Research Report

EEE40011/RME40005

SEMESTER-1, 2017

Supervisor: Jagdish Patra

Project title: Artificial Neural network based prediction of stock prices

Group ID: 22

Student name: Ngoc Khanh Nguyen

Student ID: 4957059

May 29, 2017

Table of Contents

Abstract: Predicting stock prices movement has been attracting a huge amount of effort from not only investors and economists but also researchers in data science [1]. This report investigates the problem of predicting stock prices using artificial neural network. In this report, experiments were conducted to evaluate the performance of Functional link neural network (FLANN) on forecasting IBM stock prices. The network was fine-tuned by setting up with two different configurations to select the best model. The methods took advantages of past stocks price, past stocks' attributes, financial technical indicators and market index in order to generate prediction. The results of the experiments help answer the research questions mentioned in the previous research plan: How to predict future stock closing prices of some particular business sectors companies with high accuracy using artificial neural networks.

1. Motivation

Nowadays, studying behavior of stock markets is a pivotal task of monitoring any nations' economy. It is undeniable that stock markets witness the greatest flow of capital exchange all over the world[2]. Stock markets also plays an important role in attracting and directing liquidity and savings in order to allocate scarce financial resources optimally to increase productivity of industry.

In stock market, the sole purpose of investors is to make profit by analyzing market relating information[3]. People can create incredible lead in trading if they have rational business acumen and be able to predict the future movement and status of the market. Consequently, forecasting and predicting information regarding market and company which offer stocks is one of the most important concern in the stock market. A reliable and trustworthy report on company's future financial status not only provides investors necessary insight as a base for their investment decision but also boost their confidence in gaining more profit[4].

Unsurprisingly, predicting stock price is not a new topic in finance because of its importance and obvious payoff for success. There are three major approaches to predict stock prices: fundamental analysis, technical analysis and machine learning and data mining[5]. Fundamental analysis method evaluates company's performance and credibility as well as market status to determine stock prices. On the other hand, technical analysis method does not concern with any of those information, instead, it attempts to predict future price based solely on past price. The last approach, machine learning and data mining combines both company and market status with past stock prices to produce analysis and prediction. In this report, we will study more about Artificial neural network which is a subset in machine learning and data mining methods.

2. Research objectives and timelines

2.1. Research objectives

- **Goal**: construct a neural network based model to predict future closing price of companies in retail and technology sector over specific time horizons with high accuracy.
- **Objective 1**: Understand the characteristics of the target stocks and which factors may impact the stock price movement.
- **Objective 2**: Construct the appropriate neural network architectures
- **Objective 3**: Specify evaluation metrics to determine the performance of the model.

2.2. Research timelines

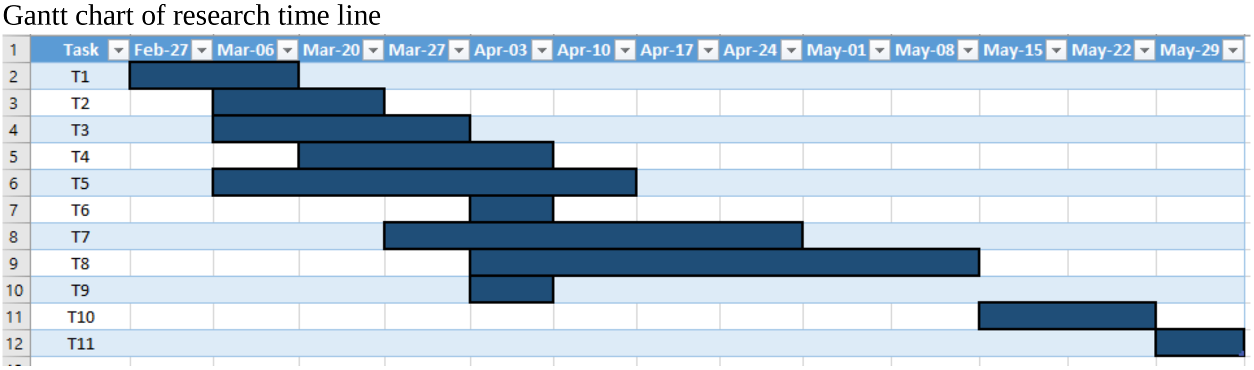
These tasks are created to with the intention to successfully achieve research objectives

Tasks	Description	Duration
T1	Selecting topic	2 weeks
T2	Understanding the selected topic	2 weeks
T3	Exploring literature	3 weeks
T4	Establishing research questions	3 weeks
T5	Finding related learning and background materials	5 weeks
T6	Forming objectives and expected goal	1 week
T7	Learning the target companies and business sectors	5 weeks
T8	Establishing initial design ideas	6 weeks
T9	Research Plan	1 week
T10	Presentation 1	2 weeks
T11	Report 1	1 week

The timeline is created in order to fulfill research's goal and main objectives as well as minor objectives arose as the research progresses

Objectives	Task	Activity	Expected completion date
	T1	Select research topic from proposal list	Completed
Learn how to carry out a research	T2	Research information related to topics of interest	Completed
Gain background knowledge about the research topic	T3	Research up-to-date literature related to stock price prediction in general and stock price prediction using ANNs in particular	Completed
Learn how to carry out a research	T4	Form a research questions based on proposed research topic and collected literatures	Completed
Research Objective 1	T5	Study the target stocks and neural network	Competed
	T6	Forming objectives and expected goal	Completed
Research Objective 2	T8	Create a draft design based on learnt information about stocks and proposed techniques	Completed
Learn how to carry out a research	T9	Establish a research plan	Completed
	T10	Presentation	Completed
	T11	Report	Completed

Gantt chart of research time line



Task	Activity	Outcomes
T1	Select research topic from proposal list	Submitted topic proposal
T2	Research information related to topics of interest	Section 3, 5.1, 5.2, 5.3
T3	Research up-to-date literature related to stock price prediction in general and stock price prediction using ANNs in particular	Section 5.3
T4	Form a research questions based on proposed research topic and collected literatures	Section 5.1, 5.2, 5.3
T5	Study the target stocks and neural network	Section 5.3
T6	Forming objectives and expected goal	Section 2.1, 3, 4
T8	Create a draft design based on learnt information about stocks and proposed techniques	Section 6, Section 7
T9	Establish a research plan	Submitted research plan
T10	Presentation	Submitted presentation
T11	Report	Submitted research report

Research progress is currently on-track to meet the all major and minor research goals and objectives specified in the Gantt chart. The results of the initial design could be used as a baseline for future researches and improvements.

3. Constraint

Stock price prediction is regarded as one of the most widely studied and challenging problems. Stock price is a notable time-series which is famous for its non-linearity and non-stationary characteristic[6]. Besides, stock price is heavily affected by many factors such as investor sentiments, industry performance, and economic factors[7]. One force can have different influence on different type of stock. The volatile nature of the stock market creates many difficulties for available prediction techniques such as time series forecasting. It could be fairly simple to explain why the stock price had this movement and behavior after it had happened, but not very easy to provide insights prior to the event in order to take advantage of it. Many economists and researchers believe that stock prices change accordingly to a random walk which makes it nearly impossible to be predicted[8]. Their belief formed the Random Walk Hypothesis which stated that the distribution of stock prices variations is similar and independent of each other[9]. Consequently, the past movement and historical prices of a stock or market would not have any power to predict its future movements[9]. Even though the hypothesis is still controversial topic between economist, it confirms the difficulty and challenge in predicting stock price.

4. Scope

The research topic covers both finance and data science field which is extremely vast and broad. As a result, in order to successfully undertake this research, the idea is to mainly emphasize on data science aspect which is manipulating data and designing an appropriate neural network while the study in finance field which is the complex behavior of the target stock in stock market will be lessen. The study on neural network will focus on examining different neural network architects such as functional link neural network (FLANN), multilayer perceptron (MLP) or recurrent neural network (RNN). Not only that, it is essential investigate and determine the best setting for each neural network to improve their prediction. More than that, as the input data to predict future stock price may contain noise which may have negative impact on built predictive model, it is pivotal pre-process input data in order to maximize networks' performance.

5. Research questions

The main research questions proposed in the initial research plan was how to predict future stock closing prices of some particular business sectors companies with high accuracy using artificial neural networks. This report shows attempts to identify potential problems and answer the research questions in details.

The research question explicitly addresses three critical problems of the research which are problem of predicting future stock prices, selecting target stocks for research and implementation of neural network for predicting stock prices.

5.1. The problem of predicting future stock prices:

5.1.1. Fundamental analysis

The traditional way in finance to predict stock price movement is Fundamental Analysis. It measures the actual value of a company or an asset by studying any related economic, financial and other qualitative and quantitative factors[10]. Fundamental analysts attempt to study any aspects which could potentially have impacts on the true value of the company. The ultimate goal of fundamental analysis is to produce a quantitative value which an investor can interpret and relate with current security's price to determine their trading strategy.

Fundamental analysis takes into account many factors such as revenues, earnings, future growth, profit margins and any other data related to company's performance in the market and industry to determine its stock value and evaluate potential for growth in the future[11]. One of the most well-known and successful investor who utilize the power of fundamental analysis is the famous billionaire Warren Buffett.

In financial context, the term fundamentals refer to qualitative and quantitative data which represent the health and stability of a business or an asset. The types data can be divided into two categories which are macroeconomic and microeconomic. Macroeconomic fundamentals are factors which affect an economy or a business at large scale. Some Macroeconomic fundamental instances are government policies on trading and statistics regarding currency exchange rate, oil prices and gold prices. On the other hand, microeconomic fundamentals denote factors which have smaller impact coverage such as supply and demand within a specific market or industry. These fundamentals are analyzed to provide analysts and investors insight whether the stocks are worth investing in or not.

Some factors which influence stock prices are:

- Demand and supply: if demand of a security is higher than its supply, for example the number of buyers is more than sellers), prices of the security will increase without a doubt. If the supply exceeds the demand, the prices will decrease.
- Interest rate: the increase in interest rate will lower the demand for funds to support manufacturing and producing goods; therefore, reduce the demands for stocks. When the interest rate decreases, the demand for funds will increase which results a high demand for stocks
- Market player: stocks prices are influenced by investor sentiments. If the investors are optimistic and believes in company's performance, the stocks prices will increase. On the other hand, if the investors are pessimistic and think the economy will possibly fall to recession, the stock prices will decreases
- Dividend announcements: dividend is a distribution of a portion of a company's earning to its shareholders. If the rate of dividend announced by company's board of directors does not match with investors expectation, share prices would decline, whereas if they are up to what investors expect, share prices would increase.

- Management profile: Executive members of a company's have important influences on how it functions and operates. As a result, if the executive team appears educated, experienced professional to investors, it is likely that company's share prices would increase.
- Trade cycle: Trade cycles denotes repeated and cyclical fluctuations in economic activity. Generally, a fluctuation is categorized as three phases. During a recession, stock price would fall dramatically. It would regain price during recovery phase and increase significantly during boom conditions.
- Speculation: High speculation in market or in a stock could lead to unpredictable and strong fluctuations in stock price. When speculation is low, the degree of fluctuation in stock price is reduced
- Political factors: Political factors such as ideology of governors, both national and international policies executed by the government and stability of government have strong impacts on the operations of stock market and investors' sentiments.
- Industrial relations: the relationships between workers and management team of a company reflects its performance. Good industrial relationships result productivity and growth which raise company's stock price. When the relationships are poor, company's performance will decrease which leads to fall of stock price.
- General market sentiments: stock price will raise when investors and traders are optimistic about the performance and growth of the market. When investors and traders and pessimistic, they tend to sell more stocks instead of buying which makes the stock price decline.
- Level of foreign investment: A high level of foreign investment leads to more buying commands being executed than selling. As a result, it raises stock's demand and price. When the level of foreign investment declines, more stocks are sold by foreign investors who left the market. The increases in selling reduces stocks' price

5.1.2. Technical analysis

Technical analysis is a method to assess stock price and forecast its movements by analyzing statistics collected from past trading activity such as stock price trend and volume. Technical analysis attempts to comprehend market status and sentiment based on price trends rather than analyzing stock's fundamental factors like fundamental analysis[12].

The foundation of technical analysis is based on three assumptions:

- The market discounts everything: this assumption is derived from Efficient Market Hypothesis (Fama, 1965) which states that stocks are already being traded at fair price and their price reflect all information regarding the company and market status. Therefore, it is unnecessary to analyze any fundamentals factors of company and market.
- Price moves in trends: technical analysts believe that stock prices have tendency to maintain their past trend instead of moving erratically. There are three types of stock trend: short, medium and long term trend.
- History tends to repeat itself: in technical analysis, it is believed that stock price movements are likely to repeat itself. This statement is due to market psychology, which depends on predictable emotions of investors like fear or excitement.

Over the years, many statistical technical indicators have been developed to support technical analysis.

- Williams %R: this indicator shows the relation between present closing level and highest and lowest price given a period.

$$\%R = \frac{Highest\ High - Closing\ Price}{Highest\ High - Lowest\ Low} * (-100)$$

- Simple moving average (MA): this indicator smooths price data to reveal and highlight stock price movement trend and filter out noises.
  n-day moving average

$$ma_n = \frac{\sum price}{n}$$

- Exponential moving average (EMA): this indicator is similar to simple moving average which has the same purpose to smooths price data to reveal and highlight stock price movement trend and filter out noises. It reduces lagging in simple moving average method by adding more weights to recent prices.
- Moving average convergence divergence: this indicator shows the relationship between two exponential moving average series of closing prices which are usually 12-day EMA and 26-day EMA.

$$MACD\ Line = \ (12 - day\ EMA) - (26 - day\ EMA)$$

- Bollinger bands: Bollinger bands consists of a simple moving average series of closing prices and two series which are calculated by subtracting and adding its standard deviation. This indicator measures stock's volatility which is how the stock price could deviate from its true value.
  Middle band: 20-day simple moving average
  Upper band: 20-day simple moving average + (20-day standard deviation * 2)
  Lower band: 20-day simple moving average - (20-day standard deviation * 2)
- On-balance volume (OBV): this indicator measures the changes in trading volume flow to interpret stock price movement.
  If the closing price is above the prior close price, then: Current OBV = Previous OBV + Current Volume
  If the closing price is below the prior close price, then: Current OBV = Previous OBV - Current Volume
  If the closing prices equals the prior close price, then: Current OBV = Previous OBV (no change)

5.1.3. Artificial neural network

Artificial neural network is a mathematical and computational model which is conceptually and structurally inspired by human nervous systems[13]. The common structure of a neural network is a system with one or multiple layers consisted of interconnected neuron units. Inputs fed into each layer are computed in each neuron unit. The results then propagated to the next layer and the process is repeated until they reach the output layer[13]. The network is trained by receiving examples. Based on how these examples are given to the network, there are two methods to train a network which are unsupervised and supervised learning. In supervised learning, model is given both input variables and corresponding output variables. Its goal is to estimate or approximate a function which maps the input to the respective output[14]. On the other hand, unsupervised learning model is not provided expected output. Unsupervised learning is used to find the structure or relationships between different inputs[14].

As the skyrocketing of computing power and the availability of finance data in the 21st century, there has been rising interest in applying neural network to problems in finance. Since neural network has the ability to extract hidden pattern and identify hidden nonlinear relationships from data with a minimal involvement of domain knowledge and human intervention, this technique attracts researchers as an alternative to existing approaches such as fundamental analysis which requires expertise in finance[15].

Over the years, many studies have proved that neural network is a potential candidate to tackle the non-linearity and non-stationary of the complex stocks price data[16]. Enke and Thawornwong showed that trading strategy built buy ANNs classification model was possible to gain more profit compared to other learning models.[16]. Mostafa and Mohamed constructed two neural networks and generalized regression neural network to forecast Kuwait stock market[17]. Kumar Chandar et al. suggested a model combined wavelet transform and ANNs[18]. Tsai and Hsiao applied 8 financial rations and 16 macroeconomic indicators on a backpropagation neural network to predict stock return in Taiwan stock market[19]. G. Tingwei et al. showed a superior performance of deep belief networks (DBNs) on predicting stock closing price with stock daily information and stock technical indicators input preprocessed by principal component analysis (PCA)[20]. F A. de Oliveira et al. used a combination of 15 technical indicators and 11 fundamental indexes in an attempt to predict stocks movement in Petrobras. Their resulted neural network achieved the accuracy of 87.50%[21]. These studies support the use of ANNs for financial forecasting. The table 1 below shows some recent studies about application of neural network in stock market prediction[15].

