
                 Loss                  88.57639591910213             87.95511150831624
  5            Accuracy                0.9393203883495146            0.9320388349514563
                 Loss                  88.2529770373236              88.36276328882136
  6            Accuracy                0.9393203883495146            0.941747572815534
                 Loss                  87.73983369757175             89.5597429797786
```

7	Accuracy	0.9393203883495146	0.9368932038834952
	Loss	87.98172275063814	89.23600560528098
8	Accuracy	0.9393203883495146	0.9393203883495146
	Loss	87.74959056275424	87.68454603005736
9	Accuracy	0.941747572815534	0.9538834951456311
	Loss	87.53045393803983	87.42143100439796
10	Accuracy	0.941747572815534	0.9393203883495146
	Loss	86.67717340291662	86.09580377264706
All	Avg Accuracy	0.935922	0.937621
All	Avg Loss	90.787941	89.999245

**********10 random 70:30 train test splits with learning rate 0.01**********

Model weights(Gradient Descent): [-2.11812586 -1.17670507 -1.36818947 -0.16477877]

Model weights(Stochastic Gradient Descent): [-2.20424345 -1.19872891 -1.45566998 -0.10017886]

No.	Stat Name	Gradient Descent	Stochastic Gradient Descent
1	Accuracy	0.9781553398058253	0.9878640776699029
	Loss	38.67451541685887	32.507876356706774
2	Accuracy	0.9878640776699029	0.9805825242718447
	Loss	30.761994088431763	33.770599099403306
3	Accuracy	0.9878640776699029	0.9805825242718447
	Loss	30.385496698425968	34.477831970935924
4	Accuracy	0.9878640776699029	0.9733009708737864
	Loss	30.158415550471283	37.02445226661268
5	Accuracy	0.9878640776699029	0.9878640776699029
	Loss	29.92496627457698	28.044883057340495
6	Accuracy	0.9878640776699029	0.9878640776699029
	Loss	29.834083858656424	33.194958285387244
