International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

Trend Forecasting in Financial Time Series with Indicator System

Seth Gyamerah[1], Andrews Awuah[2]

Department of Computer Science, Regentropfen College of Applied Science, Namoo-Bolgatanga, Ghana

Abstract: ***Trend forecasting of financial time series is a complex task, many researchers have use statistical models, time series models and machine learning model in this filed of study. The advancement in computational technologies has made it easier, deep learning models is capable of handle time series data. In this research, our objective is to design a system architecture trading strategy algorithm to track the trends movement of prices and to verify if the principal component as input variables would increase the accuracy of Deep learning models in stock market predictions task. We used S&P500 stock exchange price data in our experiment conducted. In the first stage, we preprocess the dataset where we use wavelength transform to denoise the noise in the dataset and created a new factor differences asinput variables, (the differences between Low price, High prices, etc). In the second stage, we use principal component as input variables to investigate if it will increase the model accuracy. We also selected technical indicators as input variables.The final stage, we trained and tested our design system Architecture by using LSTM model to forecasting exact future trend of the prices. We compare the result to a baseline machine learning model, Random Forest and Logistic regression model. Thorough empirical studies based on S&P500 dataset; our design system LSTM model demonstrate a high accuracy that outperform state of art- methods for trend forecasting of financial time series.***

Keywords: LSTM, S&P500, Financial time series, Technical Indicators, Principal component analysis and Wavelet transform

1. Introduction

1.1 Background to the Study

Recently most scholars have conducted research on financial time series especially in the area of stock market trading system. The interest of most individual is that if the trend of the prices of stocks is successfully predicted investors may be better guided. The profit in investment and trading by individual, institution in stock market to a large extend depends on predictability and If any program is developed and can continuously predict the trends of the dynamic stock markets would be wealthy. More over the predicted trends of financial time series especially the stock market will help the market analyst in making correct measures.Another concern for researching in this field is that it have many theoretical and experimental challenges. The most important is the Efficient Market Hypothesis (EMH:EUGENE FAMAS 1970) .The Hypothesis explained that stock prices is not an exact detail information about the market and its constituents' and that every opportunity of earning more profit ceases to exist. So is assume that no system is expected to outperform the market predictability.Hence modeling any markets under Expected market Hypothesis, the assumption is that it only possible on the speculative, stochastic components but not on the changes in value. There are a lot of argument on volatility of the Efficient market Hypothesis and random walk theory.The advance in computational and financial intelligent, behavior finance and economist have formulated opposite hypothesis called inefficient market hypothesis which states financial markets are not always affected, the markets is not always in a random walk(Pang Heping 2003).However, many researcher have agreed to the point that the stocks markets is non-linear which only appears as random walk because of it irregularities in nature. The system is sensitive to the initial conditions of the systems. The systems are dynamic, complicated to deal with normal analytical methods

In this paper indicator system is used to predict the trend, one common approach is to use Machine Learning systems to learn from the prices history data to predict the trend movement of prices. This paper goes in such direction but application of a specific design system using recurrent neural networks. Such networks have a short term memory capability and hypothesis to explore in that features can be profit in terms of result when compared to others more traditional approaches in machine learning systems.

The algorithm of choice is LSTMs (Long-Short-Term memory) use for training and testing of our design system architecture . LSTM is a type of recurrent network that has provided successful on a number of problems given its capability to separate between recent and early examples by giving different weights for each, while forgetting the memory it considers irrelevant to predict the next output.

In such instance, it is able to handle long sequences of input when compared to otherrecurrent neural networks that are able to memory short sequences.Historical price data of S&P500, stock exchange markets is used as a source of information for the network, alongside with this data a selected Technical indicators systems will also be generated to feed the network model .Principal component analysis which is commonly know for features selection, in this paper is use as input variables because it variance is capable to increase the model accuracy. In all our input variables will be TA+PCA +new created factors differences features which is generated to feed the network as input variables.Upon this dataset the model is trained, evaluated and will attempt to predict the trend movement of S&P500 stock prices by a look-back of 15hrs (900MINUTES) to forecast the next 30minutes, if the price of stock will be up or down The rest of the paper is organized as follow as: literature review, the methodology, experimental result and discussion and conclusion and Future recommendation work.

