Dalya Adams
MSDS 422 section 59
MNST Classification

Up until recently, the computing power needed to process large datasets was extremely expensive. This meant that if an organization wanted to analyze large amounts of data at a reasonable price, the dataset needed to be trimmed to decrease the time used to analyze the dataset. As computing power has become cheaper and more accessible, the reemergence of neural networks and deep learning has changed the machine learning game. These models are highly accurate and when processed using GPUS, can provide insight in a timely manner. When accuracy is paramount, neural networks are often the best solution to the problem.

The financial industry is one that is often faced with a lot of uncertainty. The number of different micro and macro factors which impact the market are innumerable. While machine learning, and specifically neural nets, might not be the best solution to predicting what a complex market might do, it can help with ensuring that less ambiguous parts of the sector are highly accurate. In this project, comparing neural nets for optical character recognition provides acts as a great gateway to utilizing character recognition and image processing in the financial sector. An issue that is encountered when utilizing neural nets is the lack of interpretability, which makes the usage ideal of areas where the reasoning behind the prediction is not as necessary. This makes image classification a perfect example of neural nets in action for a business setting.

The purpose of this research is to determine if the use of neural nets decreases the time to actionable insight on a large dataset while increasing the accuracy of predictions. In reviewing the dataset, it is composed of 70,000 observations and 784 variables. These variables are numerical attributes of each pixel making up the 28x 28 image for each observation. Each observation is a handwritten number between 0 and 9, with some being very clear representations of the number written and others being less clear, as would be expected when looking at the handwriting of different people. When viewing the descriptive stats of a subset of

the 784 variable, we immediately notice that the scale of the variables differs greatly. This extreme scale, ranging from a max value of 0 to 255, can decrease the effectiveness of certain machine learning models and deep learning algorithms.

In order to compare the processing time and accuracy of different neural nets and hyperparameter settings, we first utilize a Deep Neural Net (DNN) Classifier. After testing the impact on accuracy and processing time of different numbers of layers and nodes per layer, the data was scaled using a standard scaler and the DNN which balanced accuracy and processing time was reevaluated using the scaled data. Finally, a Convolution Neural Net (CNN) was attempted. This model has a large number of different layers, to include: an input layer, 2 convergence layers, 2 pooling layers, 2 normalization layers, 2 drop layers and 3 connected layers. The processing time and accuracy of the models attempted can be found in the Appendix of this paper.

For this dataset, the use of DNNs is superior to the use of CNNs, as judged by both accuracy and the processing time. The first DNN had 2 layers with 300 nodes on the first layer and 100 nodes on the second layer. This model was the model which balanced accuracy and processing time the best. This model has an accuracy of .9828 and a processing time of 137.6 seconds. By scaling the data prior to running the 2-layer DNN, accuracy increased from .9828 to .983. Processing time also increased to 157.65. The most accurate model was a DNN with 4 layers, with 500, 300, 100, 50 nodes, which took 350 seconds to run. The CNN had a decreased accuracy in comparison to the DNN (.9648) as well as a processing time 10x longer than the longest running DNN. The use of DNNs on this dataset, with a decreasing number of nodes by layer, appears to be the most accurate. Increasing the number of layers appears to increase the accuracy of the model as well, but also the processing time.

Dalya Adams
MSDS 422 section 59
MNST Classification

Appendix

Model	Number of Layers	Nodes per Layer	Processing Time	Training Set Accuracy	Test Set Accuracy
DNN	2	300,100	137.61	1.0	.9828
DNN	2	300,300	262.65	1.0	.9829
DNN	3	100,100,100	109.00	1.0	.9815
DNN	4	500,300,100,50	349.95	1.0	.9846
Scaled DNN	2	300,100	157.65	1.0	.983
CNN	12	32, 64, 128, 256	3174.75	.9637	.9648