Week 6 Proposal

For the Final Research Proposal I have chosen Example 1 to Evaluate the design, modify, experiment with, interpret, assess, and refine a complex neural network. I have chosen the DNN from this article: *https://towardsdatascience.com/neural-networks-to-predict-the-market-c4861b649371*. This DNN created by Vivek Palaniappan, was created to use neural networks to predict the stock market, particularly focusing on the price of stocks or indexes. Vivek stated he chose finance because it is "highly nonlinear and sometimes stock price data can even seem completely random". He further states "The main issue arises in implementing these models in a live trading system, as there is no guarantee of stationarity as new data is added. This is combated by using neural networks, which do not require any stationarity to be used". Thus, the issue he is trying to focus on is how to predict future data patterns, when the given dataset can be so random and volatile.

Vivek chose to answer this question by creating neural networks because of natural effectiveness in finding relationships between data and using those relationships to predict future data. However, Vivek chose to create a Multilayer Perceptron (MLP) and a Long Short Term Model (LSTM). An issue arises with his latter selection of the LTSM model, and the topic that I have chosen to cover for the Final Project. As noted in the article https://towardsdatascience.com/the-fall-of-rnn-lstm-2d1594c74ce0, entitled "The fall of RNN / LSTM", the author Eugenio Culurciello notes that LSTMs are note the ideal model for creating and applying a DNN. He notes that although LSTM were able to remove some of the vanishing gradient problem that RNN had created, it was not able to remove all of them. The paths from point a to point b had become more complicated "because it has additive and forget branches attached to it". Instead, he proposes a 2D Convolutional Neural Networks for Sequence-to-Sequence Prediction and a model known as the Transformer. This is what my proposal is: to take the data and use case created by Vivek, and experiment with it using more relevant DNN mainly the 2D convulsion network and the Transformer.

In viewing the five categories for grading from the rubric, data preparation and exploration will likely be minimal since this is based upon a previous project. However, if outputs are inconsistent or needing manipulation, I will edit the pre-processing of the data. This process will all be part of evaluating the design of the model but using a new model to train upon the data. Further analysis and experimentation will be done with the number of hidden layers and nodes which are needed to produce the best results. I will expand the training data to include noise and variants and assess the classes and re-examine hidden nodes. Lastly, I will characterize the role and specific functions of the different hidden layers.

Both the 2d and transformer models are based upon literature written based upon experiments and thus I will focus on these scholarly articles in reviewing research designs and modeling. I will also examine other us3es cases to see how best to create the model and optimize results.

Once the correct models have been put in place, I can review results and make further evaluations of the models. I can further experiment with noise and hidden layers, adjusting both the data and the models itself. This will affect the implementation and programming, as a deeper understanding of how the algorithms are performing will be required to optimize results.

Lastly, I will discuss the problems with using the LSTM models, highlight the vanishing gradient problem, and discuss how the two models I created were able to alleviate or reduce this problem and increase accuracy. This will lead to a deeper analysis of sequential data, and a discussion of how the proposed models affect an examination of financial stock data. The final analysis will discuss how volatile and unpredictable data relates to the models used to predict, and which models created better or more accurate results. Although there is no issue with the data, as Vivek noted DNNs are suited for non-stationary data, it will be interesting to note if the

LSTM is not the best model to select, and if other models can overcome the vanishing gradient problem!