       Machine Learning

Assignment Report (2A)

Team members:

Nipun Wahi         Hrithik Kulkarni     Mir Ameen Mohideen

2018A7PS0966H     2018A7PS0278H     2018A7PS0487H

Logistic Regression:

The data was divided into training and testing data set with a 70:30 split. Then a weight vector was created and initialised to zeros. The last element in the weight vector represents the bias.

Then Gradient descent and Stochastic Gradient descent functions were called.

The only difference between the two was in terms of the error function which was computed with respect to one random element in SGD and all elements in case of GD

Results:

The most important feature was feature 1 which had the maximum absolute weight in SGD as well as GD with all learning rates.

GD with 10 independent random splits:

Weights:

[[-4.3066789 ]

[-2.27748499]

[-2.85149757]

[-0.22494296]

[ 3.9865394 ]]

Training Stats

Accuracy = 0.9896961498439125

Precision = 0.9962192816635159

Recall = 0.9924670433145011

Loss = 27.258281230021755

Fscore = 0.9943396226415094

Test Stats :

Accuracy = 0.9854014598540145

Precision = 0.9787234042553191

Recall = 0.9956709956709956

Loss = 15.767876281412578

Fscore = 0.9871244635193133

SGD with 10 independent random splits:

Weights:

[[-0.68105689]

[-0.31800028]

[-0.27639552]

[-0.14857716]

[ 0.09813082]]

Training Stats

Accuracy = 0.9025062434963579

Precision = 0.9251561218222154

Recall = 0.9822975517890773

Loss = 271.97328569825464

Fscore = 0.9528675318878553

Test Stats :

Accuracy = 0.9479318734793185

Precision = 0.9215785081529309

Recall = 0.9917748917748919

Loss = 120.18821340807922

Fscore = 0.9553842736293967

Gradient Descent:

Learning rate=0.00001

[[-1.44042871]

[-0.77531945]

[-0.84329198]

[-0.25605462]

[ 0.85423847]]

Training Stats

Accuracy = 0.9646706555671176

Precision = 0.9794007490636704

Recall = 0.9849340866290018

Loss = 10.438668029973066

Fscore = 0.9821596244131456

Test Stats :

Accuracy = 0.9781021897810219

Precision = 0.9703389830508474

Recall = 0.9913419913419913

Loss = 47.199074591349826

Fscore = 0.9807280513918629

Accuracy vs epochs :

Shape, square

Description automatically generated

# Learning rate=0.0001

[[-2.67005971]

[-1.47583862]

[-1.78289737]

[-0.14389337]

[ 2.62778512]]

Training Stats

Accuracy = 0.9864089490114464

Precision = 0.9961904761904762

Recall = 0.9849340866290018

Loss = 40.12560760209089

Fscore = 0.990530303030303

Test Stats :

Accuracy = 0.9854014598540146

Precision = 0.982832618025751

Recall = 0.9913419913419913

Loss = 19.018831560117324

Fscore = 0.9870689655172413

Accuracy vs epochs

Shape, square

Description automatically generated

# Learning Rate=0.001

[[-5.28860649]

[-2.77051614]

[-3.49635806]

[-0.33452235]

[ 4.70634546]]

Training Stats

Accuracy = 0.9917315296566076

Precision = 0.996219281663516

Recall = 0.992467043314501

Loss = 24.111983684310894

Fscore = 0.9943396226415094

Test Stats :

Accuracy = 0.9854014598540146

Precision = 0.9787234042553192

Recall = 0.9956709956709957

Loss = 15.842389845079894

Fscore = 0.9871244635193133

Shape, square

Description automatically generated

# SGD

# Learning rate=0.00001

# Weights:

[[-0.08589993]

[-0.08546586]

[-0.00641872]

[-0.00490893]

[-0.00216998]]

Training Stats

Accuracy = 0.7460978147762747

Precision = 0.7222222222222222

Recall = 0.8796992481203008

Loss = 744.0516689327236

Fscore = 0.7932203389830509

Test stats

Accuracy = 0.8004866180048662

Precision = 0.7761194029850746

Recall = 0.9043478260869565

Loss = 308.6360343054742

Fscore = 0.8353413654618473

Shape

Description automatically generated

# Learning Rate=0.0001

Weights

[[-0.16157106]

[-0.13198572]

[-0.00941518]

[-0.02081083]

[-0.00026101]]

Training Stats

Accuracy = 0.7458678459937566

Precision = 0.7775974025974026

Recall = 0.9020715630885122

Loss = 63.407708397728506

Fscore = 0.8352223190932868

Test Stats :

Accuracy = 0.8029197080291971

Precision = 0.7737226277372264

Recall = 0.9177489177489177

Loss = 27.231904005903157

Fscore = 0.8396039603960395

Chart, line chart

Description automatically generated

# Learning rate=0.001

Weights:

[[-0.68294396]

[-0.33670299]

[-0.27966735]

[-0.16135218]

[ 0.09446525]]

Training Stats

Accuracy = 0.9034167533818939

Precision = 0.9253996447602132

Recall = 0.9811676082862524

Loss = 27.009093689870184

Fscore = 0.9524680073126143

Test Stats :

Accuracy = 0.9464720194647201

Precision = 0.9196787148594378

Recall = 0.9913419913419913

Loss = 11.917713988300948

Fscore = 0.9541666666666668

Line chart

Description automatically generated