7	Accuracy	0.9878640776699029	0.9805825242718447
	Loss	30.196803603758323	29.953711573922025
8	Accuracy	0.9878640776699029	0.9854368932038835
	Loss	29.95504240096974	29.202384669556167
9	Accuracy	0.9878640776699029	0.9757281553398058
	Loss	30.010092374501667	37.48456924104304
10	Accuracy	0.9878640776699029	0.9830097087378641
	Loss	30.274579236006844	28.117326122572038
All	Avg Accuracy	0.986893	0.982282
All	Avg Loss	31.017599	32.377859

**********10 random 70:30 train test splits with learning rate 0.1**********

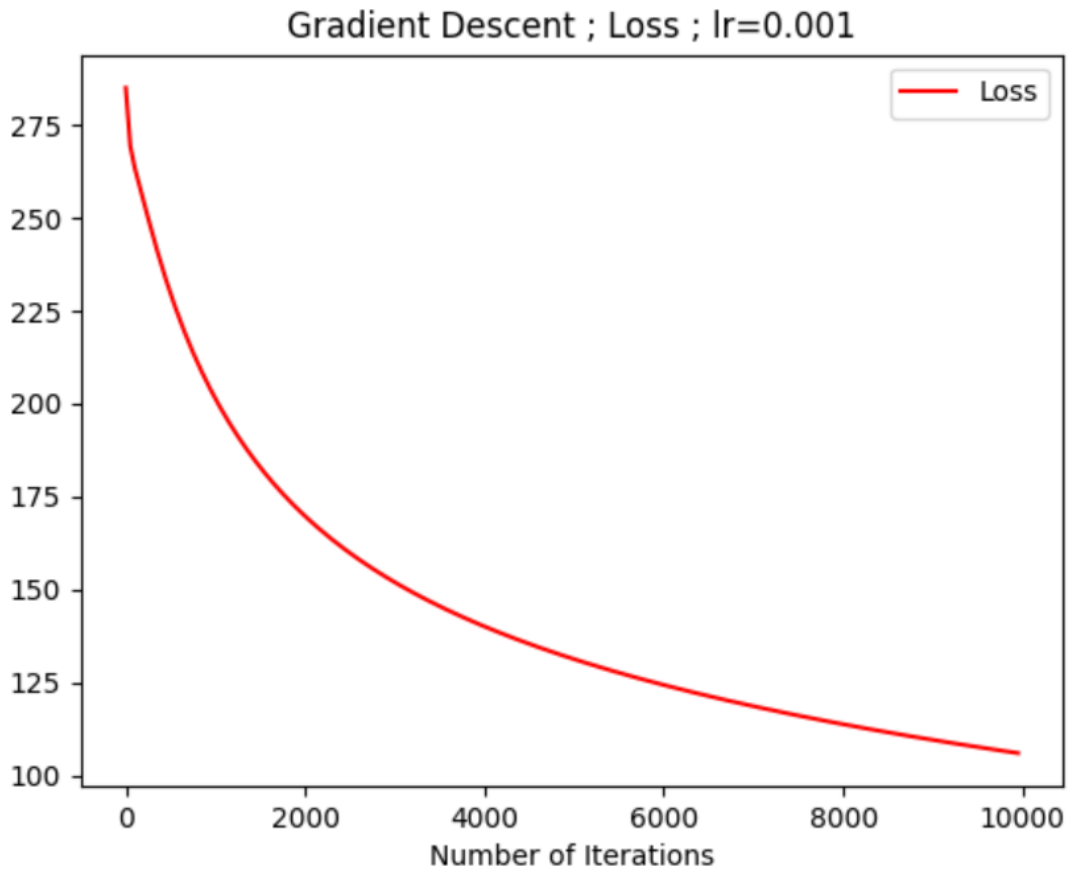
Model weights(Gradient Descent): [-3.90184941 -2.06698732 -2.5540207 -0.08621851]

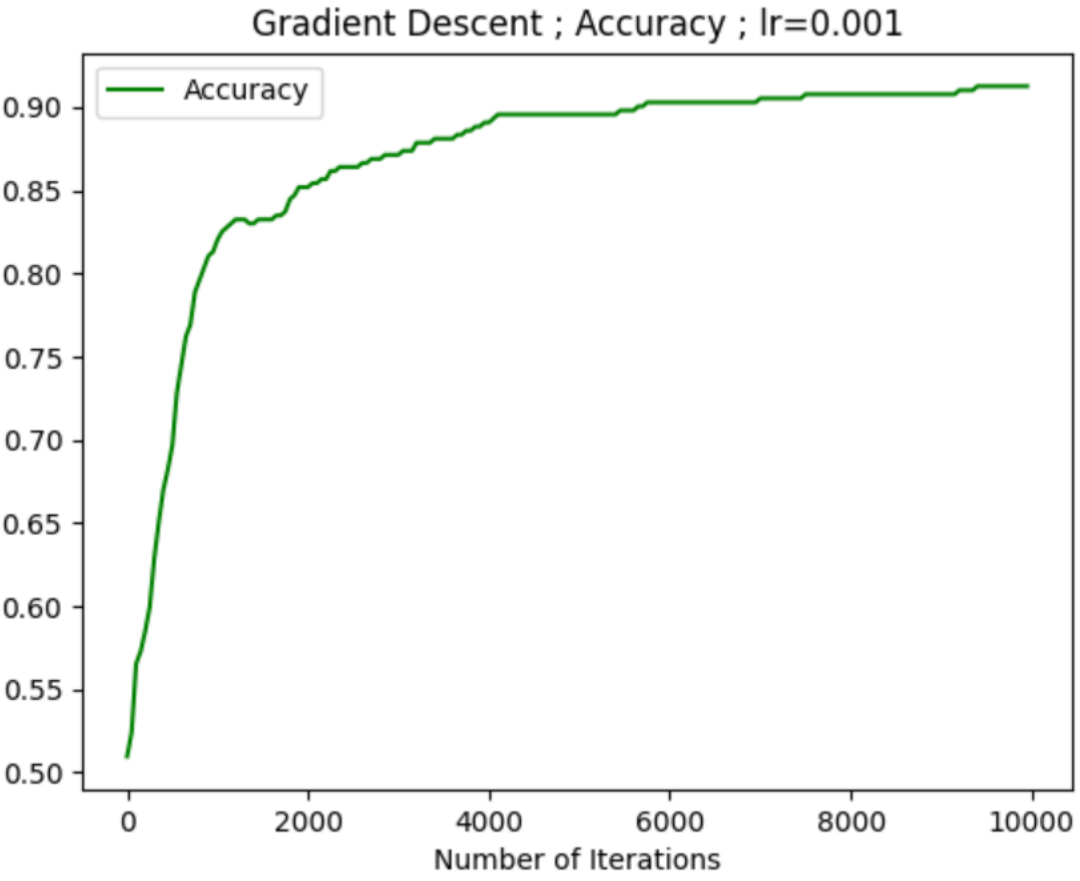
Model weights(Stochastic Gradient Descent): [-4.23018508e+00 -2.12585863e+00 -2.99690504e+00
3.25044314e-03]

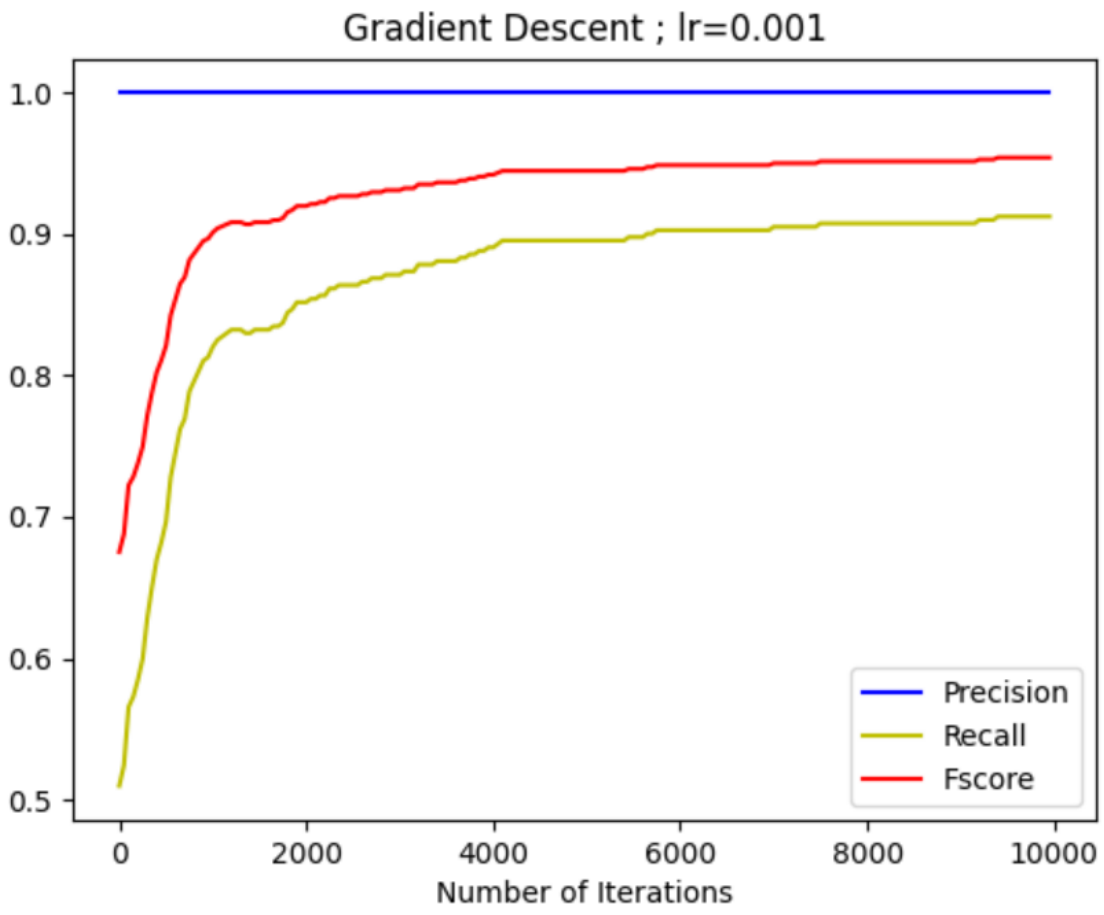
