Stock Price Prediction Using Machine Learning

BARIŞ CEM BEKTAŞ

Prof. Dr. SÜLEYMAN TOSUN
Proje Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

Bilişim Enstitüsünün

Yazılım Mühendisliği için

Öngördüğü DÖNEM PROJESİ olarak

hazırlanmıştır.

Haziran, 2021

Barış Cem Bektaş'ın hazırladığı **"Makine Öğrenmesi Teknikleri Kullanarak Hisse Senedi Fiyatlarının Tahmin Edilmesi"** adlı bu çalışma **Prof. Dr. Süleyman Tosun** tarafından **Bilişim Enstitüsü Yazılım Mühendisliği**'nde Dönem Projesi olarak kabul edilmiştir.

Prof. Dr. SÜLEYMAN TOSUN
Proje Danışmanı

BİLDİRİM

Hacettepe Üniversitesi Bilişim Enstitüsü, dönem projesi yazım kurallarına uygun olarak hazırladığım bu tez çalışmasında,

- Dönem projesi içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,

ve bu dönem projesinin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez/dönem projesi çalışması olarak sunmadığımı beyan ederim.

20.05.2021

Barış Cem Bektaş

Stock Price Prediction Using Machine Learning

Barış Cem Bektaş, June 2021, 43 Page

Özet: Borsa tahmini, pekçok ekonomist ve araştırmacının üzerinde çalışmış olduğu önemli bir alandır.Etkin Piyasalar Hipotezi ve Rastlantısal Yürüyüş Kuramı gibi geçmiş ekonomi teorileri, borsa fiyatlarını tahmin etmenin mümkün olmadığını savunsa da, daha sonraki araştırmalar, hisse senedi fiyatlarının bazı teknikler ve matematiksel formüller kullanılarak doğru bir şekilde öngörülebileceğini göstermektedir. İçinde bulunduğumuz son yüzyılda borsa tahmininde temel analiz ve teknik analiz olmak üzere iki ana yaklaşım kullanılırken, bilgisayar bilimi alanındaki gelişmeler sayesinde 1990'lı yıllardan itibaren borsa fiyatlarının tahmininde makine öğrenmesi teknikleri kullanılmaya başlanmıştır. Yapay Sinir Ağlarının hisse senedi fiyatlarını yüksek doğruluk oranlarında tahmin edebildiğini gösteren birçok araştırma bulunmaktadır. Literatürdeki araştırmaların bir çoğu NASDAQ, S&P 500 veya BIST 30 gibi borsa endekslerinin tahmin edilmesine yöneliktir. Buna ek olarak alandaki mevcut araştırmaların çoğunda teknik indikatörler tahmin modeline dahil edilmemiştir. Bu projede fiyat tahmini için dört büyük (Aselsan, Türk Hava Yolları, Arçelik, Ereğli Demir Çelik) Türk şirketinin hisse senetleri seçilmiş olup sonuçların doğruluğunu artırmak adına tahmin modeline bazı teknik indikatörler eklenmiştir. Tahmin

tekniği olarak Yapay Sinir Ağının gelişmiş bir versiyonu olan Uzun Kısa Süreli Bellekli Sinir Ağı kullanılmış. Girdi veri seti olarak 22 Mayıs 2006 ile 18 Mayıs 2021 tarihleri arasındaki fiyat ve hacim verileri kullanılmıştır.
Fiyatları tahmin etmek için mimari açıdan farklılık gösteren iki farklı model geliştirildi. Sonuçlar, şirketlerin Türk borsasındaki hisse senedi fiyatlarının yüksek

doğruluk oranlarında tahmin edilebileceğini ve teknik indikatörlerin temel tahmin modeline uygulanmasıyla tahmin doğruluğunun arttırıldığını göstermektedir.

Abstract: Stock market forecasting is an important field on many economists and researchers have been studying. Although previous economical theories such as Efficient Market Hypothesis and Random Walk Theory argue that prediction that predict stock market prices is not possible, further researches show that stock prices can be foreseen accurately by using some techniques and mathematical formulas. While two main approaches which are fundamental analysis and technical analysis have been used for the prediction of stock market for last century, thanks to developments in computer science area, machine learning techniques have been used since 1990's in prediction of stock market prices. There are many researches show that Artificial Neuron Networks can predict stock prices in high accuracy rates. Most of researches in literature regarding with predicting stock market indexes such as NASDAQ, S&P 500 or BIST 30. In addition, in most of the existence researches on the field technical indicators are not utilized in their prediction models. In this project it was chosen four big (Aselsan, Turkish Airline, Arcelik, Eregli Demir Celik) Turkish companies' stocks for price prediction and some technical indicators were added to prediction model to get increase accuracy of results. Long Short Term Memory Neural Network which is advanced version of Artificial Neuron Network was utilized as prediction technique. Price and volume data which between the period of May 22, 2006 and May 18, 2021 are utilized as input datasets. Two different models that differ in their architecture were developed to predict prices. The results show that companies' stock prices in Turkish stock market can be predicted in high accuracy

ratios as well as when technical indicators are applied to the base prediction model, prediction accuracy was increased.

Key Words

stock price prediction, technical analysis, LSTM, artificial neural network, prediction of Turkish stock prices, stock price prediction using deep learn

Table of Content

Table of Content

1. **Introduction**
2. **Related Works**
3. **Technical Analysis**
   - 3.1. Technical Indicators
     - 3.1.2. Ema
     - 3.1.3. Sma
     - 3.1.4. Macd
4. **Machine Learning**
   - 5.1. Supervised Learning
   - 5.2. Unsupervised Learning
   - 5.3. Support Vector Machines
   - 5.4. Deep Learning and Artificial Neural Networks
   - 5.5. Recurrent Neural Network
   - 5.6. Long-Short Term Memory
5. **Approach and Application**
   - 5.1. Data Acquisition
   - 5.2. Data Preparation
     - 5.2.1. Cleaning Inappropriate Input
     - 5.2.2. Deleting Inadequate Columns
     - 5.2.3. Data Normalization
   - 5.3. Applying Model
     - 5.3.1. Hyper-parameters
6. **Experimental Results**
   - 6.1. Evaluation Metrics
     - 6.1.1. Mean Squared Error
     - 6.1.2. (Scaled) Mean Squared Error
   - 6.2. Graphs and Evaluation

6.2.1. Aselsan - ASELS
6.2.2. Türk Hava Yolları - THYAO
6.2.3. Arçelik - ARCELK
6.2.4. Ereğli Demir Çelik - EREGL
7. Conclusion
8. References

1. Introduction

The Stock market which is one of the key elements of financial system gives people facility of buying, holding or selling stocks or wide derivative of financial instruments. People who have capital, mostly prefer to buy company's shares from stock market by hoping with their short or long term returns. Besides of this understandable expectation, many investors who are especially trading in short periods cannot be successful on making profit from the market. High volatility and non-linear price changes as well as various economical impacts and political situations make stock market prices very hard to predictable.

Two different approaches has been coming to fore in regarding with prediction stock market prices for almost a century. Fundamental Analysis and Technical Analysis. These theories and their backgrounds are mentioned on below.

Fundamental Analysis

Fundamental Analysis is based on measuring intrinsic value of a stock. Analysts use this method for determining a stock’s fair value. They consider all external and internal factors that effect stock’s long term profitability and sustainability. These are macroeconomic factors such as state of the economy and industry condition or micro economic factors such as revenues of stock, return of expenses, liabilities, cash flows etc.[1]

Technical Analysis

Another approach that is used for future price prediction is technical analysis. Technical analysis is study of historical price fluctuations and technical analysts use price data in calculations to to predict future price of stock.[2] Technical analysts argues that a stock’s future price is hidden of its historical price. In other words, past price movements repeat their-self again and again. Hence, if one can understand and analyze these price movements, he can predict future price movement accurately. In fact, this thinking is based on the theory of history repeats itself.

The difference between fundamental analysis and technical analysis approaches

shows itself in time periods of valuation. Whereas, fundamental analysis considers long term periods (three months to ten years), technical analysis focuses on short term prediction. (couple minute to three months)

Conventional methods have been used for many years by analysts. On the other hand, improving technology, hardware and software systems which are developing in exponential speed led to using machine learning techniques for financial market as well as for stock market. Even though there are variations of machine learning techniques such as k-nearest neighbors, support vector machines and random forest which are used for developing prediction models, LSTM (Long Short-Term Memory) has proven its succeed in prediction of stock market price. [3]

In this paper, we will first examine technical analysis and its some indicators, then we will explain methodology of our deep learning model. After these processes we will apply our model and discuss our findings.

2. Related Work

One of the first researches about stock market price prediction is Efficient Market Hypothesis suggested by Fama, E.F. (1965). According to the theory, current prices already contain all exist information and there is no possibility consistently to make abnormal profit by buying and selling. Hence, one cannot make profit higher than profit of buy and hold strategy. It also claims that stocks are always on the price of that they should be. No one can can get chance of buying a stock from undervalued price. If price is low that means the risk is high and vice versa.[4] Another theory which was presented by Horne, J. C., & Parker, G. G. (1967) is Random Walk Theory. It argues that changing stock prices in different periods are independent from each others and they occur randomly. Hence, past stock prices cannot be used for prediction of future prices and it is not possible to predict a stock's future price accurately.[5] These two theories are accepted as highly controversial. There are studies show that stock prices can be predicted by using different approaches. One of the first studies about technical analysis is proposed by Dooley and Shafer(1983).They develop an approach in order to make possible to get profit from foreign exchange's interest rates.[6] After few years of their researches,

Schulmeister(1988) observed some foreign exchange rates are following a sequential model in specific time periods.[7] Booth et al.(1994) developed an econometric model called ARIMA. The model was occurred six variables considered macroeconomic factors to predict future prices. Model was promising since it increased accuracy of prediction of future prices. [8] Leigh (2002) showed profitable of using of technical analysis in his work by using combination of multiple technical indicators.[9] Besides of the conventional analysis researches, there have been many other researches conduct in machine learning field. One of the earliest researches in predicting stock market prices with machine learning was conducted by White(1988) He made a neural network model to predict IBM stock daily prices.[10] After his work, many scientist started to be interested with this developing field. Wang et al.(2003) used artificial neuron network(ANN) for prediction of stock market prices. They emphasize volume data of stocks for better accuracy rate of prediction.[11]

Nelson et al. (2017) applied LSTM model on different technical indicators to estimate future stock prices. His work showed that LSTM model has better performance to make prediction in time series model than other machine learning models.[12] Maknickash and Maknickiene(2019) focused on RNN(Recursive Neural Network) to develop model for prediction stock prices. They made research to find optimal number of neurons and number of hidden layer to model success.[13] KIM T. and KIM(2019) H used stock time series and stock trend graphs in their LSTM-CNN model and they showed that LSTM-CNN approach was better than any other machine learning models. [14]

While technical analysis models and machine learning techniques are separately good on prediction of stock prices, it is hard to use them in combined due to the limitations of conventional machine learning. Conventional machine learning models can use only one time period data as input. [15]. On the other hand, thanks to deep learning which has recently developed branch of machine learning, time series data and multi period data can be given to the model as input. Thus, it can be used in combination with econometric data and technical analysis models.[15]

3. Technical Analysis

Technical analysis is the approach which measures the past price action of a stock in order to predict that stock's future price movements. Technical analysis roots are reaching out “The Dow Theory” suggested by Charles Dow. He thought that expectation about national economy determines long run market price in advance of actual economic developments. He computed averages of past prices to predict future prices on the market.[16] Technical Analysis in today’s form include three main assumptions.