Table 1: Recent studies about application of neural network in stock market prediction

Authors (Years)	Data type (No. of input features x lagged times)	Target outputs	Number of samples (Training: Validation : Test)	Sampling period	Method	Performance Measure
Enke and Mehdiyev (2013)	US S&P500 index (20x1)	Stock price	361	Jan-1980 to Jan-2010 (daily)	Feature selection + fuzzy clustering + fuzzy NN	RMSE
Niaki and Hoseinzade (2013)	Korea KOSPI200 index (27x1)	Market direction (up or down)	3650 (8:1:1)	1-Mar-1994 to 30-Jun-2008 (daily)	Feature selection + ANN	Statistical tests
Cervello-Royo et al. (2015)	US Dow Jones index (1x10)	Market trend (bull/bear-flag)	91307	22-May-2000 to 29-Nov-2013 (15-min)	Template matching	Trading simulation
Patel, Shah, Thakkar, and Kotecha (2015)	India CNX and BSE indices (10x1)	Stock price	2393*	Jan-2003 to Dec-2012 (daily)	SVR + {ANN, RF, SVR}	MAPE, MAE, rRMSE, MSE
T.-L Chen and Chen (2016)	Taiwan TAIEX[a] and US NASDAQ[b] indices (27x20)	Market trend (bull-flag)	3818[a], 3412[b] (7:0:1)	7-Jan-1989 to 24-Mar-2004 (daily)	Dimension reduction + Template matching	Trading simulation

Chiang, Enke, Wu, and Wang (2016)	World 22 stock market indicies ({3~5}x1)	Trading signal (stock price)	756 (2:0:1)	Jan-2008 to Dec-2010 (daily)	Particle swarm optimization + ANN	Trading simulation
Chourmouziadis and Chatzoglou (2016)	Greece ASE general index (8x1)	Portfolio composition (cash:stock)	3907*	15-Nov-1996 to 5-Jun-2012 (daily)	Fuzzy system	Trading simulation
Qiu, Song, and Akagi (2016)	Japan Nikkei 225 index (71x1)	Stock return	237(7:0:3)	Nov-1993 to Jul-2013 (monthly)	ANN + {genetic algorithm, simulated annealing} Deep NN	MSE
Arevalo, Nino, Hernandez, and Sandoval (2016)	US Apple stock (3x{2~15}+2)	Stock price	19109 (17:0:3)	2-Sep-2008 to 7-Nov-2008 (1-minute)	Deep NN	MSE, directional accuracy
Zhong and Enke (2017)	US SPDR S&P500 ETF (SPY) (60x1)	Market direction (up or down)	2518 (14:3:3)	1-Jun-2003 to 31-May-2013 (daily)	Dimension reduction + ANN	Trading simulation, statistical tests

5.2. Target stock

In this report, IBM stock will be studied in details. IBM is an American multinational technology company which headquarters in Armonk, New York, United States. The full name of the company is International Business Machines Corporation. IBM manufactures and produces computer hardware, middle ware and software. It offers a diverse range of service from cloud computing, cognitive computing, commerce, data and analytics, Internet of Things, IT infrastructure, mobile and security[22]. IBM business model is business-to-business (B2B) in which its target customers are focused on manager of a business or government organization – someone concerned everyday with saving time, labor, or money to improve the productivity of the firm[23]. IBM is also famous for its dedication for research work. Its first research facility was the Watson Scientific Computing Laboratory located in Columbia University in New York. IBM Research department is known as the largest industrial research organization in the world. It consists of 12 labs which are spread across 6 continents[24]. IBM employees have been awarded five Nobel Prizes and six Turing Awards. They are also honored by National Medals of Technology and five National Medals of Science which were bestowed by the President of the United States.

IBM stock is listed and traded on the New Your Stock Exchange (NYSE). The company is included in the Dow Jones Industrial Average index which is the second-oldest US market index. Its stock is also included in S&P 100 and S&P 500 index which represents well established companies in US stock market.

To successfully predict IBM stocks, it is essential to understand the company's stock quotes. A stock quote includes this set of information:

- Week High and Low: the highest and lowest prices of a stock over the previous 52 weeks in total.

- Company name and type of stock: name of the company who offers stocks followed by special symbols or letters to indicate classes of shares. If there is nothing followed the company's name, it is a common stock
- Ticker symbol: a unique name which identifies the stock. Ticker of IBM stock is IBM.
- Dividend per Share: the percentage return on the distribution of company's earning to shareholder.
- Price/Earnings Ratio: The P/E ratio is the ratio between the market price of a company's stock and its earning. This information can be compared between different companies to determine whether a company's stock is overvalued or undervalued.
- Trading Volume: this figures denotes the total number of stocks traded for the day.
- Day High and Low: the highest and lowest prices of a stock during the trading day.
- Close: this is the last trading price of a stock when the market closed on the day.
- Net change: the difference between stock price compared to the previous day's closing price.

5.3. Implementation of neural network

5.3.1. Neurons

A neuron is a basic computational unit of a neural network. It processed input fed from other neurons or from an external source to produce an output. Each neuron is associated a set of parameters including a weight (w) and a bias (b). There are many selections for a neuron's activation function such as sigmoid, tanh, and ReLu unit. The activation function takes a n-dimension input vector and results in a scalar activation a which is also neuron's output. An output of a neuron with sigmoid activation function is computed as the weighted sum of its input as shown below:

$$a = h_{W,b}(x) = \frac{1}{1 + \exp(-(w^T x + b))}$$

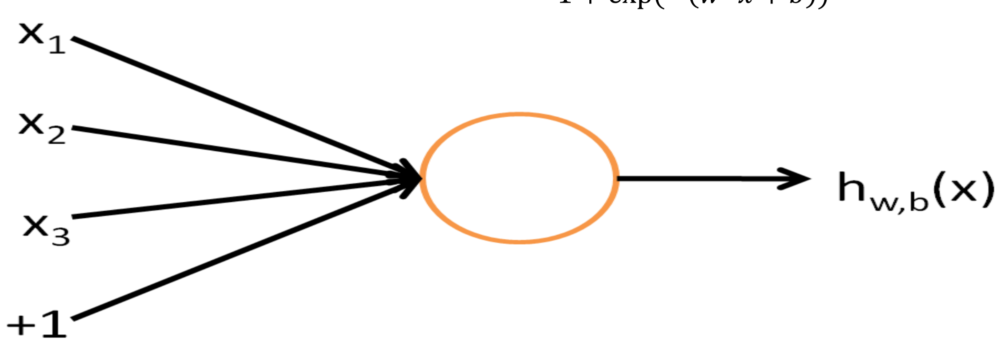


Figure 1: A single neuron[25]

5.3.2. A single layer of neurons

A sing layer of neurons contains multiple neurons described above. Consequently, in a single layer of neurons, there are different neurons' weight $\{w^{(1)}, w^{(2)}, \dots, w^{(m)}\}$, biases $\{b_1, b_2, \dots, b_m\}$ and respective activations $\{a_1, a_2, \dots, a_m\}$:

$$a_1 = \frac{1}{1 + \exp\left(-\left({w^{(1)}}^T x + b_1\right)\right)}$$

$$\dots$$

$$a_m = \frac{1}{1 + \exp\left(-\left({w^{(m)}}^T x + b_m\right)\right)}$$

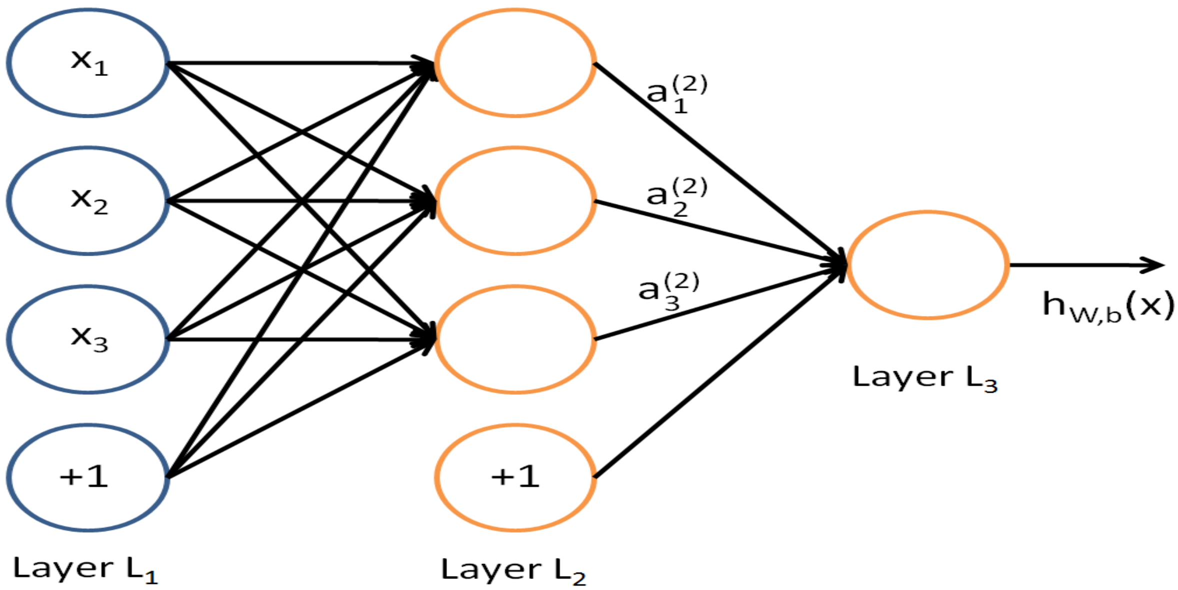


Figure 2: One layer neural network[25]

5.3.3. **Functional link neural network (FLANN)**

Functional Link Neural Network was introduced by Klassen and Pao for pattern recognition [26]. The architect of this network is much straightforward compared to multilayer perceptron as it consists of 1 input, 1 output layer and no hidden layer. As a result, there is only a single-layer trainable weights and bias. Despite of its simplicity, it is still capable of capture the non-linear relationships of input data. FLANN has been successfully employed in many applications such as system identification [27, 28], classification[29, 30] and prediction[31, 32]. Figure 3 shows a general structure of a FLANN.

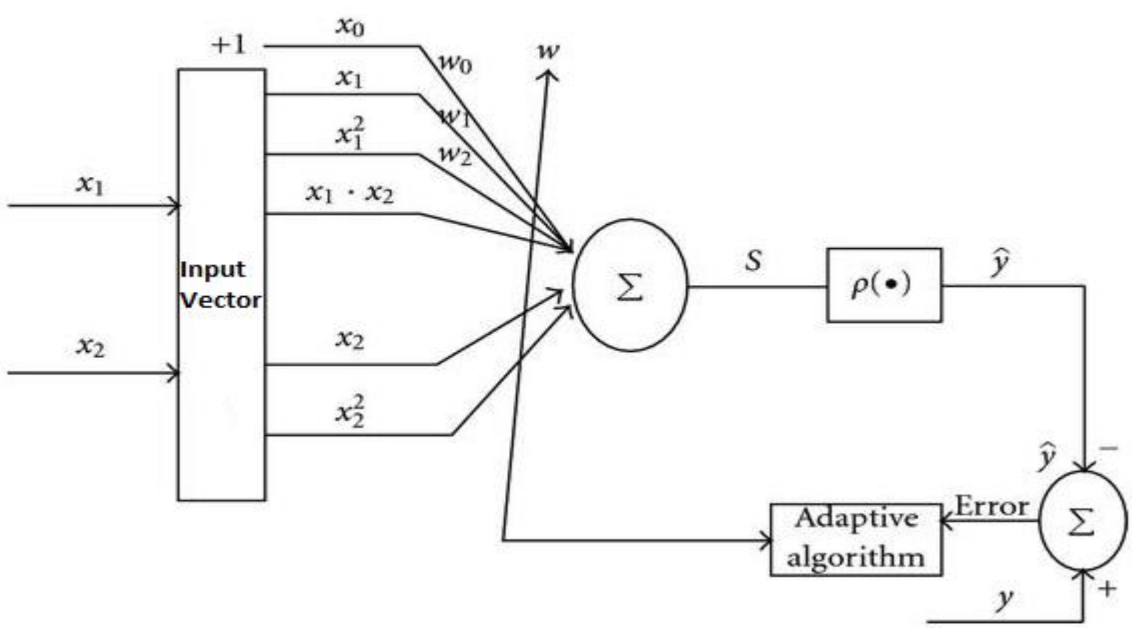


Figure 3: FLANN structure

For stock price prediction, J. C. Patra et al. proposed a 2-layer neural network architect named functional link artificial neural network (FLANN) to predict the next-day's closing price of Exxon Mobil (energy sector), Citigroup (banking sector) and IBM (technology sector)[33]. The data was collected on daily basis from Yahoo Finance between Jan 1998 and Apr 2008. It contained 2478 samples in total. The model predicted next-day's closing price based on 14 different features which consisted of not only open and closing prices but also industrial indices and technical indicators for higher accuracy. The structure of the FLANN network had two parts. The first part was a trigonometric functional expansion component to map the input data from low dimensional space to a higher dimensional space. The second part was a single neuron with tanh activation function. The experiment yielded a reasonably low average percentage error (APE) 1.4% to 4.5% with acceptable hit rate of 0.62.

5.3.4.Multilayer perceptron (MLP).

A multilayer perceptron contains one or more hidden layers (apart from one input and one output layer). Multiple neurons in different layers are hooked together so that output of a neuron in a layer can be input of another in the next layer. Figure 4 shows a general structure of MLP.

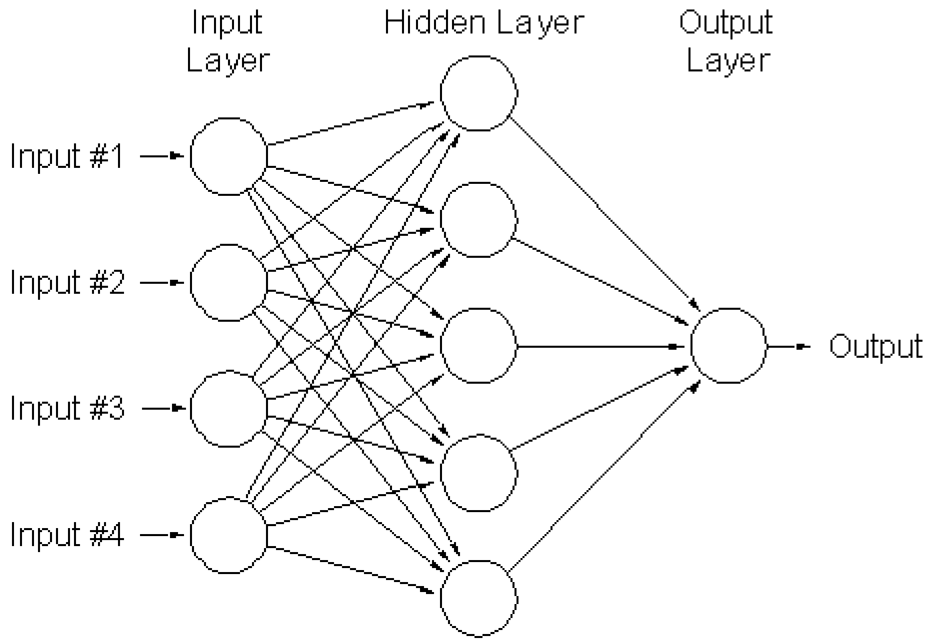


Figure 4: MLP structure

F. A. d. Oliveira et al. used supervised neural network to predict the future closing price over different time horizons and possible future behavior of stock PETR4 which belong to Petrobras (energy sector) traded on the Sao Paulo Stock Exchange (BM&FBOVESPA)[34]. The data set contained 2384 data samples between 04/01/2000 and 18/08/2009. Each samples included 22 features which were categorized as stock quotes, macroeconomics series and technical indicators such as opening, closing price and 5, 22, 200-day exponential moving average. The structure of the neural network consisted of three layers which were 1 input layer, 1 hidden layer and 1 output layer. The number of input unit was the prediction window size multiplied by the number of sample set series. The hidden layer had 22 hidden units which was the number of features. The number of neurons in the output layer varied by the prediction horizons. The best performance of the model was obtained with 5-day window and a prediction horizon of 1 day with root mean square (RMS) error of 0.00129.

M. Qiu et al. introduced a new set of input features and applied Global search techniques, a genetic algorithm (GA) and simulated annealing to improve prediction ability of the ANN models to forecast the return of the Nikkei 225 index[35]. The data covered the period from Nov 1993 to Jul 2013. In total, 71 features which contained financial indicators and macroeconomic data were employed to predict the return of the Nikkei 225 index. The neural network consisted of three layers which were 1 input layer, 1 hidden layer and 1 output layer. To improve the performance of the model, genetic algorithm (GA) and simulated annealing was employed to optimize the weights and bias of the ANN.

5.4. SWOT analysis

Based on the understanding of research question, a SWOT analysis is conducted.

Strengths	Weakness
Data availability: there are plenty of available and public stocks related data which can be easily mined from reliable sources such as Yahoo Finance or Google Finance **Data science skills**: strong skills in linear algebra, statistics and python programming.	**Lack of domain knowledge in finance**: studying finance to grasp the fundamentals of stock market in an arduous and challenging task. **Lack of computing power**: Improving training time and computing accuracy of neural networks requires at least a decent set up of computer.
Opportunities	Threats
Stock evaluation: this study could be applied to analyses IBM stock behavior in stock market. **Trading strategy**: prediction produced by this study could be used to form a strategy for investors to decide whether to buy, sell or hold their stock	**Accuracy of models**: small error between actual and predicted stock price value could potentially lead to great loss in stock market **Prediction speed**: predicting speed and prediction horizons must match with each other. The model would be pointless if it takes a week predicts next 5-day closing prices.