No.	Stat Name	Gradient Descent	Stochastic Gradient Descent
1	Accuracy	0.9878640776699029	0.9563106796116505
	Loss	16.18865842209692	50.20916102857387
2	Accuracy	0.9878640776699029	0.9781553398058253
	Loss	15.587521472276558	26.73742233042657
3	Accuracy	0.9878640776699029	0.9757281553398058
	Loss	15.59005615300998	24.500348125690618
4	Accuracy	0.9878640776699029	0.9878640776699029
	Loss	15.568440136919317	12.576364071841613
5	Accuracy	0.9878640776699029	0.9490291262135923
	Loss	15.561210409809227	59.155850160473484
6	Accuracy	0.9878640776699029	0.9854368932038835
	Loss	15.57570636356912	16.163749094428944

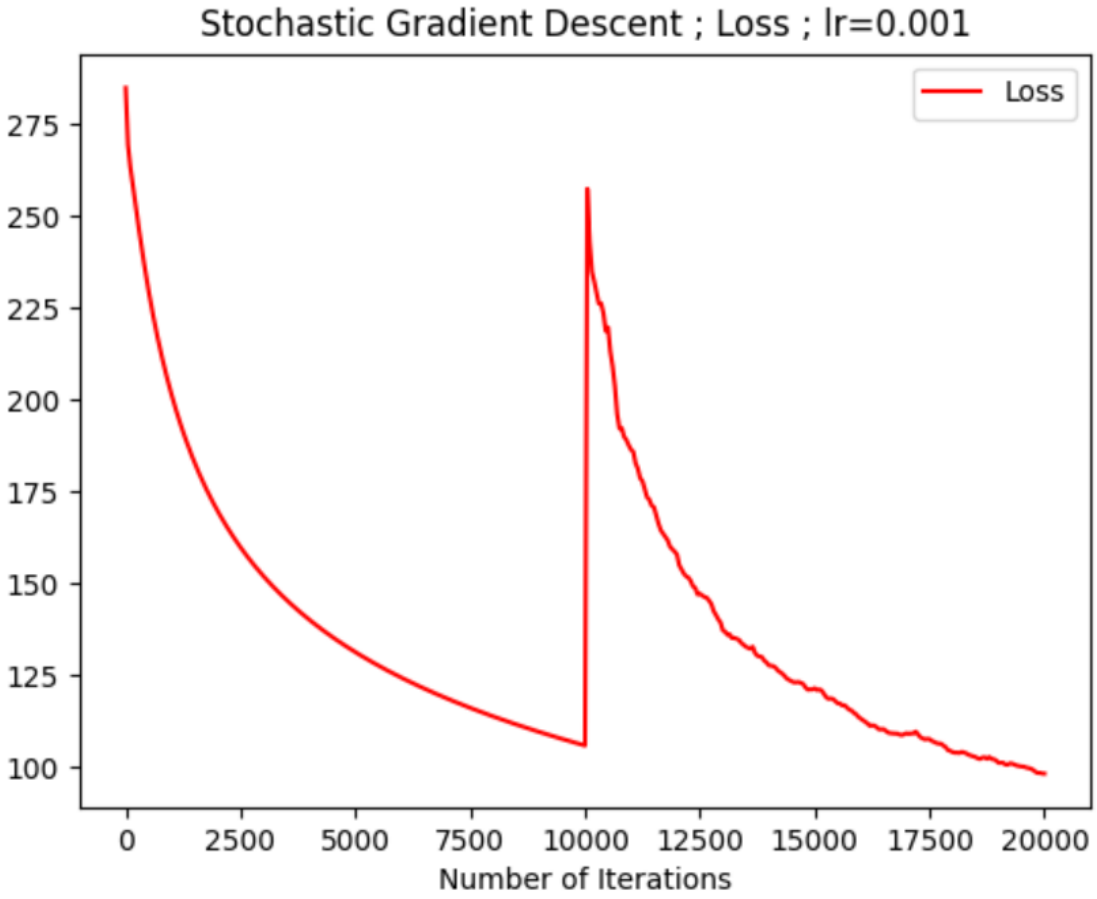
7	Accuracy	0.9878640776699029	0.9757281553398058
	Loss	15.523095716661201	20.150024659096264
8	Accuracy	0.9878640776699029	0.9805825242718447
	Loss	15.578512143922362	22.504402881760956
9	Accuracy	0.9878640776699029	0.970873786407767