The market discounts everything

The price of the stock discounts all information which are macroeconomic and micro economic existing in any given time. Technical analysts do not need any fundamental information about stock, sector or national economy. Because of the theory says past movements of price consider all these information, analysts just measure previous price and volume data of the stock.

The history tends to repeat itself

The theory assumes that the price movements have repetitive model and repeat itself in historical trends. It argues that because of human psychology, investors tend to give similar reaction when similar market condition happens. Thus, repetitive pattern of price movements make market and investors behaviors predictable. If a technical analyst detects repetitive pattern, he can make profit by acting through this model.

Prices move in trends:

According to the theory, price movements are happening in a trend. When a trend occurs, price of the stock will keep going to follow this direction in specific time period. The trend can occur in horizontal, upwards (bull market) or downwards (bear market).

3.1. Technical Indicators

Technical analysts developed mathematical models and calculations to predict

future price of a stock in higher accuracy. These models are called as indicators and, oscillators. They are applied on historical price and volume data to extract “hidden information” which is exist inside of these data. They are guiding to analyst where to buy or sell the stock. Even though there are hundreds of technical indicators are used for technical analysis, we utilized three of the most important indicators for our project.

3.1.1. Simple Moving Average – SMA

Simple moving average is a formula that calculates average of prices, mostly closing prices, by the given time periods. 14 closing prices are general accepted to calculate. Prices are being added first, then divided into total time period (trading day). It is the simplest technical indicator that is used by analyst. The formula of SMA:

$$SMA = \frac{\sum_{i=0}^{n} C}{n}$$

Figure 1: Formula of SMA

3.2.2. Exponential Moving Average – EMA

Exponential Moving Average (EMA) is a derivative from Simple Moving Average (SMA) that measuring average of price stock in a given period of time. The difference EMA from SMA is that EMA applies increasing weights to more current prices. Through given weights, the model will follow price more closely. Hence it can reflect more current prices and trends. EMA’s rate of accuracy in prediction better than SMA’s one in theoretically

$$EMA_t = (C * K) + (EMA_y * (1 - K))$$
$$K = \frac{2}{(N+1)}$$

Figure 2: Formula of EMA

3.2.3. Moving Average Convergence Divergence - MACD

Moving Average Convergence Divergence is a trend focusing technical indicator that shows mathematical relationship between two moving averages prices. It is calculated by subtracting the 26 period of EMA from the 12 period of EMA. Technical analysts are accepted nine day EMA of the MACD as “signal line” which is guiding analysts to buy or sell decision. If calculated MACD is above its signal line then buy signal occurs, or if MACD crosses below its signal line, the model gives sell signal.

$$MACD = (EMA_{26} - EMA_{12})$$

Figure 3: Formula of MACD

4. Machine Learning

Machine learning is the study of computer science provide computer to learn automatically from data by its experience.[17] Machine learning algorithm builds a model by itself for extracting information from the training data which is given to it. Machine learning approaches are divided into two which are supervised learning and unsupervised learning.

4.1. Supervised Learning

Supervised learning is the method of training machine learning algorithm through

example inputs and their related labels. The algorithm's purpose is to increase accuracy of giving label to input data as much as possible. The more we feed algorithm with training data, the better results our model gets. According to their tasks and output variable, supervised learning algorithm are categorized in two subset. These are classification and regression. Output variables of classification tasks should be categorized such as brand of a car, color of a pencil, marriage status of a person etc. Its output should be limited with a set of values. On the other hand, output of regression tasks must be continuous variable such as degree of air temperature, price of a bicycle, height of a person etc. In essence, supervised learning algorithms are used to detect relationship and difference of degree between two object.

4.2.Unsupervised Learning

In unsupervised learning we give input data which is unlabeled and categorized to our algorithm and we expect from algorithm to build a pattern and discover relationships between them. There are two sub categories in unsupervised learning by their tasks. These are Clustering task and association task. Association tasks are mostly about discover patterns and rules related with input data. For instance, prediction of political tendency of people. However, in clustering tasks, algorithm tries to make group with our data by its specific features.

4.3. Support Vector Machines

Support Vector Machines (SVM) are part of supervised learning algorithms that grouped data into two which divided by a linear boundary. SVM can be used for both classification and regression analysis. It builds a model and put training data on space according to its corresponding feature. When new data comes, SVM predicts their belonging group or category and mapped them into this space.

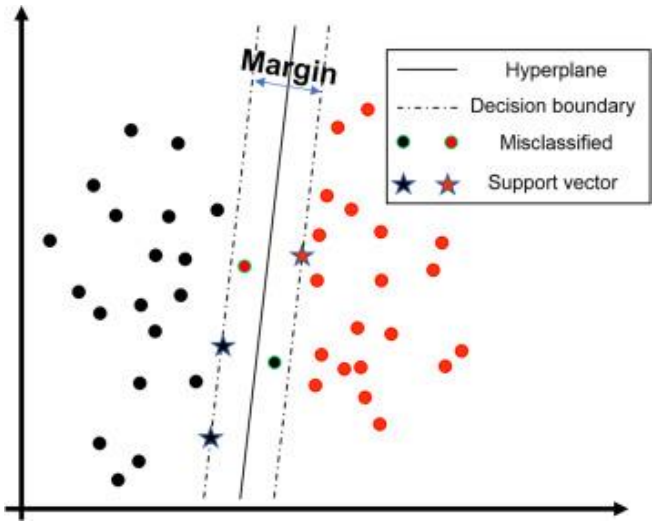

Figure 4: Support Vector Machines

Although SVM gives good results in classification or regression problems in some degree, it has limitations. When we start to work with large scale of data or we need to get pattern recognition with data contains huge mathematical values, the dimension space of SVM is getting bigger and it starts to have exponential complexity. On account of this problem, the algorithm's system requirement to run is increased hugely. Hence, traditional SVM algorithms are not scalable. [18]

4.4. Deep Learning and Artificial Neuron Network

Deep learning is one of the subcategory of Machine learning in AI which is capable of unsupervised learning with unlabeled data by using its multi layer neuron units. Deep learning algorithms can extract information from data automatically through its neuron network which is called as Artificial Neuron Network.

Artificial Neuron Network (ANN) is the imitation of the natural neuron network of the human nervous system. In ANN, neurons are connected with each other on network form. likewise in the natural neuron network. ANN units called as neurons or nodes. Each node has dendrites and axons to connect with adjacent nodes. These units receive input through their dendrites, performs it and send the processed information to other nodes through axons.

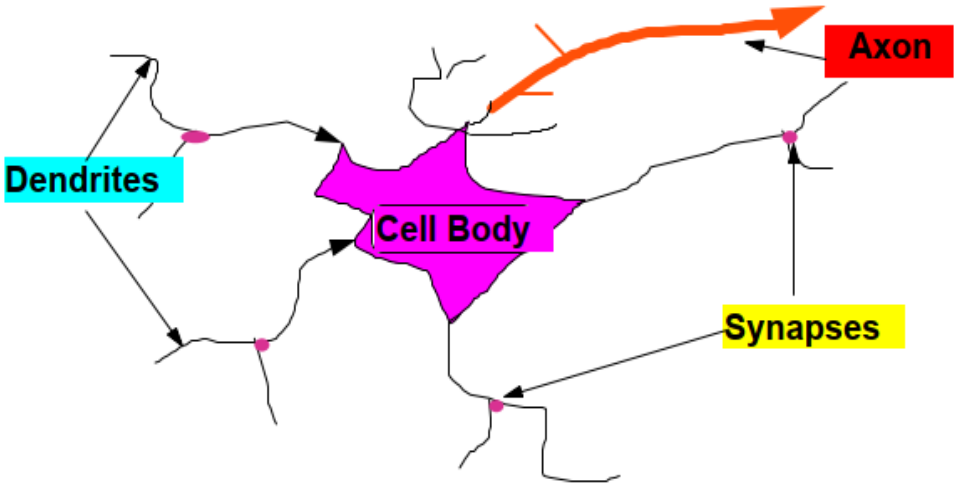

Figure 5: A Biological Neuron, Source

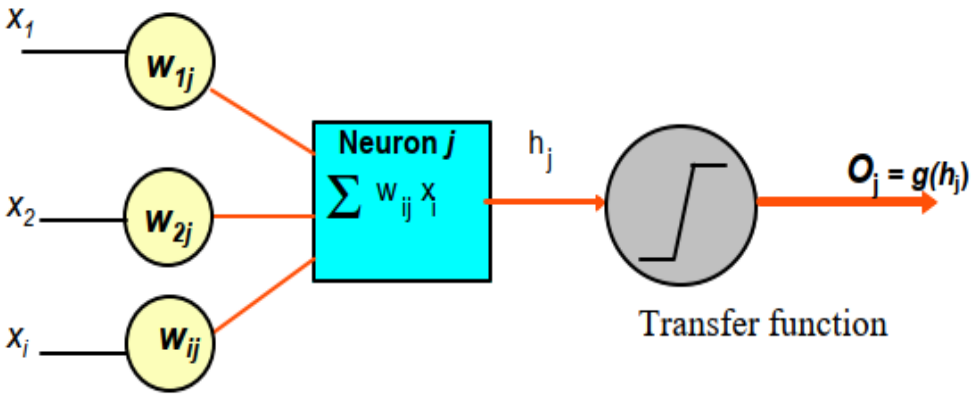

Figure 6: An Artificial Neuron, Source

In ANN each input has a specific weight which can be either positive or negative. Whereas positive values activate the node, negative values inhibit it. The neuron, sums received values of signals by multiplied them by their weight. The output of summing passed through to transfer function (activation function) which is mostly a logistic function to process and in the final output is sent to other neurons.[19]

There is back-propagation function was developed for ANN structure. The aim of the back-propagation algorithm is to minimize error in feed forward neural network. Back propagation algorithm takes output value from the network output layer and gives back it to the network input layer again to get decrease error rate which comes from previous process. In each iteration, weights of connections get determined

again until output error rate is acceptable.

In traditional ANN there is three layer which are formed by group of neurons. These are input layer, hidden layer and output layer. This earliest form of ANN model is sufficient for most of the calculations such as prediction house prices or classifying objects.

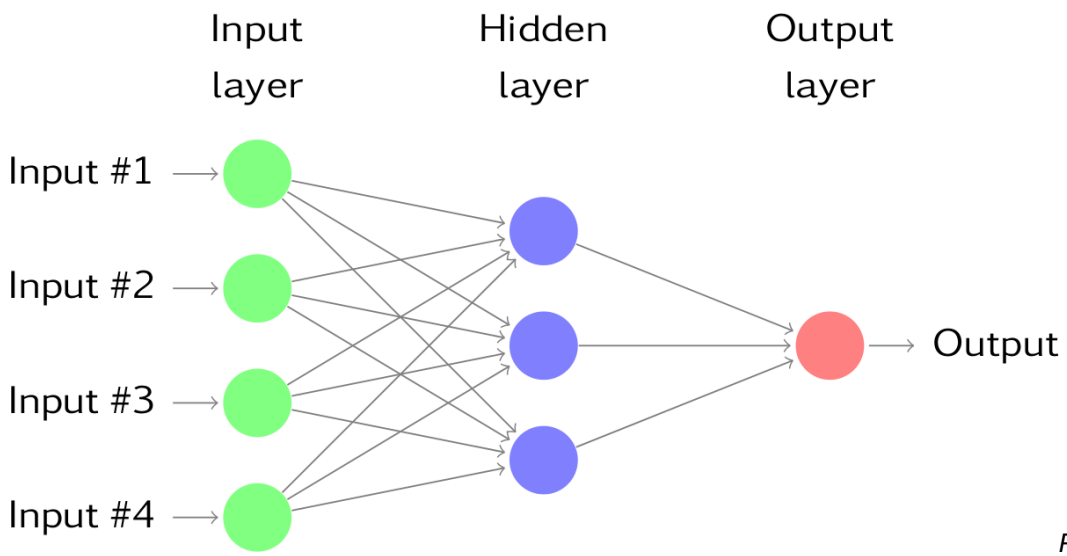

Figure 7: Simple form of ANN

On the other hand, when it needs to learn complex model, to process thousands of values or to make future prediction of a time series, simple form of ANN starts to be incapable. For processing this type of complex problems, Deep Neuron Network (Deep Learning) has been developed. Deep Neuron Network (DNN) in essentially is advanced form of ANN. The “deep” word in it is coming from multiple hidden layers it has. When the compare with ANN which has single hidden layer, DNN may has two to hundred hidden layers. Because of DNN has more hidden layers and connections that ANN has, it is capable of processing complex problems.

