

-Submitted by Abhishek Kumar Shastry

Learning rate is a parameter to control how fast the model learns the data. A high learning rate might make the model extremely sensitive and can even fail to converge during gradient descent. On the other hand a low learning rate will result in painfully slow training of the model and also consuming more resources.

With learning rate of .1 logistic regression learns the data fairly well in less than 10 epochs.

'Plasma glucose concentration a 2hrs in an oral glucose tolerance test' is the best predictor of diabetes.

'No. of times pregnant' is the worst predictor of diabetes.

Both of the above were found by checking the weights of the trained model whose features had been normalized before training. In a normalized model a feature with high weight magnitude will saturate the sigmoid function to either 0 or 1 more quickly than others and hence is a good predictor of either class.

One advantage of the perceptron is that it is faster than logistic regression to train since it does not use the sigmoid function.

Normalize function has been implemented in main.py Feature normalization helps to bring all the feature values to the same comparable range. If features are not normalized then the feature having higher values will dominate(at least in the beginning of training) and model prediction might be biased toward that feature.