	Loss	15.593627226221829	28.708323752080982
10	Accuracy	0.9878640776699029	0.9053398058252428
	Loss	15.619810755261373	131.65290939184922
All	Avg Accuracy	0.987864	0.966505
All	Avg Loss	15.638664	39.235856

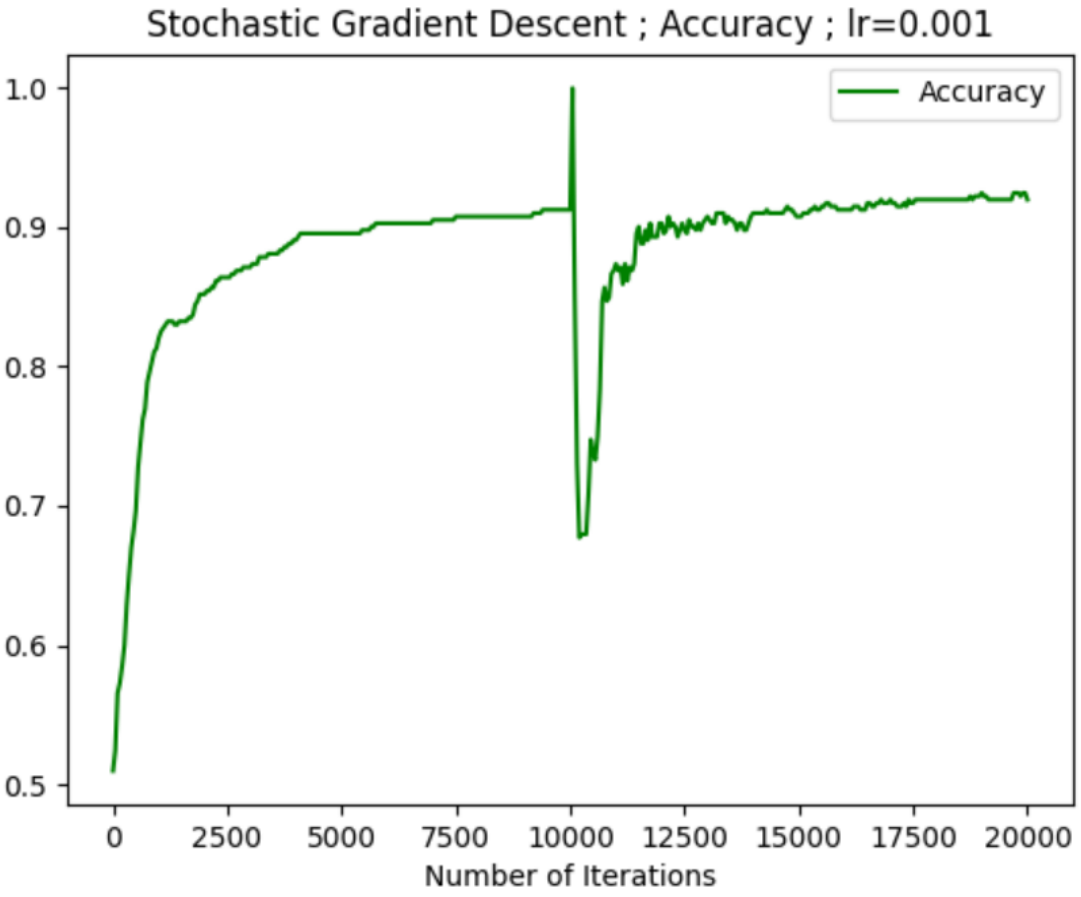
4) Plots of accuracy for three different learning rates using GD and SGD, i.e. three plots for GD with varying learning rate (say η1, η2, η3) and three plots for SGD with the same set of learning rate on anyone data split.

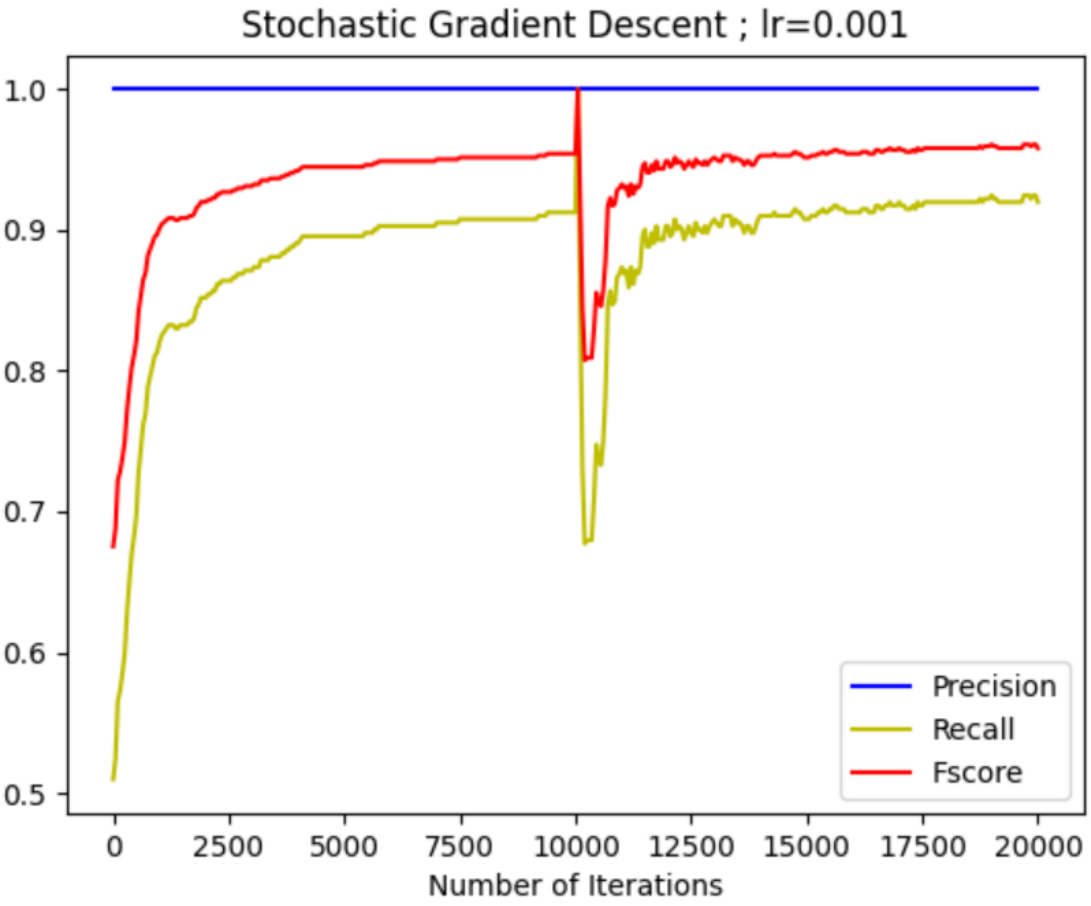

Gradient Descent ; Loss ; lr=0.001
Loss
275
250
225
200
175
150
125
100
0
2000
4000
6000
8000
10000
Number of Iterations


Gradient Descent ; Accuracy ; lr=0.001
Accuracy
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0
2000
4000
6000
8000
10000
Number of Iterations

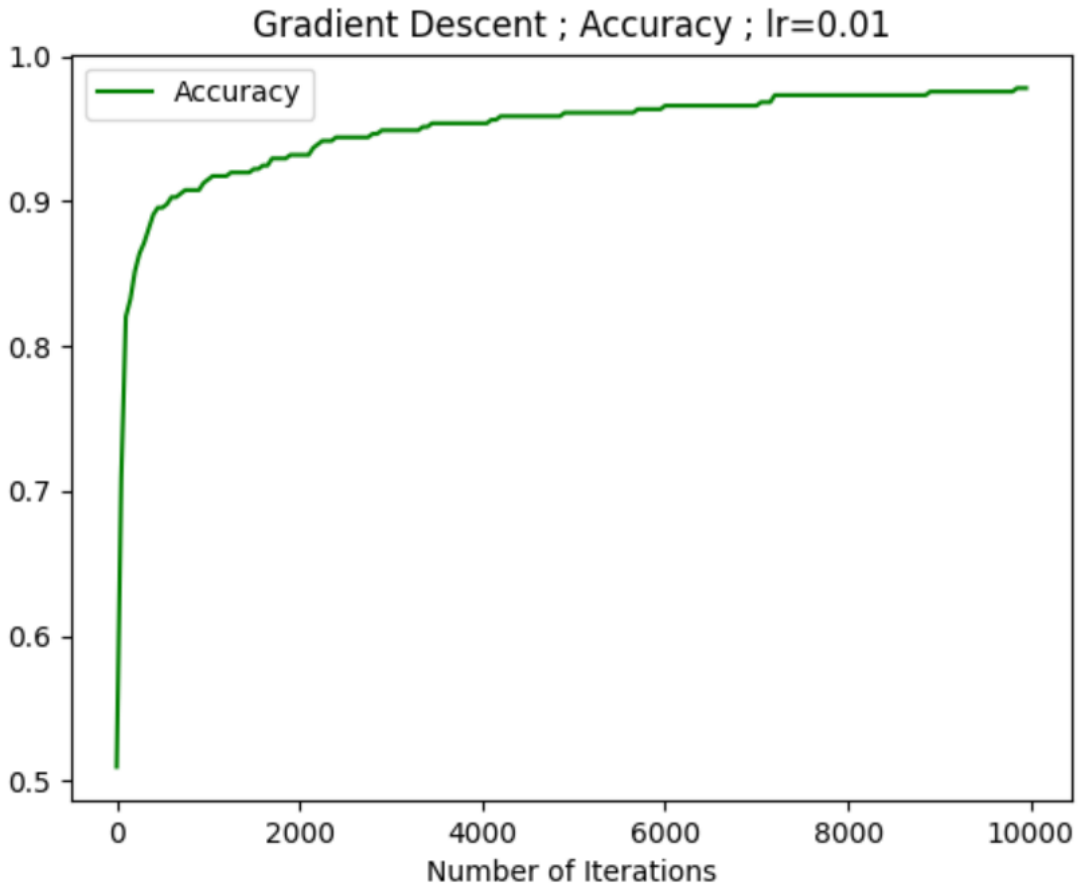

