Stock Price Forecast using Wavelet Transformations in Multiple Time Windows and Neural Networks

Ajla Kulaglic
Graduate School of Science Engineering and Technology,
Istanbul Technical University,
Istanbul, Turkey
kulaglic@itu.edu.tr

Burak Berk Üstündağ
Graduate School of Science Engineering and Technology,
Istanbul Technical University,
Istanbul, Turkey
berk@berk.tc

***Abstract*— This paper presents a highly reliable and accurate stock-price prediction model. We aim to anticipate the stock price with respect to multiple patterns in different time scales. The stock price time-series are decomposed, using discrete wavelet transform (DWT), into temporal resolution of varying scales. Then, each subseries is used to predict the stock price using two types of neural network (NN) models with one and two hidden layers. Results show that having multiple time windows in input datasets together with DWT decrease the RMSE of NN models below 10%.**

Keywords*— *discrete wavelet transform, neural networks, stock price forecasting, Apple Stock Prices, financial time series data

I. Introduction

Financial market is a complex system involving huge amount of information to produce an output. Those information ranging from fundamental information to socio-political events and news to investor's behavior [1]. Due to a complex market dynamics and a large-scale structural shifts in economy these modeling approaches failure to accurately represent and forecast big movements in national and financial markets. In the literature two approaches for forecasting the movements of stock prices could be found: fundamental and technical analysis [1]. Fundamental analysis includes fundamental indicators of a market, such as Return on Equity (ROE), Earning per Share (EPS) and Price to Earnings (PE) ratio. Technical analysis evaluate stock prices changes based on the behavior of previous stock price values.

A stock market price represents the movement average of many individual stocks; a price reflects mainly market movement rather than movement of a stock [2]. In the context of financial modeling and forecasting the movements of stock prices machine learning techniques are used.

Machine Learning (ML) mainly uses artificial algorithms to learn the pattern and derive a model with ability to predict new data with the same structure [3]. The capability of various machine learning techniques to capture and model non-stationary and non-linear data cause their wide applicability in time series forecasting [4-10].

Several different Neural Network (NN) models have been adopted to forecast the financial time series, especially stock prices [5-10]. The Recurrent Neural Network (RNN) are mostly used due to their ability to represent certain computational structures in a parsimonious mode [5]. Dynamic RNN used in [5] provide five-day ahead forecasting. On the other hand, in [6] using internal states the temporal relationships between time-series data have been constructed. Incorporating autoregressive filters into a full RNN structure in [7] they obtained a temporal dimension.

The NN models due to their limitations suffer from the problem of overfitting and getting stuck in the local optima. Overcoming these issues and accomplish accuracy can be done by improving the algorithms, preprocessing the data or both [4]. The data preprocessing allows us to transform the time-series data into a format that reveals certain characteristics. For that purpose, in the recent times, the signal processing techniques have been widely used. Different variations of transformations are used for decomposition of time-series data into time-frequency and time-domain. Fourier Transform (FT) decomposes the signal only into frequency domain where the information related to occurrence of frequency is not captured. The FT eliminates the time resolution. On the other hand, Wavelet Transform (WT) gives a good local representation of a signal in both time and frequency domain simultaneously.

Many practical applications of WT in finance and economics can be found in the literature [3-10]. The WT decompose the data into time and frequency domain capturing the information related to occurrence of frequency and hold the time resolution [3]. Due to limitations that WT is limited with the dyadic length constraint, some studies try to modified the WT and overcome this requirement. Modified Discrete WT (MDWT) is circular shift invariant and is not limited with the dyadic length constraint [4].

In this study we used Discrete Wavelet Transformation (DWT) for decomposition approach. Thereafter obtaining the wavelet coefficients we construct two NN models and analyze the error rate due to different periodicity of time-series data. The Apple Stock prices for the period May 2008 until May 2018 are used for forecasting purpose. From the study it can be seen that using multiple time windows data as input, together with preprocessing mechanism gives better accuracy.

The rest of the paper is organized as follows: Section 2 presents the detailed information about methodology and dataset used in this work, together with the information about decomposition of time-series data and construction of NN models. In sections 3 we discuss the results depicts the different level of decompositions and different NN models. The section 4 contains the conclusion of the presented study.

