BASIC IMAGE CLASSIFICATION – MNIST DATASET

(Of handwritten numbers)

Background

The main problem that we are trying to solve here would be to identify handwritten digits. This can be made used in banks for a machine to read through a cheque to avoid the necessity for human presence to verify the transaction, this way the process will also be fast and not require any additional costs for the bank. It could also be used by police to get the number plate digits of fast paced cars.

Problem Definition

Handwritten digits can be easily identified by the human brain. Being able to identify numbers of different handwritings is a natural thing for us humans. Handwritten digits can be present on cheques, letters, to name a few. A machine cannot directly recognize handwritten numbers as it can recognize typed out numbers. Because handwritten digits looks different and can vary from person to person based in their handwriting. This is a linear search problem where once the shape of the digit is recognized it searches through what it’s learnt to give an output. Optimization would be done on the data by normalizing it.

Data: An image of a handwritten number

The dataset which we will be using is the famous MNIST (Modified National Institute of Standards and Technology) dataset. It is a dataset of handwritten digits that is commonly used for training various image processing systems. This dataset contains over 60000 image data of handwritten numbers. Each image is of the dimension 28x28.

State: State space search is the process in which consecutive states (configurations) of an instance are considered in order to find the goal state with the desired property.

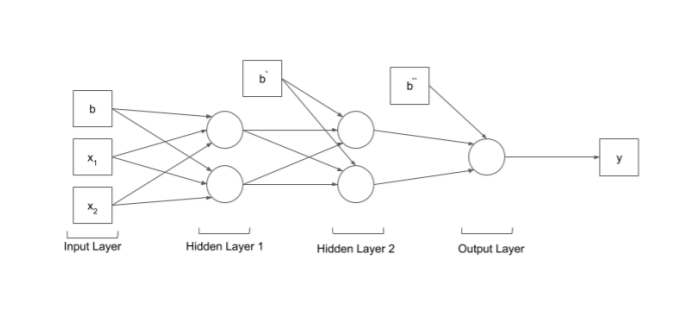
* Initial State: The image data of the handwritten data
* Goal test: The goal is just the model identifying the handwritten digit

Complexity

The state space Neural Network comprises two main advantages: Being a neural model, it has the flexibility to represent any linear or non-linear function. Being a state space model, the number of outside connections is minimum.  
Here in this project, the size is linearly dependent on N which is the number of images in the training dataset.