2. Literature Review

The most popular technique used in financial time series analysis to detect the trend and seasonality is Autoregressive integrated moving average. ARIMA [1]. The current advancement in technology has improved modern techniques in forecasting of financial time series data. Deep LSTM, Shallow LSTM, 1-CNN and machine learning models were used to predict stock market data [2]. An attention LSTM model is also used in prediction of financial time series data [3]. In[4] ARIMA model is used in prediction of the closing prices of time series data.Va Yassine, Khadija and Faddoul in [5] formulated classification model designed based on LSTM network to predict the probability of investing or not. Gauting, Yatong and Takahiro in [6] implemented a deep learning models as a strategy for trading. Manuel R.Vergas et al in [7] use LSTM model and Technical indicators to forecast S&P500 index. In [8] a hybrid model based on LSTM network and several GARCH model is used in to forecast the stock prices volatility. Abe,M. Nakayama, applied deep learning model for forecasting stock returns in the cross-section.

3. Methodology

This section is about method used in analyzing the trend of S&P500 stocks, Wherewe design a system Architecture of stock market price prediction which are explained and its associated process flow.

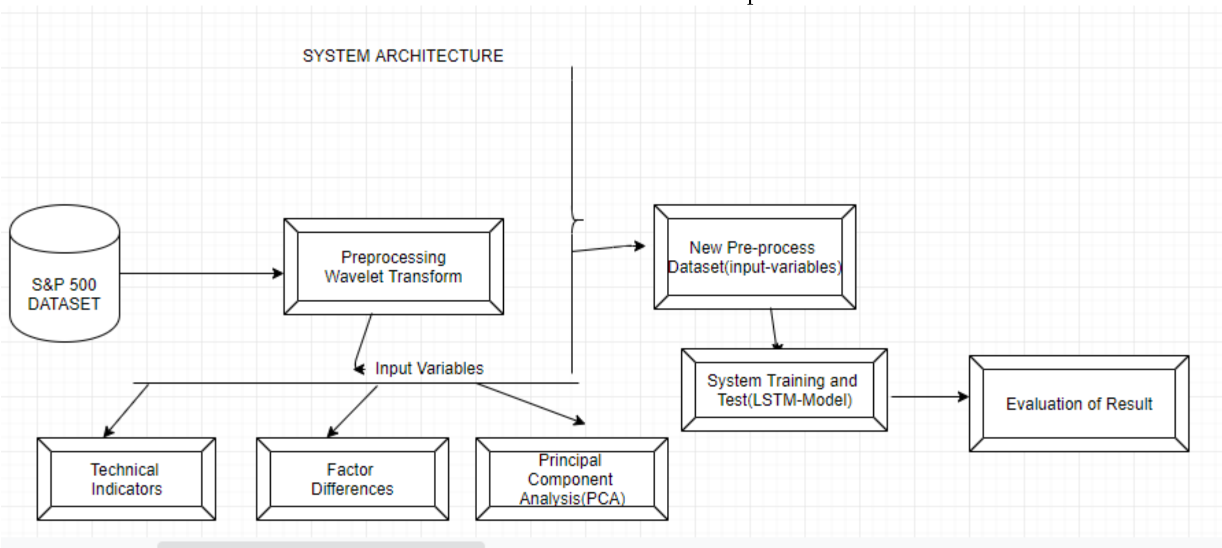

Figure 1

3.1 Dataset Description

Historical prices data for a particular, stocks are gathered in the format of time series of candles (Open, Close, High, Low and Volume) in a granularity of 15minutes. For this task, data was collected from finam.ru, an open stock trading firm where every individual has equal chances of having access to historical prices of different stocks data. The S&P500 was our sample of data gathered from 1999 to 2003

3.2 Preprocessing

In order to work with this dataset, we preprocess it, where we first visualize it to illustrate it outmost importance. This makes it possible to identified trends, missing values, noise and outliers as shown in figure 2 below.