II. Materials and Methods

A. Apple Stock Prices

Apple Stock Prices (ticker symbol AAPL) data provided by National Association of Securities Dealers Automated

978-1-5386-7893-0/18/$31.00 ©2018 IEEE

Fig.1 Apple Stock Prices for the period May 2008 – May 2018

Quotations (NASDAQ) [13] are examined in this study. NASDAQ is an American Stock Exchange and it is the second-largest exchange in the world by market capilization.

Apple is one of the world's leading consumer electronics and personal computer companies established in 1977 in USA [14]. In this research we consider last ten years of AAPL stock prices collected between May 2008 and May 2018, excluding weekends and holidays (Fig.1). From the figure it can be seen that the prices are in constant growth with exception of second quarters 2013 and 2016. In those periods the falling stock prices are detected. Although our dataset contains open prices, close prices, high and low prices and trading volume of the day, here we use only closing prices of a day. The total dataset includes 2520 points.

For purpose of this study, the time-series data are organized into two different datasets. The unit period is "day's resolution" with weekly shift. The first input dataset contains 8 business days, while second dataset takes 4 business days with 4 weekly average values of one-month week resolution. The output dataset consists of 5 business days with weekly shift. In the Fig. 2 and Fig. 3 the arrangement of time-series data for first and second datasets are shown, respectively.

The first dataset always starts with Wednesday and ends with Friday. The second dataset starts with Tuesday and ends with Friday containing also 4 average values of current week, last week and week before last week. The forecasting is done for next week's business days (Monday, Tuesday, Wednesday, Thursday and Friday), after the last Friday from input datasets.

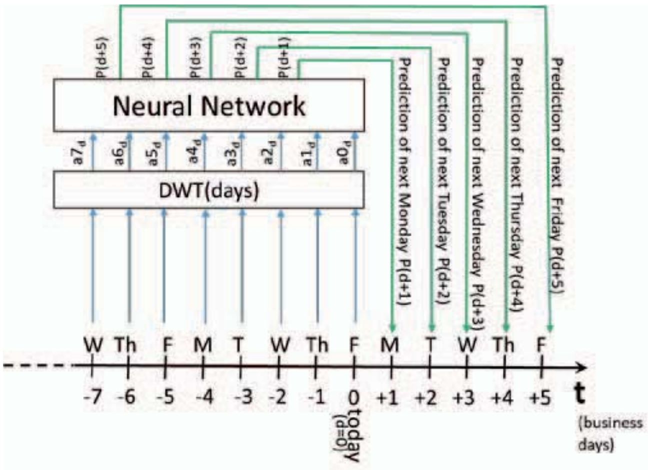

Fig. 2. Arrangement of time-series data for the first dataset

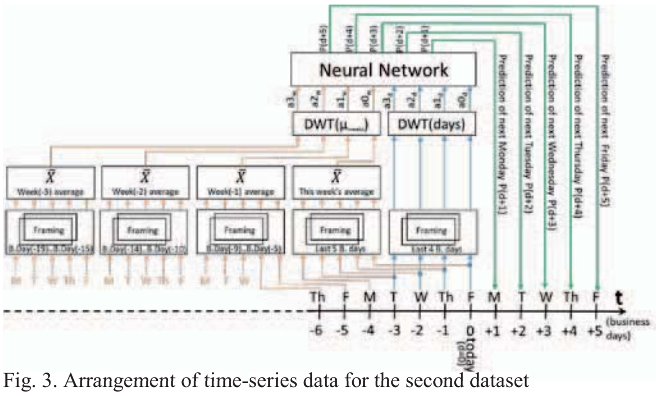

Fig. 3. Arrangement of time-series data for the second dataset

B. Methods

This work presents a learning stock prices forecasting framework using two different NN models. Back Propagation (BP) feed-forward NNs with differed number of layers are utilized. The first model contains one (Fig. 4), whilst second model contains two hidden layers (Fig. 5). The first network model structure is 5x16x6 which means that we have 5 inputs in the input layer, 16 neurons in the first hidden layer, and 5 outputs. The size of hidden layer is chosen with respect to the number of input variables.