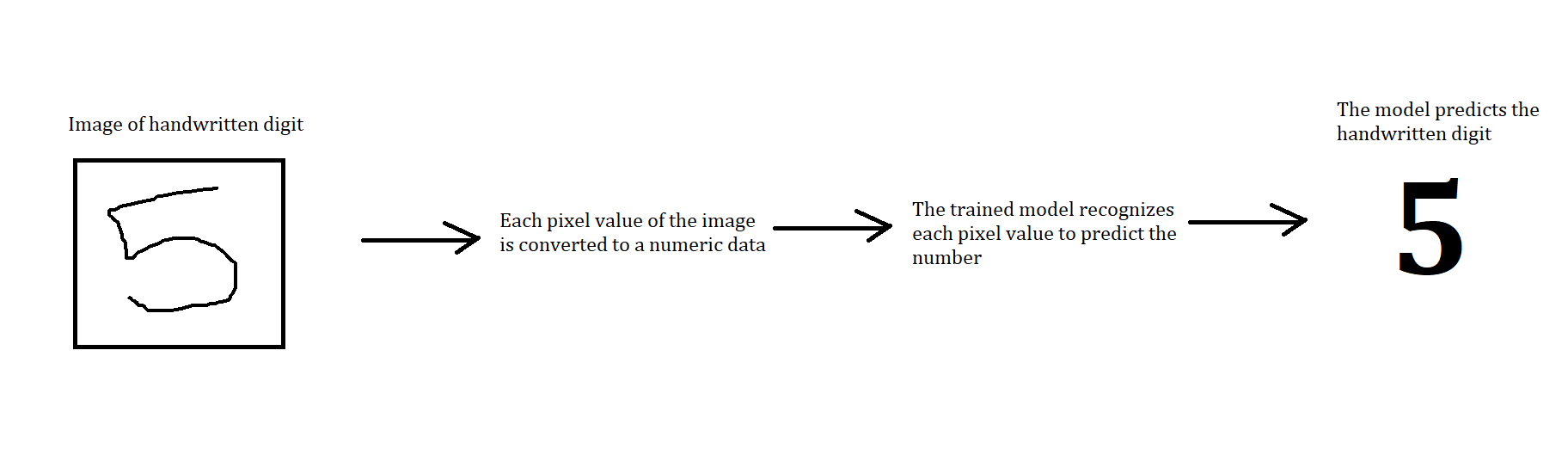
Algorithm

The algorithm that will best be suited for this particular study would be neural networks. The main aim of a neural network is to recognize relationships in data. Much like how the human nervous system works which is a set of neurons along with the brain in place to study data. Neural network finds a correlation between the given input and output. Here we will be cascading the functions together, which are the hidden layer and provide an activation function to the linear output. The purpose of an activation functions is to help the neural network find non-linear patterns in the data. The activation function gives the model flexibility.

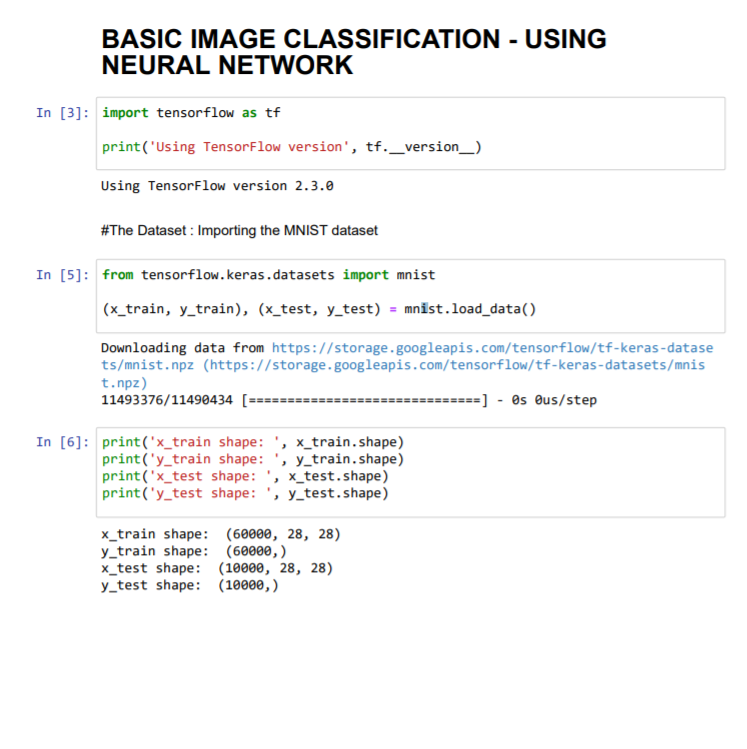
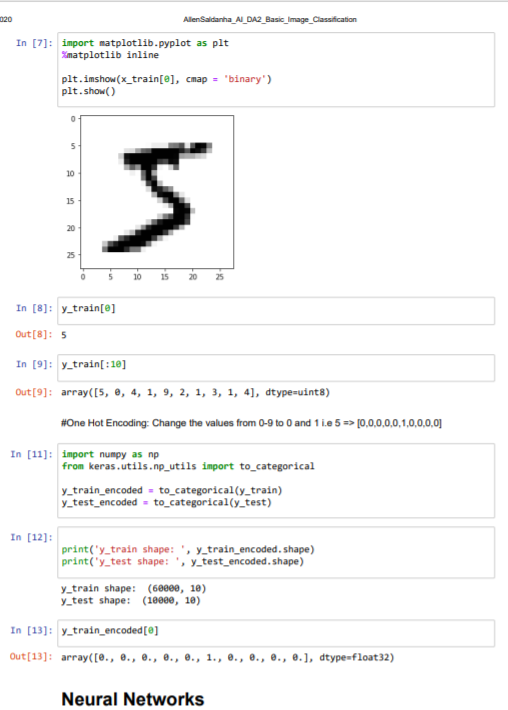