Due to the complexity of financial stock exchange environment, stock prices contains a lot of noise that make it very difficulty to achieved a good model accuracy when trying to forecast it trend movement. Since is the interest of investors to achieve good forecasting model, then there is a need to reduce the noise in the dataset because stock environment is contains noise from the news articles and other source media information. A wavelettransform, a mathematical function introduce in the preprocessing stage by denoise S&P500 dataset and present it trend and structure of the dataset. We transform the dataset using wavelet mathematical transform function:

$$X_{\omega}(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} x(t)\psi\left(\frac{t-b}{a}\right)dt$$

After transformation we drop the co-efficient with more standard deviation away from the co-efficient and inversely transform the new co-efficient to get the denoise S&P 500 dataset.

Input Variables

4.1 Factor Differences

We proposed new factors as differences between the prices of data, the differences add more informative to our LSTM network. This is also used as initial to detect the trend differences in the prices movement. We were more interested in the candles (Open, High, Low, Close), so we drop the volume prices column at the initial stage.

The factors differences in OHLC was defined as:

- OH $= (O_p - H_p)$, HL = $(H_p - L_p)$, OL = $(O_p - L_p)$,
- HC = $(H_p - C_p)$, LC = $(L_p - C_p)$, OC = $\left(O_p - C_p\right)$

Where O, H, L , C is open high, low and close prices , and it differences between the prices The new created features is added to the input variables of our dataset.

4.2 Principal Component Analysis

The Principal Component Analysis is mainly use for feature extraction when there are large amount of input features which is used to find the features with more importance but In this work it is used as input variables to check if it will increase the model accuracy. We implement the principal component on the dataset by the use PCA() class in Scikit Learn. The fit() function is used on the class of dataset and a chosen number of dimension with the transform() function. The PCA class contains explained -variables which return the variance caused by each of the component.

4.3 Technical Indicators

A selected technical indicator were used initial to perform predictions on the dataset and it was finally included as input variables. Due to non-stationary of stocks data, technical indicator help in monitoring the upward and downward of prices and also use to reduce the noise in dataset. Below are the few technical indicators use in initial prediction of trend where we use time interval of 30, 45 and 60minutes predictions.

Table 2

Name	Technical Indicators (Definition/Implication
MACD	Moving average convergence divergence display the trend following the characteristic of market prices.
CCI	Consumer channel index identify prices reversals, price extreme and the strength of rise prices
ROC	Rate of change is a momentum indicator that measures the percentage change in prices from one period to the next.
RSI	Relative strength index is a momentum oscillator that detect the speed and changes of price movements
Momentum	Momentum is an indicator that measure the velocity of changes in prices in time.
OBV	Open balance volume is a momentum indicator that use the volume prices to predict changes of prices.
MA	Moving average is an indicator that smooth out the price by removing the noise of price changes over short period of time.
EMA	Exponential moving average is a moving average that have a greater weight importance on the recent prices of data. The exponential has more advantage than other moving average because it faster to response to prices fluctuation in market
ADX	Average directional index is an indicator use to determine the strength trend of prices movement. The trend can either be up or down base on the two indicators, positive directional and negative directional index

5. System Training and Testing

Given the recent success of RNNs in sequential data, we decided to apply such network, the LSTM model on our new process S&P500 data. We trained and test our network to determine if the prices of S&P500 stocks will up or down option.