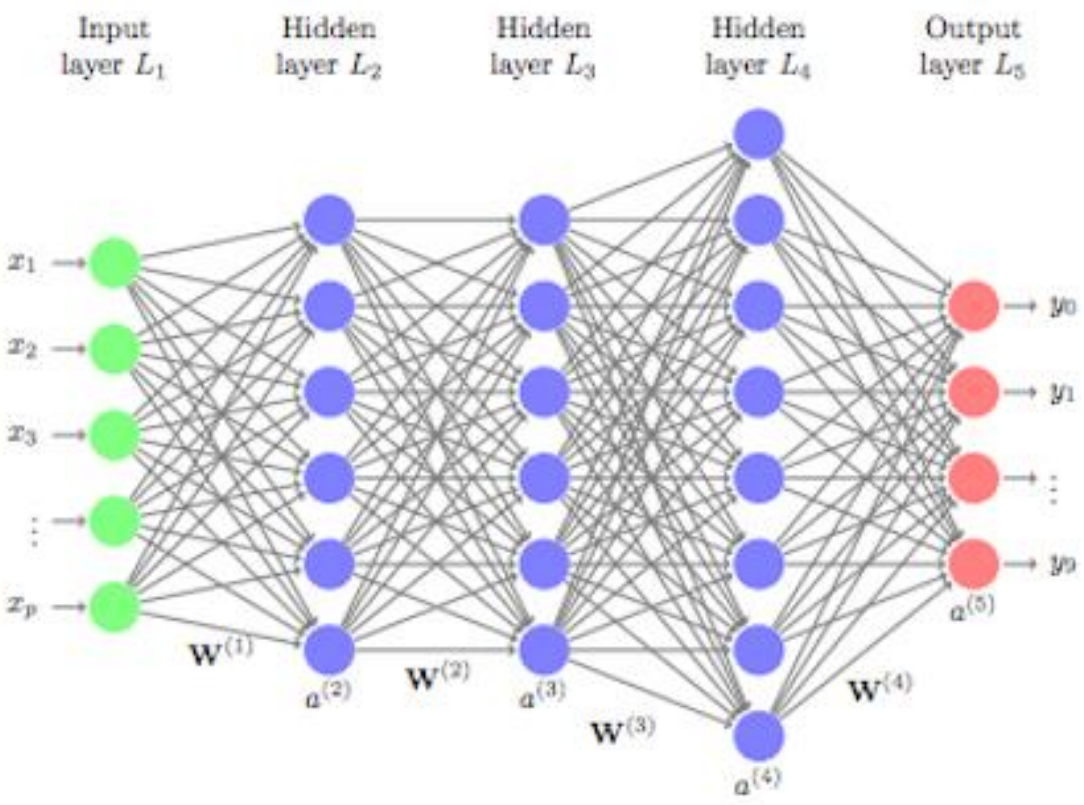

Figure 8: Deep Neuron Network

The first developing type of ANN is Feed Forward Neural Network (FFNN). In FFNN, connections between neurons are in one direction. In other words, layers which occurs from neurons receive signal from their previous layers and send output signal to their forward layers. Signal does not process in cycle but it flows in one direction through the neuron layers. FFNNs are capable of solving complex problems however, in some cases we want to hold information belongs previous iterations on current state. Because of this requirement, there is need for sophisticated structure than early ANN's has.

4.5. Recurrent Neural Network

In sequential data such as time series problem required to process historical values for predicting to future data. Hence, Artificial Neuron Network should hold historical data values and build patterns between them to predict future data points. Because of traditional ANN's cannot be capable of processing data belongs previous step's,

different techniques we need. Recurrent Neural Networks is remarkable development to solve this temporal effect.

Recurrent Neural Network (RNN) which has neurons capable of self-loop is an advanced subclass of ANN's. Neurons in hidden layers of RNN can hold information which belongs previous iterations. This internal state works as a type of memory for neurons and for whole network structure. Thanks to this memory function, RNN can perform complex tasks such as weather forecasting, face recognition, voice recognition etc.

In early version of ANN's neurons are independent in terms of performing information. As a consequence of it, sequential input such as time series data cannot be processed efficiently. In RNN model each iteration of network generates a hidden state which is self-loop state.

Researches show that RNN have higher accuracy ratios compare with classic(vanilla) ANN's on times series data prediction.[20]

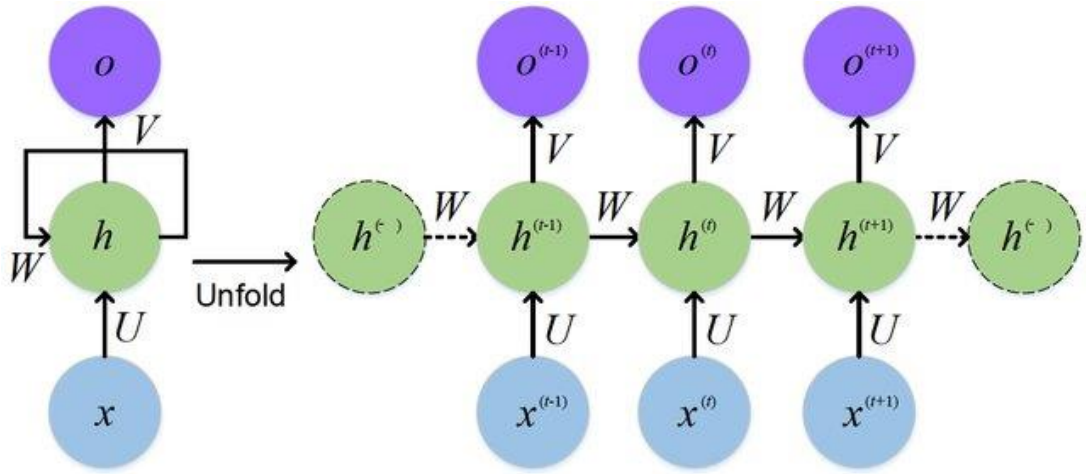

Figure 9: Recurrent Neuron Network

4.6.Long-Short Term Memory (LSTM)

Classic RNN's are capable of holding long term data for complex processing however, back-propagation function of RNN suffer from a significant problem which is called as *vanishing or exploding gradient problem. This* problem causes when the

network either stops or continues to learn until it reaches over-learning point. Hence it cannot converge on minimum error point. Long-Short Term Memory (LSMT) which is a sophisticated version of RNN architecture has solved vanishing gradient problem by forgetting intelligently irrelevant previous information[21] LSTM is highly suitable to process long term historical data due to their efficient memory capacity.

Classic (vanilla) RNN's has architecture chain like and they have looping function to hold previous information.

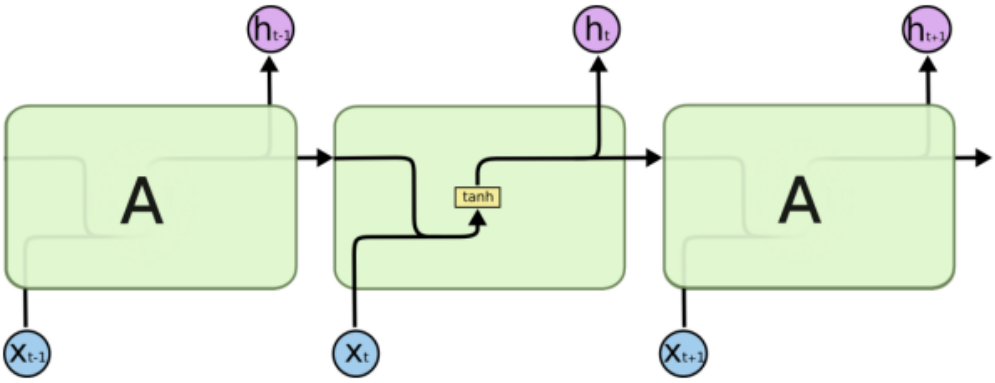

Figure 10: Classic (Vanilla) RNN module

In LSTM, there are three types of gates which are input gates, output gates and forget gates. The input gate receives new information and represents current state of neuron. The forget gate intelligently selects relevant information with current data out of past information and throwaways unused or irrelevant ones. Thus, the neuron are not overflow with unnecessary previous information. Past irrelevant information from the forget gate and new information from the input gate are processed through cell state vector.[21] After this step, the output gate produces the prediction value for the model forecasting and sends it to the next node

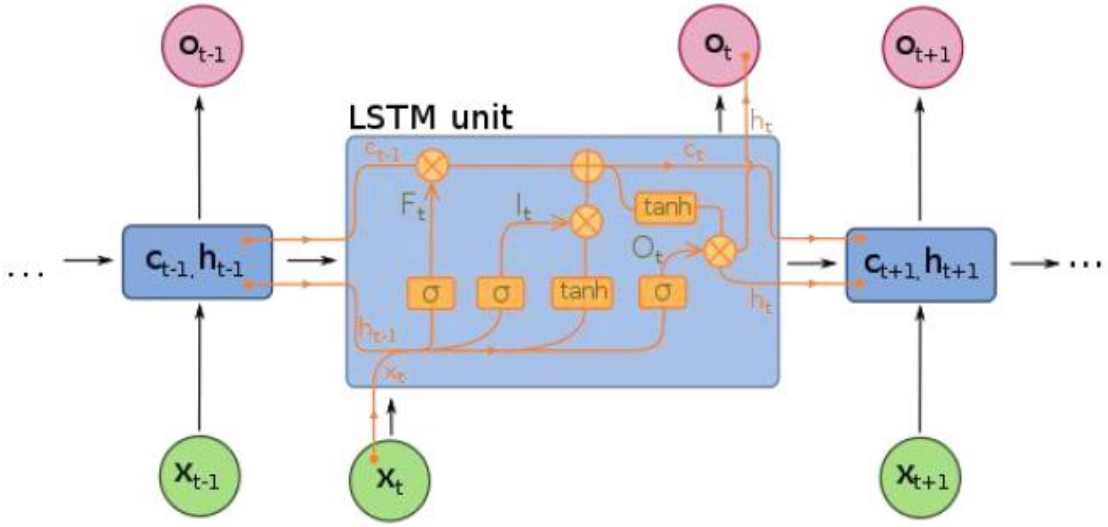

Figure 11: Long short-term memory unit (LSTM)

LSTM uses backpropagation through time (BPTT) algorithm to learn data which is flowing from network and the algorithm makes LSTM powerful on calculating univariate and multivariate time series.[21]

Researches show that LSTM gets better result compare with other Artificial Neuron Network methods in stock price prediction. [22] [3] [23]

4.6.1. Sliding Window Method

Sliding window method in LSTM is the basis of turning time series datasets into a supervised learning problem. Previous time data are used for the prediction of next time data step by step in sliding window method. The number of previous time frames called as window width or size of the lag.[25]

