The problem that I am trying to solve is to create a concurrent real time system that uses the concepts of Artificial Intelligence to make predictions about the stock market. This report follows an algorithm design type report, thus many previous research papers will be used in order to make the algorithm correct and effective. The main problem is that these days trying to predict the market is quite impossible, this is a statement to which I partially agree with but through the use of artificial intelligence I believe that a system can be developed that would be able to predict whether a stock would go up or whether the stock would go down in the next day. From my knowledge, many of the existing techniques cannot effectively tackle this problem because there is a lot of noise that an individual must clear in order to find some information that is of value, in addition due to the fact that there are countless numbers of factors that are untraceable can majorly affect the stock. The solution to this problem will be discussed in the next section. The intuition behind this technique is that if one can make a system that will be able to process important information concurrently, and is real time then one can ultimately be able to make a system that is able to make predictions about the stock with at least a success rate that is above 50%.

A research paper titled “Stock Market Prediction System with Modular Neural Networks” by Takashi Kimoto and Kazuo Asakawa, is focused around the creation of a buying and selling timing prediction system for stocks on the Tokyo Stock Exchange. The system achieved an accurate prediction rate, and the simulation on the stock trading showed an excellent profit. The architecture of the system that has been created in the paper claims that it is made up of several neural networks that learned the relationship between various technical and economical indexes and the timing for when to buy or sell the stock. Their focus was to predict the best time to buy and sell for one month in the future. The results of the neural network had achieved a successful prediction rate that was above 50%. Another research paper titled “Stock Prediction using Artificial Neural Networks” by Abhishek Kar, dives into prediction indices of the stock market by the use of Neural Networks. In their architecture of the neural network there are a number of various activation functions that are implemented along with options for cross validation sets. Their neural network on had achieved a best case accuracy of 96% on test data. There architecture employs the use of concurrency to produce faster results along with having a multi-layer neural network. The last research paper that is investigated is titled “Time Series Forecasting using Neural Networks” by Bogdan Oancea, and Stefan Cristian Ciucu. Their work compares the performances of different feed forward and recurrent neural networks and training algorithms for predicting the exchange rate of EUR/RON and USD/RON. They use data series with daily exchange rates starting from 2005 to 2013. After investigating these research papers, it was clear that Neural Networks would be one of the tools that would be needed in order to create a prediction making system.

The technique that I have developed is created in 3 pieces. The first piece is setting up a database. This needs to be working for all the other steps to be working properly. The database is essential, because not only will this store massive amounts of data but it should also be able to update the data on a daily basis so that the Neural Network itself achieves “learning”. That being said, to dive into the specifics a MYSQL database should be set up with working stored procedures that should scrape data from websites that include relevant information about the stock. The stored procedures also need to be able to process all that information and output back a matrix that contains all the information about the data for the second step of this technique. The second step of this technique is to be able to process the data. This means that I will employ a decision tree to be able to pick out the first 1000 most relevant factors that are required for an accurate prediction. The decision tree will follow a process that is something similar to the tasks given to us in assignment 3 for this course. This assignment made us create a decision tree that would be able to classify documents, if they either belonged to the atheism category or the graphics category. Although it achieved a prediction rate of at least 80%, the most important concept that was taken out from this endeavor was the need to be able to pick important words from that distinguished where the article belonged. Similarly, in the second step of this technique, I need to figure out what factors are more important than others. By using decision tree, those factors could be determined. After having at least a 1000 factors that are relevant to the stock, they need to be formatted correctly. This means that a matrix must be developed that has at least a 1000 columns these columns would represent the factors that are associated with the stock and the rows of the matrix would show numerical data about how that specific factor is either rising on falling on that specific date. The third step of this would be to send the data matrix into the multilayered recurrent neural network. The neural network would be trained with at least 10 000 data sets and then would be investigated for its correctness. The output of this neural network should be a Boolean answer along with the probability of confidence. The real time system is also concurrent, meaning that as the factors that are identified will be updated every 30 seconds in the database, so while the neural network is processing the information, the database will be updating the data so everything will be up to date. A technique to test the whether or not the system is making correct predictions would be test it out on real live data, this means that the system will be trained with all the important information up till today, and then it will make a prediction about whether or not the stock will rise or fall tomorrow. By doing this several times and recording it we can find the level of correctness for the system.

This approach would be processing intensive, meaning that the computer will have to do many operations for this techniques to work. Given that limitation, I believe that this new technique could effectively tackle the problem of predicting stocks. Some future research recommendation that I would like is finding an optimized way of figuring out what are the most important associated factors. At the moment a decision tree is being used to find those factors, but I believe there is a better way for a prediction.