In order to feed the input variable to our network, we created new function that is dataset NumPy array and set a look-back which is the number of previous time steps to be used as input variable to predict the next time prediction. In this work our look-back is used as default to create a dataset, where X_t the input variable and is defined as: $\left(x_{1,}x_2 \dots\dots..x_t\right)$, X_t is the prices of data along with other input variables and Y is the next time prediction $(t+1)$. We trained and test our system architecture using LSTM network on the input variable to forecast 30 minutes prices movement into the future and the output should be in the range of [0,1] The LSTM network model was build on Kera's Library which comprise of input layers that takes both the Technical indicators, Price data and others as input and feed on output layer using sigmoid function. The input layer has a dimensionality of 20 features. And 500 mini-batches over 150 epochs were chosen a dropout of 0.2, metrics is accuracy and for our model Optimizer is Adam.

5.1 Baseline approach

For comparison, the baseline approach is chosen based on classical machine learning algorithm. The machine learning algorithm are traditional method that is widely used but less complex than the model used in this work. Using the same input variables trained and test our system architecture

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

predictions on S&P500 new process data. we chose the Logistic regression and Random Forest model. We implemented the two baseline models using Scikit -Learn library

6. Evaluation of Result

$$y_{prob} = \begin{cases} 0, & \hat{y}_{model} < 0.5 \\ 1, & otherwise \end{cases}$$

We further move on to evaluate the performance of our prediction model, the selected performance metrics: Recall, Precision and F1 is chosen. We define the chosen metrics as positive class (true positive), negative class (true negative).

$$\text{Precision} = \frac{truepositive}{truepositive\ + falsepositive},$$

$$\text{Recall} \quad = \frac{truepositive}{truepositive\ + falsenegative}$$

The first objective is to find out if the PCA will increase or not increased the accuracy of the model. According to the LSTM model on our system Architecture, the PCA as input variables, achieved training accuracy of 0.798% and validation accuracy of 0.687% and when not used as input variables the training accuracy was 0.786% and validation accuracy of 0.586%. And it is concluded on the first objective that the principal component analysis as input variables improved the accuracy of our design system architecture model prediction.

Below is the result of our LSTM and the baseline models:

Table 2: Result of Deep Learning (LSTM) and Baseline models

Metrics	Deep learning Model (or LSTM model) (%)	Baseline (random forest model) (%)	Baseline (logistic regression model) (%)
Accuracy	0.687%	0.56%	0.551
Precision	0.538	0.514	0.594
Recall	0.697	0.412	0.236
F1-Score	0.607	0.458	0.354

Source: Results of Data Analysis

The main objective is to achieved a good algorithm in terms of accuracy for tracking the trend movement of prices. From the result above, LSTM model used in training and testing the system architecture give clear indication of the stock market prices movement. This can be used as trading strategy algorithm since is hardly difficult to achieve better accuracy of stock market data. Deep learning model at times is unablesignificantly to overfit the training dataset, which suggest that there may not be enough signals in pure market data to forecast even the trend movement of the prices. However in our experiment, conducted on S & P500 our LSTM model network trained and tested on system architecture show promise by achieving 0.768% training accuracy and validation accuracy 0.687% with the baseline model also perform comparable to LSTM model. We can observe that in general the design system architecture (LSTM network model) outperform the baseline approach trained and tested on the design system Architecture.

7. Conclusion and Recommendation

In this work, it is showed that our design system architecture performance on trend forecasting of financial time series data. We used LSTM network model in training and testing our system. The noise in the stock market environment make it arguable in achieving a good prediction accuracy, is it always assume to be probability of 50% in terms of accuracy that is Up or Down options trend movement.We conducted an experiment on S&P500 stock exchange data, where our design system architecture model chosen perform slightly differences to the baseline models. However, our problem statement in this study has been achieved, where we were able to detect the Up and Down trend in the prices movement, the system architecture design modelpredicted more upwardsthan downwards S&P500 stock exchange market prices. Our design system architecture LSTM network model is used in other stock exchange market 1minutes data points and it achieved a better result. but there are more works to be done to improve the accuracy of the model on financial time series data.Some area that will be needed to achieved better accuracy in our future work, Some scholars are concerning the used of sentiment analysis on twitter to predict stock market trend movement, Beside the social media, other qualitative indicators like news, internet and domestic policy changes can also be used as input to predict the trend of stocks. Another important concept is Elliot waves principles which can perform better on stock market trend prediction.