The structure of second model is 5x16x8x5 with 16 neurons in first and 8 neurons in second hidden layer. Rectified Linear Unit (ReLu) activation function is used for both models (Eq. 1). ReLU is more practical and compared to the widely used activation functions (sigmoid and hyperbolic tangent) enables better training [16, 17].

$$f(x)=x^{+}=max(0,x) \tag{1}$$

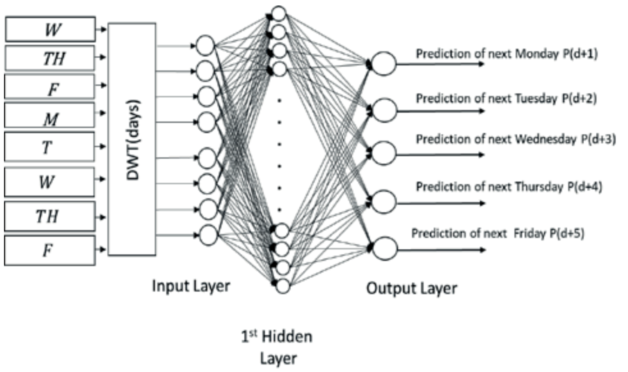

Fig. 4. One hidden layer Neural Network Model

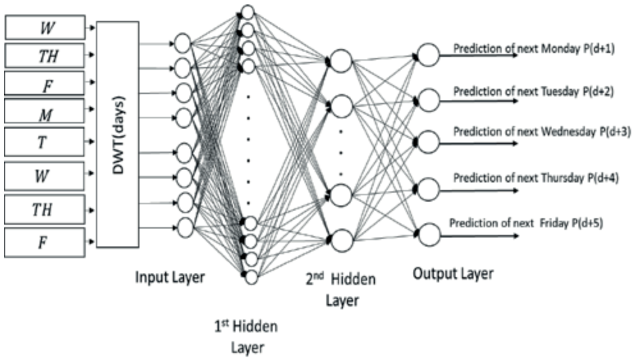

Fig. 5. Neural Network Model with Two Hidden Layers

TABLE I. RMSE FOR TWO DIFFERENT NN MODELS WITH RESPECT TO THE APPLIED DATA PREPROCESSING MECHANISM

Network architecture		8 business days			4days+4 average weeks		
		Training RMSE	Test RMSE	R	Training RMSE	Test RMSE	R
NN {8-16-5}		1.87	3.53	0.95	1.93	3.02	0.98
NN {8-16-8-5}		1.99	3.89	0.94	2.21	3.37	0.97
NN {5-16-5}	With DWT	1.83	3.35	0.96	2.32	3.60	0.97
NN {5-16-8-5}	With DWT	1.80	3.29	0.95	2.53	4.35	0.95

Considering that the price movements of stocks are extremely noisy [18], we decompose the stock price time series data to eliminate noise. The Wavelet Transform is used since its ability to handle the non-stationary of financial time series data [19]. The time-series data are decomposed into discrete wavelets using Discrete Wavelet Transformation (DWT). We represent weekly information of the stock prices using at least one low-frequency or average values of wavelet coefficients from different levels.

This research applies the Haar function as the wavelet basis function. The main reason for choosing this function is due it reduce the processing time significantly [19]. The decomposed data are used as input elements to a neural network to capture valuable information during a training processes.

III. RESULTS

The proposed work is carried out into two parts. The first part is arrangement of input dataset and decomposition of previously explained datasets. The decomposition is done from 1st to 3rd level. Appropriate coefficients from each level are used as inputs to the NN models. The Python programming language is used for the implementation of described work.

Root Mean Square Error (RMSE) is computed and used as measure of the difference between values predicted by a model and obtained values (Eq. 2)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \widehat{X_i})^2}{n}} \quad (2)$$

Where $\widehat{X}_i$ and X_i denote the model and target output, respectively.

