

A TUTORIAL ON APPLYING ARTIFICIAL NEURAL NETWORKS AND GEOMETRIC BROWNIAN MOTION TO PREDICT A STOCHASTIC TIME SERIES

Suresh Venugopal¹

¹ Application Engineer, Houston, Texas, USA
mpsureshnair@gmail.com

Abstract

Several challenges in the engineering or financial world can be resolved with a proper handle on data. Amongst other applications in engineering, system identification and parameter estimation are widely used in developing control strategies for automation. In this domain, there would be requirements to design an adaptive control system. In order to design an adaptive control system, an adaptive model needs to be estimated or identified. This is generally done by studying the data and creating a transfer function. In the process, regression, artificial neural networks (ANN), random walk theory and Markov chain estimates are used to understand a time series and create a model. While some of these processes are stationary, some are non-stationary. These methods are chosen based on the nature and availability of historical data. One of the issues that always remain is which method is appropriate for a certain application. The objective of this tutorial is to illustrate how artificial neural network and Geometric Brownian motion can be used in this regard. An attempt is made to predict the future price of a stock of a corporation. Stock prices are an example for a stochastic time series. Initially, an artificial neural network is used to predict the stock price. The network is designed as a Multi layer Back propagation type network. Profit over earnings and S&P are used as inputs. Thereafter, Geometric Brownian motion is explained and used on the same dataset to come up with its predictions. The results from both neural network and geometric Brownian motion are compared.

Key Words: Artificial Neural Network, Geometric Brownian Motion, Stochastic time series, and stock price prediction

1. INTRODUCTION

Broadly speaking, there are two forms of data. They are

1. Deterministic: As the name implies, this form of data can be completely and accurately estimated. The relationship between input and output is conclusively determined.

2. Stochastic: This form of data contains inherent randomness. By definition, a stochastic process is an infinite collection of random variables defined on a common probability model. These random variables are usually indexed by an integer or a real number such as time. Stochastic Process [7] is very important and has applications in engineering, economics and science etc. It is used to model various phenomena where the subject of study varies discretely or continuously with time in a non-predictable fashion. A wide variety of engineering measurements are stochastic. In order to develop a control system for a system, which follows a stochastic process, the system needs to be appropriately modeled. The same logic applies in finance. A stock market follows a stochastic process. In order to effectively predict the path of a stock price, stochastic time series needs to be studied and effectively modeled.

In this tutorial, artificial neural network (ANN) and Geometric Brownian Motion (GBM) are used on stock price data. The goal is to illustrate the different steps involved in

setting up a neural network or a geometric Brownian motion model.

1.1 Artificial Neural Network

Artificial Neural Network is a very well studied branch of Mathematics. ANNs are used for system identification, classification and prediction and is widely popular in the industry. The idea is to create a network that can learn from the past data and use it to predict future data. The success or failure of the network relies on the efficiency of the network to use the information that it has learned.

An artificial neural network [5][3] is a network composed of carefully crafted layers of neurons that relate inputs of a system to its outputs. This pattern matching or training of the network is performed in a supervised learning environment. Once trained, the network is used to predict outputs. In that sense, all neural network does is to estimate the transfer function between input and output. As long as the transfer function and its parameters are valid, prediction will be reliable. Conversely, if the network is used in real time for control purposes in a rapidly changing environment, the network will have to be trained more frequently.

Inputs are combined with weights using an activation function to connect to the intermediate layer. The intermediate layer is then connected to the output layer

using another activation function and the weights for that connection. Initially, these weights are chosen at random. Then, as the iteration proceeds, these weights are adjusted to minimize the error function. Figure 1 shows an example of the network.

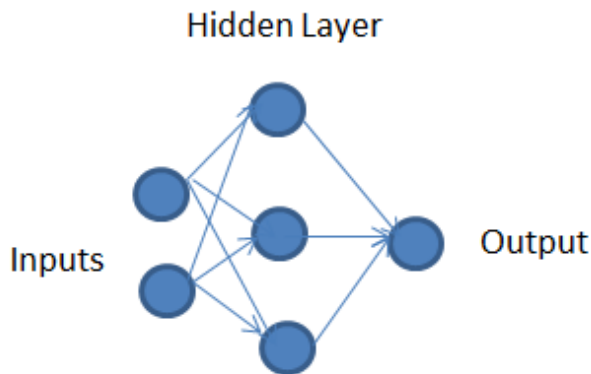


Fig-1: Multilayer Neural Network With One Hidden Layer, Two Input Neurons And One Output Neuron

1.2 Geometric Brownian Motion

Geometric Brownian [4][6] is used to study stochastic data. A stochastic process X_t is said to follow Geometric Brownian Motion only if it satisfies the following differential equation

$$dX_t = uX_t dt + \sigma X_t dW_t$$

Where u is the drift term, σ is the volatility term and W_t is a Wiener process. Thus if drift and volatility terms are known, geometric Brownian motion solution can be produced throughout the time interval. The degree of randomness is decided by the volatility. After a model based on GBM is setup, it can then generate random trials out of this model. Monte Carlo simulation is done at this point to predict the future value many times. This sample of future values can help us understand where the stock price is headed.