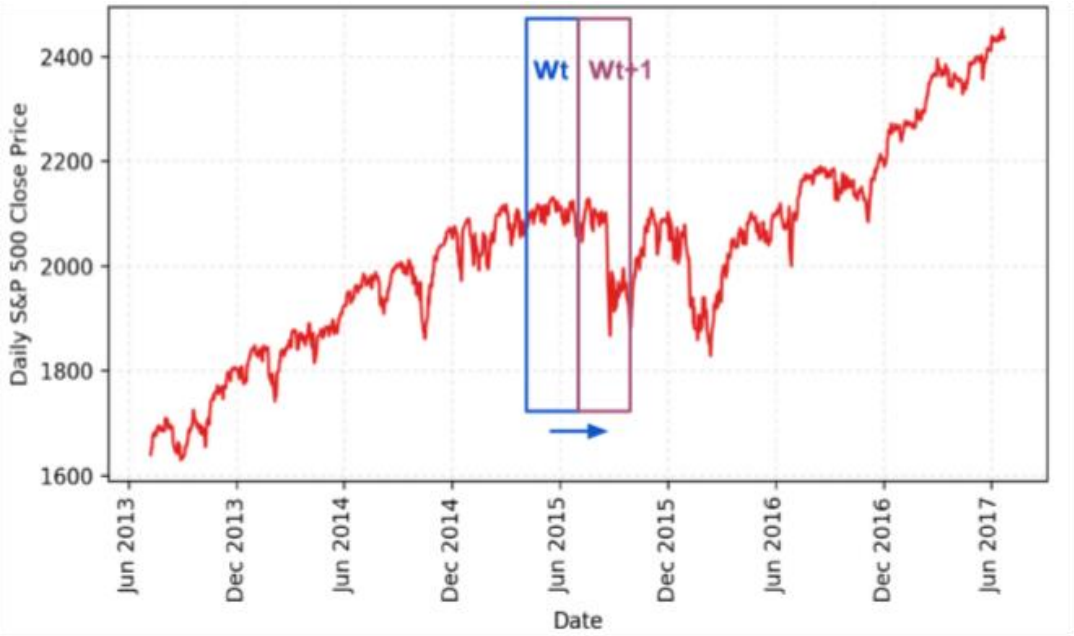

Figure 12: Presentation of Sliding Window

It can be seen how sliding window method works above. The “Wt” values(previous time series data) are used to predict “Wt+1” values (next time series). Through this step by step approach our model can predict assigned datasets which is %10 of our total data. In our project, as it is mentioned in chapter 5.4.1.Tuning Hyper parameters, “history points” hyper parameter controls our window width. We assigned 50 history points as window width. It means that our prediction model uses 50 previous days to predict just the next day coming after these 50 days.

5. Approach and Application

In this project we used historical time-series data of the chosen stock market companies to predict their future prices in determined time frame. For non-linear time series problems LSTM networks are the one of the most preferable algorithms as emphasized in chapter 4 on this report. Thus, we used LSTM network architecture for the prediction model.

We use Python 3.7 with libraries Keras API that run on top of TensorFlow. TensorFlow is an open sourced library developed by Google’ Brain Team. It is used for deep learning applications and used for processing with high scale of data.

5.1. Data Acquisition

We chose four huge company of Turkey (Aselsan-ASELS, Arcelik-ARCLK, Turk Hava Yollari-THYAO, Eregli Demir Celik-EREGL) by their market cap to predict their future prices. We acquired historical price data of the selected companies from yahoo.finance.com which is one of the most visited website in finance field of the world. Historical price data cover 3874 number of days between 22/05/2006 to 18/05/2021 for all companies. Datasets contains “Date”, “Open”, “High” “Low”, “Close” “Adj Close” and “Volume” columns for each day.

1	Date,Open,High,Low,Close,Adj Close,Volume
2	2006-05-22,0.433701,0.433701,0.389595,0.411648,0.321640,6269397
3	2006-05-23,0.389595,0.393271,0.374894,0.382245,0.298666,18216970
4	2006-05-24,0.385920,0.385920,0.374894,0.378569,0.295794,5534654
5	2006-05-25,0.382245,0.389595,0.378569,0.385920,0.301537,6971491
6	2006-05-26,0.400622,0.418999,0.400622,0.415323,0.324511,17650506
7	2006-05-29,0.393271,0.441051,0.393271,0.426350,0.333127,7527956
8	2006-05-30,0.426350,0.426350,0.393271,0.396946,0.310153,9781092

Figure 13: Sample of datasets

5.2. Data Preparation

Before we process our datasets with our model we need to get our data cleaned from noises, unsuitable figures or inputs etc to make sure our LSTM algorithm working smoothly and efficiently.

5.2.1. Data Cleaning

In our datasets, there are some trading days that have columns are null. Thus, we need to clean out these inappropriateness manually before we put our datasets to our process environment.

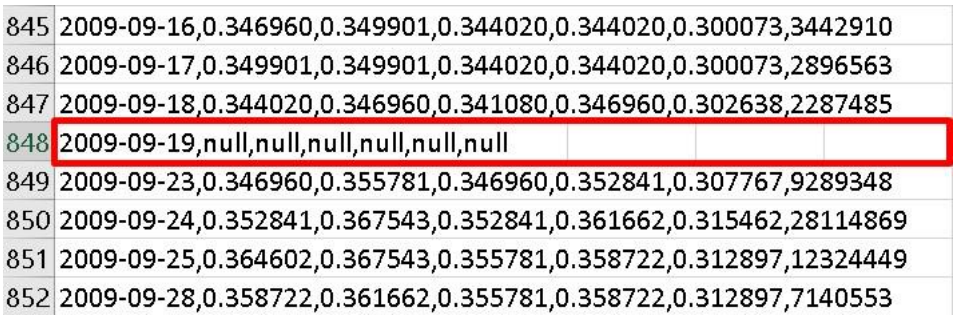

Figure 14: Cleaning null values out from the datasets

5.2.2 Dropping Inadequate Columns

We have seven columns in our datasets include “Date” column. However, our model do not need to know date because it does not effects output of algorithm.

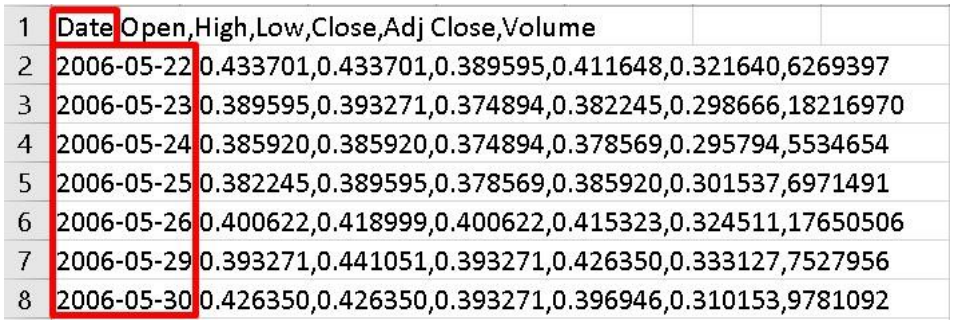

Figure 15: Dropping date column from datasets

There are two columns in regarding close price of the stock on the datasets. “Close” and “Adj Close” (Adjusted Close). Whereas “Close” column indicating close price figure which seems on stock market screens on that date, “Adjusted Close Price” represents the value after applied all divided distributions and splits to close price of that day. Using “Adj Close” prices instead of “Close” prices may seem more reasonable at first sight. However, because of inconsistency Adj Close prices with Open prices in regarding daily basis, we used “Close” prices instead of “Adj Close” prices to train our model. Thus, we should drop “Adj Close” column from the datasets.

1	Date,Open,High,Low,Close Adj Close Volume
2	2006-05-22,0.433701,0.433701,0.389595,0.411648 0.321640 6269397
3	2006-05-23,0.389595,0.393271,0.374894,0.382245 0.298666 18216970
4	2006-05-24,0.385920,0.385920,0.374894,0.378569 0.295794 5534654
5	2006-05-25,0.382245,0.389595,0.378569,0.385920 0.301537 6971491
6	2006-05-26,0.400622,0.418999,0.400622,0.415323 0.324511 17650506
7	2006-05-29,0.393271,0.441051,0.393271,0.426350 0.333127 7527956
8	2006-05-30,0.426350,0.426350,0.393271,0.396946 0.310153 9781092

Figure 16: Dropping Adj Close column from the datasets

Since “date” and “Adj close” columns are inseparable from other columns in excel table, they could not be deleted as manually. We used a python code to delete it from datasets.

```
def csv_to_dataset(ASELS):
    data = pd.read_csv(ASELS)
    data = data.drop('Date', axis=1)
    data = data.drop('Adj Close', axis=1)
    data = data.drop(0, axis=0)
```

Figure 17: Dropping date and adj close columns from the datasets (Python Code)

After “null” row and “Date” and “Adj Close” columns are cleaned up from our datasets, we have 3863 rows for ARCLK, 3868 for ASELS, 3862 for EREGL, 3849 for THYAO which are trading days and 5 columns which are “Open”, “Low”, “High”, “Close”, “Volume” as our datasets.

5.3. Normalization

After we delete and clean inadequate and unnecessary columns and values from our datasets, we need to define a scaler to normalize data. Normalization is process that scales data in the value between “0” and “1”. It is important because in time series problems different time periods of data have different value ranges. If normalization is not applied, the earlier data can be close to 0 while new ranges of data added to neural network and earlier data cannot add much value to the learning process. MinMaxScaler provided by scikit learn framework is a function that we use for normalization step on our work.

$$X_{new} = \frac{X_i - \min(X)}{\max(x) - \min(X)}$$

Figure 18: Formula of MinMaxScaler

```
16      data = data.values
17
18      data_normaliser = preprocessing.MinMaxScaler()
19      data_normalised = data_normaliser.fit_transform(data)
20
```

Figure 19: Data normalization python code

5.4. Applying Model

Adam Optimizer is used as optimization algorithm to make sure the best learning rate is provided.

5.4.1. Tuning Hyper parameters

In our work, for determining the optimal parameters, we experienced different combination of batch size, epoch size, dropout rate, train test split ratio and history points. Optimal parameters for our model are mentioned below.

Train/Test Split

We split our datasets into two as training and test data. %90 part of our datasets is determined as training data and %10 part of it as test data. Therefore, as an example, for Aselsan Stock we have 3482 value for training to our model and 386 value for testing our results.

Epoch Size

Epoch number is the value used to determine how much time training data backpropagated to neuron network. The optimum rate of epoch size may change project to project in regarding the size of datasets and other hyper parameters. For this project we tried 30-40-50-60-70 epoch sizes and we reached the result that 50 epoch size is the optimum for our datasets. Thus we used 50 epoch size.

Batch Size

The batch size indicates the number of samples that is propagated through the network. Batch normalization was created to speed up learning process of ANN[24] If batch sizes are holded so small, accurate rate of gradient estimate decreases. Using big batch sizes have some advantages. However it should be determined by the size of datasets. Advantages of using big batch sizes:

Decrease system memory requirements: When datasets is split into small parts and the network trained with these small parts, memory that system needs to process decreases

Increase the learning rate: Neural network weights are updated after every propagation. In other words, each batch processed by network, parameters is changed. Network update rate can increases when batch size increases. It results with increased learning rate of the ANN.

Dropout Rate

Dropout is a method decreases over-fitting problem in ANN. Dropout function is randomly eliminating units derived from previous batches. Through this elimination process effectiveness of the ANN increases. The optimal dropout rate are accepted as 0.1 for the input layer and between 0.5 and 0.8 for internal layers.[24]

History Points

History points a hyper parameter that we build up to control number of days of the historical price that model get uses for prediction. For instance when we determine history points as 30, it means the model will train previous 30 days to predict just the next day.