Gradient Descent ; lr=0.001
1.0
0.9
0.8
0.7
0.6
0.5
0
2000
4000
6000
8000
10000
Number of Iterations
Precision
Recall
Fscore

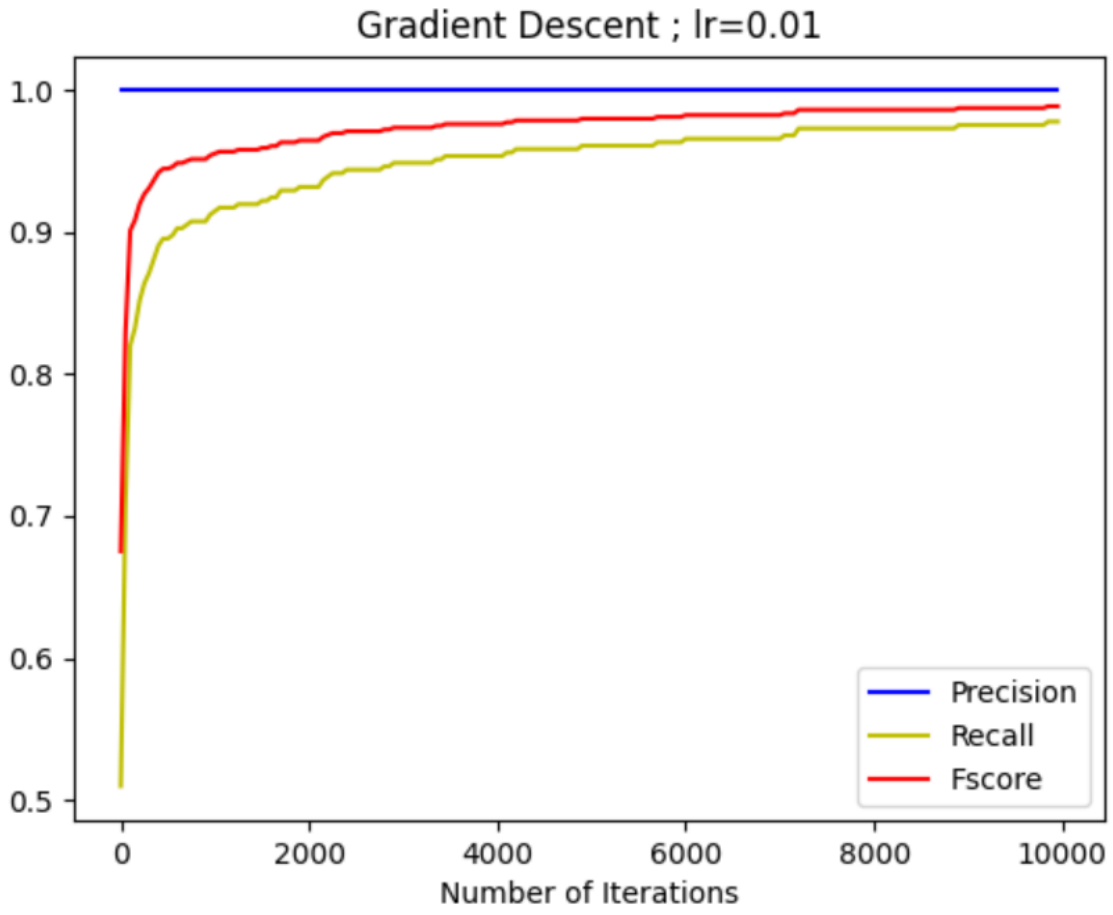

Stochastic Gradient Descent ; Loss ; lr=0.001
275
250
225
200
175
150
125
100
0
2500
5000
7500
10000
12500
15000
17500
20000
Number of Iterations
Loss

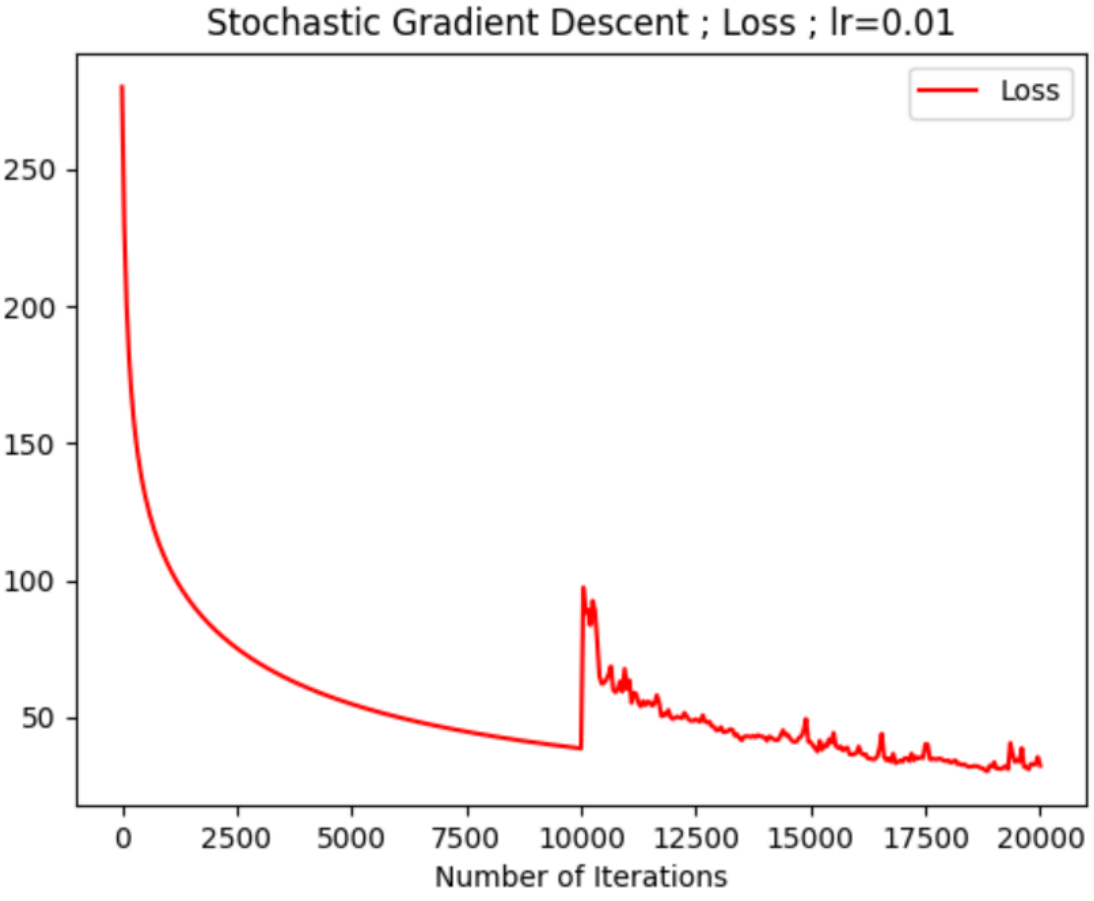

Stochastic Gradient Descent ; Accuracy ; lr=0.001
Accuracy
1.0
0.9
0.8
0.7
0.6
0.5
0
2500
5000
7500
10000
12500
15000
17500
20000
Number of Iterations

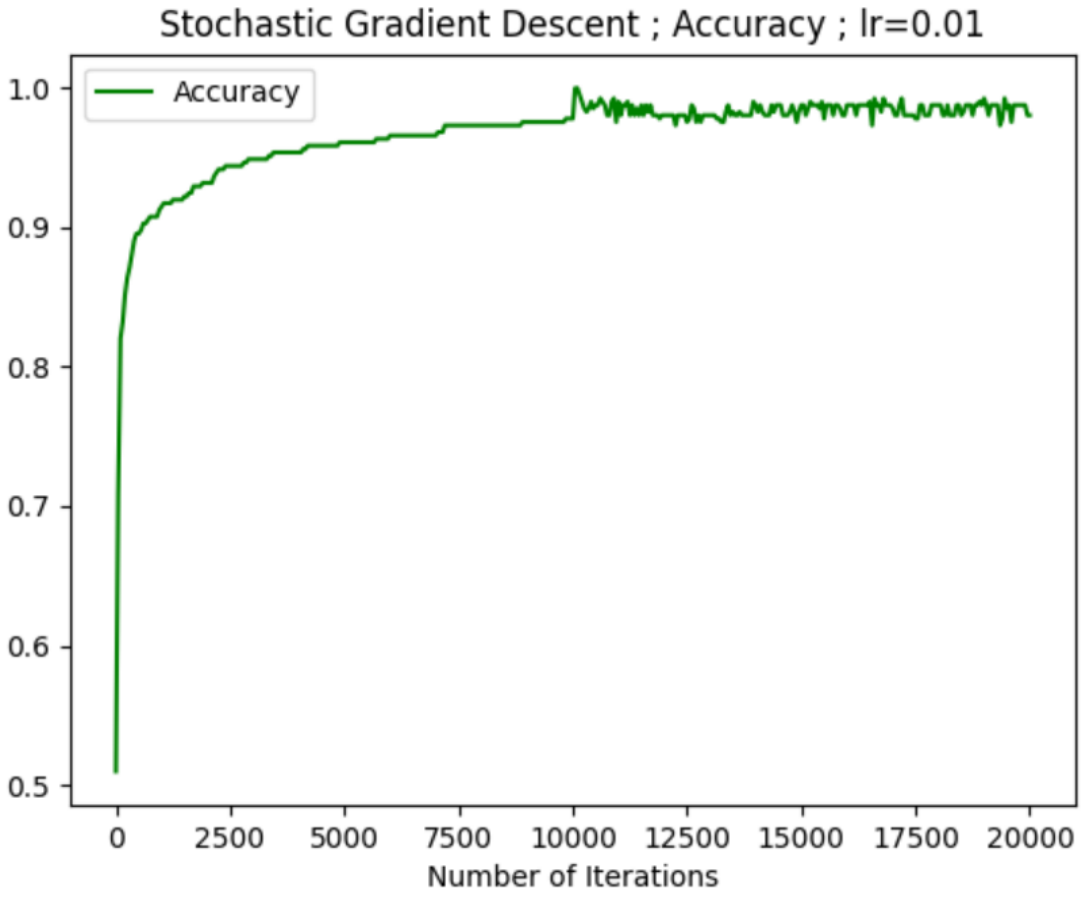

Stochastic Gradient Descent ; lr=0.001
Precision
Recall
Fscore
1.0
0.9
0.8
0.7
0.6
0.5
0
2500
5000
7500
10000
12500
15000
17500
20000
Number of Iterations

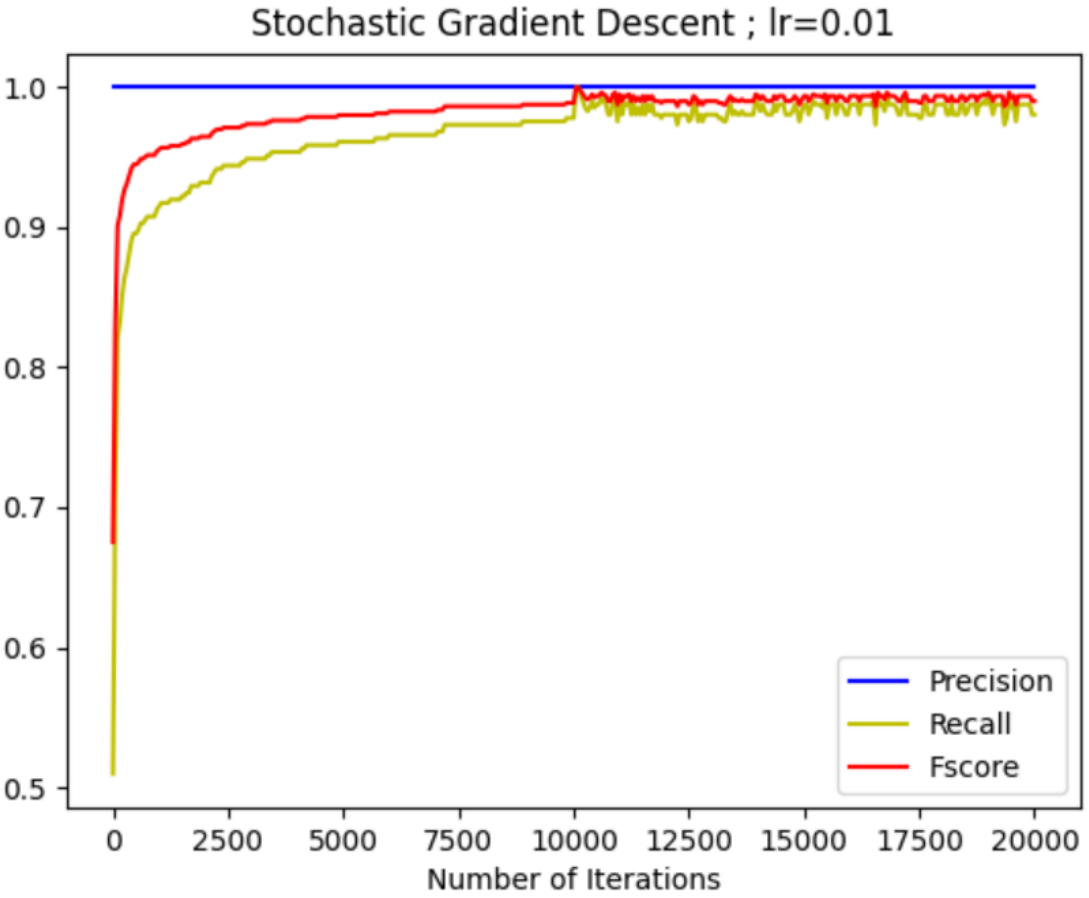

Gradient Descent ; Loss ; lr=0.01
Loss
250
200
150
100
50
0
2000
4000
6000
8000
10000
Number of Iterations