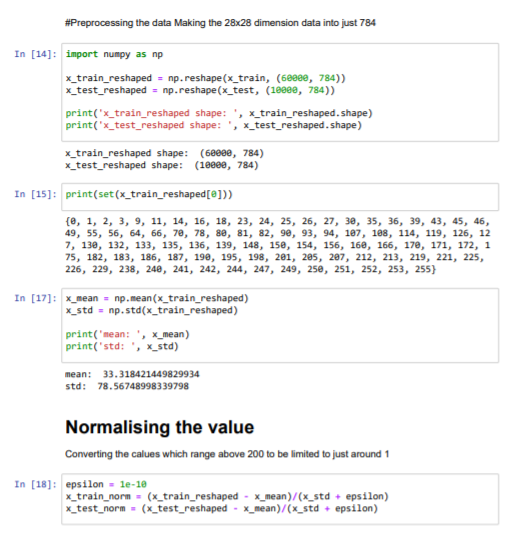
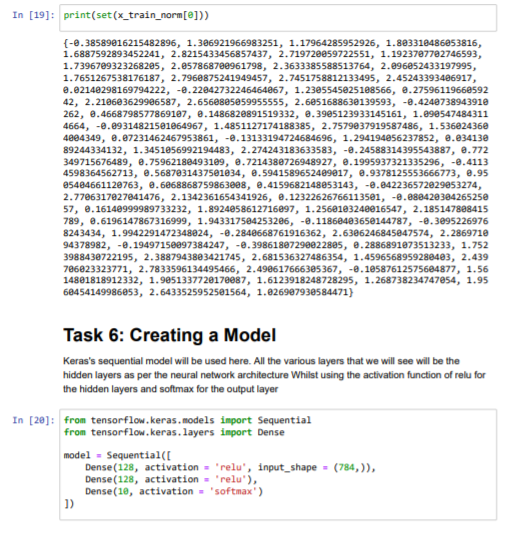
This will be the structure of our model, the number of nodes shown here will vary as the number and complexity of our data increases. A dense layer will be used which is neural network layer in which all the nodes are connected to all the nodes of the preceding layer. The model will consist of one input layer, one output layer and 2 hidden layers. We will not be modifying or adding anything extra to this algorithm. But we will be optimizing our data by normalizing it. There are no issues in implementing this algorithm.