6. Methodology

6.1. Toolbox

In this research, the neural network will be constructed by Tensorflow. It is a python based open source software library for numerical computation using data flow graphs. Tensorflow was created by Google Brain Team to conduct machine learning and deep neural networks research. It provides an extensive suite of functions and classes that allow users to build various models from scratch. Tensorflow was launched in Nov 2015. Since then, its popularity has increased considerably which has created a large support community. It is estimated that more than 10000 commits and 3000 Tensorflow-related repositories have been made in one year. Some companies which employ Tensorflow are Google, OpenAi, Snap Inc., Uber and eBay[36].

Python programming language v3.5	Computation and analysis	NumPy v1.13 Matplotlib v2.0.02
	Neural network	Tensorflow CPU v1.1
Hardware	CPU: Intel® Core™ i5-4690 CPU @ 3.50GHz, 4 physical cores, 4 logical processors Memory: DDR3 2133MHz 8.00GB	

6.2. Proposed solutions

In this report, the neural network based prediction system to predict next-day closing price of IBM stock implemented FLANN as the core architect of the predictive model. Prediction of the model was made based on historical closing prices, market fundamentals and technical indicators. To capture changes in economic environment and stock trend, 14 features were used[33]:

Features	Ticker	Description
Dow Jones Industrial Average	DJIA (x_1)	Macroeconomics series (see section 3.2)
Historical close price of past 3 days	Close_1 (x_2), Close_2 (x_3), Close_3 (x_4)	Stocks quotes
7-day moving average, for past 4 periods	$ma7_1$ (x_5), $ma7_2$ (x_6), $ma7_3$ (x_7), $ma7_4$ (x_8)	Technical indicator (see section 3.1.2)
30-day moving average, for past 4 periods	$ma30_1$ (x_9), $ma30_2$ (x_{10}), $ma30_3$ (x_{11}), $ma30_4$ (x_{12})	Technical indicator (see section 3.1.2)
Standard deviation for the past 30 days	Std_30 (x_{13})	Technical indicator
Open price	Open (x_{14})	Stocks quotes

The data was collected from Yahoo finance website at http://finance.yahoo.com/ on daily basis for the period from Jan 1st 2000 to Dec 31st 2016. The data set consisted of 4277 data samples. After the historical closing prices had been collected, the simple moving average and standard deviations were calculated using formula showed in section 5.1.2.

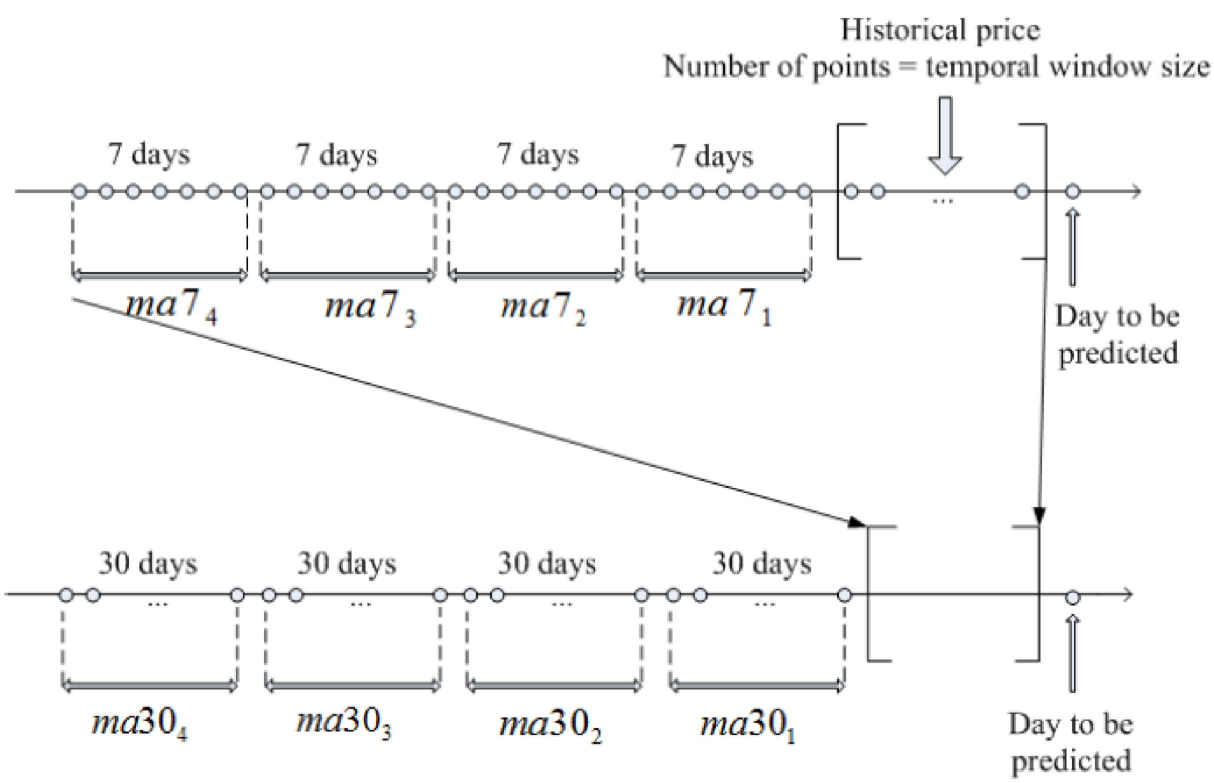


Figure 5: Moving average and historical prices features usage

Figure 5 illustrates how the historical prices and moving averages series were used in the predictive model. The closing prices of 3 days prior to the predicted day were used. The 7-day moving averages were taken prior to the 3-day period and the 30-day moving averages were taken prior to the last 7-day moving averages. By applying this idea, exact value of closing price of samples which is closer to the predicted day are considered to have bigger impact than the historical prices which is further. The moving averages series were used to capture the overall trend of the past 4 weeks and the past 4 months.

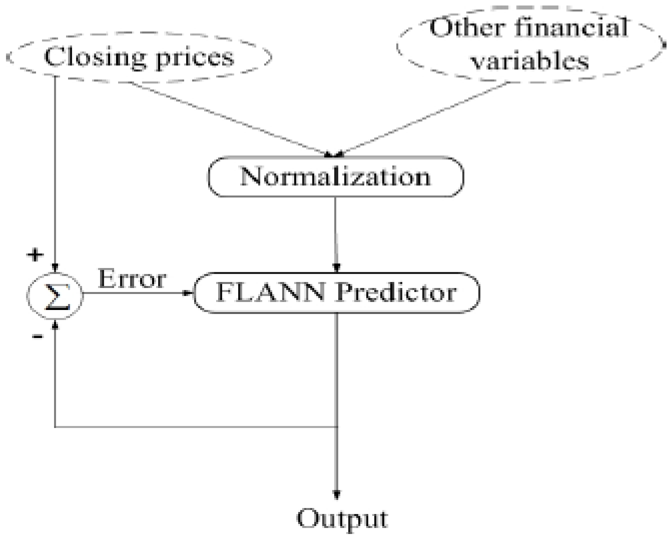


Figure 6: Structure of predictive model

Fig 6 shows the basic structure of the predictive model. The model was inputted with data consisted of the above 14 features. Data was then normalized and its dimension was expanded before being fed into the FLANN. The FLANN was trained using backpropagation algorithm. Its cost function was minimized using gradient descent.

6.3. Data overview

The data set original had 4277 samples. After calculating simple moving average series, there remained 4129 samples. The data set was divided into two separated set which are training set and testing set. The training set consisted of 3373 samples was used to train the predictive model. The testing set includes 756 samples which was used to evaluate the performance of the model.

Table 2:Input data description

	IBM (unprocessed)	IBM (processed)	Training	Testing
Start	01/01/2000	08/03/2000	08/03/2000	01/01/2014
End	31/12/2016	31/12/2016	31/12/2013	31/12/2016
Number of samples	4277	4129	3373	756

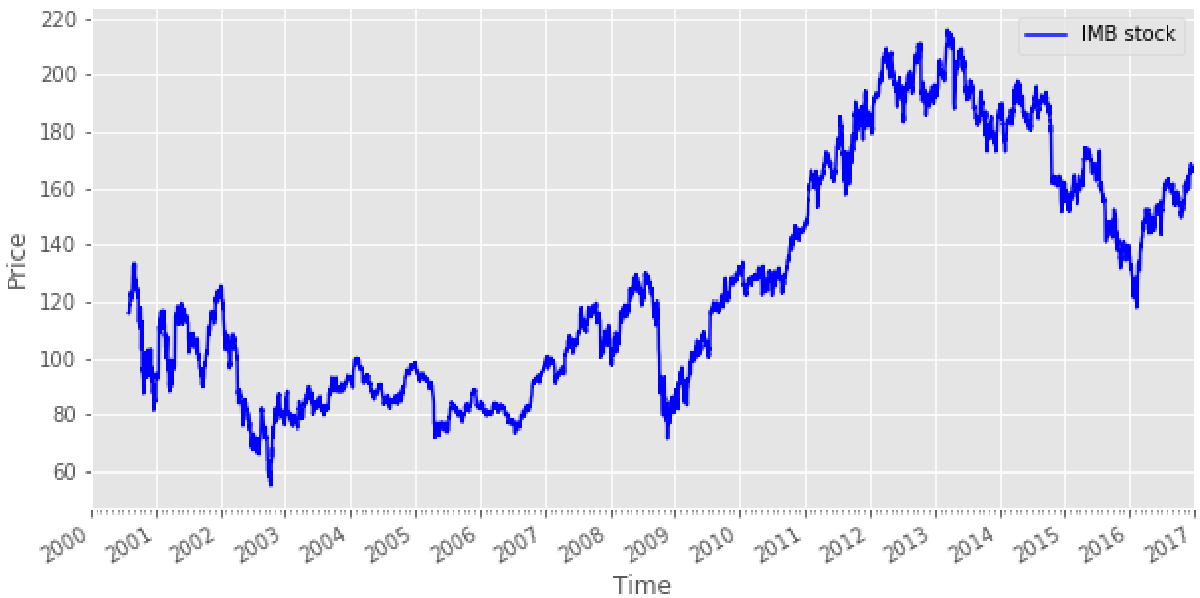


Figure 7: IBM stock price movement from 01/01/2000 to 31/12/2016

Figure 7 shows the stock price movement of IBM from 08/03/2000 to 31/12/2016. Between 2000 and 2002, IBM stock price kept fluctuating then dropped dramatically in the beginning of 2002. The price bottomed at $55.07 USD in the end of 2002. The instability and decrease of the stock price during this period could be possibly a result of the early 2000s recession which hit the US in 2002 and 2003. After 2003, IBM stock price started to recover as the price slightly increased. However, due to the United States housing bubble occurred in 2008, 2009, the price swiftly declined. Nevertheless, between 2009 and 2014, IBM stock price rapidly increased and peaked at $215.80 USD in 2013. After that, the price showed gradual downward trends but started to recover again at the beginning of 2016.

Table 3: Statistic summary of features Open, DJIA, Close_1, Close_2, Close_3

	Open	DJIA	Close_1	Close_2	Close_3
count	4129.00	4129.00	4129.00	4129.00	4129.00
mean	127.01	12268.74	127.08	127.07	127.06
std	41.13	2980.20	41.14	41.14	41.14
min	54.65	6547.05	55.07	55.07	55.07
0.25	90.50	10241.02	90.55	90.55	90.55
0.50	117.14	11239.77	117.40	117.38	117.36
0.75	162.34	13712.21	162.30	162.28	162.28
max	215.38	19974.62	215.80	215.80	215.80

Table 4: Statistic summary of features std30, ma7_1, ma7_2, ma7_3, ma7_4, ma30_1, ma30_2, ma30_3, ma30_4

	std30	ma7_1	ma7_2	ma7_3	ma7_4	ma30_1	ma30_2	ma30_3	ma30_4
count	4129.00	4129.00	4129.00	4129.00	4129.00	4129.00	4129.00	4129.00	4129.00
mean	3.52	126.90	126.81	126.72	126.65	126.32	126.01	125.75	125.48
std	2.05	41.05	41.03	41.01	40.99	40.95	40.87	40.82	40.75
min	0.59	57.54	57.54	57.54	57.54	57.54	57.54	57.54	57.54
0.25	2.05	90.58	90.58	90.58	90.58	90.58	90.58	90.58	90.58
0.50	3.05	116.84	116.71	116.58	116.51	115.95	115.59	115.09	115.01
0.75	4.53	162.19	162.12	161.91	161.91	161.91	161.91	161.81	161.81
max	13.56	213.82	213.82	213.82	213.82	213.82	213.82	213.82	213.82

Table 3 and 4 show the summary statistic of each features in the processed data set which includes number of samples, mean, standard deviation, lower percentile (25%), median (50%), upper percentile (75%) and maximum value. From the table, it can be seen that the statistics figures of closing price and moving average series are approximately the same as moving average series were calculated from the closing price series. It can also be observed that the value range of DJIA features is significantly bigger than other features. On the other hand, value range of std30 feature is extremely low compared to the others.

6.4. Data preparation:

Due to different in value range of features, it is necessary to normalize them to guarantee the performance of the model. As the activation function of FLANN is tan hyperbolic function (Tanh) which mapped input value to output domain of [-1, 1], the normalization method used in this solution is the linear interpolation which normalize input data into range [-1, 1].

Cover theorem stated that "The probability that classes are linearly separable increases when the feature are nonlinearly mapped to a higher dimensional feature space"[33]. Based on this theorem, the input data with 14 features was expanded into a higher dimensional space using three different methods: trigonometric expansions, Chebyshev expansion and Legendre expansion. The number of expansion term using in this report was 8 for three expansion method as it produced the best result in a study on predicting stock price using FLANN by J. C. Patra et al mentioned in section 3.3.1 [33]. Three expansion methods were examined separately with similar configuration of FLANN to determine the method which could result in the best predictive performance.

Table 5: Data preprocessing methods

Method	Description	Formula
Normalization	linear interpolation target range $[-1,1]$	$X_{norm} = 2 * \frac{X - \min(X)}{\max(X) - \min(X)} - 1$
Expansion function	Trigonometric function	Input vector X is expanded into 8 terms as follow: $X_{expand} = [X \ \sin \pi X \ \cos \pi X \ \sin 2\pi X \ \cos 2\pi X \dots \ \sin 4\pi X \ \cos 4\pi X]$
	Chebyshev polynomials	Input vector X is expanded into 8 terms as follow $T_0(x) = 1$ $T_1(x) = x$ $T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x) \ for \ 8 \geq n \geq 2$
	Legendre polynomials	Input vector X is expanded into n terms as follow $L_0(x) = 1$ $L_1(x) = x$ $L_n(x) = \frac{(2n-1)xL_{n-1}(x) - (n-1)L_{n-2}(x)}{n} \ for \ 8 \geq n \geq 2$

Table 5 shows an overview of normalization and dimension expansion method used in the proposed solution. Let X be a 14-dimensional input vector given by $X = [x_1, x_2, \dots, x_{14}]$. Each element $x_i (i = 1 \dots 14)$ was expanded into 8 more terms shown in table 4. As a result, the processed input data consisted of 126 features in total.

6.5. Training model

Instead of training only a single predictive model, an ensemble of 10 models was used to obtain better predictive performance than one single model could achieve. In this method, 10 FLANN models were created with different initial weights.

Each model was trained using mini-batch gradient descent algorithm. This algorithm scans through the entire training set, each time it takes a number of training example (a training batch) and updates the parameters according to the gradient of the error with respect to that group of training examples. As a result, it can immediately start making progress and get close to the minimum quickly. However, since it only learns from a small group in the whole training set, it may never converge exactly to the minimum. Instead, it may oscillate between points around the minimum value which can be reasonably good approximation to the true minimum[20]. This process is repeated for multiple iterations.

The performance of a model on testing set was check in every iteration to select the settings which could work well on testing set. In an iteration, the results of all model in the ensemble were average to obtain the final prediction at that iteration. After finishing training, the best prediction results produced by each model in the ensemble were averaged to obtain the ultimate prediction.

6.6. Evaluation metrics:

The performance of the model can be asses using these metrics:

- Root mean Squared error (RMSE): it measures the different between predicted values and actual values

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Where:
$n \text{ is the number of predictions}$
$y_i \text{ is the } i \text{ predicted value}$
$\hat{y}_i \text{ is the } i \text{ actual value}$

- Mean absolute error (MAE): it measures how close forecasts or predictions are to the predicted outcomes

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Where:
$n \text{ is the number of predictions}$
$y_i \text{ is the } i \text{ predicted value}$
$\hat{y}_i \text{ is the } i \text{ actual value}$

- Hit rate: it measures the accuracy of the prediction of tomorrow's closing price compared to today. It values in the range between 0 and 1 where 1 denotes all the trends predicted correctly, 0 denotes all the trends are predicted incorrectly.

$$RMSE = \frac{1}{n}\sum_{i=1}^{n}u[(\hat{y}_i - y_{i-1})(y_i - y_{i-1})]$$

Where:
$n \text{ is the number of predictions,}$
$y_i \text{ is the } i \text{ predicted value}$
$\hat{y}_i \text{ is the } i \text{ actual value}$

6.7. Network configuration

In machine learning, it is observed that the training of predictive model heavily depends on the learning rate. Therefore, two network configurations with different learning rate was set up to construct the proposed solution. The details of two configurations are described in table 6.

