Project 5 – Handwritten Digits

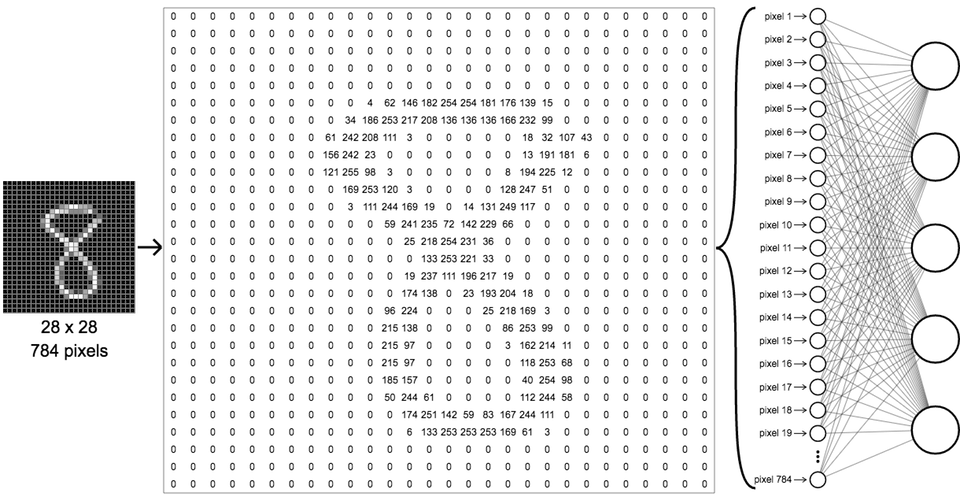
# MNIST

MNIST ("Modified National Institute of Standards and Technology") is the de facto “hello world” dataset of computer vision. Since its release in 1999, this classic dataset of handwritten images has served as the basis for benchmarking classification algorithms. As new machine learning techniques emerge, MNIST remains a reliable resource for researchers and learners alike.

In this project, your goal is to correctly identify digits from a dataset of tens of thousands of handwritten images. This week your job is to identify hand written digits from 0 – 9, this is a multi-classification problem. This should be a less arduous challenge than previous weeks however, you’ll need to be diligent in your coding, and like always there is a bar to meet on Kaggle.

You have been provided two CSV files:

* Digits\_train.csv
* Digits\_holdout.csv



## Tasks

Your job is to build and tune **three** different models at least one needs to be a neural network MLP others can be XGboost, RF or other method of your choice; document their performance using accuracy as our metric; and visualize examples where they disagree with your prediction on your test set. Unlike our other projects you’ll need to craft a document comparing the three different model’s performance and what you did to tune the hyper parameters and have examples of each model’s failure to predict.

***This project is worth 20 points***

Exceeding the bar

* Report should be professionally written and visually appealing
* ~~You’ve used K-Fold CV, documented grid tuning of hyperparameters for each of all three models and have enough experiments to justify exceeding the bar.~~
* ~~Organized so that it’s easy to compare the performance of the models.~~
* You have used a function or two to evaluate your model and produce and visualize a handful of examples of where your model failed to predict on test set.
* ~~Performance of your model significantly exceeds the benchmark~~

Meeting the bar -10 to -5

* Report should be professionally written and visually appealing
* Organized so that it’s easy to compare the performance of all three models
* Examples of each model’s failure to predict on test set
* Exceeds the benchmark

Below the bar

* Report should be professionally written
* Examples of each model’s failure to predict test set
* meets the benchmark

## What to turn in

* Model Report – I leave it up to you and your creativity the key is making it easy to compare the differences in models and explain how you chose your hyper parameters, if you didn’t use k-fold grid search (which is fine) can you compare accuracy impact on different parameters?
* Code, did you write a function or two?
* Kaggle submission

Here is an example of the predictions being wrong, 1 predicted as an 8 a 2 predicted as a 3 etc.