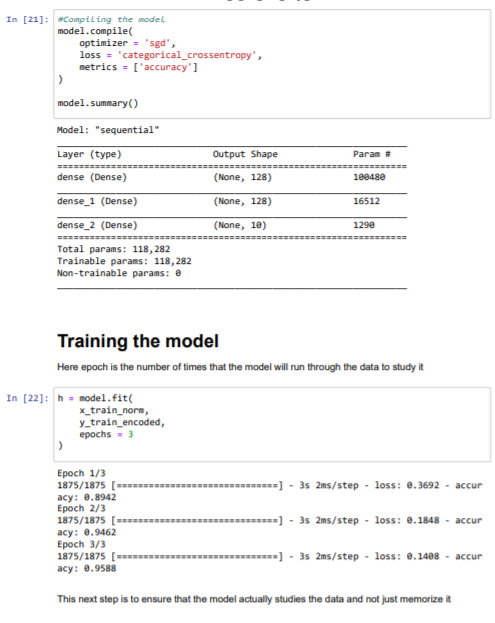
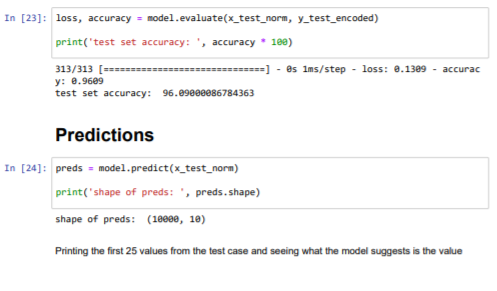
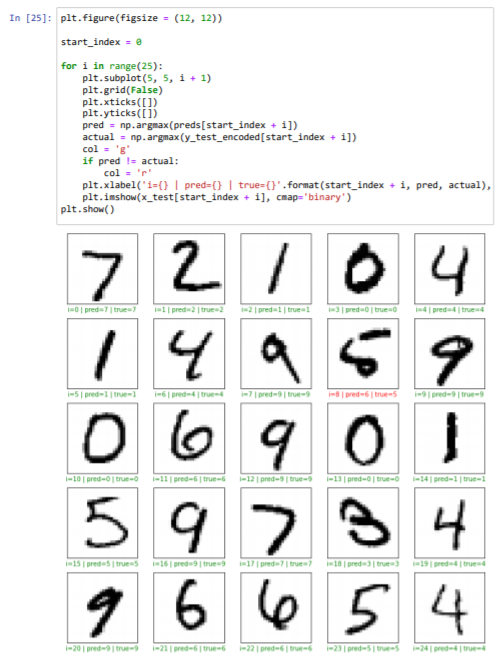
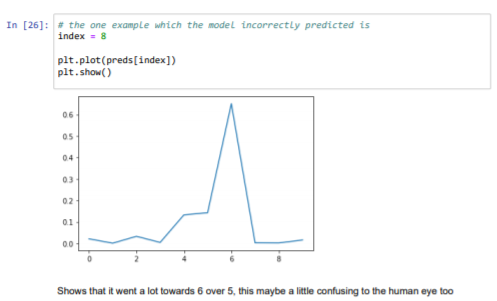
Worked Example



Code

Summary

The conclusions that we got were very positive. The model provides an accuracy of over 95% which is pretty good as it shows the efficiency of the model too. The model is feasible as long as the user input of the handwritten number is legible and understandable by humans as well. As we see from the above 25 examples that we tried out we might not always get the right answer. Creating the model was the hard part, as determining the number of layers required would vary at times, and hence to make the model fast and accurate we need to get a feasible number for the number of layers. And understanding the dataset before creating the model was the most important insight that was required for this project.

References

1. [Behnke, 2003] Sven Behnke. Hierarchical Neural Networks for Image Interpretation, volume 2766 of Lecture
2. Notes in Computer Science. Springer, 2003MNIST classifier, "Classify MNIST digits using Convolutional Neural Networks"
3. Romanuke, Vadim. "The single convolutional neural network best performance in 18 epochs on the expanded training data at Parallel Computing Center, Khmelnitskiy,
4. [Chellapilla et al., 2006] Kumar Chellapilla, Sidd Puri, and
5. Patrice Simard. High performance convolutional neuralnetworks for document processing. In International Workshop on Frontiers in Handwriting Recognition, 2006.
6. [Ciresan et al., 2010] Dan C. Ciresan, Ueli Meier, Luca M.Gambardella, and Jurgen Schmidhuber. Deep big simple ¨neural nets for handwritten digit recognition. Neural Computation, 22(12):3207–3220, 2010.
7. [Coates et al., 2010] Adam Coates, Honglak Lee, and Andrew Ng. An analysis of single-layer networks in unsupervised feature learning. In Advances in Neural Information Processing Systems, 2010.
8. [Olshausen and Field, 1997] Bruno A. Olshausen and David J. Field. Sparse coding with an overcomplete basis set: A strategy employed by V1? Vision Research, 37(23):3311–3325, December 1997.