2. PROBLEM INTRODUCTION

In this study, the objective is to predict the stock price of a corporation and also provide the steps involved in the process. As discussed in the previous section, Artificial Neural Network and Geometric Brownian Motion concepts would be independently tested. Available data include

1. Stock Price Data of the corporation from May 2013 to Dec 2014
2. Price Over Earnings of the corporation from May 2013 to Dec 2014
3. S&P from May 2013 to Dec2014

Next section sets up the problem with the ANN approach. Thereafter, the result of ANN approach is discussed. After

that, the problem is setup to use Geometric Brownian Motion. Results and discussions are followed after that.

2.1 Artificial Neural Network Approach

The problem is to attempt to predict stock prices of an XYZ corporation using historical information available. Figures 2,3 and 4 show the training data. P/E and S&P data are going to be the inputs and stock price is going to be the actual target data. The choice of those inputs comes from the fact that P/E reflects the company performance and S&P reflects general economy.

A back propagation multi-layer type network [1] is used in the study. The network has 2 inputs, 1 hidden layer with 6 neurons and 1 output neuron. A sigmoid is used as the activation function. The choices for above mentioned parameters were based on trial and error. The steps [2] [3] for successful creation of the network and training include:

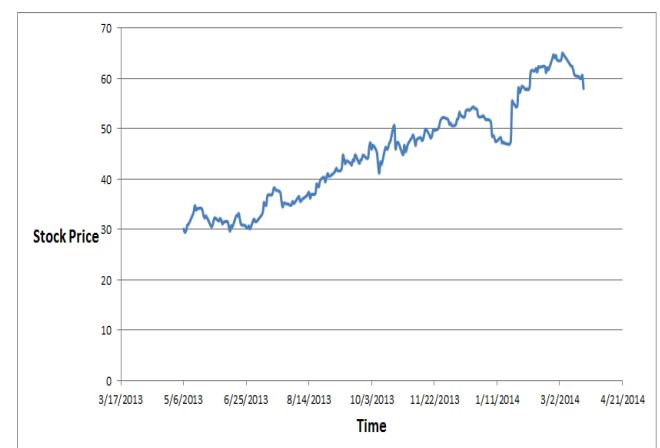


Fig- 2: Historical Stock Price

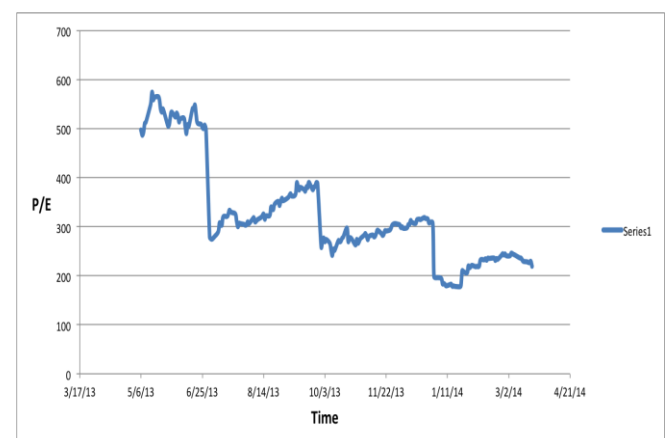


Fig-3: Historical P/E

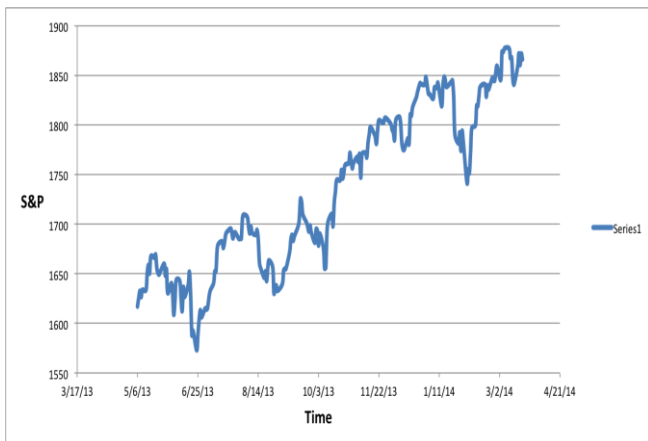


Fig-4: Historical S&P

1. Normalization of the data
Data is normalized between 0 and 1. This helps the network to perform better when every input and output is mapped in the same range.
2. Initialize weights
At first, weights are selected at random. In this study, there are 18 weights. 12 weights maps inputs to the hidden layer and then 6 weights map hidden layer to the output.

