FML : Assignment 01 Report

Department of Computer Science

Indian Institute of Technology Bombay

Guided by

Prof. Sunita Sarawagi

Foundation of Machine Learning

CS 725

Prepared By

Bharat Patidar
23M0761

Rohit Singh Yadav
23M0773

Report :

In this Exercise/Assignment we have learnt the training of Logistic Regression Model and Linear Classification Model.

The optimal value of accuracy depends on the following three parameters.

1. Learning Rate

2. Epochs

3. Momentum

The below table states the different parameter value for optimal accuracy :

	Logistic Regression	Linear Classification
Learning Rate	0.1	0.1
Epochs	1000	250
Momentum	0.9	0
Accuracy(%)	85.33%	96%
Loss	4.35	7.65

Report of Logistic Regression

In the Logistic Regression Model the Dataset is taken from https://github.com/ashutoshbsathe/cs725-hw . In this model the binary dataset is used where the number of classes were 2 also the number of features of dataset is also 2.

We have implemented functions : calculate_loss(), calculate_gradient(), update_weights() and get_prediction() functions. Also implemeted preprocess() and sigmoid() functions for model training purpose.

In this report we have to analyse the effect of the three parameters on the model's accuracy and loss. As the parameters are :

1. Learning Rate

2. Epochs

3. Momentum

Effect of Learning Rate on Loss and Accuracy :

As the Learning rate of the model increases the step size or the gradient loss increases hence , the rate of change of learning rate directly affects the accuracy . Also the epochs is also a important parameter while finding the effeicient value of learning rate.

As in the calculation we have taken the epochs as 100 initially for calculating learning rate.

Epochs = 100

Momentum = 0

Learning Rate	Accuracy (%)	Loss
1e-1	82.67%	2.25
1e-2	77.33%	0.91
1e-4	77.33%	0.69
1e-6	77.33%	0.69

Conclusion :

As the Accuracy is most optimal when we use Learning Rate = 1e-1 = 0.1 . As we decrease the size of the Learning Rate alpha the accuracy decreases . But with the decrease in learning rate the loss also decreases.

Hence , the most efficient Learning Rate is 0.1 as it gives the accuracy 82.67%.

Effect of # of Epochs on Loss and Accuracy :

As the Epochs of the model increases the number of iterations increases . hence , the rate of change of Epochs directly affects the accuracy . # of Epochs states that how much time the model will search for the minima or in search of valley the model will how much effort. Also the epochs is also a important parameter while finding the effeicient value of learning rate.

As in the calculation we have taken the Learning Rate as .0.1 for calculating # of Epochs.

 Learning Rate = .0.1

Momentum = 0

# of Epochs	Accuracy (%)	Loss
100	82.67%	2.25
250	84%	3.03
500	84%	3.66
1000	85.33%	4.35

Conclusion :

As the Accuracy is most optimal when we use Learning Rate = 1e-1 = 0.1 and use the # of Epochs = 1000 . As we increase the number of the Epochs the accuracy increases . But the increase in the Epochs causes the increase in loss.

Hence , the most efficient Learning Rate is 0.1 and Number of Epochs is 1000.

Effect of Momentum on Loss and Accuracy :

As the Momentum states that gradients accumulated from past iterations will push the cost further to move around a saddle point even when the current gradient is negligible or zero. It is better when we use gradient descent .

As in the calculation we have taken the Learning Rate as 0.1 And # of Epochs 1000 for calculating # of Epochs.

Learning Rate = 0.1

# of Epochs = 1000

Momentum	Accuracy(%)	Loss
0	85.33%	4.35
0.9	85.33%	4.35

Conclusion :