References

[1] Kelvin Li and Glen Yu(2000), Deep Learning for stock price forecasting. Journal of machine learning.

[2] Takaya Shimtate and Lukas Pichi(2019), Trend prediction classification for high frequency Bitcoin Time series with Deep learning .Journal of risk and Financial management.

[3] Dr MakramZardi, Najiran university (2016). Forecasting stock market trend by Logistic regression and Neural networks evidence of KSA stocks. Journal of Economics and finance.

[4] MingyueQiu and Yu song(2016), predicting the direction of stock market index movement using an optimized Artificial neural network model. Journal PMC 4873195.

[5] GautingChen,Yatongchenand Takahiro Fusgimi(2017).Application of Deep learning to algorithmic trading. Thesis work at Stanford university computational and mechanical engineering department.

[6] Hyan woo Byun, SeanghoBeek and Jae Joon.Ah(2015), using principal component analysis for multi-currencies -trading in the foreign exchange market. Journal of Financial and Quantitative Analysis.

[7] Yassine Touzani, Khadija Douzi and Faddoul Khoukhi (August 2018).Stock forecasting : new model for uptrend and downtrend detection anticipating based on Long-term -memory. Journal of operational research.

[8] Ruey .S.Tay university of Chicago(2002) Analysis of financial time series :wiley -inter-science publication .

[9] S.AL wadi, Mohammed Almasarweti and Ahmed AtallahAlsaraireh(March 2018). Predicting closed price

time series data using ARIMA model. Journal of statistic and Econometrics.

[10] Manuel R.Vargas, Carlos E.M. dosAnjos, Gustavo L.G.Bichara, Alexandre G.Evsulkoff(July, 2018) , Deep learning for stock market predictions using Technical indicators and Financial New articles .International Joint conference on neural networks.

[11] Felipe Barboza oriani and Guilherme Palermo Coelho (December, 2016). Evaluating the impact of technical indicators on stock forecasting. International conference 2016, symposium series on computational intelligence.

[12] ErbanBeyaz, FiratTekiner , Xiao-jun Zeng, john A. Keane (2018) comparing technical and fundamental indicators in stock price forecasting. International conference on High performance computing and communication, IEE 16th conference on smart city.

[13] Rajashree, Dash and Pradipta Kishone Dash (March,2016). A hybrid stock trading framework integrating technical analysis with machine learning techniques. The Journal of finance and Data science 2(2016) 42-57.

[14] Francisco J.Ruiz, Alberta Jama, German Sanchez, Jose A. Sanabria and NunaAgell(2011) An interval technical indicators for financial time series forecasting .Journal of statistical software, vol.

[15] Weihong Huang and Yu Zhang(June 31,2014) Asymmetry Index of stock price fluctuations .Journal of Global Economics

[16] BasitJanvir, khan MomanJareletc (August 2017) Evolving Technical Trading strategies using Genetic Algorithms .A case study of Pakistan stock exchange. Conference Paper: IDEAL 2017

[17] Fiu Feng, Xiangnan He, Xiang Wnagetc (March, 2018) Temporal Relational Banking for stock prediction. Journal of HCM transactions information system 37(2) 1-30

[18] Tobias Schadler(Dec,25,2018) Measuring irrationality in Financial markets. Achieves Business Research Vol.6 No.12

[19] Gary R Weckmen, Siriam Lakshminarayanan (May, 2004) Identify technical indicators for stock market prediction with Neural Network. Conferences: proceedings of the IIE Annual conference.

[20] David Acheme, Olufurene Rebecca etc (July 2014). A prediction stock market technical analysis using Fuzzy Logic. The Journal of computer and information science.