Table 6: Network configurations for simulation

	Configuration 1	Configuration 2
Network	FLANN	FLANN
Input data	IBM stock data	IBM stock data
Input range	[-1,1]	[-1,1]
No. training samples (m)	3373	3373
Batch size	30	30
No. Training batch $\left(\frac{m}{Batch\ size}\right)$	113	113
No. Epochs (N)	20000	10000
Input layer	126 units	126 units
Hidden layer	No hidden layer	No hidden layer
Output layer	1 unit tanh activation function	1 unit tanh activation function
Learning Rate	Fix learning rate: 10^{-3}	Decay learning rate Initial learning rate: $\alpha_0 = 10^{-3}$ After 3000 epochs $\alpha_t = \alpha_0\left(1 - \frac{n_t - 3000}{N}\right)$ Where n_t is the current epoch
Regularization	No regularization	No regularization
Cost function	$\frac{1}{m}\sum_{1}^{m}\left(Y_{true} - Y_{predicted}\right)^2$	$\frac{1}{m}\sum_{1}^{m}\left(Y_{true} - Y_{predicted}\right)^2$
Objectives	Minimize the cost function	Minimize the cost function

7. Simulation results

7.1. Configuration 1

- **Performance on training set**

Table 7: Cost value of training models with configuration 1

Epoch	Cost value		
	Trigonometric	Chebyshev	Legendre
0	0.2662	0.2419	0.3904
500	0.0450	0.0991	0.0250
1000	0.0467	0.0810	0.0313
1500	0.0417	0.0805	0.0288
2000	0.0394	0.0615	0.0251
2500	0.0453	0.0619	0.0291
3000	0.0394	0.0684	0.0233
3500	0.0506	0.0889	0.0338
4000	0.0390	0.0848	0.0212
4500	0.0525	0.0751	0.0307
5000	0.0471	0.0547	0.0231
5500	0.0387	0.0582	0.0318
6000	0.0485	0.0896	0.0216
6500	0.0426	0.0952	0.0298
7000	0.0488	0.1270	0.0227
7500	0.0470	0.1016	0.0308
8000	0.0359	0.0731	0.0208
8500	0.0479	0.0726	0.0287
9000	0.0426	0.1001	0.0226
9500	0.0498	0.0926	0.0299
10000	0.0425	0.0679	0.0204
10500	0.0403	0.0889	0.0267
11000	0.0463	0.0777	0.0233
11500	0.0395	0.0964	0.0279
12000	0.0513	0.0588	0.0220
12500	0.0394	0.0986	0.0235
13000	0.0530	0.0613	0.0251
13500	0.0473	0.0903	0.0246
14000	0.0390	0.0866	0.0244
14500	0.0487	0.0968	0.0215
15000	0.0428	0.0738	0.0267
15500	0.0489	0.0930	0.0211
16000	0.0471	0.0771	0.0245
16500	0.0360	0.0770	0.0215
17000	0.0481	0.0819	0.0259
17500	0.0428	0.0889	0.0209
18000	0.0500	0.0968	0.0223
18500	0.0426	0.0583	0.0236
19000	0.0404	0.0988	0.0223
19500	0.0465	0.0605	0.0235
20000	0.0397	0.0902	0.0211

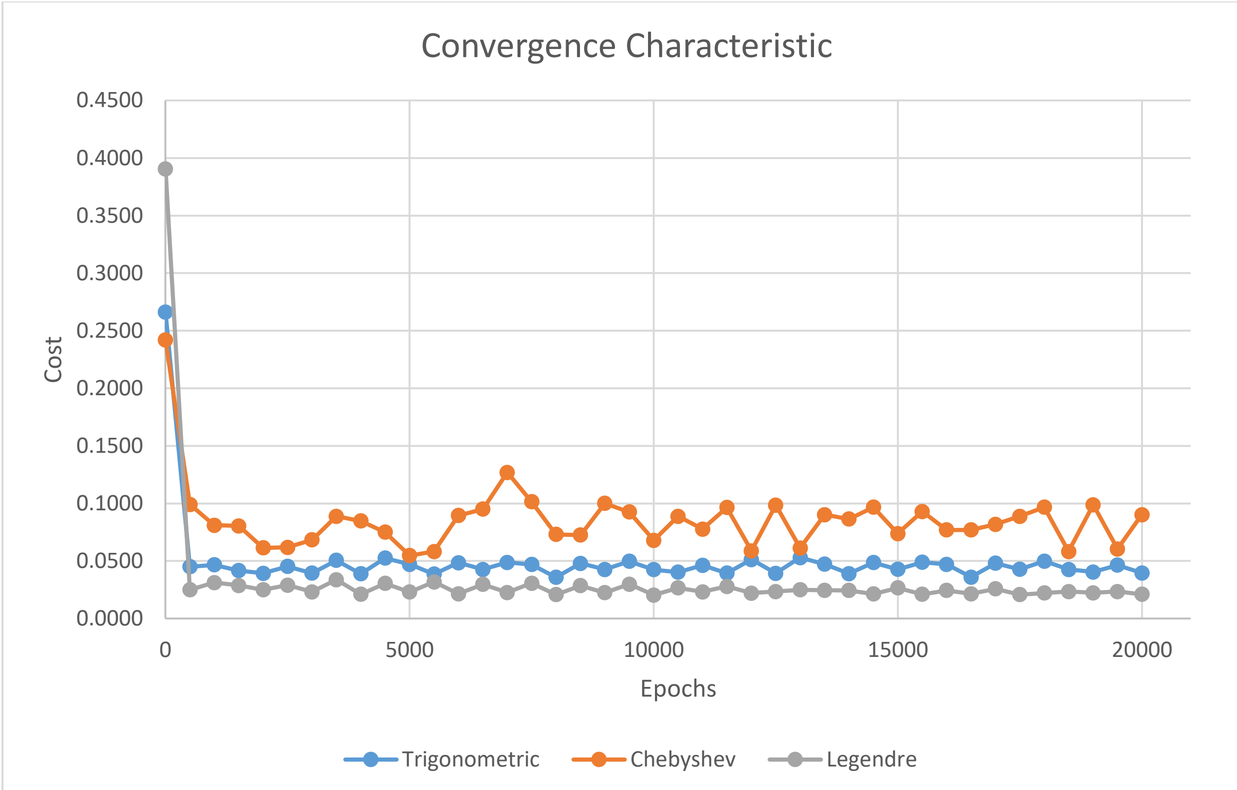


Figure 8: Convergence characteristic of three models with different expansion methods for 2000 epochs

Figure 8 shows the convergence characteristic of three models with different expansion functions. For the first configuration, the overall trend is a reduction of cost value for the first 10000 epochs. With the learning rate of 10^{-3}, the figure and the table data shows a quick learning progress of all three models as the cost values declined significantly for the first 500 epochs. After that, the cost values leveled off. However, from 10000 epochs onwards, the cost values fluctuated and did not show any sight of decreasing anymore. Among three models, the model with Legendre expansion method had in the lowest cost value while the one with Chebyshev expansion method not only resulted in higher cost values but also unstable oscillation in cost values. Figure 9 – 14 shows the details performance of model with Legendre expansion function at 0^{th}, 10000^{th} and 20000^{th} epoch.

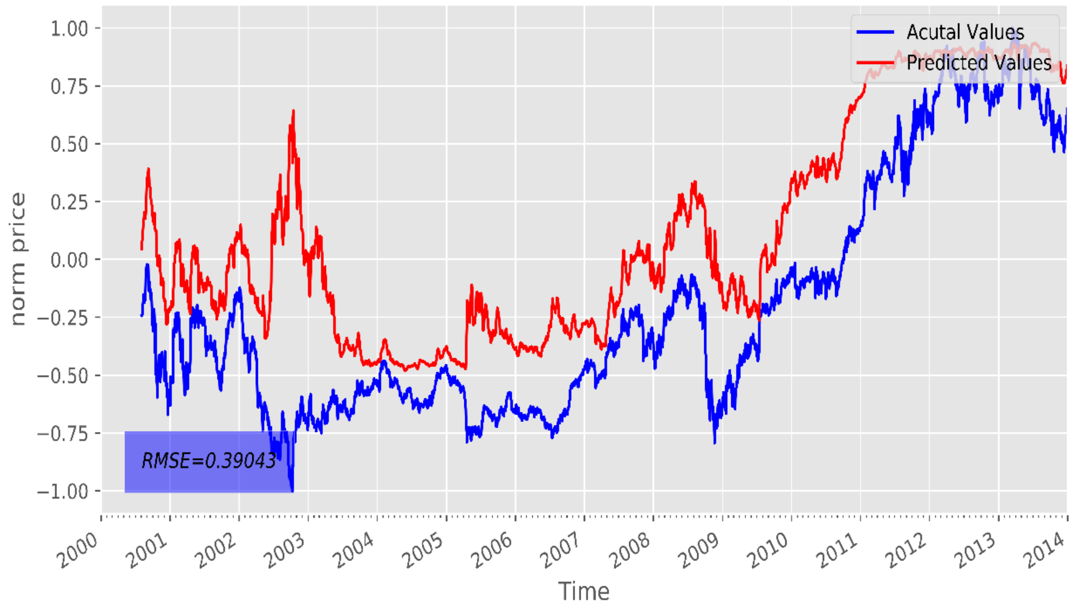


Figure 9: Actual values and predicted values of training set at 0th epoch of Legendre model in configuration 1

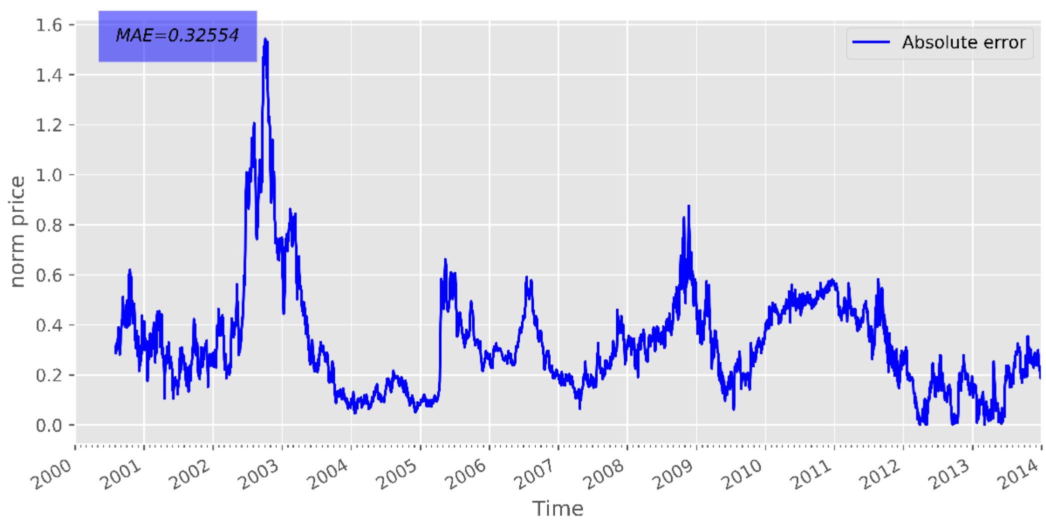


Figure 10: Mean absolute errors between actual value and predicted value of training set at 0th epoch of Legendre model configuration 1

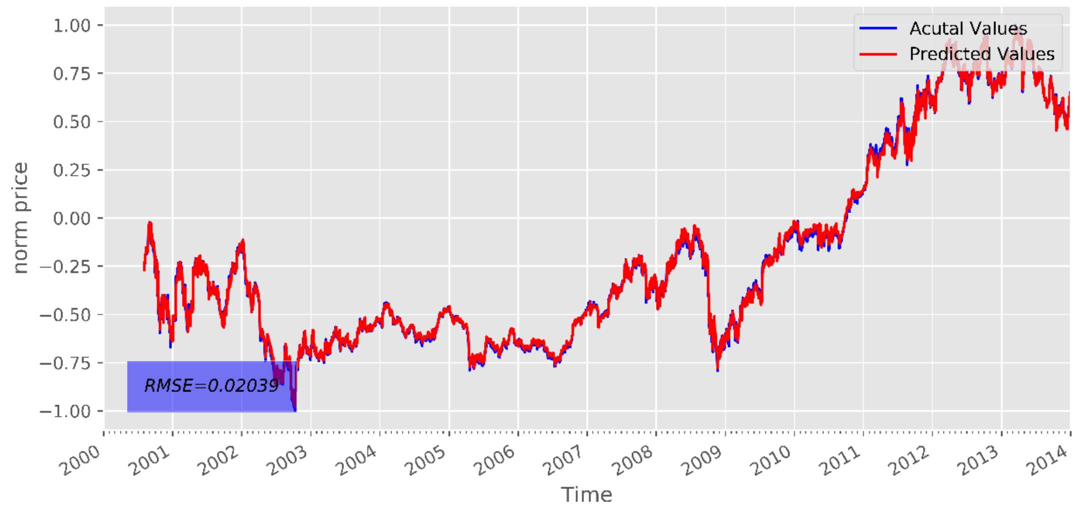


Figure 11: Actual values and Predicted values of training set at 10000th epoch of Legendre model configuration 1

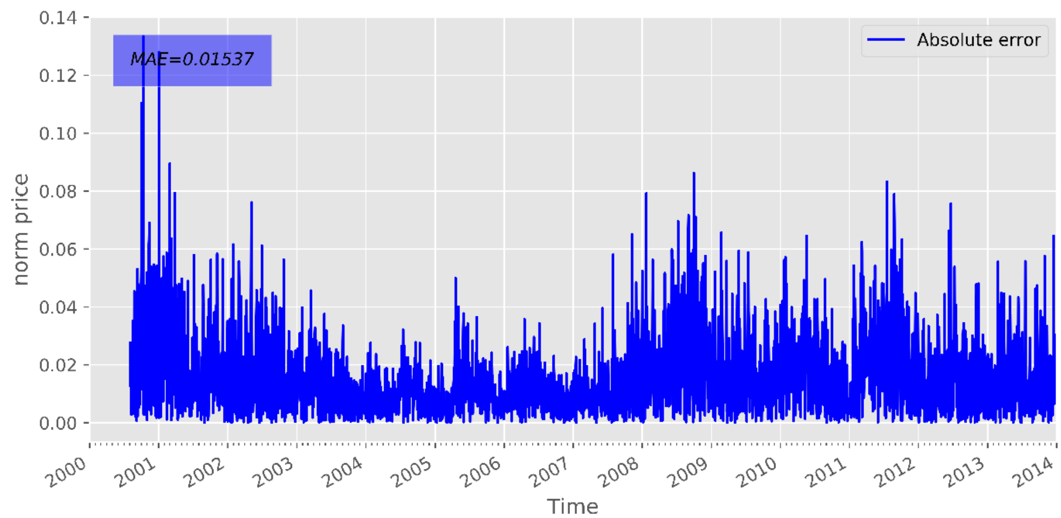


Figure 12:Mean absolute errors between actual value and predicted value of training set at 10000th epoch of Legendre model configuration 1

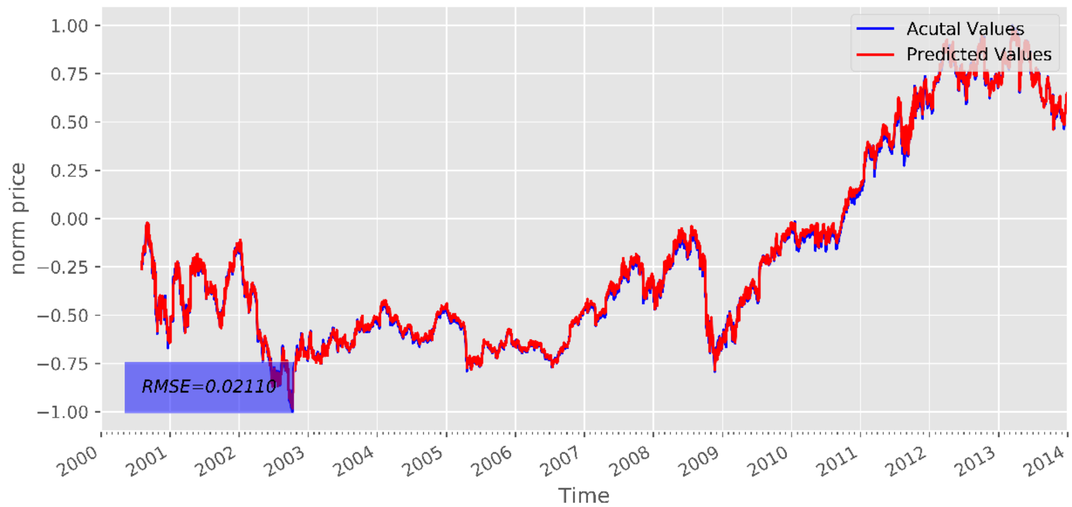


Figure 13: Actual values and Predicted values of training set at 20000th epoch of Legendre model configuration 1