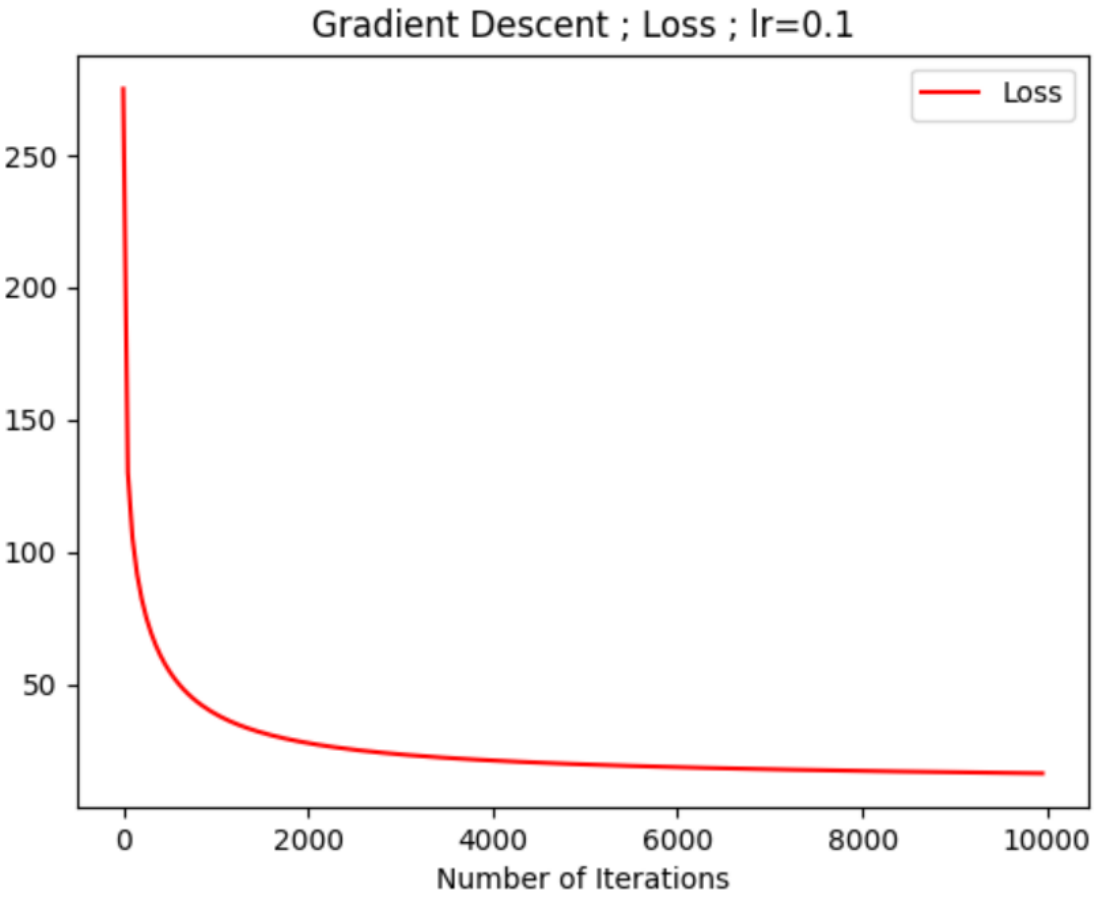

Gradient Descent ; Accuracy ; lr=0.01
Accuracy
1.0
0.9
0.8
0.7
0.6
0.5
0
2000
4000
6000
8000
10000
Number of Iterations

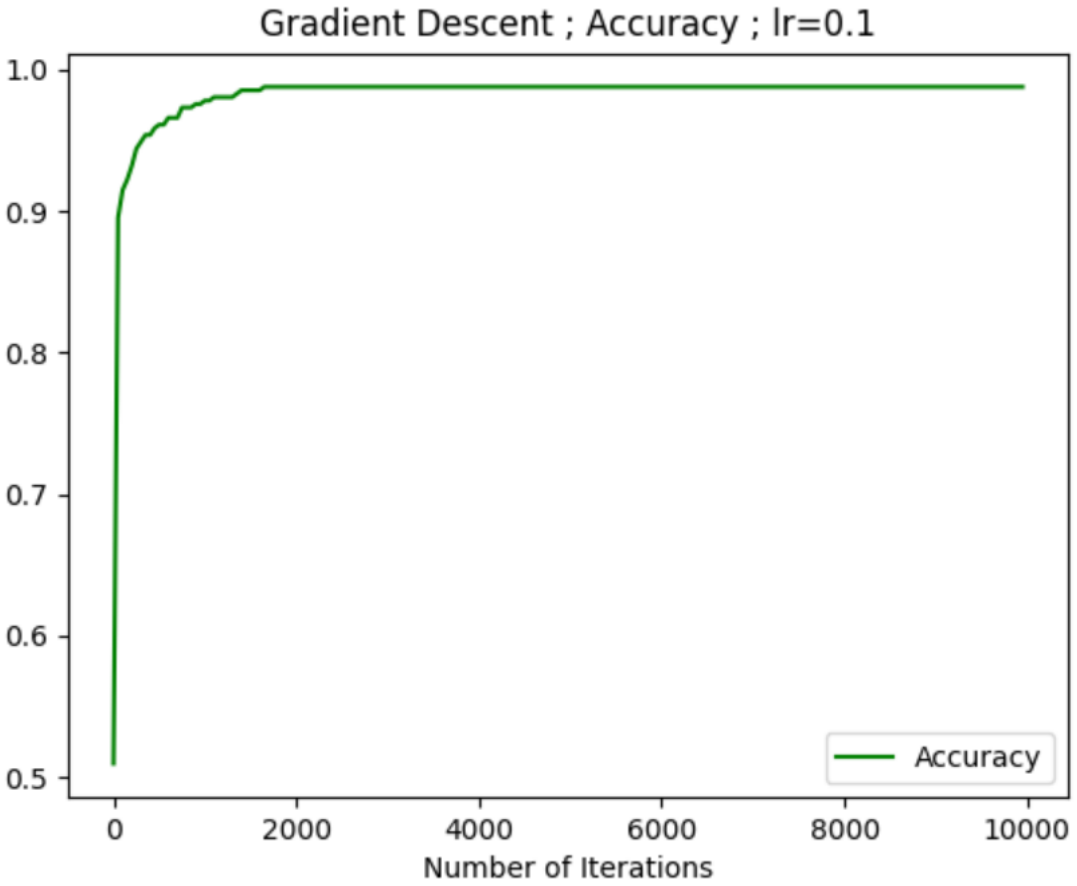

Gradient Descent ; lr=0.01
1.0
0.9
0.8
0.7
0.6
0.5
0
2000
4000
6000
8000
10000
Number of Iterations
Precision
Recall
Fscore

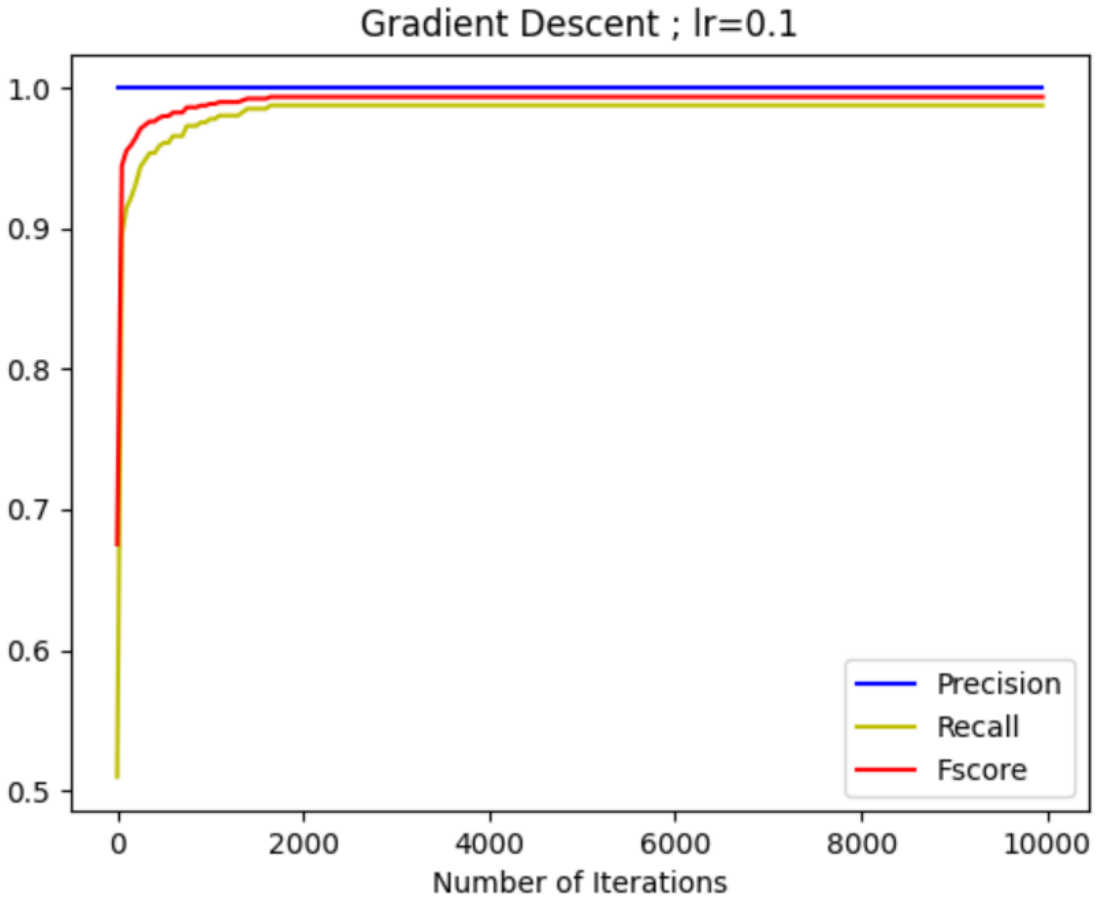

Stochastic Gradient Descent ; Loss ; lr=0.01
250
200
150
100
50
0
2500
5000
7500
10000
12500
15000
17500
20000
Number of Iterations
Loss

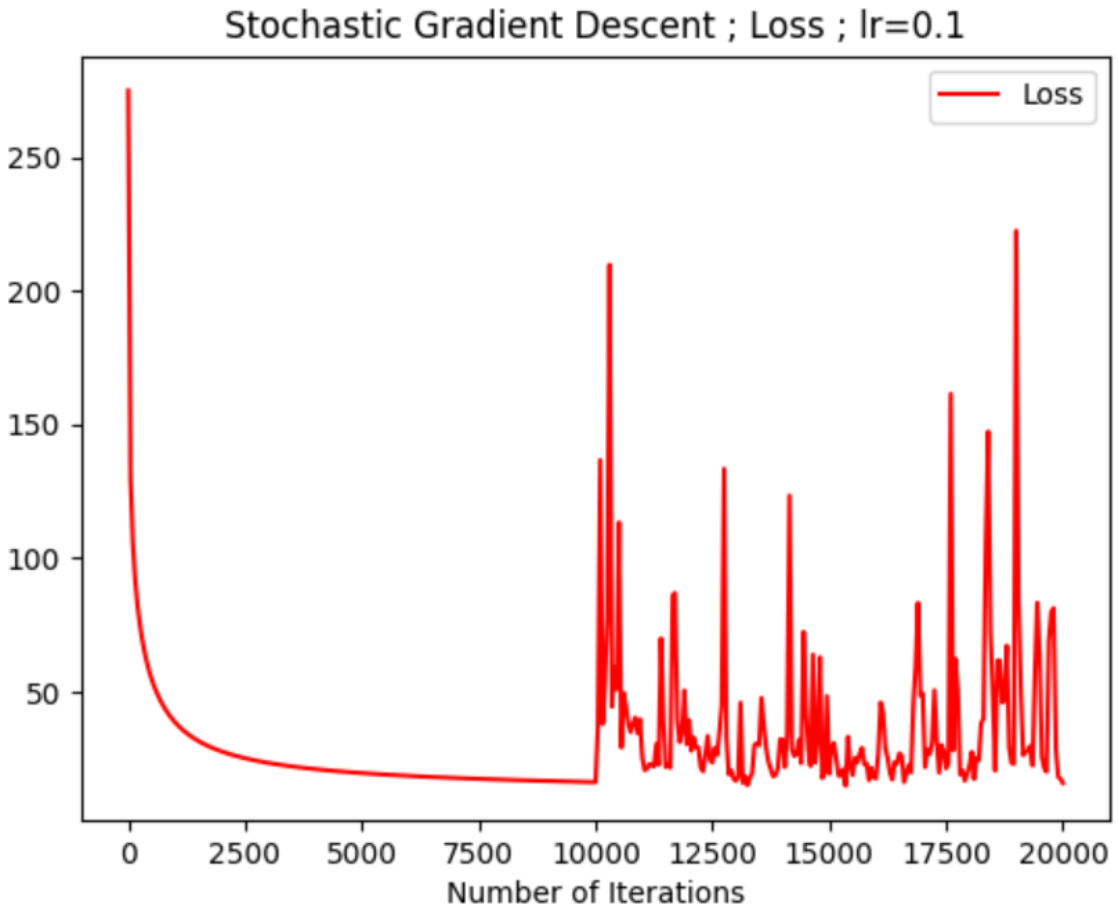

Stochastic Gradient Descent ; Accuracy ; lr=0.01
Accuracy
1.0
0.9
0.8
0.7
0.6
0.5
0
2500
5000
7500
10000
12500
15000
17500
20000
Number of Iterations

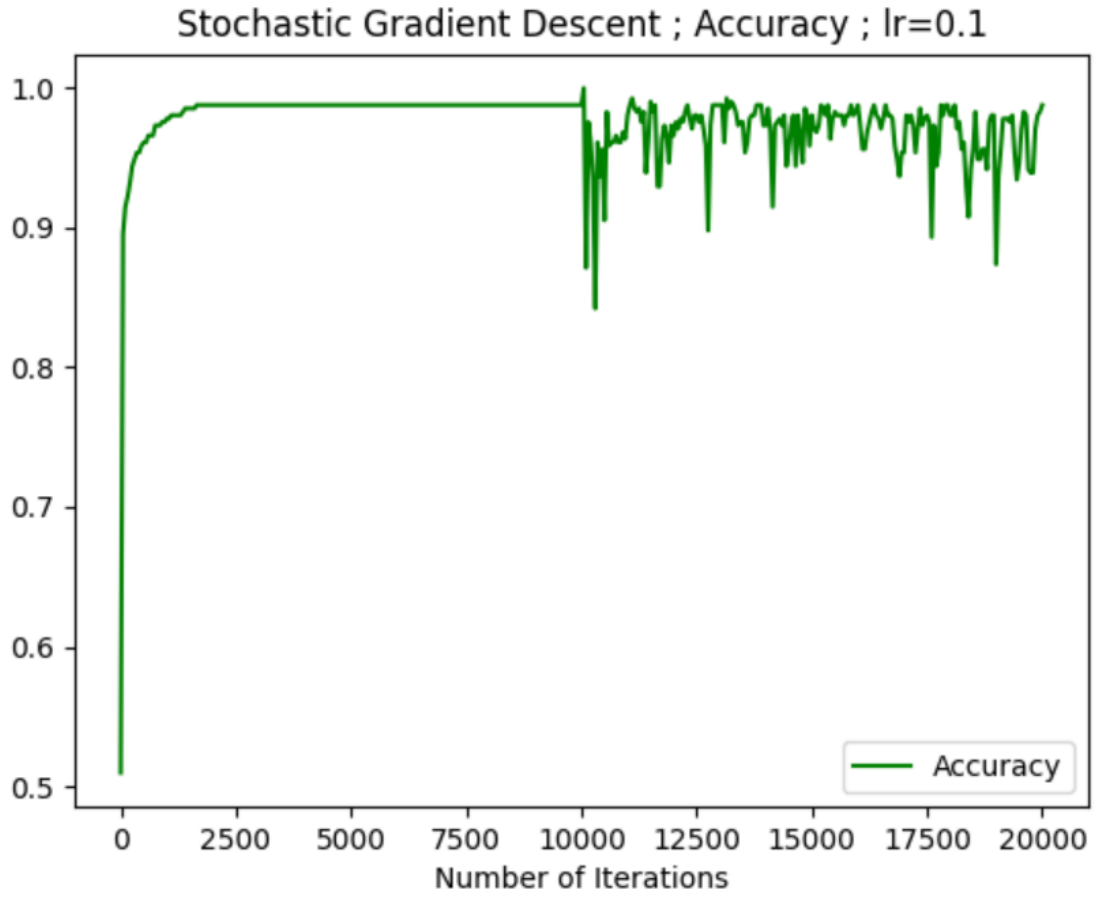

Stochastic Gradient Descent ; lr=0.01
1.0
0.9
0.8
0.7
0.6
0.5
Precision
Recall
Fscore
0
2500
5000
7500
10000
12500
15000
17500