Epoch Size	Batch Size	Dropout Rate	Train/Test Split Ratio	History Points
50	32	0.2	0.9	50

Table 20: Hyper parameters of our model

Neurons

Number of neurons is one of the key points of an ANN architecture.

We determined 50 neurons in LSTM layer, 64 neurons in each three of dense layers and 1 neurons in output layer in basic model

In technical model, we used two branch of ANN in order to combine prediction of Lstm model and prediction of technical indicator model. The first branch has 50 neurons in LSTM layer and second branch has 20 neurons in Dense layer. These ANN's output flows to third ANN which provide combination of them. Third ANN model has 64 neurons in Dense layer and 1 neuron in output layer.

Neurons of Basic Model					**Neurons of Technical Model**			
LSTM	DENSE	DENSE	DENSE	OUTPUT	LSTM	DENSE		OUTPUT
50	64	64	64	1	50	20	64	1

Table 2: Number of neurons on basic model and technical model

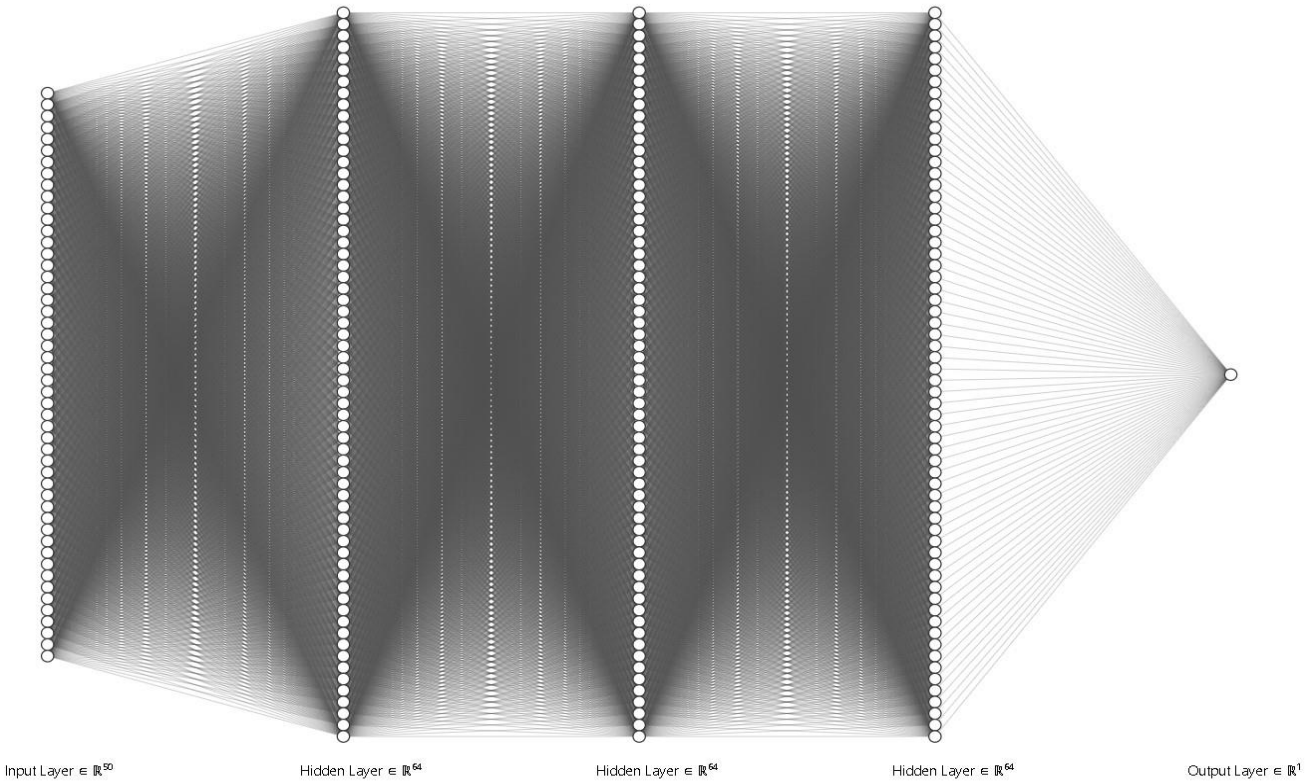

Figure 21: Neuron Network of our Basic Model

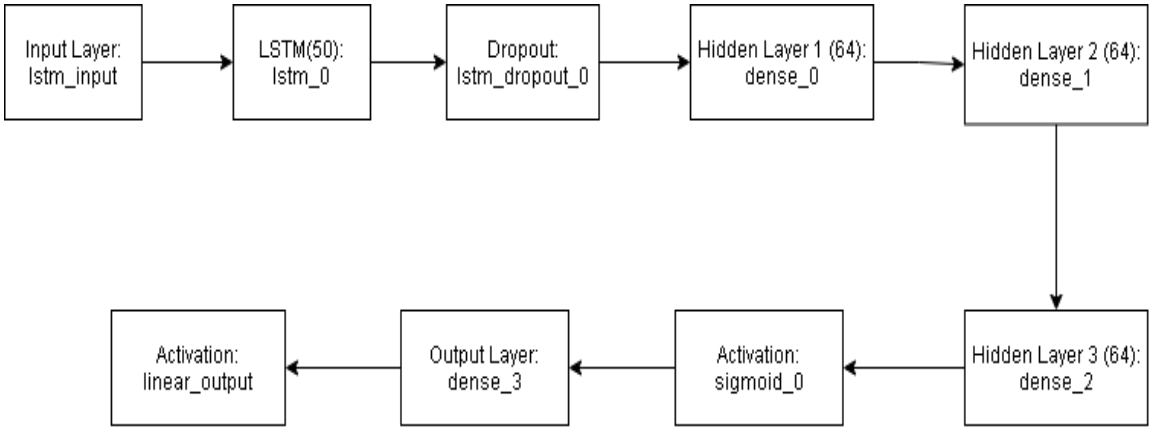

Figure 22: Basic Model Architecture

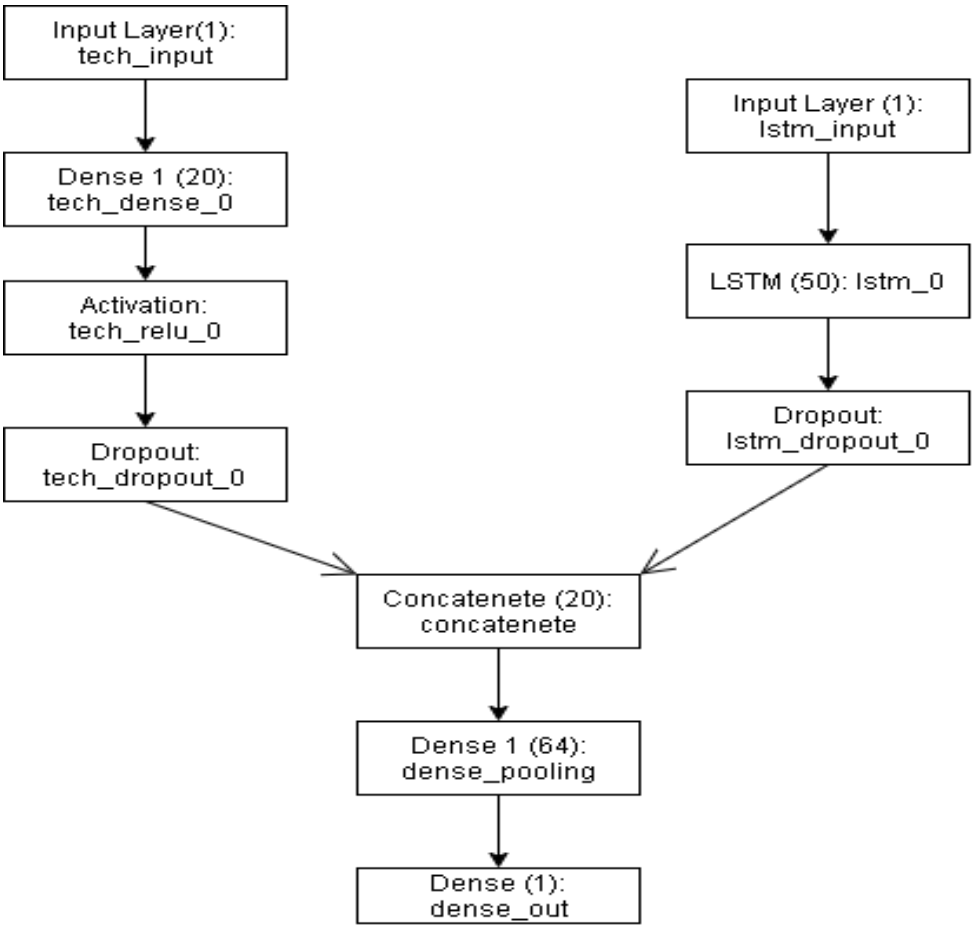

Figure 23: Technical Model Architecture

6. Experimental Results

We trained our model and we got prediction results by giving raw historical price data that are belongs to our companies to our LSTM Neuron Network. Therewith, we applied our technical indicators which are SMA and MACD to same datasets separately and compared the results of them. First model called as basic model and second model called as technical model.

6.1. Evaluation Metrics

We need some metrics to evaluate our results. On below, it can be seen the metrics we used to evaluate correctness ratio of our prediction models.

6.1.2. Mean Squared Error

Mean Squared Error(MSE) is an estimator that is used for measuring averages of the errors in the set of values. It shows the difference between actual values and the estimated values. The calculation formula of mean squared error is below.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2.$$

Figure 24: Formula of Mean Squared Error

Yi shows the actual values whereas Y^i shows the predicted values on the formula. Mean Squared Error formula is mostly used for calculating error of continuous values.

6.1.3. (Scaled) Mean Squared Error – (Real) Mean Squared Error

Since we normalized our data between 0 and 1 by MinMaxScaler function, our MSE of results will be smaller than it would be. Hence, results need to rescale before measuring MSE in healthy. Whereas MSE before inverse transformation named as scaled MSE, after rescaled it called as Real MSE or original MSE.

$$mse_{\text{original}} = \frac{1}{m}\sum_{n}|y_n - p_n|^2 = (max - min)^2 mse_{\text{scaled}}.$$

Figure 25: Measuring of original(real) MSE

6.2. Graphs and Evaluation

In our results graphs orange graph shows the prediction, and the blue one shows the actual value.

6.2.1. Aselsan – ASELS

Aselsan is the biggest Turkish Defence Company by its 34.4 billion dollar market cap. In addition it is 48th largest defense company in the world in terms of revenue [25][26] We got Aselsan's historical stock data covers from 22 May 2006 to 18 May, 2021 from yahoo.finance.com for utilizing in our model. The dataset had contained 3869 days before we deleted the null row which date is 19 August,2009. After deletion we had 3868 days to use. We splited our dataset into two as train(%90) and test(%10) data. After splitting we have 3482 value for training our model and 386 value for testing our results. Graphical demonstration and MSE results of our model has showed below.