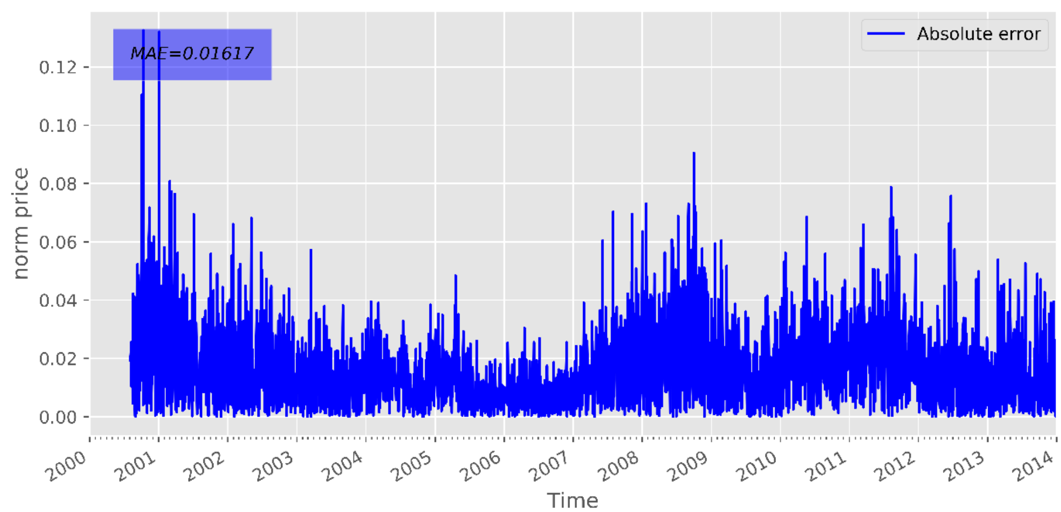


Figure 14: Mean absolute errors between actual value and predicted value of training set at 20000th epoch of Legendre model configuration 1

- **Performance on test set**

Table 8: Performance of models with different expansion methods on test set in configuration 1

Epoch	Root Mean Square Error (RMS)			Mean Absolute Error (MAE)			Hit Rate		
	Tri	Che	Leg	Trig	Che	Leg	Trig	Che	Leg
0	0.5394	0.4507	0.5518	0.5127	0.4269	0.5116	0.4967	0.4967	0.4967
500	0.4772	0.3578	0.3034	0.4269	0.3164	0.2663	0.5099	0.5020	0.5033
1000	0.3869	0.2671	0.2350	0.3431	0.2300	0.1989	0.5099	0.5139	0.5073
1500	0.3029	0.2297	0.2006	0.2651	0.1958	0.1640	0.5073	0.5033	0.5219
2000	0.2363	0.2425	0.2068	0.2044	0.2052	0.1693	0.5113	0.5033	0.5219
2500	0.1789	0.2372	0.2235	0.1512	0.1993	0.1846	0.5179	0.5007	0.5152
3000	0.1269	0.1887	0.2144	0.1023	0.1518	0.1752	0.5126	0.5060	0.5139
3500	0.1024	0.1463	0.1871	0.0811	0.1132	0.1505	0.5152	0.5245	0.5126
4000	0.0774	0.2194	0.1858	0.0586	0.1816	0.1490	0.5298	0.4980	0.5285
4500	0.0676	0.1133	0.2090	0.0503	0.0845	0.1711	0.5245	0.5285	0.5179
5000	0.0625	0.1796	0.2087	0.0465	0.1454	0.1702	0.5417	0.5073	0.5152
5500	0.0644	0.1066	0.1836	0.0471	0.0784	0.1476	0.5616	0.5430	0.5099
6000	0.0821	0.1123	0.1769	0.0624	0.0851	0.1410	0.5073	0.5179	0.5325
6500	0.0852	0.1306	0.1998	0.0654	0.0991	0.1630	0.5232	0.5192	0.5205
7000	0.0829	0.1007	0.2008	0.0631	0.0820	0.1633	0.5086	0.4927	0.5152
7500	0.1138	0.1449	0.1758	0.0896	0.1147	0.1406	0.5126	0.5073	0.5126
8000	0.1181	0.0860	0.1712	0.0929	0.0699	0.1361	0.5179	0.4980	0.5298
8500	0.1291	0.0857	0.1941	0.1050	0.0696	0.1581	0.5232	0.4980	0.5219
9000	0.1310	0.0890	0.1912	0.1049	0.0726	0.1547	0.5232	0.5126	0.5166
9500	0.1432	0.1199	0.1661	0.1172	0.1001	0.1318	0.5099	0.4808	0.5139
10000	0.1562	0.0695	0.1682	0.1303	0.0573	0.1338	0.5113	0.4967	0.5298
10500	0.1547	0.0744	0.1894	0.1296	0.0586	0.1541	0.5086	0.4967	0.5205
11000	0.1643	0.1445	0.1800	0.1367	0.1201	0.1447	0.5166	0.4834	0.5179
11500	0.1827	0.1510	0.1569	0.1544	0.1267	0.1235	0.5046	0.4781	0.5166
12000	0.1740	0.0747	0.1674	0.1446	0.0603	0.1337	0.5126	0.4887	0.5258
12500	0.1819	0.0901	0.1832	0.1528	0.0680	0.1484	0.5126	0.4940	0.5166
13000	0.1856	0.1957	0.1671	0.1551	0.1582	0.1331	0.5166	0.4954	0.5152
13500	0.2015	0.2161	0.1500	0.1684	0.1769	0.1172	0.5113	0.4795	0.5232
14000	0.1941	0.1402	0.1677	0.1638	0.1129	0.1345	0.5086	0.4768	0.5232
14500	0.2163	0.1219	0.1747	0.1803	0.0920	0.1404	0.5060	0.4715	0.5219
15000	0.2206	0.2145	0.1536	0.1846	0.1667	0.1208	0.5086	0.4914	0.5139
15500	0.2106	0.2717	0.1469	0.1740	0.2202	0.1147	0.5152	0.4781	0.5298
16000	0.2425	0.2443	0.1672	0.2027	0.1953	0.1344	0.5099	0.4768	0.5232
16500	0.2444	0.1824	0.1643	0.2036	0.1361	0.1306	0.5126	0.4768	0.5166
17000	0.2587	0.2412	0.1421	0.2145	0.1830	0.1101	0.5113	0.4795	0.5192
17500	0.2525	0.3339	0.1476	0.2107	0.2637	0.1158	0.5073	0.4927	0.5285
18000	0.2681	0.3323	0.1637	0.2218	0.2639	0.1312	0.5073	0.4808	0.5152
18500	0.2820	0.2280	0.1514	0.2345	0.1752	0.1189	0.5046	0.4834	0.5166
19000	0.2797	0.2615	0.1359	0.2316	0.1946	0.1044	0.5073	0.4861	0.5298
19500	0.2857	0.3835	0.1502	0.2368	0.2983	0.1188	0.5099	0.4914	0.5232
20000	0.3064	0.3970	0.1556	0.2547	0.3115	0.1232	0.5060	0.4821	0.5219

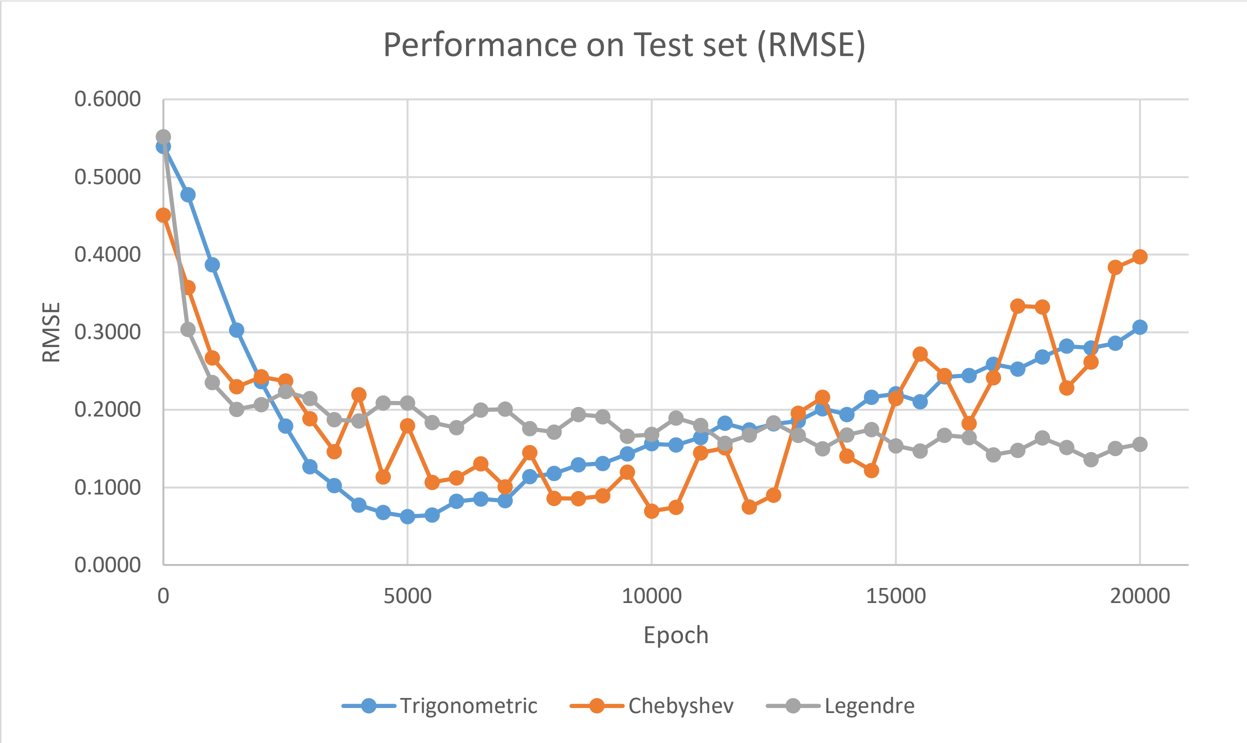


Figure 15: Root mean square errors between actual and predicted values of test set throughout training process produced from models with three different expansion function in configuration 1

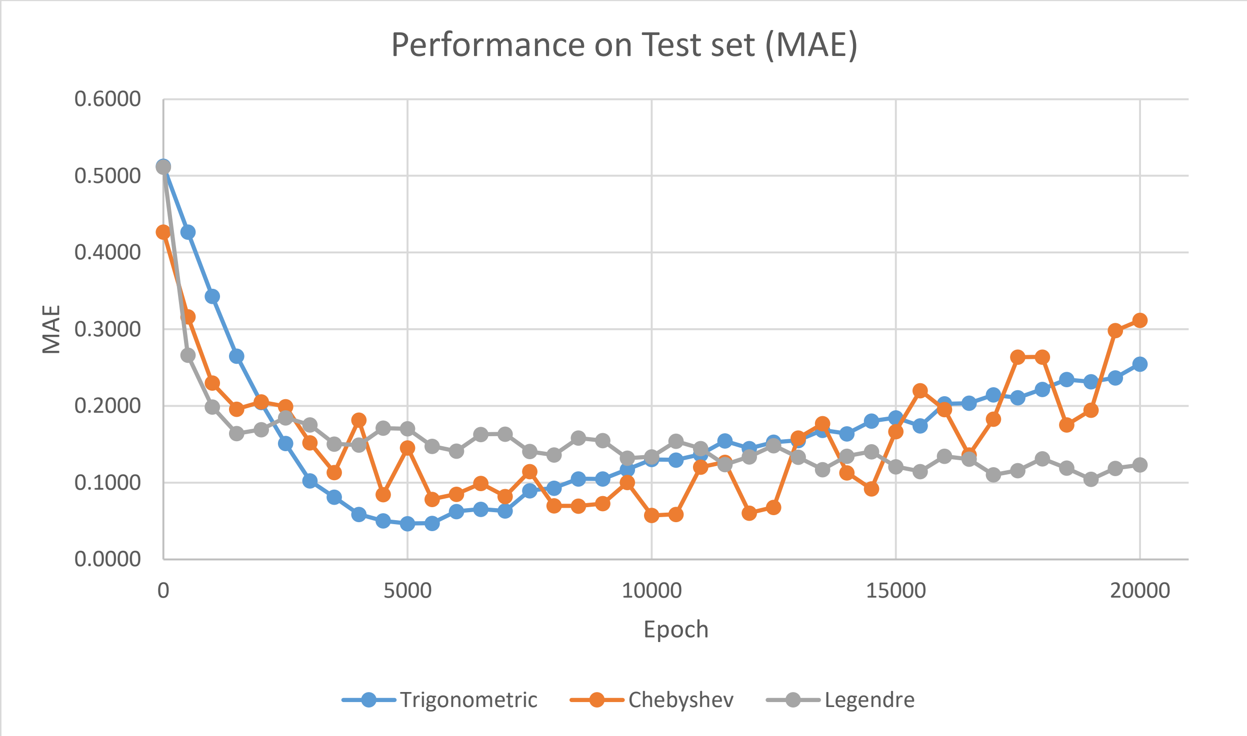


Figure 16: Mean square errors between actual and predicted values of test set throughout training process produced from models with three different expansion function in configuration 1

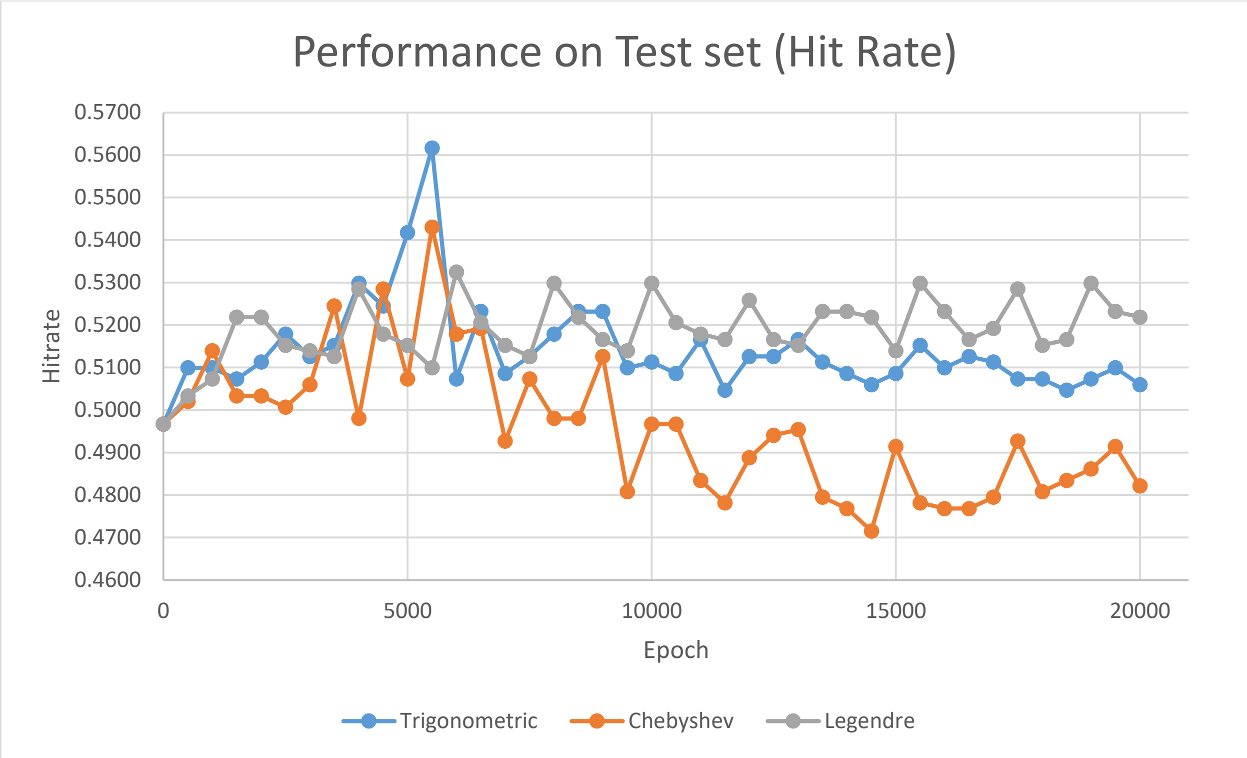


Figure 17: Hit rate between actual and predicted values of test set throughout training process produced from models with three different expansion function in configuration 1

Figures 15 – 17 shows the performance of three models on the test set which including RMSE, MAE and Hit rate respectively throughout training process. For model with trigonometric expansion, the RMSE and MAE reached the lowest error value at around 5000th epoch. The lowest RMSE and MAE values produced by this model also lower compared to the lowest RMSE and MAE produced by other models. The hit rate of trigonometric model was also better than the other two. Its hit rate peaked approximately at the same iteration with the lowest RMSE and MAE. However, after 5000 epochs, the RMSE and MAE values started to increase and the hit rate started to decrease. Despite of having the best performance on training set, the model with Legendre expansion resulted in much higher RMSE and MAE than the trigonometric one. This is clearly due to overfitting as the model overfit the training set and failed to generalize on the test set. For the fix learning rate configuration, the model with Chebyshev expansion had the poorest performance as all of its evaluation metrics results are significantly worse than the other two models.