2.2 Results And Analysis Of ANN Approach

A back propagation algorithm was used for training the network. As shown in Figure 2, 3 and 4, data from May 6th 2013 to March 21, 2014 were used for training. This was about 221 data points for each input. Goal of the training was to minimize the mean square error between the actual stock price and the output of the model. Microsoft Excel was used to perform the calculations. In this particular study, training was stopped when MSE reached 0.0008. At that point, the model was tested. 197 data points from March 24th to Dec 26th 2014 were used to test the model. Test inputs P/E and S&P are shown in Figure 5 and Figure 6. Figure 7 shows the results of training and testing. Red and Blue charts on the left show training. Red is the actual stock price and Blue is the result of training. Green and purple show testing. Green is the actual test prices and purple is the prediction. It shows that the predicted output follows the trend of the actual data. Quantitatively, however, 48% of the predicted output for the rest of the year was within 5 dollars of the actual stock value. This was calculated by taking the output of the program and subtracting it from the actual stock price. The percentage of results that were below 5 dollars was calculated. This result can be improved by changing the network by choosing independently or a combination of additional hidden layers, different activation functions, additional inputs that can add value to the already existing information.

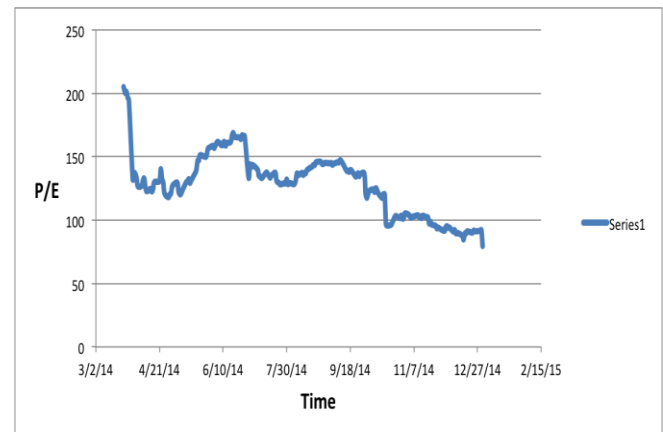


Fig-5 : Test Price Over Earning Input

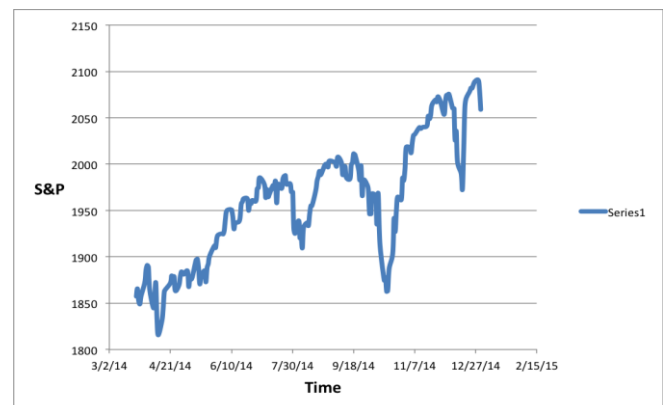


Fig-6: Test S&P Input

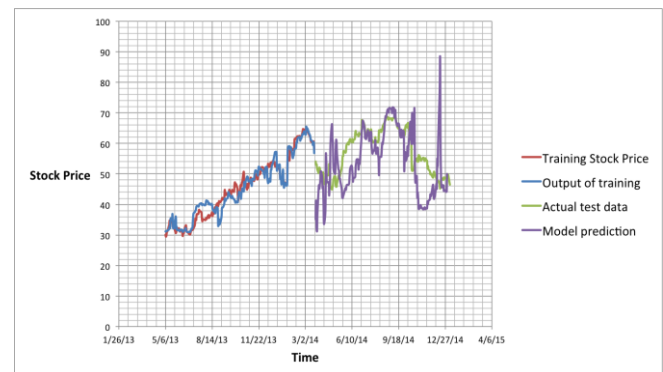


Fig-7: Training And Testing

2.3 Geometric Brownian Motion Approach

GBM modeling requires the knowledge of drift and volatility of the data. If this is not known, this should be calculated from past data. In this study, these parameters are calculated from the historical stock prices given in Figure 2. The steps to estimate drift and volatility from the past data are as follows:

1. Calculate the difference between today's closing stock value and the previous day's closing stock value.
2. Divide this difference by the today's closing stock value. The steps 1 and 2 together help to identify the relative change in the price from the previous day.