As the Accuracy is most optimal when we use Learning Rate = 1e-1 = 0.1 and use the # of Epochs = 1000 . As we increase the momentum the accuracy increases . But the increase in the Momentum causes the increase in loss.
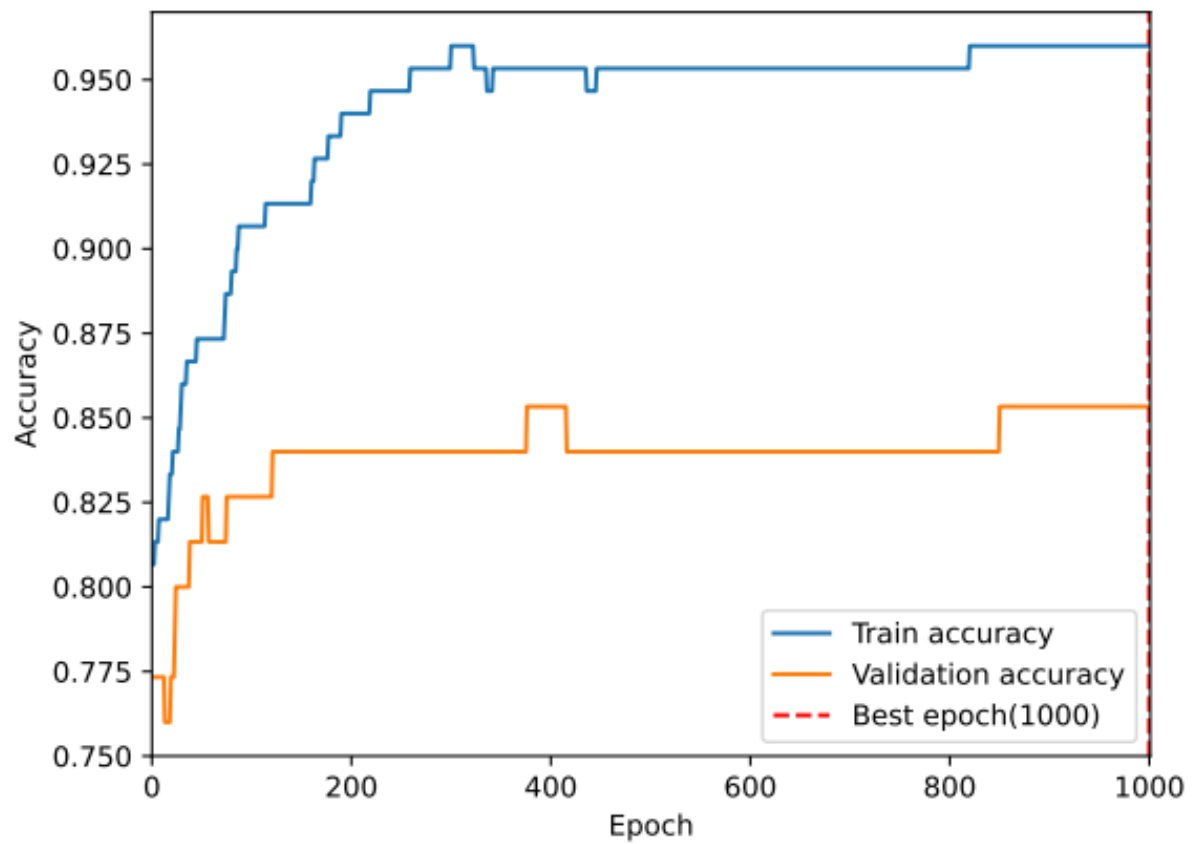
Hence , the most efficient Learning Rate is 0.1 and Number of Epochs is 1000 and Momentum is 0.9 .
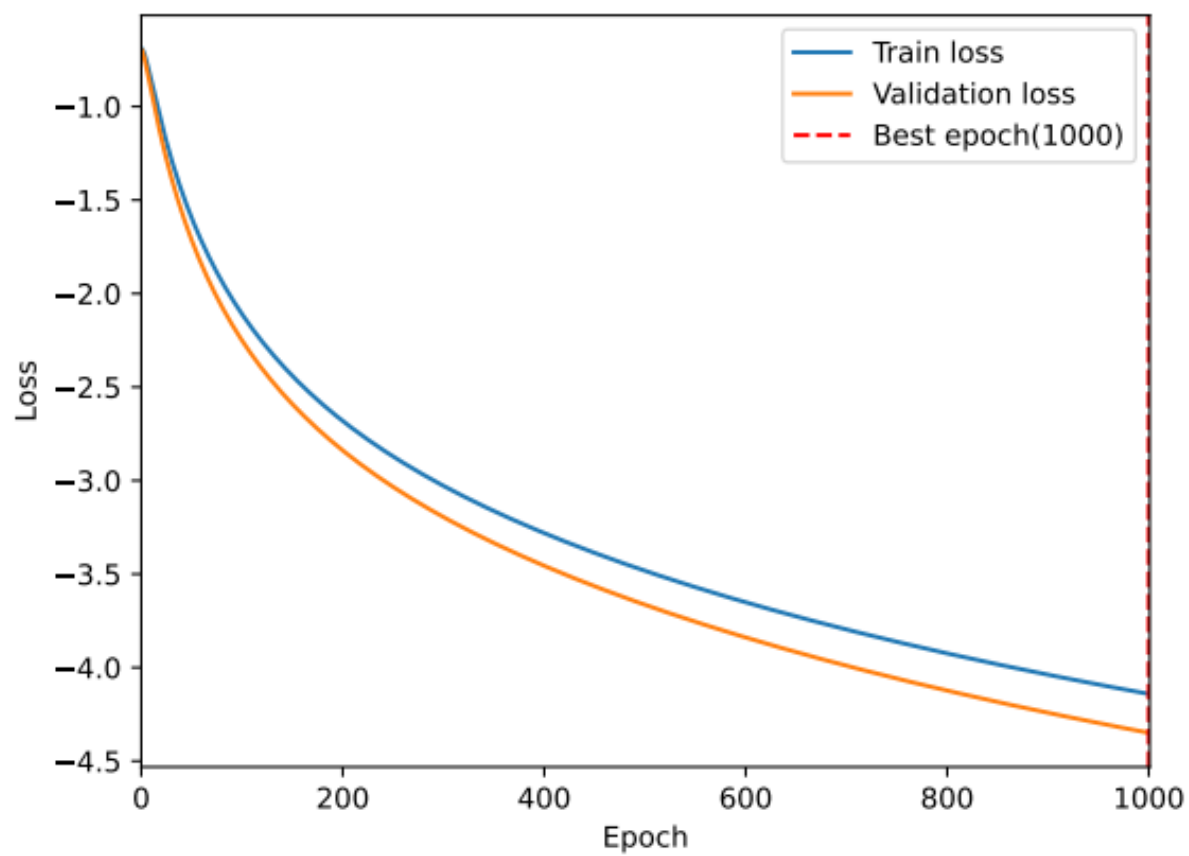
Graph :

