

CSE 6363 Machine Learning Assignment 2

Tirumala Manukonda(1001662386)
Department of Computer Science
University of Texas Arlington
Texas, TX 76019
tirumalareddy.manukonda@mavs.uta.edu

November 5, 2020

1 Problem 1 :- Implementing logistic regression using sigmoid function and cross entropy

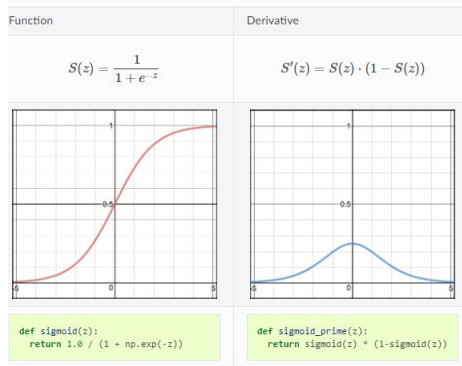
1.1 Sigmoid and Loss Function

Logistic regression model using sigmoid function and logloss function
Logistic Regression is a Machine Learning classification algorithm that is used to predict the probability of a categorical dependent variable. In logistic regression, the dependent variable is a binary variable that contains data coded as 1 (yes, success, etc.) or 0 (no, failure, etc.). In other words, the logistic regression model predicts $P(Y=1)$ as a function of X .

a). Sigmoid Function:-

Sigmoid takes a real value as input and outputs another value between 0 and 1

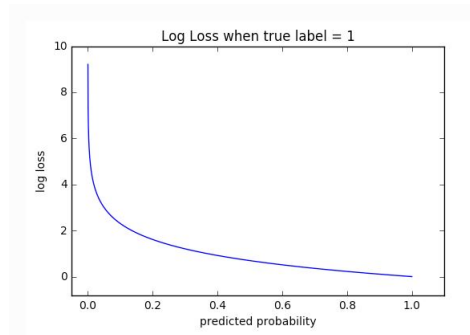
$$S(z) = 1 / (1 + e^{-z}) \quad (1)$$



b). Logloss Function:-

Cross-entropy loss, or log loss, measures the performance of a classification model whose output is a probability value between 0 and 1. Cross-entropy loss increases as the predicted probability diverges from the actual label

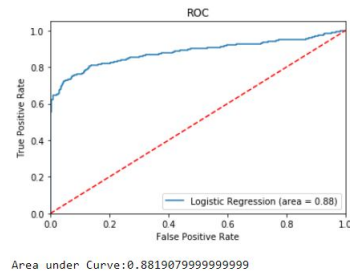
$$(y \log(p) + (1-y) \log(1-p)) \quad (2)$$



1.2 Analysis from the Model Implementation

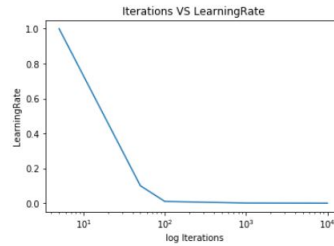
a). The Accuracy of the classification is around 0.84, which is says that the model is doing good predictions even-though its not the best.

b). ROC is the graphical plot between True Positive Rate(TPR) and False Positive Rate (FPR), my Roc Curve has AUC area under curve which is around 88 percent that says that the model has 0.88 probability of predictions as correct are under it. When its 1 we say it as the ideal model and anything less than 0.5 we can say its a model. I could say this is a better model even-though its not best.



c). The graph between the no of iterations and learning rate shows that as the no of iterations increases the learning rate decreases and both of them are

inversely proportional.

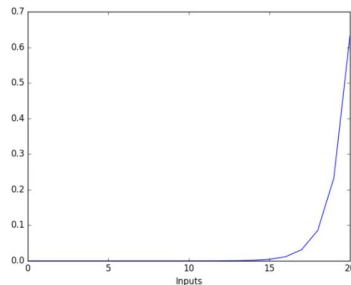


2 Problem 2 :- Implementation multi-class logistic regression using a soft-max function and cross entropy

Softmax Regression is a form of logistic regression that normalizes an input value into a vector of values that follows a probability distribution whose total sums up to 1.

2.1 Softmax Function and Cross Entropy

a). Softmax Function:- Softmax function calculates the probabilities distribution of the event over 'n' different events, this function will calculate the probabilities of each target class over all possible target classes. Later the calculated probabilities will be helpful for determining the target class for the given inputs. The main advantage of using Softmax is the output probabilities range. The range will 0 to 1, and the sum of all the probabilities will be equal to one.



b). Cross Entropy:-

If the classifier outputs 1 for the training example, then the cost is zero. The cost increases exponentially as the classifier's output decreases towards 0.

2.2 Analysis from the Model Implementation

My Multiclass classification looking good with the good predictions of individual class labels 0,1,2,3,4 and overall good accuracy of 97 , precision and recall also 97 (percent). Which i could say 97 percent of the over classifications of the model are correct.

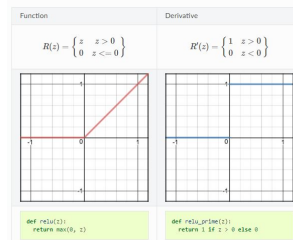
	precision	recall	f1-score	support
0	0.99	0.99	0.99	980
1	0.99	0.99	0.99	1135
2	0.96	0.94	0.95	1032
3	0.95	0.97	0.96	1010
4	0.98	0.97	0.98	982
accuracy			0.97	5139
macro avg	0.97	0.97	0.97	5139
weighted avg	0.97	0.97	0.97	5139

3 Problem 3 :- Implementation of Artificial Neural Network Classifier using existing libraries to classify the MNIST dataset

A multilayer perceptron (MLP) is a feedforward artificial neural network that generates a set of outputs from a set of inputs. An MLP is characterized by several layers of input nodes connected as a directed graph between the input and output layers. MLP uses backpropagation for training the network

3.1 RELU Activation Function and Sigmoid Activation

a). RELU Activation Function:- The formula is deceptively simple: $\max(0, z)$



b). Sigmoid Function has discussed in 1.1.a

3.2 Analysis from the Model Implementation

:-

a).Model Implementation using RELU :- From the classification report, it gives Accuracy of 97 and precision and recall as 97, which says that my model is doing good with RELU Activation Function

	precision	recall	f1-score	support
0	0.98	0.98	0.98	980
1	0.98	0.98	0.98	1135
2	0.96	0.94	0.95	1032
3	0.97	0.97	0.97	1010
4	0.97	0.98	0.97	982
accuracy			0.97	5139
macro avg	0.97	0.97	0.97	5139
weighted avg	0.97	0.97	0.97	5139

b).Model Implementation using Sigmoid :-

From the above classification report, it gives Accuracy of 99 and precision and recall as 99, which says that my model is doing good and better than the RELU Activation Function

	precision	recall	f1-score	support
0	0.99	0.99	0.99	980
1	1.00	0.99	0.99	1135
2	0.98	0.98	0.98	1032
3	0.99	0.99	0.99	1010
4	0.99	0.99	0.99	982
accuracy			0.99	5139
macro avg	0.99	0.99	0.99	5139
weighted avg	0.99	0.99	0.99	5139

4 Comparing Problem 2 and Problem 3

Looking at the Model classification reports of Multi-class logistic Regression and ANN Report, Overall ANN does better comparing that it has 99 accuracy with the Sigmoid activation function, Even-though it has same accuracy with RELU, but still ANN is better at predicting the models and its faster also, Softmax regression is slow with large amount of data.

References

I have referred the following documents

- [1] Activation Functions - From ML Glossary
- [2] Loss Functions - From ML Glossary