20000
Number of Iterations

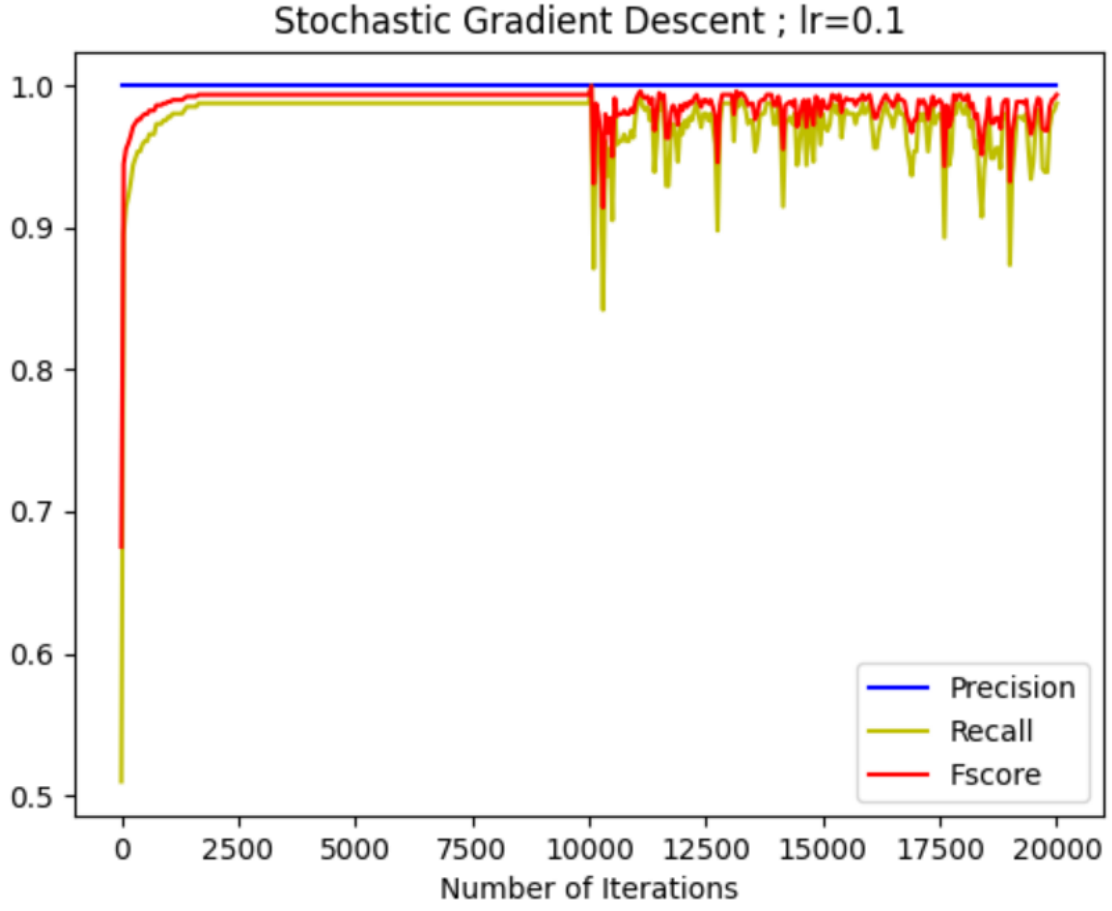

Gradient Descent ; Loss ; lr=0.1
Loss
250
200
150
100
50
0
2000
4000
6000
8000
10000
Number of Iterations


Gradient Descent ; Accuracy ; lr=0.1
1.0
0.9
0.8
0.7
0.6
0.5
Accuracy
0
2000
4000
6000
8000
10000
Number of Iterations


Gradient Descent ; lr=0.1
1.0
0.9
0.8
0.7
0.6
0.5
0
2000
4000
6000
8000
10000
Number of Iterations
Precision
Recall
Fscore


Stochastic Gradient Descent ; Loss ; lr=0.1
250
200
150
100
50
0
2500
5000
7500
10000
12500
15000
17500
20000
Number of Iterations
Loss


Stochastic Gradient Descent ; Accuracy ; lr=0.1
1.0
0.9
0.8
0.7
0.6
0.5
0
2500
5000
7500
10000
12500
15000
17500
20000
Number of Iterations
Accuracy


Stochastic Gradient Descent ; lr=0.1
1.0
0.9
0.8
0.7
0.6
0.5
0
2500
5000
7500
10000
12500
15000
17500
20000
Number of Iterations
Precision
Recall
Fscore