3. Do this calculation for the rest of the data.
4. Calculate the mean and standard deviation of this percentage difference in closing values. This mean and standard deviation are used as drift and volatility respectively.

In the case of this problem, rate of change of stock prices for the data shown in Figure 2 was calculated using the following equation

$$\text{Change \%} = \frac{X_t - X_{t-1}}{X_t} * 100$$

2.4 Results And Analysis Of GBM Approach

Drift and Volatility were calculated on the change percentage. Drift came out to be 0.16% and Daily Volatility came to be 2.73%. Since there are 252 trading days, the delta t (dt) was calculated to be 0.003. Stock price path was predicted from March 24th to Dec 26th 2014. Figure 9 shows the results for the prediction.

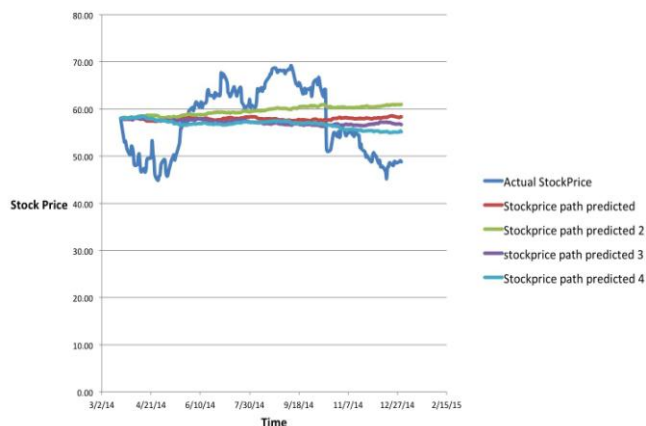


Fig-8: Stock Price Path Prediction Using GBM

Blue curve shows the actual data for the stock. Remaining colors show different paths that the stock could take. Several iterations of path could be generated using Monte Carlo simulations. In this study, only 4 paths for the stock price were generated. It appears that they are not capturing the dynamics of the actual stock price. Generating several hundred more trials of stock price path using Monte Carlo Simulations may solve this problem. Alternatively, is to find the data from the past that is more inclusive of the trends in the actual data from March 24th to Dec 26th 2014 to calculate drift and volatility. A mean across the sample would be a better estimate of what to expect.

3. CONCLUSIONS

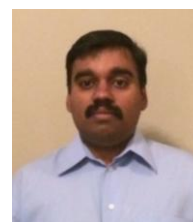
Neural Network and Geometric Brownian Motion were used on the same dataset to predict stock prices from March 24th to Dec 2014 based on the historical data from March 2013 to March 2014. Based on the results of this study, Neural Network is slightly better in predicting stock price movements when compared to GBM. Neural Network model was able to capture the trend and got accuracy of 48% within 5 dollars for a period of 9 months. This can be

vastly improved by diligent selection of inputs, number of hidden layers, and its activation functions. In that sense, it is more of a supervised learning. In contrast, Geometric Brownian motion did not capture the trend. This is partly because the drift and volatility were calculated based of a data that may not be truly inclusive of the dynamics of the time from March to Dec 2014. Also, a larger set of Monte Carlo trials would result in a better set of predictions for GBM. In conclusion, both ANN and GBM can be used for stock price prediction. However, it is understood from this study that they both require different levels of dynamics in their past data to successfully predict future data.

REFERENCES

- [1] Ali, A., Hammad, A., & Al-Ghandoor, A. (2011). Modeling Stock Exchange Prices Using Artificial Neural Network: A Study of Amman Stock Exchange. *Jordan Journal of Mechanical and Industrial Engineering*, 439-446.
- [2] Dunne, R. (2007). *A Statistical Approach to Neural Networks for Pattern Recognition*. John Wiley & Sons, Inc.
- [3] M.Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Singapore: Springer.
- [4] [Morters, P., & Peres, Y. (2010). *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge: Cambridge University Press.
- [5] Neenwi, S., Asagba, P., & Kabari, L. (2013). Predicting the nIGERIAN sTOCK mARKET USING Artificial NeuralNetwork. *European Journal of Computer Science and Information*, 30-39.
- [6] Oksendal, B. (2013). *Stochastic Differential Equations: An Introduction with Applications*. Heidelberg: Springer.
- [7] Zitkovic, G. (2010, December 24). Introduction to Stochastic Process-Lecture Notes. Austin, TX, USA.

BIOGRAPHIES



Suresh Venugopal: Dr. Venugopal holds a PhD in Engineering Mechanics. He did his PhD from The University of Alabama, Tuscaloosa. His research interests include Vibration of mechanical systems and how they can be controlled. Currently, he works for an Oil & Gas Company as an

Applications Engineer.