Learning Rate = 0.1

# of Epochs = 250

Momentum = 0

Report of Linear Classifier

In the Linear Classifier Model the Dataset is taken from https://github.com/ashutoshbsathe/cs725-hw . In this model the Iris dataset is used where the number of classes were 3 also the number of features of dataset is also 4.

We have implemented functions : calculate_loss(), calculate_gradient(), update_weights() and get_prediction() functions. Also implemeted preprocess() and sigmoid() functions for model training purpose.

In this report we have to analyse the effect of the three parameters on the model's accuracy and loss. As the parameters are :

1. Learning Rate

2. Epochs

3. Momentum

Effect of Learning Rate on Loss and Accuracy :

As the Learning rate of the model increases the step size or the gradient loss increases hence , the rate of change of learning rate directly affects the accuracy . Also the epochs is also a important parameter while finding the effeicient value of learning rate.

As in the calculation we have taken the epochs as 100 initially for calculating learning rate.

Epochs = 100

Momentum = 0

Learning Rate	Accuracy (%)	Loss
1e-1	96%	6.23
1e-2	80%	4.10
1e-4	40%	2.56
1e-6	40%	2.41

Conclusion :

As the Accuracy is most optimal when we use Learning Rate = 1e-1 = 0.1 . As we decrease the size of the Learning Rate alpha the accuracy decreases . But with the decrease in learning rate the loss also decreases.

Hence , the most efficient Learning Rate is **0.1**

Effect of # of Epochs on Loss and Accuracy :

As the Epochs of the model increases the number of iterations increases . hence , the rate of change of Epochs directly affects the accuracy . # of Epochs states that how much time the model will search for the minima or in search of valley the model will how much effort. Also the epochs is also a important parameter while finding the effeicient value of learning rate.

As in the calculation we have taken the Learning Rate as 0.1 for calculating # of Epochs.

 Learning Rate = 0.1

Momentum = 0

# of Epochs	Accuracy (%)	Loss
100	96%	6.23
250	100%	7.65
500	100%	11.23
1000	100%	13.26

Conclusion :

As the Accuracy is most optimal when we use Learning Rate = 1e-1 = 0.1 and use the # of Epochs = 250 . As we increase the number of the Epochs the accuracy increases . But after a

thresold value the accuracy remains constant. But the increase in the Epochs causes the increase in loss.

Hence , the most efficient Learning Rate is 0.1 and Number of Epochs is 250.

Effect of Momentum on Loss and Accuracy :

As the Momentum states that gradients accumulated from past iterations will push the cost further to move around a saddle point even when the current gradient is negligible or zero. It is better when we use gradient descent .

As in the calculation we have taken the Learning Rate as 0.1 And # of Epochs 250 for calculating Momentum.

Learning Rate = 0.1

# of Epochs = 250

Momentum	Accuracy(%)	Loss
0	100%	7.65
0.9	100%	13.53

Conclusion :

As the Accuracy is most optimal when we use Learning Rate = 1e-1 = 0.1 and use the # of Epochs = 250 . As we increase the momentum the accuracy increases . But the increase in the Momentum causes the increase in loss.

Hence , the most efficient Learning Rate is 0.1 and Number of Epochs is 250 and Momentum is 0.

Graph :

Learning Rate = 0.1

# of Epochs = 250

Momentum = 0