As the training reached above 10000 iterations, the performance on test set of all models decreased significantly. As the training progresses, the models not only learn the details but also noise in the training data to the extent that it becomes a detriment to the capability of generalization on test set which consists of unseen data to the models[37]. Therefore, in the next simulation, to prevent overfitting, the model was only trained for 10000 epochs.

- **The best performance**

The average prediction of the best selected prediction from each model is shown in table 8. Based on the results displayed in table 8, it is clearly that model with trigonometric expansion had the best performance on test set as its RMSE and MAE between predicted and actual values of test set was respectively 0.0609 and 0.0762 which were remarkably higher than other two models. Therefore, the best model settings of configuration 1 is the model with trigonometric expansion.

Figure 18 shows the actual and predicted values produced by the model with trigonometric expansion.

Figure 19 shows the mean square errors between actual and predicted values produced by the model with trigonometric expansion.

Table 9: The best performance of each model

Epoch	Root Mean Square Error (RMS)			Mean Absolute Error (MAE)			Hit Rate		
	Tri	Che	Leg	Trig	Che	Leg	Trig	Che	Leg
Best performance	**0.0609**	**0.0762**	**0.0931**	**0.0499**	**0.0621**	**0.0601**	**0.5126**	**0.4954**	**0.5497**

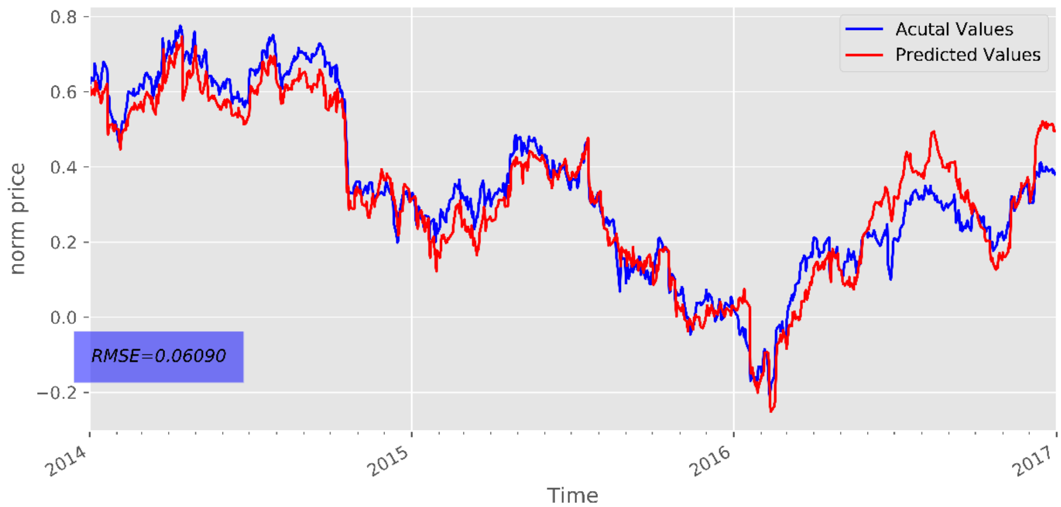


Figure 18: Best prediction generated by models with trigonometric function expansion configuration 1

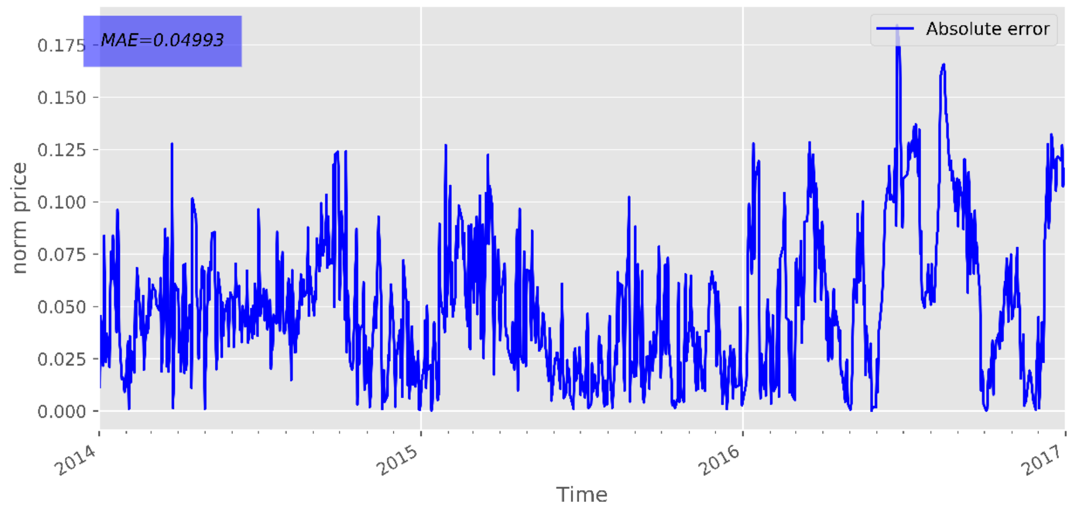


Figure 19: the mean square errors between actual and predicted values produced by the model with trigonometric expansion configuration 1

7.2. Configuration 2

➢ **Performance on training set**

Table 10: Cost value of training models with configuration 2

Epoch	Root Mean Square Error (RMS)		
	Tri	Che	Leg
0	0.4991	0.3013	0.5520
500	0.5075	0.0871	0.2892
1000	0.4047	0.0638	0.2442
1500	0.3145	0.0988	0.2499
2000	0.2625	0.0543	0.2384
2500	0.2212	0.0816	0.2170
3000	0.1686	0.0704	0.2109
3500	0.1068	0.0837	0.2262
4000	0.1082	0.0500	0.2135
4500	0.0863	0.0562	0.2214
5000	0.0747	0.0475	0.2103
5500	0.0820	0.0530	0.2117
6000	0.0563	0.0322	0.2139
6500	0.0666	0.0333	0.2138
7000	0.0630	0.0323	0.2120
7500	0.0788	0.0342	0.2119
8000	0.0759	0.0361	0.2135
8500	0.0598	0.0414	0.2151
9000	0.0824	0.0219	0.2137
9500	0.0862	0.0229	0.2153
10000	0.0806	0.0359	0.2152

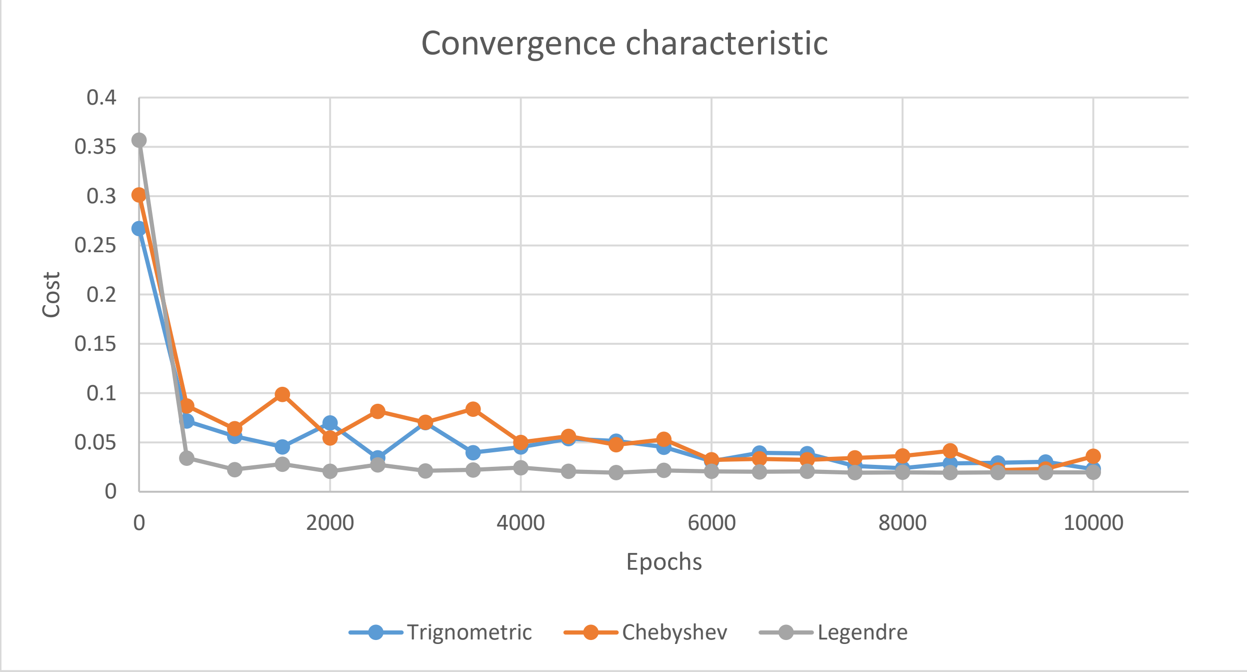


Figure 20: Convergence characteristic of three models with different expansion methods for 2000 epochs

Figure 20 shows the convergence characteristic of three models with different expansion functions in configuration 2. For the second configuration, the overall trend is a remarkable decrease of cost values throughout the whole training processes. With the decay learning rate, the figure and the table data shows a quick learning progress of all three models as the cost values declined significantly for the first 500 epochs. After that, the cost values continued to decrease. Among three models, again, the model with Legendre expansion method had in the lowest cost value and converged faster for the first 6000 epochs. However, as the training progressed, there was not a difference between convergence speed between all three models compared to the fix learning rate configuration. Figure 9 – 14 shows the details performance of model with Legendre expansion function at 0th, 10000th and 20000th epoch.

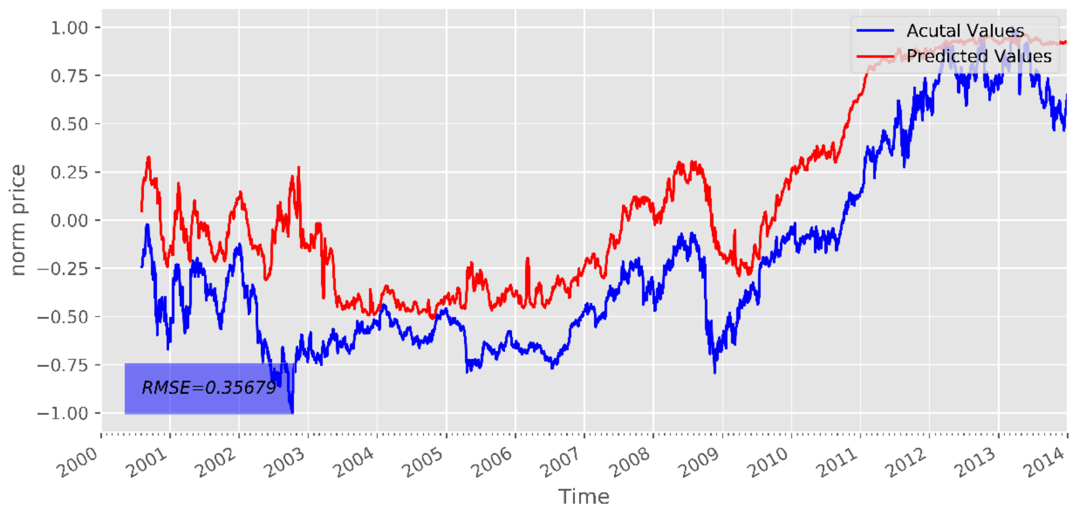


Figure 21:Actual values and predicted values of training set at 0th epoch of Legendre model in configuration 2

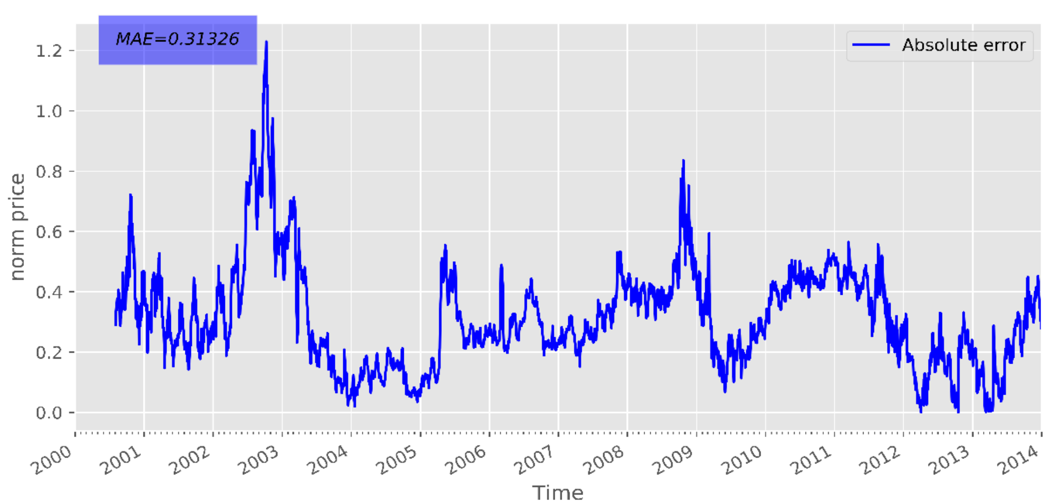


Figure 22: Mean absolute errors between actual value and predicted value of training set at 0th epoch of Legendre model configuration 2

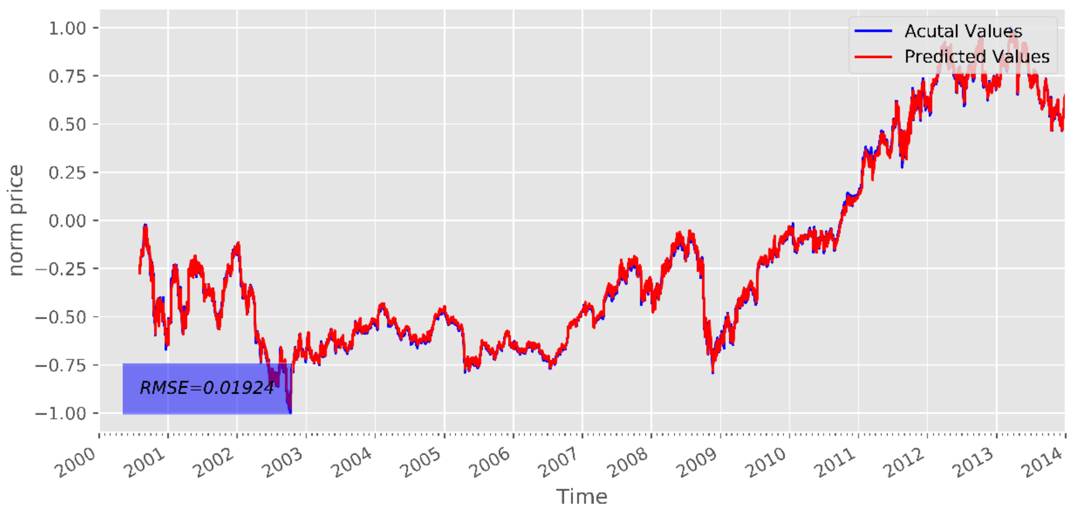


Figure 23: Actual values and predicted values of training set at 10000th epoch of Legendre model in configuration 2

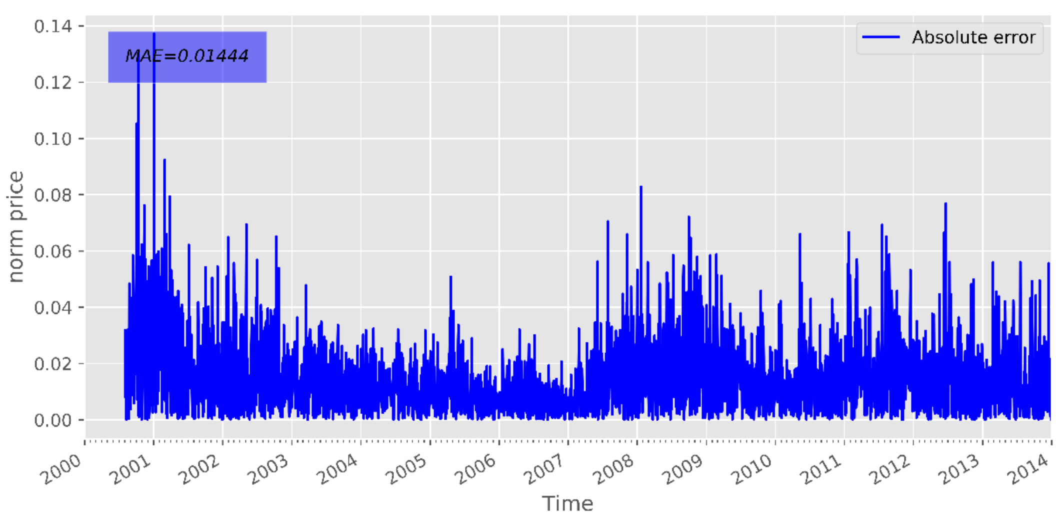


Figure 24:Mean absolute errors between actual value and predicted value of training set at 10000th epoch of Legendre model configuration 2