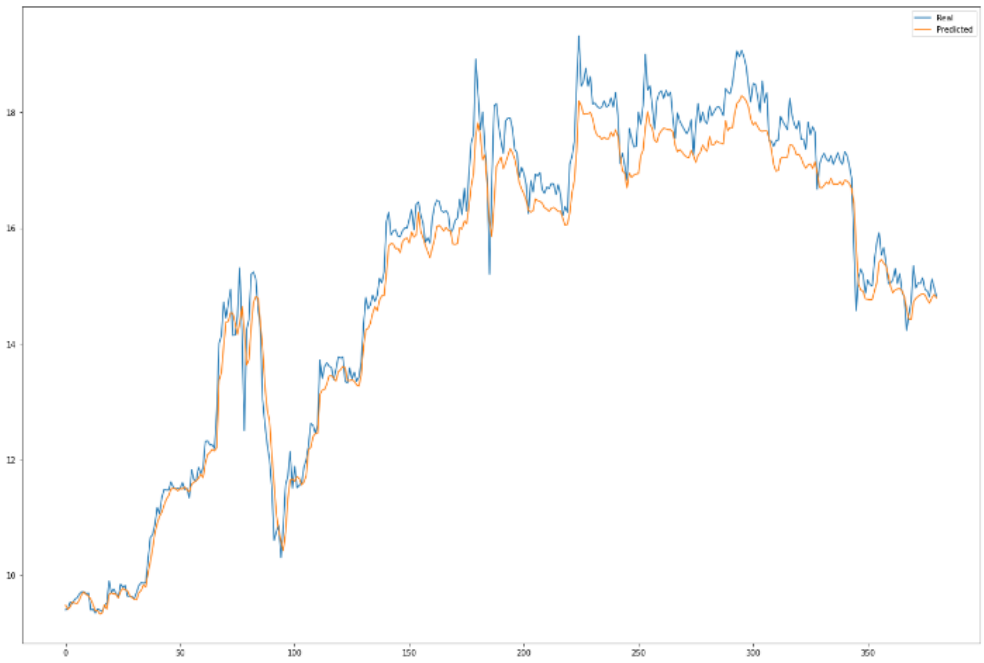

Figure 26: Basic Model

Real MSE	0.2154771115839
Scaled MSE	2.160171544701

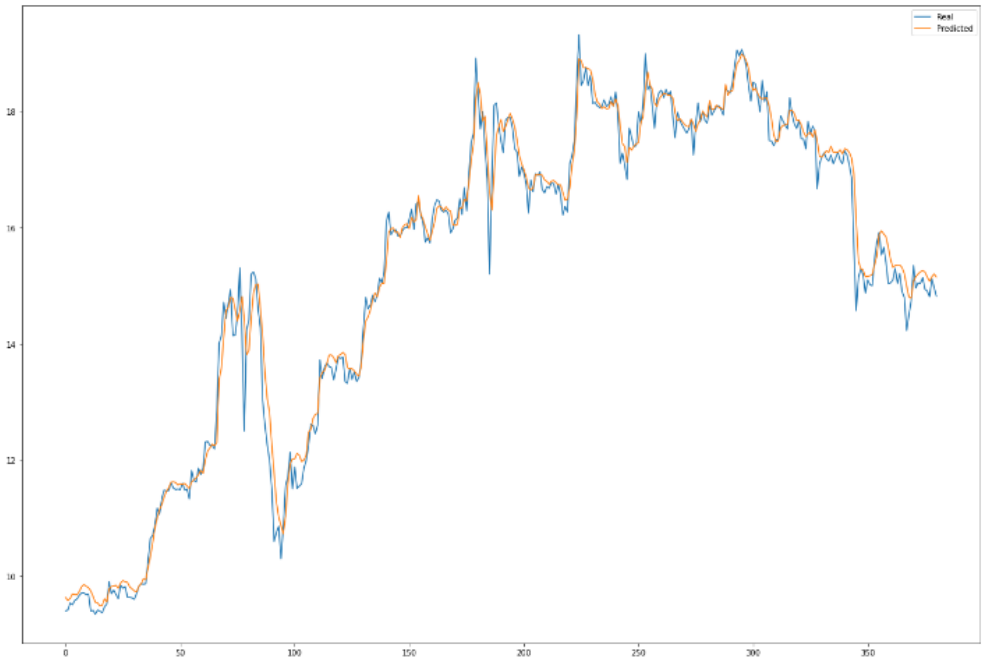

Figure 27 Technical Model (SMA)

Real MSE	0.10429301820763
Scaled MSE	1.0455440421818

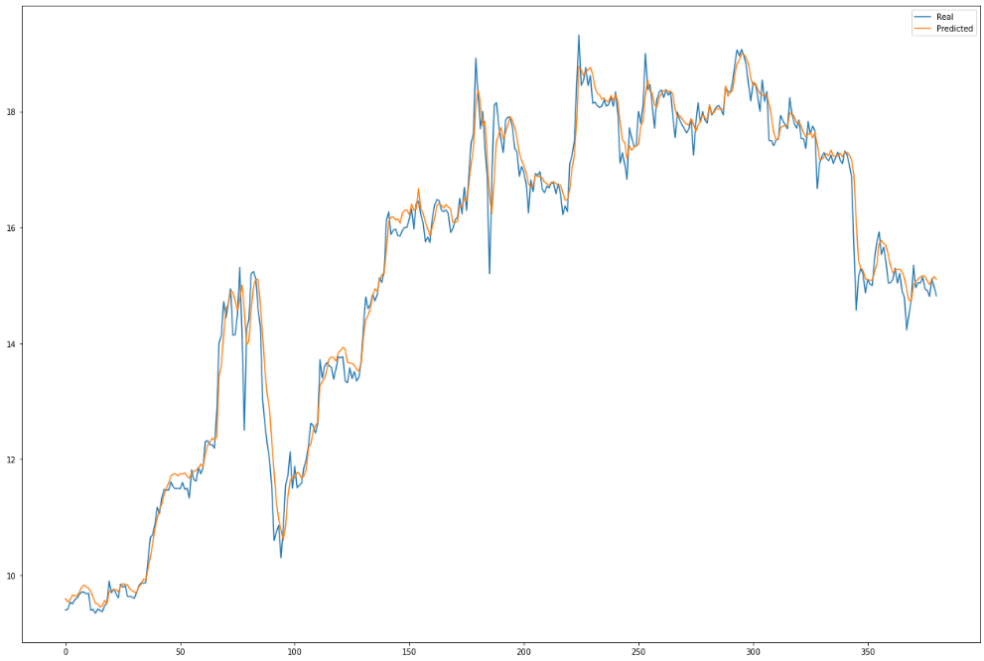

Figure 28: Technical Model (MACD)

Real MSE	0.10616482548413
Scaled MSE	1.0643090274099

As it can be seen in results, in technical model thanks to using technical indicators which are SMA and MACD, error rate (MSE) smaller compare with the basic model erro rates. In Aselsan stock example, SMA as a technical indicator gives the best results in our model with 1.04 scaled MSE ratio.

Evaluation	**Basic Model Results**	**Technical Model Results**	
		SMA	**MACD**
Real MSE	0.2154771115839	0.10429301820763	0.10616482548413
Scaled MSE	2.160171544701	**1.0455440421818**	1.0643090274099

6.2.2. Turk Hava Yollari – THYAO

Turk Hava Yollari (Turkish Airlines) is the national flag carrier airline company of Turkey and one of the biggest airline companies in the world.[28] We took stock data of THYAO dates between 18 May, 2006 and 18 May, 2021. Before we

deleted null rows, we had 3869 trading day. However, we detected 20 null rows in dataset and cleaned these rows to provide seamless of our model. After deletion process, 3849 rows last. We split dataset as test(%10) and training(%90) hence our dataset was splitted as 384 of test values and 3465 of train values. Resuts are below:

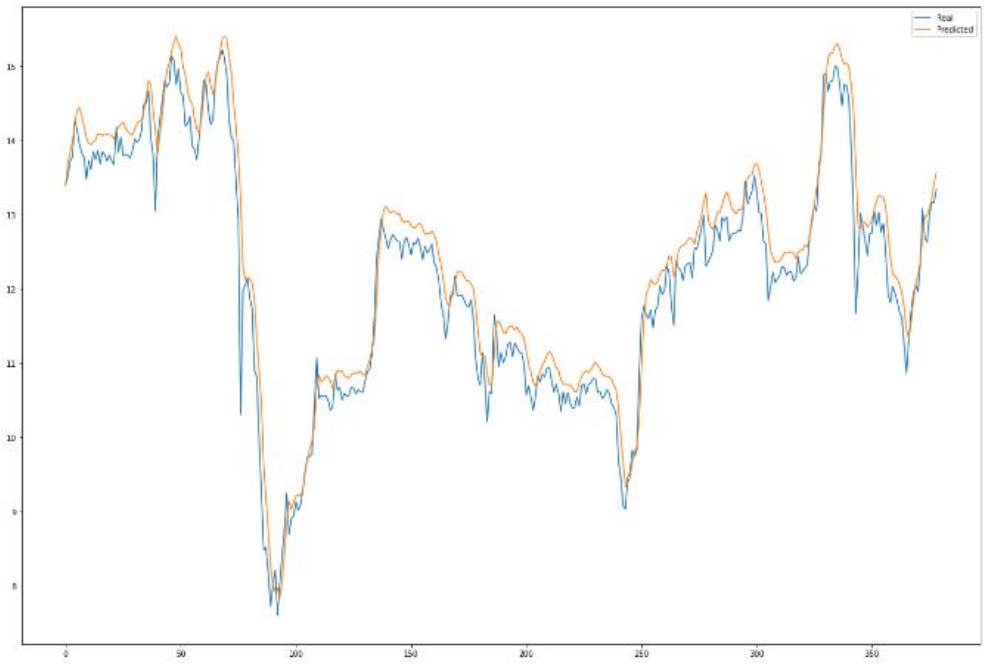

Figure 29: Basic Model

Real MSE	0.172198842014186
Scaled MSE	2.259827323020819

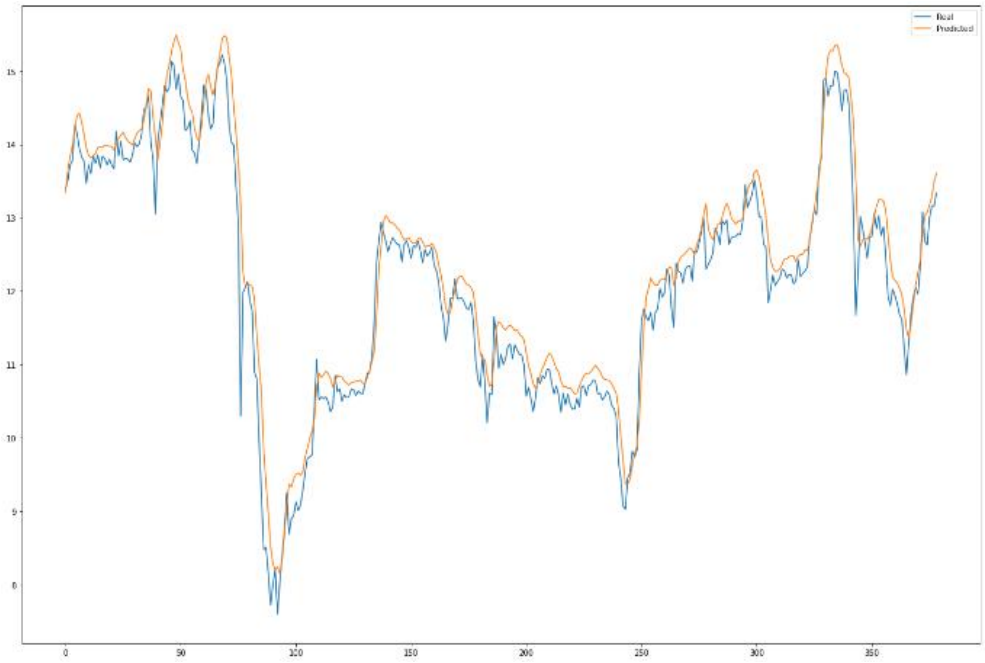

Figure 30: Technical Model (SMA)