The datasets are randomly divided into training and test set (70% for training set, 30% for test set). The first dataset, 8 business days with weekly shift time interval after arrangement contains 382 data samples. Training set contains 250 data samples, while test set contains 132.

Concurrently, the second dataset after arrangement into 4 business days with average weekly values contains 501 data samples, with 350 data samples in training and 151 data samples in test dataset.

The results are presented in the Table 1 for both NN models. The RMSE for the dataset with 8-business days is relatively high, compared with one of different multiple time windows. In addition, the obtained results for the second dataset with 4-business days and 4 average weekly data considerably reduces the error.

Different size of time window and shift rate significantly reduce the errors, likewise the use of DWT as preprocessing mechanism improves our results. Similarly, increase in the number of hidden layers decrease the RMSE error and gives better accuracy.

In Table 2 and 3, the RMSE errors with respect to days in a week are presented. According to the literature, the most common patterns are the “Monday Effect”, as well as the “January Effect” [11,12]. Presented results indicate the smaller errors for Mondays as starting week and existence of Wednesday and Friday effect with higher RMSE in each of presented models.

TABLE II. DAILY RMSE FOR FIRST DATASET WITH RESPECT TO TWO DIFFERENT NN MODELS AND THE APPLIED DATA PREPROCESSING MECHANISM

NN model:		**M**	**T**	**W**	**T**	**F**
(8x16x5) ReLu	**train**	1.33	1.68	1.75	2.20	2.24
	test	2.43	2.92	3.34	3.85	4.08
NN model:		**M**	**T**	**W**	**T**	**F**
(8x16x8x5) ReLU	**train**	1.22	1.53	1.84	2.42	2.58
	test	2.46	2.96	3.53	4.79	5.03
NN model:		**M**	**T**	**W**	**T**	**F**
(5x16x5) ReLu, DWT	**train**	1.10	1.47	1.78	2.31	2.22
	test	2.18	2.72	3.29	4.10	4.05
NN model:		**M**	**T**	**W**	**T**	**F**
(5x16x8x5), ReLu, DWT	**train**	0.95	1.52	1.99	2.04	2.21
	test	1.87	3.19	4.17	3.87	3.98

TABLE III. DAILY RMSE FOR SECOND DATASET WITH RESPECT TO TWO DIFFERENT NN MODELS AND THE APPLIED DATA PREPROCESSING MECHANISM

NN model:		**M**	**T**	**W**	**T**	**F**
(8x16x5) ReLu	**train**	1.25	1.58	1.86	2.23	2.48
	test	2.05	2.57	3.10	3.24	3.83
NN model:		**M**	**T**	**W**	**T**	**F**
(8x16x8x5) ReLU	**train**	1.44	1.93	2.10	2.58	2.75
	test	1.95	2.76	3.39	3.87	4.34
NN model:		**M**	**T**	**W**	**T**	**F**
(5x16x5) ReLu, DWT	**train**	1.68	2.11	2.27	2.68	2.69
	test	2.47	3.17	3.72	4.12	4.22
NN model:		**M**	**T**	**W**	**T**	**F**
(5x16x8x5), ReLu, DWT	**train**	2.25	2.31	2.38	2.78	2.84
	test	4.07	4.13	4.25	4.63	4.64

IV. CONCLUSION

In this study, we used wavelet coefficients of time series stock price patterns in multiple time windows for Apple Inc. stock price forecast.

In order to demonstrate the performance difference of wavelet preprocessing in multiple time windows, we prepared three test cases. In all cases we used 8 inputs for training and testing the forecasts. We evaluated the forecast performance for next 5 business days with 5 business days (weekly) shift for all cases. For this reason, Mondays stock value forecast by the end of Fridays' is clearly having least error rate since it is the first day after known latest closing value. On the other hand, using multiple time windows provides dramatically much higher forecast performance than single time window. When 8 business days' values in weekly shift (5 business days) is directly applied to the neural network, Mondays' root means square error (RMSE) is computed as 3.53. In the second case, when we use 4 business days and last 4 weekly average values as 8 inputs to the neural network, RMSE error is reduced down to 3.02. In the third case, we additionally apply discrete wavelet transform to the 8 business days' values as preprocessing before the neural network then the RMSE error is decreased below 3.29. Training set errors are as low as 25% of the testing data set errors.