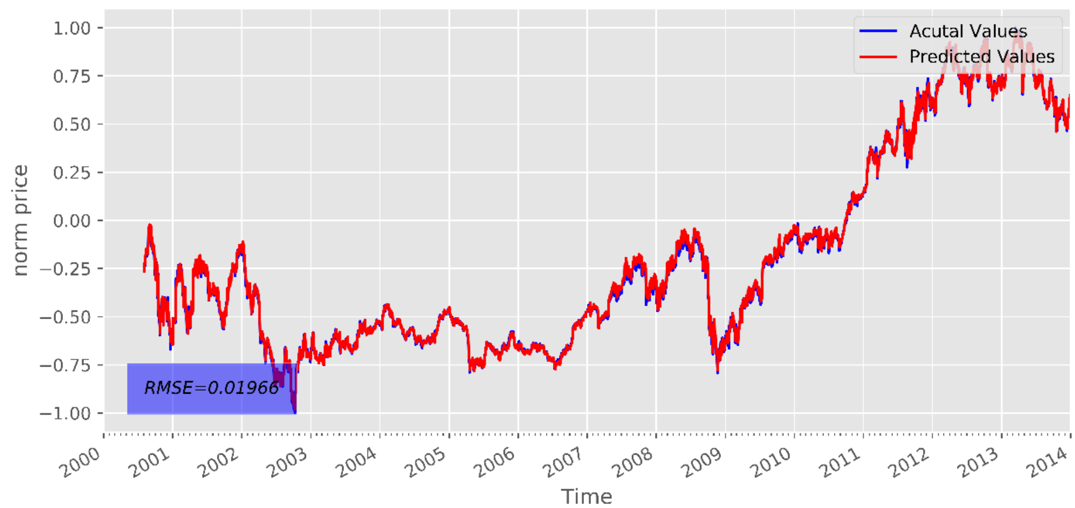


Figure 25: Actual values and predicted values of training set at 20000th epoch of Legendre model in configuration 2

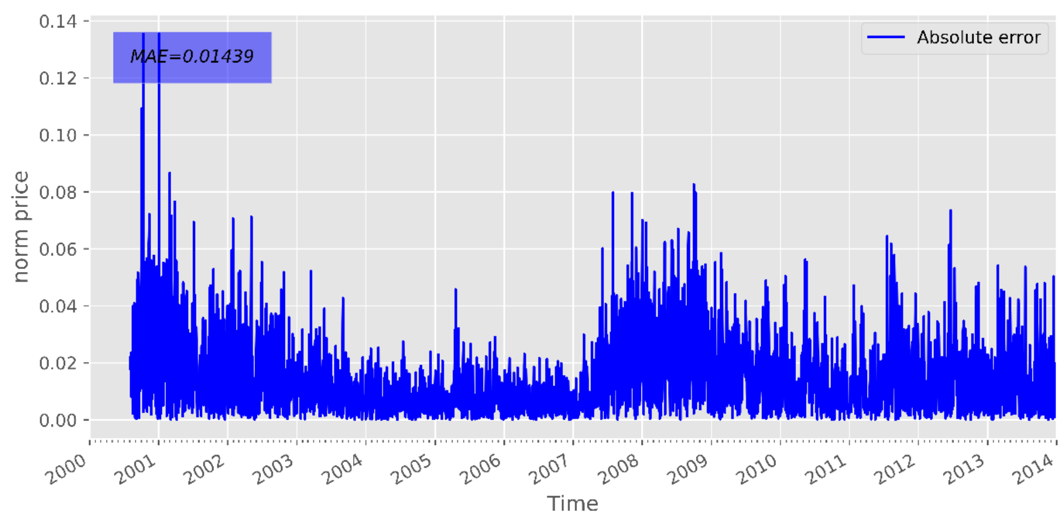


Figure 26: Mean absolute errors between actual value and predicted value of training set at 20000th epoch of Legendre model configuration 2

- **Performance on Test set**

Table 11: Performance of models with different expansion methods on test set in configuration 2

Epoch	Root Mean Square Error (RMS)			Mean Absolute Error (MAE)			Hit Rate		
	Tri	Che	Leg	Trig	Che	Leg	Trig	Che	Leg
0	0.4991	0.4251	0.5520	0.4709	0.3946	0.5211	0.4967	0.4967	0.4967
500	0.5075	0.3475	0.2892	0.4527	0.3054	0.2517	0.5099	0.5007	0.5033
1000	0.4047	0.2618	0.2442	0.3579	0.2269	0.2047	0.5113	0.4993	0.5205
1500	0.3145	0.1956	0.2499	0.2741	0.1616	0.2079	0.5073	0.5099	0.5166
2000	0.2625	0.2279	0.2384	0.2279	0.1906	0.1968	0.5099	0.5033	0.5099
2500	0.2212	0.2450	0.2170	0.1859	0.2034	0.1759	0.5073	0.5113	0.5205
3000	0.1686	0.1844	0.2109	0.1402	0.1487	0.1715	0.4980	0.5073	0.5139
3500	0.1068	0.1936	0.2262	0.0821	0.1574	0.1840	0.5245	0.5033	0.5219
4000	0.1082	0.1941	0.2135	0.0833	0.1560	0.1724	0.5139	0.5046	0.5232
4500	0.0863	0.1510	0.2214	0.0634	0.1180	0.1793	0.5338	0.5020	0.5298
5000	0.0747	0.1367	0.2103	0.0536	0.1059	0.1702	0.5417	0.5152	0.5258
5500	0.0820	0.1042	0.2117	0.0616	0.0753	0.1705	0.5258	0.5470	0.5325
6000	0.0563	0.1128	0.2139	0.0400	0.0857	0.1727	0.5656	0.5192	0.5245
6500	0.0666	0.1352	0.2138	0.0491	0.1082	0.1726	0.5377	0.5046	0.5245
7000	0.0630	0.1290	0.2120	0.0464	0.1034	0.1705	0.5616	0.5060	0.5338
7500	0.0788	0.1175	0.2119	0.0608	0.0949	0.1704	0.5325	0.5179	0.5311
8000	0.0759	0.1204	0.2135	0.0592	0.0983	0.1714	0.5470	0.5219	0.5364
8500	0.0598	0.1038	0.2151	0.0447	0.0819	0.1731	0.5762	0.5205	0.5311
9000	0.0824	0.1015	0.2137	0.0667	0.0816	0.1714	0.5497	0.5245	0.5364
9500	0.0862	0.1007	0.2153	0.0708	0.0811	0.1726	0.5483	0.5258	0.5338
10000	0.0806	0.1348	0.2152	0.0656	0.1125	0.1722	0.5338	0.5258	0.5351

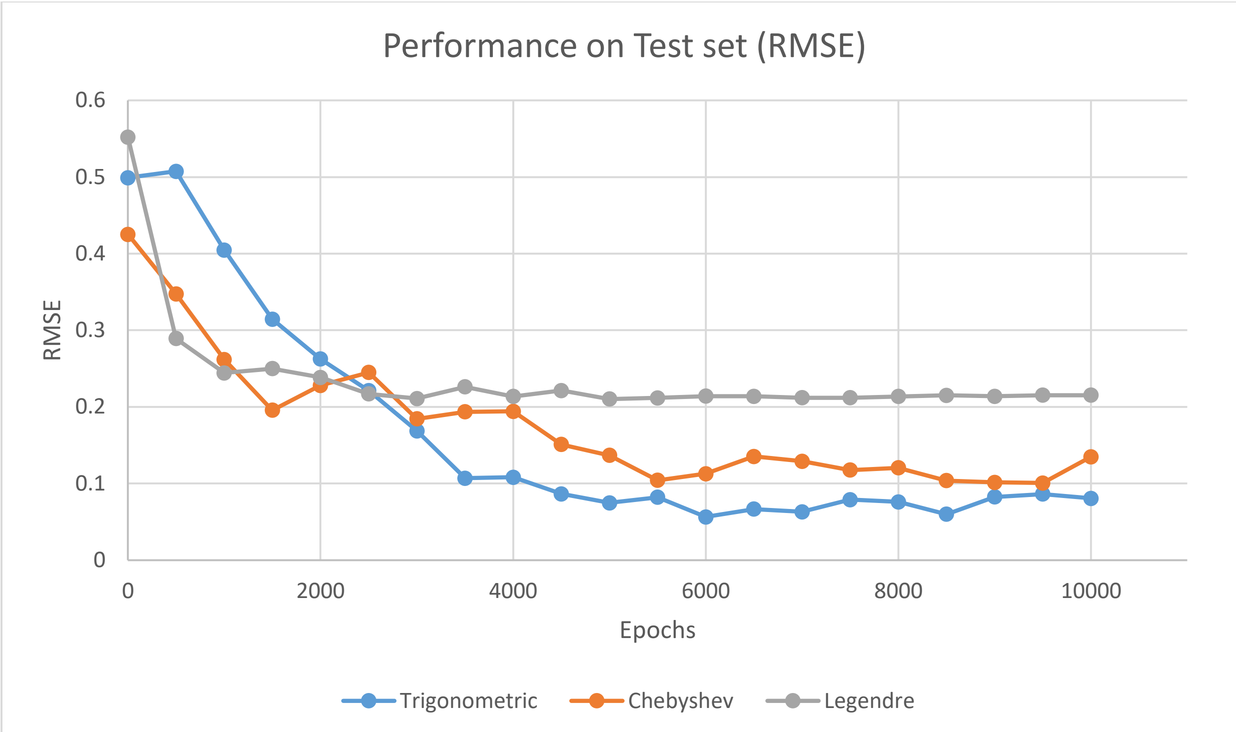


Figure 27: Root mean square errors between actual and predicted values of test set throughout training process produced from models with three different expansion function in configuration 2

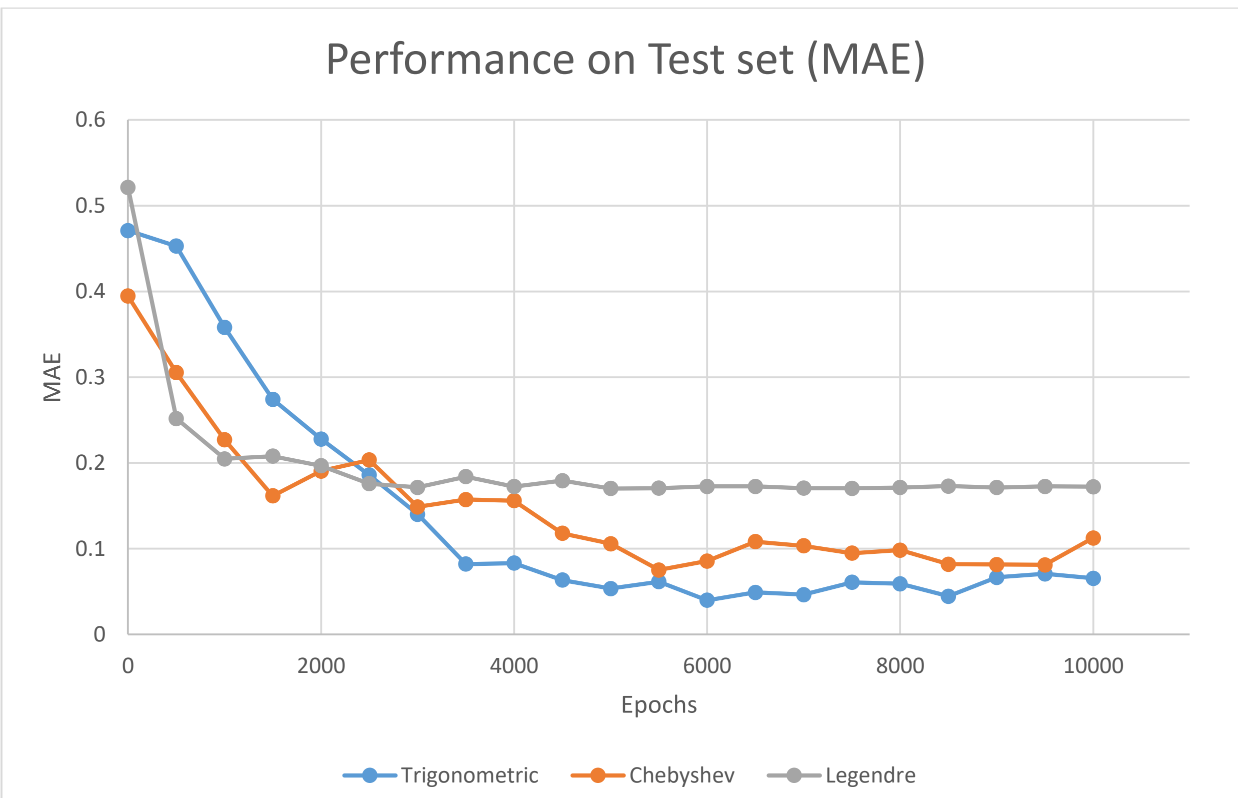


Figure 28: Mean absolute errors between actual and predicted values of test set throughout training process produced from models with three different expansion function in configuration 2

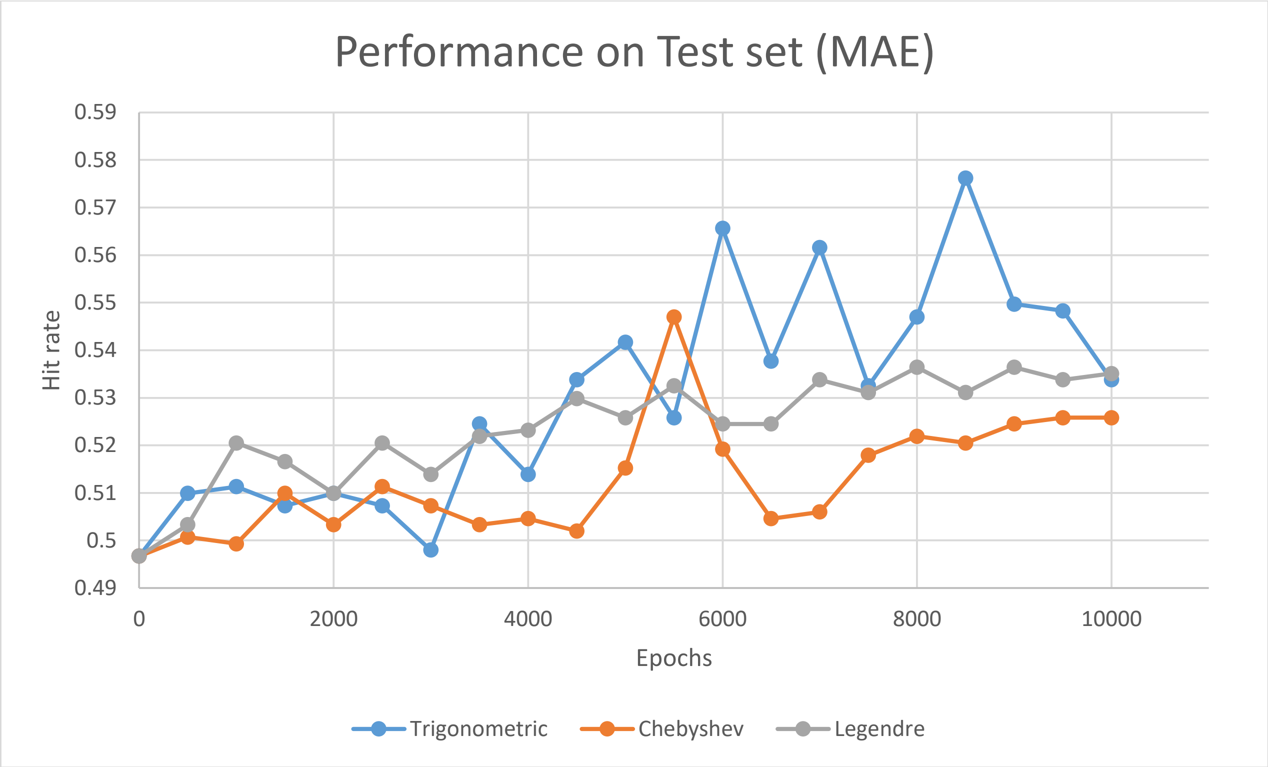


Figure 29: Hit rate between actual and predicted values of test set throughout training process produced from models with three different expansion function in configuration 2

Figures 27 – 29 show the performance of three models on the test set which including RMSE, MAE and Hit rate respectively throughout training process. The values of performance metrics resulted from the second configuration stopped getting worse as the training progressed compared to the first configuration. It confirms the benefits of decay learning rate and the overfitting prevention by reducing training iterations.

For model with trigonometric expansion, the RMSE and MAE bottomed at around 6000th epoch. The RMSE and MAE values resulted from this model also remarkably lower compared to the other two. The hit rate of trigonometric model was also significantly higher than the Chebyshev and Legendre model. Its hit rate peaked approximately at the same iteration with the lowest RMSE and MAE. In this configuration, the same phenomenon happened again with Legendre model as it was overfitted again. The performance of the Chebyshev model was in between the trigonometric and Legendre models.