Real MSE	0.1671594019901359
Scaled MSE	2.1936929395030957

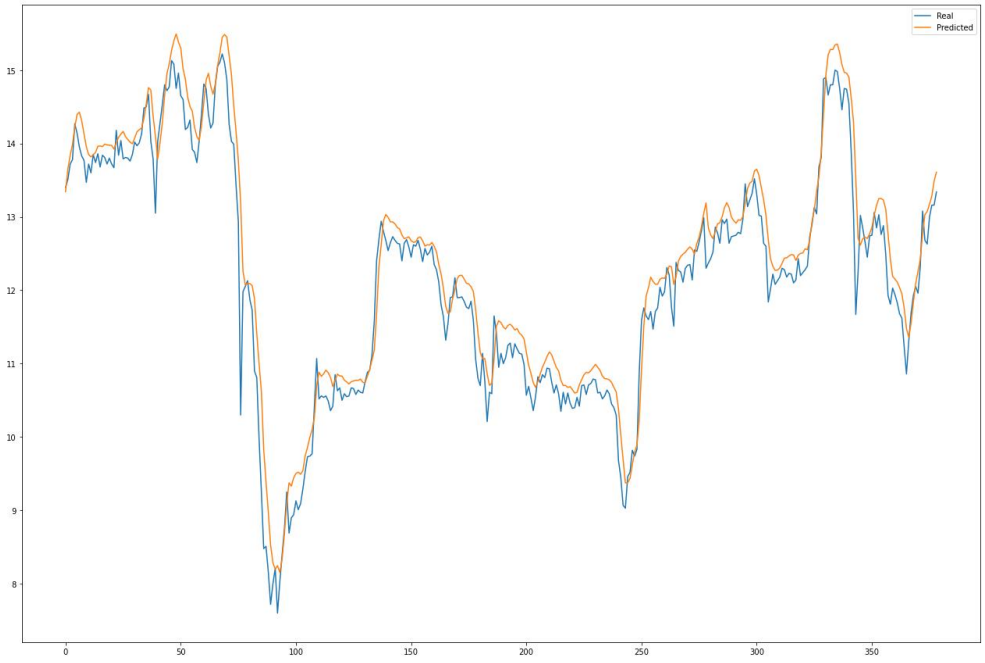

Figure 31: Technical Model (MACD)

Real MSE	0.14831880397998595

Scaled MSE	1.9464409971126764

In THYAO testing, our technical model which is using MACD one gives the best results. THYAO testing shows that technical model which is using technical indicators gives better accuracy rate than basic model gives. Concantrated results are below.

Evaluation	**Basic Model Results**	**Technical Model Results**	
		SMA	**MACD**
Real MSE	0.172198842014186	0.167159401990135	0.1483188039799859
Scaled MSE	2.259827323020819	2.193692939503095	1.9464409971126764

6.2.3. Arcelik – ARCLK

Arcelik is is a Turkish multination household applianced manufacturer. Its market cap 21.32 billion dollar and it is one of the biggest company in Turkey.[29] We acquired Arcelik's historical stock data covers from 22 May 2006 to 17 May, 2021 from yahoo.finance.com for using in our model. The dataset was contained 3874 days before we deleted the null rows. However after deletion process it drops 3863 rows. We determined as %90-%10 of train to test ratio. Thus we have 3477 train and 386 test values after seperation. Results are showed below as graphs.

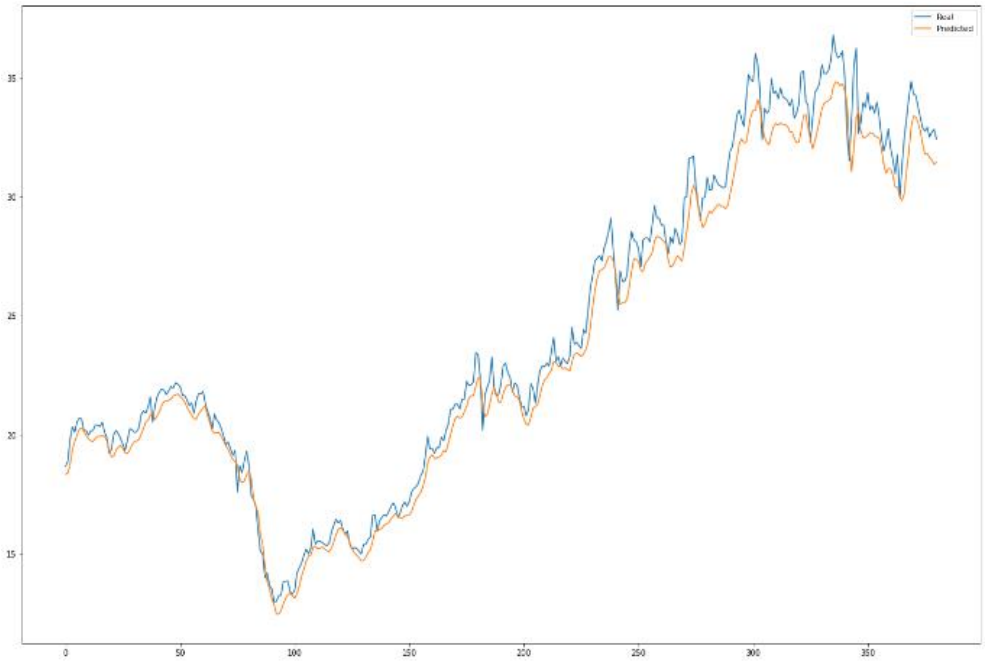

Figure 29: Basic Model

Real MSE	0.909514539116
Scaled MSE	3.82148982072

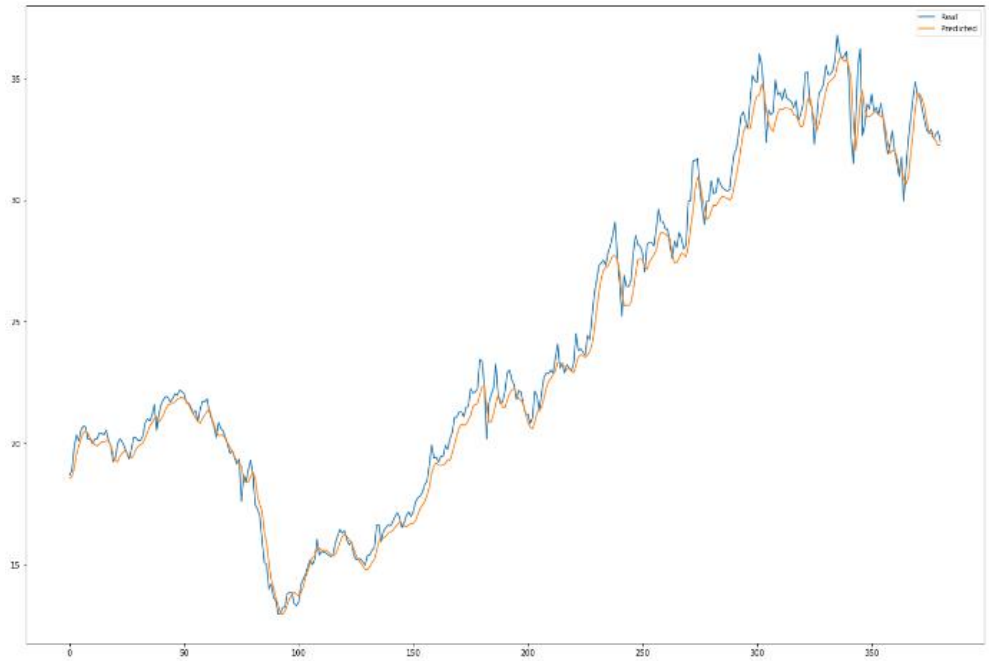

Figure 30: Technical Model (SMA)

Real MSE	0.488505528626
Scaled MSE	2.052544324

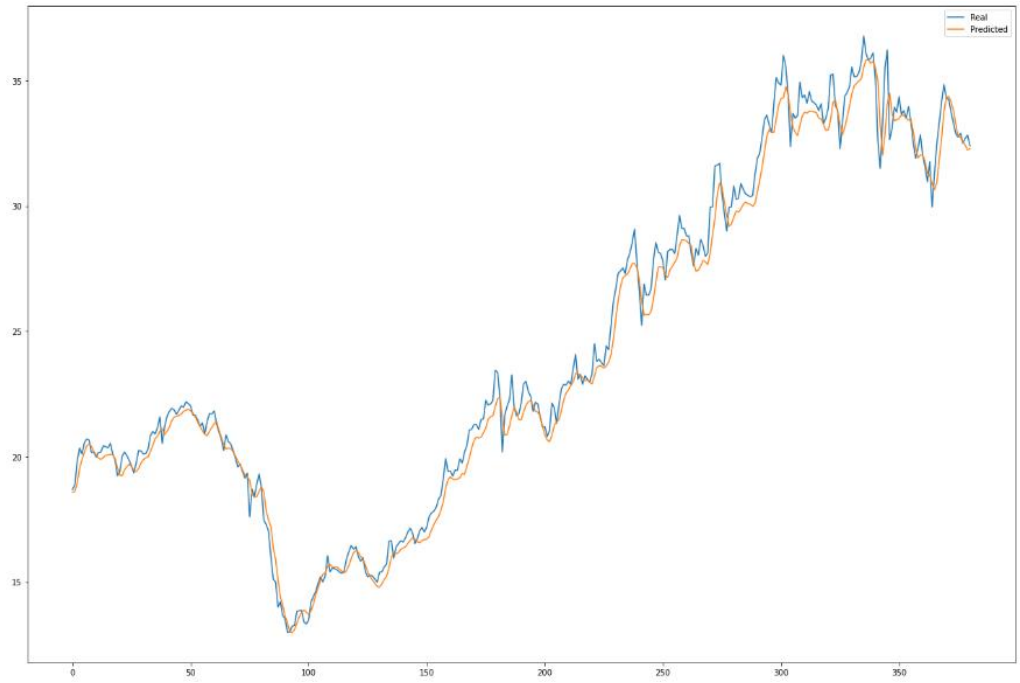

Figure 31: Technical Model (MACD)

Real MSE	0.40344375175787
Scaled MSE	1.69514188533315

In Arcelik testing we got 1.69 best error rate between prediction and real values in our MACD technical model. Arcelik testing show using technical indicators which especially MACD increased results. The concantrated results are below.

Evaluation	**Basic Model Results**	**Technical Model Results**	
		SMA	**MACD**
Real MSE	0.909514539116	0.488505528626	0.40344375175787
Scaled MSE	3.82148982072	2.052544324	1.69514188533315

6.2.4. Eregli Demir Celik – EREGL

Eregli Demir Celik company is one of the biggest industrial companies in Turkey and it occupies the 43rd place among the largest steel companies in the world.