Results have shown that having different size of time windows in input datasets by using DWT as preprocessing mechanism notably decreases the RMSE of NN models.

REFERENCES

[1] M. G. Novak, D. Veluscek, "Prediction of stock price movement based on daily high prices", Quantitative Finance, doi:10.1080/14697688.2015.1070960, 2015

[2] T.J. Hsieh, H.F. Hsiao, W.C. Yeh, "Forecasting stock markets using wavelet transform and recurrent neural networks: An integrated system based on aritifial bee colony algorithm", Applied Soft Computing 11(2011) 2510-2525

[3] X.Y. Qian, S. Gao, "Financial Series Prediction: Comparison Between Precision of Time Series Models and Machine Learning Methods", CoRR, 1706.00948, 2017, http://arxiv.org/abs/1706.00948

[4] D. Jothimani, R. Shankar, S.S. Yadav, "Discrete Wavelet Transform-Based Prediction of Stock Index: A Study of National Stock Exchange Fifty Index", Journal of Financial Management and Analysis, 28(2), 35-49, 2015.

[5] A. Aussem, J.Campbell, F.Murtagh, "Wavelet-Based Feature Extraction and Decomposition Strategies for Financial Forecasting", Journal of Computational Intelligence in Finance, March/April, 1998.

[6] C. L. Giles, S. Lawrence, A. C. Tsoi, "Noisy Time Series Prediction using Reccurent Neural Network and Grammatical Inference", Machine Learning, Volume 44, No. ½, July/August, pp.161-183, 2001.

[7] S. Lahmiri, "Wavelet low- and high-frequency components as features for predicting stock prices with backpropagation neural networks", Journal of King Saud University – Computer and Information Sciences, pp. 218-227(26:2014)

[8] L. Ortega, K. Khashanah, "A neuro-wavelet model for the short-term forecasting of high-frequency time series of stock returns", Journal of Forecasting, pp.134-146, (33:2014)

[9] B. Al-Hnaity, M.Abbod, "A novel hybrid ensemble model to predict FTSE100 index by combining neural network and EEMD", 2015 European Control Conference (ECC), pp. 3021-3028, 2015

[10] H. Tao, "A wavelet neural network model for forecasting exchange rate integrated with genetic algorithm", IJCSNS International Journal of Computer Science and Network Security, 6:60-63, 2006.

[11] J. F. Kaashoek, H.K. van Dijk, "Neural network pruning applied to real exchange rate anaylsis", Journal of Forecasting, 21:559-577, 2002.

[12] G. M. Caporale, A. Plastun, "Price gaps: Another market anomaly?", Investment Analysts Journal, 46:4,279-293, doi: 10.1080/10293523.2017.1333563

[13] K. Wang, Y. Li and J. Erickson, "A new look at the Monday Effect", The Journal of Finance, vol LII, no.5, December, 1997

[14] https://www.nasdaq.com/symbol/aapl

[15] "Monthly Reports". World-Exchanges.org. World Federation of Exchanges. Archived from the original on August 17, 2014. Retrieved June 3, 2015.

[16] http://secfilings.nasdaq.com/edgar_conv_html%2f2017%2f11%2f03%2f0000320193-17-000070.html#FIS_BUSINESS

[17] Xavier Glorot, Antoine Bordes and Yoshua Bengio(2011). Deep sparse rectifier neural networks(PDF). AISTATS.,

[18] Yann LeCun, Leon Bottou, Genevieve B. Orr and Klaus-Robert Müller (1998). "Efficient BackProp"(PDF). In G. Orr and K. Müller. Neural Networks: Tricks of the Trade. Springer

[19] R. J. Shiller, "Do stock prices move too much to be justified by subesquent changes in dividends?" 1980

[19] W. Bao, J. Yue, Y. Rao, "A deep learning framework for financial time series using stacked autoencoders and long-short term memory", PLoS ONE 12(7): e0180944, https://doi.org/10.1371/journal.pone.0180944