- **The best performance**

The average prediction of the best selected prediction from each model is shown in table 11. Based on the results displayed in table 11, it is obvious that model with trigonometric expansion had the best performance on test set as its RMSE and MAE between predicted and actual values of test set was respectively 0.0416and 0.0319 which were significantly higher than other two models. Therefore, the best model settings of configuration 2 is the model with trigonometric expansion.

Figure 30 shows the actual and predicted values produced by the model with trigonometric expansion. Figure 31 shows the mean square errors between actual and predicted values produced by the model with trigonometric expansion.

Table 12: The best performance of each model in configuration 2

Epoch	Root Mean Square Error (RMS)			Mean Absolute Error (MAE)			Hit Rate		
	Tri	Che	Leg	Trig	Che	Leg	Trig	Che	Leg
Best performance	**0.0416**	**0.0596**	**0.1618**	**0.0319**	**0.0417**	**0.1260**	**0.5775**	**0.5550**	**0.5166**

For the fix learning rate configuration, the models with the trigonometric expansion function has the best performance.

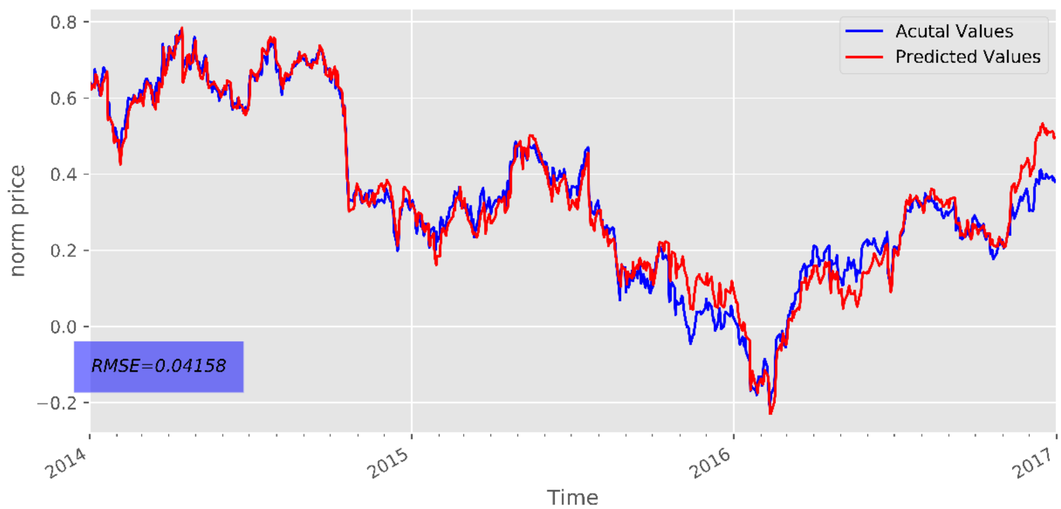


Figure 30: Best prediction generated by models with trigonometric function expansion in configuration 2

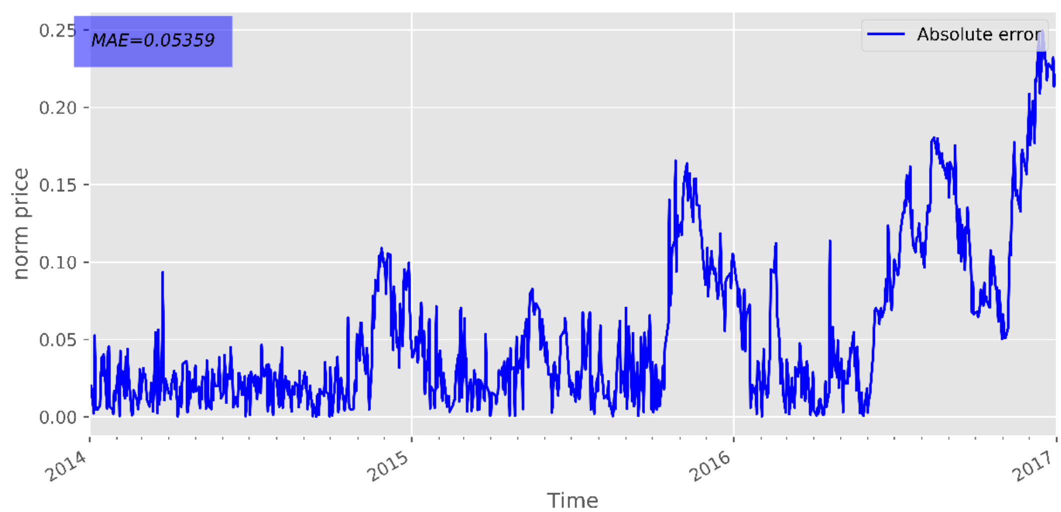


Figure 31: the mean square errors between actual and predicted values produced by the model with trigonometric expansion configuration 2

8. Summary and discussion

- The simulation used two different FLANN configurations to trained models to predict next-day closing price of IBM stock. Input data fed into models were processed by three different expansion method which were Trigonometric, Chebyshev and Legendre.
- In configuration 1, models were trained for 20000 epochs with a fix learning rate using ensemble methods while models were trained for only 10000 epochs with a decay learning rate in configuration 2. The purpose of using less iterations in configuration 2 was to prevent models to learn random noise from input data to prevent overfitting. Not only that, the decay learning rate used in configuration 2 was intended to improve the convergence of models.
- In both configurations, models were trained using ensemble method. The model performance on test set was checked every training iteration to select the best settings which resulted in the best prediction.
- Results from configuration 1 showed that the model with Trigonometric expansion function had the best performance among three dimension-expansion methods. The root means square error (RMSE) between the actual and predicted price on test set produced by the models was relatively low (0.0609).
- Results from configuration 2 confirmed the superior performance of model with Trigonometric expansion function. The performance configuration 2 showed a faster more stable convergence speed than configuration 1. Not only that, the best resulted RMSE from configuration 2 (0.0416) was moderately better than configuration 1. These results validated the importance of choosing learning rate in training models as well as the advantage of decay learning rate.
- The results verify the effectiveness of FLANN in the prediction of stock prices in the next trading day. By taking advantages of macroeconomic series (DJIA index), technical indicators (simple moving averages) besides historical prices, the FLANN model could generate reasonably good prediction (RMSE of 0.0416).
- Even though the simulations resulted in a decent prediction, the predicting horizons was only one day in advance which is not an ideal and practical predicting horizons in the stock market

9. Future work and expansions

- The results from the simulations shows a promising result in predicting IBM stock prices moving using FLANN, it could be expanded into an online prediction model which is capable of predicting stock market in real time with high speed.
- Since there are still many more features which have not been studied in this research such as William%R, MACD or exponential averages (mentioned in section 5.1.2), there are still rooms for improvements of the current models. Furthermore, as there are many variants of neural network approaches which are more complicated than the FLANN such as multilayer perceptron (MLP) and recurrent neural network (RNN), the results from the simulations could be used as a baseline for developments of more sophisticated and accurate models.

Bibliography

[1] S. Wang, Z. Tang, and B. Chai, "Exchange rate prediction model analysis based on improved artificial neural network algorithm," in *2016 International Conference on Communication and Electronics Systems (ICCES)*, 2016, pp. 1-5.

[2] M. M. Rounaghi and F. Nassir Zadeh, "Investigation of market efficiency and Financial Stability between S&P 500 and London Stock Exchange: Monthly and yearly Forecasting of Time Series Stock Returns using ARMA model," *Physica A: Statistical Mechanics and its Applications,* vol. 456, pp. 10-21, 8/15/ 2016.

[3] X. Li, H. Xie, L. Chen, J. Wang, and X. Deng, "News impact on stock price return via sentiment analysis," *Knowledge-Based Systems,* vol. 69, pp. 14-23, 10// 2014.

[4] S. Barak and M. Modarres, "Developing an approach to evaluate stocks by forecasting effective features with data mining methods," *Expert Systems with Applications,* vol. 42, no. 3, pp. 1325-1339, 2/15/ 2015.

[5] M. Ballings, D. Van den Poel, N. Hespeels, and R. Gryp, "Evaluating multiple classifiers for stock price direction prediction," *Expert Systems with Applications,* vol. 42, no. 20, pp. 7046-7056, 11/15/ 2015.

[6] N. L. D. Khoa, K. Sakakibara, and I. Nishikawa, "Stock Price Forecasting using Back Propagation Neural Networks with Time and Profit Based Adjusted Weight Factors," in *2006 SICE-ICASE International Joint Conference*, 2006, pp. 5484-5488.

[7] Y. Chen and Y. Hao, "A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction," *Expert Systems with Applications,* vol. 80, pp. 340-355, 9/1/ 2017.

[8] Investopedia. (26/05/207). *Random Walk Theory*. Available: http://www.investopedia.com/terms/r/randomwalktheory.asp

[9] P. H. Cootner, *The random character of stock market prices*. M.I.T. Press, 1964.

[10] Investopedia. (22/05). *Introduction to Fundamental Analysis*. Available: http://www.investopedia.com/university/fundamentalanalysis/

[11] Investopedia. (22/05). *Fundamental Analysis*. Available: http://www.investopedia.com/terms/f/fundamentalanalysis.asp

[12] Investopedia. (22/05). *Technical Analysis*. Available: http://www.investopedia.com/terms/t/technicalanalysis.asp

[13] H.-I. Suk, "Chapter 1 - An Introduction to Neural Networks and Deep Learning," in *Deep Learning for Medical Image Analysis*: Academic Press, 2017, pp. 3-24.

[14] F. Q. Lauzon, "An introduction to deep learning," in *2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA)*, 2012, pp. 1438-1439.

[15] E. Chong, C. Han, and F. C. Park, "Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies," *Expert Systems with Applications,* vol. 83, pp. 187-205, 10/15/ 2017.

[16] D. Enke and S. Thawornwong, "The use of data mining and neural networks for forecasting stock market returns," *Expert Systems with Applications,* vol. 29, no. 4, pp. 927-940, 11// 2005.

[17] M. M. Mostafa, "Forecasting stock exchange movements using neural networks: Empirical evidence from Kuwait," *Expert Systems with Applications,* vol. 37, no. 9, pp. 6302-6309, 9// 2010.

[18] S. Kumar Chandar, M. Sumathi, and S. N. Sivanandam, *Prediction of Stock Market Price using Hybrid of Wavelet Transform and Artificial Neural Network* (2016). 2016.

[19] C.-F. Tsai and Y.-C. Hsiao, "Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches," *Decision Support Systems,* vol. 50, no. 1, pp. 258-269, 12// 2010.

[20] G. Tingwei, L. Xiu, C. Yueting, and T. Youhua, "Deep learning with stock indicators and two-dimensional principal component analysis for closing price prediction system," in *2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, 2016, pp. 166-169.

[21] F. A. de Oliveira, C. N. Nobre, and L. E. Zárate, "Applying Artificial Neural Networks to prediction of stock price and improvement of the directional prediction index – Case study of PETR4, Petrobras, Brazil," *Expert Systems with Applications,* vol. 40, no. 18, pp. 7596-7606, 12/15/ 2013.

[22] (22/05). *IBM Products*. Available: https://www.ibm.com/products/en-us/?lnk=hmpr

[23] J. Spohrer, "IBM's service journey: A summary sketch," *Industrial Marketing Management,* vol. 60, pp. 167-172, 1// 2017.

[24] (22/05). *IBM Labs*. Available: http://www.research.ibm.com/labs/

[25] A. Ng, J. Ngiam, C. Y. Foo, Y. Mai, and C. Suen. (06-04). *Unsupervised Feature Learning and Deep Learning Tutorial*. Available: http://deeplearning.stanford.edu/wiki/index.php/UFLDL_Tutorial

[26] Y. H. Pao and Y. Takefuji, "Functional-link net computing: theory, system architecture, and functionalities," *Computer,* vol. 25, no. 5, pp. 76-79, 1992.

[27] J. C. Patra and C. Bornand, "Nonlinear dynamic system identification using Legendre neural network," in *The 2010 International Joint Conference on Neural Networks (IJCNN)*, 2010, pp. 1-7.

[28] J. C. Patra and A. C. Kot, "Nonlinear dynamic system identification using Chebyshev functional link artificial neural networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics),* vol. 32, no. 4, pp. 505-511, 2002.

[29] "A comparative study of three artificial neural networks for the detection and classification of gear faults," *International Journal of General Systems,* vol. 34, no. 3, pp. 261-277, 2005/06/01 2005.

[30] P. P. Raghu, R. Poongodi, and B. Yegnanarayana, "A combined neural network approach for texture classification," *Neural Networks,* vol. 8, no. 6, pp. 975-987, // 1995.

[31] R. Ghazali, A. J. Hussain, and P. Liatsis, "Dynamic Ridge Polynomial Neural Network: Forecasting the univariate non-stationary and stationary trading signals," *Expert Systems with Applications,* vol. 38, no. 4, pp. 3765-3776, 4// 2011.

[32] R. Majhi, G. Panda, and G. Sahoo, "Development and performance evaluation of FLANN based model for forecasting of stock markets," *Expert Systems with Applications,* vol. 36, no. 3, Part 2, pp. 6800-6808, 4// 2009.

[33] J. C. Patra, N. C. Thanh, and P. K. Meher, "Computationally efficient FLANN-based intelligent stock price prediction system," in *2009 International Joint Conference on Neural Networks*, 2009, pp. 2431-2438.

[34] F. A. d. Oliveira, L. E. Zárate, M. d. A. Reis, and C. N. Nobre, "The use of artificial neural networks in the analysis and prediction of stock prices," in *2011 IEEE International Conference on Systems, Man, and Cybernetics*, 2011, pp. 2151-2155.

[35] M. Qiu, Y. Song, and F. Akagi, "Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market," *Chaos, Solitons & Fractals,* vol. 85, pp. 1-7, 4// 2016.

[36] https://www.tensorflow.org/. (06-04). Available: https://www.tensorflow.org/

[37] J. Brownlee. (26/05). *Overfitting and Underfitting With Machine Learning Algorithms*. Available: http://machinelearningmastery.com/overfitting-and-underfitting-with-machine-learning-algorithms/

Appendix 1

Project progress

Week 1

Date: 2/3/2017
Attendance: Jagdish Patra, all undergraduate students
Topics discussed: Introduction and overview of2017-HS1-EEE40011/RME40005-Final Year Research Project 1 (BET)
Task to be completed: Selecting topics

Week 3

Date: 15/3/2017
Attendance: Jagdish Patra, Ngoc Khanh Nguyen
Topics discussed:

- Finalize topic selection
- Literature exploration

Miles stones/Task completed: select "Artificial Neural network based prediction of stock prices" which is supervised by Jagdish Patra.

Week 4

Date: 21/03/2017
Attendance: Jagdish Patra, Ngoc Khanh Nguyen
Topics discussed:

- Literature reviews
- Forming research questions
- Backgrounds materials in stock price predictions

Miles stones/Task completed: Reviews up to 5 papers related to predicting stock prices using artificial neural networks

Week 5

Date: 29/03/2017
Attendance: Jagdish Patra, Ngoc Khanh Nguyen
Topics discussed:

- Literature reviews
- Backgrounds materials in stock price predictions
- Select target stocks
- Select programming language and toolbox

Miles stones/Task completed:

- Select IBM stock as the initial target stock of the study
- Select python and Tensorflow as toolbox for the research

Week 6

Activities:

- Factors which could affects stock prices
- Study different approaches in stock price prediction
- Mine IBM stock related data from Yahoo finance
- Research plan

Miles stones/Task completed:

- Collect training and testing data for neural network models
- Research plan submission

Week 7

Attendance: Ngoc Khanh Nguyen
Activities:

- Explore IBM stocks data set
- Investigate neural networks approach
- Online modules

Miles stones/Task completed: online modules submission

Week 8

Attendance: Ngoc Khanh Nguyen
Activities:

- Study in details functional link neural network (FLANN)
- Make a design of functional link neural network (FLANN)

Miles stones/Task completed: finish a draft design of FLANN to predict next-day closing price of IBM

Week 9

Attendance: Ngoc Khanh Nguyen
Activities:

- Study in details multilayer perceptron (MLP)
- Make a design of multilayer perceptron (MLP)

Miles stones/Task completed: finish a draft design of FLANN to predict next-day closing price of IBM

Week 10

Date: 10/05/2017
Attendance: Jagdish Patra, Ngoc Khanh Nguyen
Topic discussed:

- Performance of FLANN model and how to improve it
- Performance of MLP model and how to improve it
- Different type of dimension expansion methods

Miles stones/Task completed:

Week 11

Date: 15/05/2017
Attendance: Jagdish Patra, Ngoc Khanh Nguyen
Topic discussed:

- Performance of FLANN model using different type of expansion function
- Improvement of FLANN model

Activities:

- Research report for semester 1

Miles stones/Task completed: Finalize FLANN model design

Week 12

Date: 22/05/2017
Attendance: Jagdish Patra, Ngoc Khanh Nguyen
Topic discussed:

- Performance of FLANN model using different type of expansion function
- Future works and expansion for FLANN model.

Activities:

- Research presentation for semester 1

Miles stones/Task completed:

- Presentation
- Research Report