[30] We took stock data of EREGL dates between 18 May, 2006 and 17 May, 2021. Before we deleted null rows, we had 3874 trading day. However, we detected 12 null rows in dataset and cleaned these rows to provide seamless of our model. After deletion process, 3862 rows last. We split dataset as test (%10) and training (%90) hence our dataset was splited as 386 of test values and 3476 of train values. Results are below:

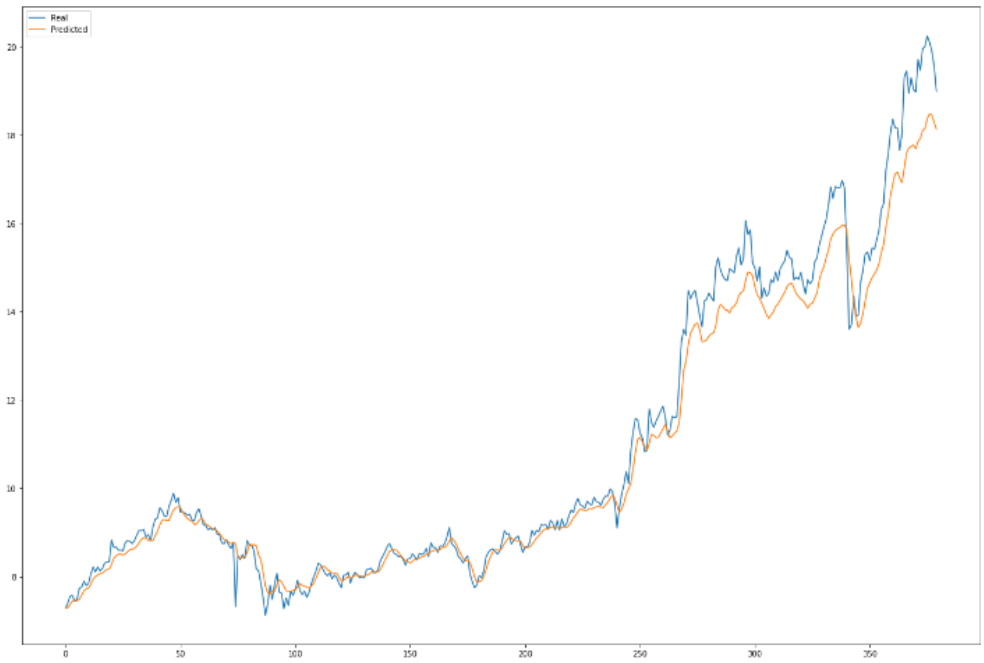

Figure 32: Basic Model

Real MSE	0.15459700629953896
Scaled MSE	1.1792296437798548

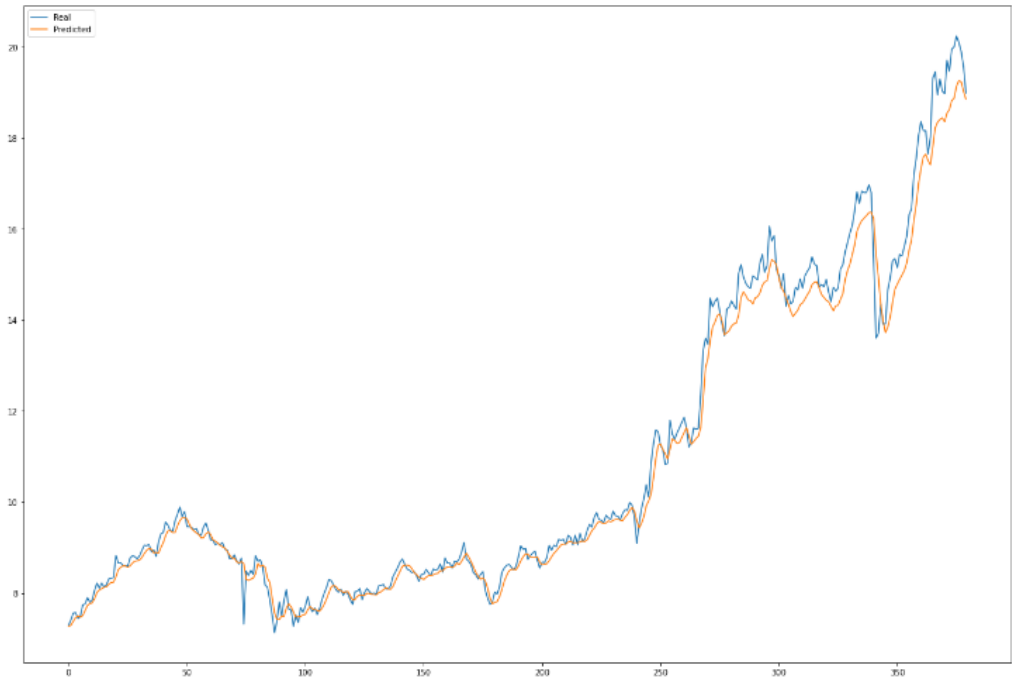

Figure 33: Technical Model (SMA)

Real MSE	0.31847537673332343
Scaled MSE	2.4292553526569294

Figure 34: Technical Model (MACD)

Real MSE	0.12452254945526049
Scaled MSE	0.9498287525191496

In EREGL testing, our technical model which is using MACD gives the best results. EREGL testing shows that technical model which is using technical indicators gives better accuracy rate than basic model gives. Concantrated results are below.

Evaluation	**Basic Model Results**	**Technical Model Results**	
		SMA	**MACD**
Real MSE	0.15459700629953	0.318475376733323	0.12452254945526
Scaled MSE	1.17922964377985	2.429255352656929	**0.94982875251914**

7. Conclusion

High volatility of stock market makes their future prices difficult to predict. Although there had been theories such as Efficient Market Hypothesis and Random Walk Theory argue that prediction stock prices is impossible, economists have developed some mathematical models and formulas to predict stock market prices since 1980's. Thanks to developments in Machine Learning techniques, Artificial Neuron Networks have begun to be used in stock market area after 1990's. In literature, most of researches are related in predicting stock index prices such as Nasdaq, S&P 500 or BIST 100 and technical indicators are not involved into the prediction model in many of them. However in this paper, we used Long Short Term Memory which is advanced type of Artificial Neuron Networks to predict stock prices of four big Turkish companies(Aselsan, Turkish Airlines, Arcelik, Eregli Demir Celik) and we utilized two significant technical indicators for our prediction model. Historical stock prices cover from 22 May 2006 to 18 May 2021 were gathered, prepared, cleaned up and optimized. After our datasets had been ready, we built up our LSTM architectures in two aproaches. In our first model, historical price and volume data were sended through the LSTM network and no adding technical indicator were used. It was named as basic model. In our second model we added two different

technical indicator which are SMA and MACD separetely to our basic prediction model. It was named as technical model. We used MSE (Mean Squared Error) to calculate our model's error rate. Results show that LSTM type of Neural Network is reasonably successful in prediction of stock prices. Our model's accuracy rate is between 3.82 and 0.94 for four companies' prediction results. Another fact that our project indicate when technical indicators were applied, the accuracy rate increased methodically in compare with the results of basic model. Thus, besides our project shows that Machine Learning is a very effective technique for prediction of stock prices , technical indicators which are MACD and SMA are also effective methods on prediction of stock market prices on their own.

8. References

[1] Anonim, Fundamental analysis, https://en.wikipedia.org/wiki/Fundamental_analysis May, **2021**

[2] T. Turner. A Beginner's Guide to Day Trading Online. Adams Media, 2nd edition, **2007**

[3] R. Akita, A. Yoshihara, T. Matsubara and K. Uehara, "Deep learning for stock prediction using numerical and textual information," 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), pp., 1-6, doi: 10.1109/ICIS.2016.7550882. **2016**

[4] Fama, E. F., Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance, 25(2), 383. doi:10.2307/2325486 **1970**

[5] Horne, J. C., & Parker, G. G. , The Random-Walk Theory: An Empirical Test. Financial Analysts Journal, 23(6), 87-92. doi:10.2469/faj.v23.n6.87 **1967**

[6] Dooley, M. P., and Shafer, J. Analysis of short run exchange rate behavior: March 1973 to November 1981. In D.Bigman and T. Taya (eds) Exchange Rate and Trade Instability: Causes, Consequences and Remedies. Cambridge, MA.: Ballinger Publishing: 43- 69. **1983**

[7] Schulmeister S., Technical Analysis and Stock Market Efficiency, pg 19-22, **1988**

[8] Booth, G.G., Martikainen, T., Sarkar, S.K., Virtanen, I. and Yliolli, P., "Nonolinear dependence in Finnish stock returns", European Journal of Operational Research, Vol. 74 No. 2, pp. 273-283, **1994**

[9] Leigh, W., Purvis, R. & Ragusa, J. M., Forecasting the NYSE Composite Index with Technical Analysis, Pattern Recognizer, Neural Network, and Genetic Algorithm: A Case Study in Romantic Decision Support. Elsevier, Decision Support Systems, 32, 361-377. **2002**

[10] White, "Economic prediction using neural networks: the case of IBM daily stock returns," IEEE 1988 International Conference on Neural Networks, pp. 451-458 vol.2, doi: 10.1109/ICNN.1988.23959. **1988**

[11] Wang X, Lin W. Stock Market Prediction Using Neural Networks: Does Trading Volume Help in Short-term Prediction, Proceedings of the International Joint Conference on Neural Networks, 2003, pp. 2438-2442 vol.4, doi: 10.1109/IJCNN.2003.1223946. **2003**

[12] Nelson, D.M.Q., Pereira, A.C.M. and de Oliveira, R.A., "Stock Market's Price Movement Prediction With LSTM Neural Networks", in 2017 International Joint Conference on Neural Networks, pp. 1419-1426. **2017**

[13] Maknickas, A. and Maknickiene, N., "Support system for trading in exchange

market by distributional forecasting model", Informatica, Vol. 30 No. 1, pp.73-90. **2019**

[14] Kim, T. and Kim, H.Y., "Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data", Plos One, Vol. 14 No. 2, p. e0212320. **2019**

[15] Baek, Y. and Kim, H.Y., "ModAugNet: a new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module", Expert Systems with Applications, Vol. 113, pp. 457-480. **2018**

[16] Wilhelm Griffllioen, Technical Analysis in Financial Markets, Faculty of Economics and Business (FEB), pg 4-5 **2003**

[17] Mitchell Tom, Machine Learning. New York: McGraw Hill. 1st Edition, **2019**

[18] Hwanjo Y, Jiong Yan, Jiawei Ha, Classifying Large Data Sets Using SVMs with Hierarchical Clusters, 3rd Edition, **2012**

[19] Tan, Clarence, Artificial Neural Networks: A Financial Tool as Applied in the Australian Market, Award date:199 pg 40-41 , **1997**

[20] Halit Apaydin, Hajar Feizi, Mohammad Taghi Sattari, Muslume Sevba Colak, Shahaboddin Shamshirband, Kwok-Wing Chau, Comparative Analysis of Recurrent Neural Network Architectures for Reservoir Inflow Forecasting, Water ,12, 1500, **2020**

[21] Sidra Mehtab, Jaydip Sen, Abhishek Dutta, Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models, Cornell University, September, **2020**

[22] Xuan Ji, Jiachen Wang, Zhijun Yan, A Stock Price Prediction Method Based On Deep Learning Technology, International Journal of Crowd Science, March, **2021**

[23] Achkar R., Elias-Sleiman F., Ezzidine H., Haidar N., "Comparison of BPA-MLP and LSTM-RNN for Stocks Prediction", in 2018 6th International Symposium on Computational and Business Intelligence, pp. 48-51, Praxis Business School, **2018**

[24] Christian Garbin, Xingquan Zhu, Oge Marques, "Dropout Vs. Batch Normalization: An Empirical Study of Their Impact to Deep Learning", Springer Science and Business Media, LLC, part of Springer Nature 2020 PG 4-5 **April 2019**

[25] Jason Brownlee, Time Series Forecasting as Supervised Learning, https://machinelearningmastery.com/time-series-forecasting-supervised-learning/ **December, 2016**

[26] Anonim, Aselsan, https://en.wikipedia.org/wiki/Aselsan , May, **2021**

[27] Anonim, Aselsan, https://finance.yahoo.com/quote/ASELS.IS/, May, **2021**

[28] Anonymous, Turkish Airlines, https://en.wikipedia.org/wiki/Turkish_Airlines, May, 2021

[29] Anonymous, Arcelik, https://en.wikipedia.org/wiki/Ar%C3%A7elik, May, **2021**

[30] Anonymous, https://en.wikipedia.org/wiki/Erdemir, May, **2021**