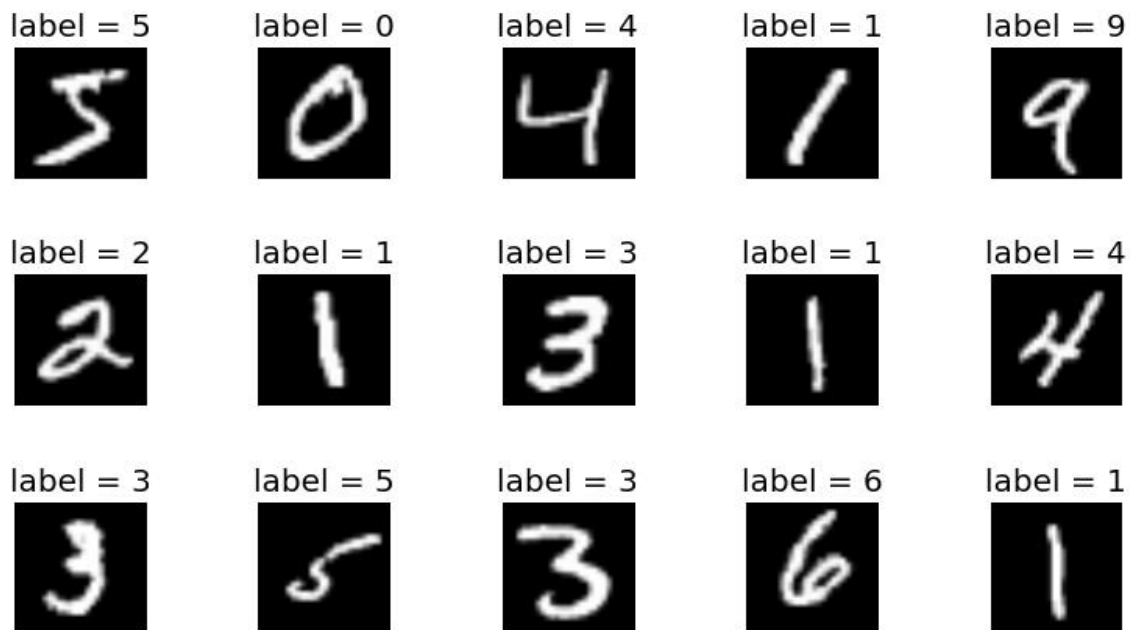


## Project HW2 - Analysis

The “mnist” data is a popular dataset for multiclass classification Problem. Which contains hand written digits. Here’s some examples of what mnist data looks like:



## Softmax Classification

Each digit is classified between 0 to 9 (10 classes). In softmax we need to convert labels into shape of “one hot vector”.

After applying softmax and getting the probabilities, weights are updated simultaneously using gradient descent.

The training data contains 60,000 examples and it is more efficient to batch data for training in this case.



Since all algorithm are written without using any ready frameworks, it might not be the most efficient version of softmax classification but it has given some good results.

The accuracy for softmax classification is **0.9092** which is quite impressive for only 50 iterations.

### One Vs All Classification

The other method however uses sigmoid classification which is for binary classifications, and it partials data for each class.

We flagged each class one time and made it positive, and the rest of the data negative. Doing it for 10 times, it gives us partial weights which should be united for prediction tests, and with choosing maximum arguments between weights of classes, the final model is ready to predict!

This method however, didn't get a higher result from softmax and the accuracy was **0.8991** and it was three times slower than softmax (under same circumstances) which makes sense since this method has to train 10 models individually though softmax only does it once.



## Summery

One vs all is easier to understand and implement. And it's a good way to start multiclass classification.

It also is independent of how many classes dataset has, while softmax performance can be effected by it.

Overall, they both are good methods of classification and they each performs better than the other under different circumstances.

## Sources

<https://ljvmiranda921.github.io/notebook/2017/08/13/softmax-and-the-negative-log-likelihood/>

<https://towardsdatascience.com/softmax-function-simplified-714068bf8156>

<https://towardsdatascience.com/multi-class-classification-one-vs-all-one-vs-one-94daed32a87b>

<https://www.coursera.org/lecture/machine-learning/multiclass-classification-one-vs-all-68Pol>

<https://regenerativetoday.com/multiclass-classification-with-logistic-regression-one-vs-all-method-from-scratch-using-python/>

<https://medium.com/data-science-bootcamp/understand-the-softmax-function-in-minutes-f3a59641e86d>

<https://www.coursera.org/lecture/deep-neural-network/training-a-softmax-classifier-LCsCH?authMode=signup&redirectTo=%2Flearn%2Fdeep-neural-network%3Faction%3Denroll>

