Karachi Stock Exchange Price Prediction using Machine Learning Regression Techniques

Mazhar Hameed, Khurum Iqbal, Rozaida Ghazali, Fawwad Hassan Jaskani, Zunaira Saman

Abstract

Accurate stock market returns are quite difficult for the company because of the unpredictable and non-linear nature of the financial stock markets. With the development of artificial intelligence and increased computer power, programmed prediction approaches have demonstrated that they are increasingly effective in predicting stock values. In this study, the Artificial Neural Network, LSTM, and LR techniques were used to predict the closing price for the following day for five companies belonging to different business sectors. In today's economy, the stock market or equity market has a profound influence. The prediction of stock prices is quite complex, chaotic, and it is a big challenge to have a dynamic environment. Behavioural finance means that investors' decision-making processes are affected by emotions and attitudes in response to particular news. In order to help investors' judgements, we have supplied a technology for the analysis of the stock exchange. The method combines historical price prediction. For predicting, LSTM (Long Short-Term Memory), ANN and LR are employed. It includes the latest information on trade and analytical indicators. Financial data: Open, high, low and close stock prices are used to build new variables needed for model input. The models are validated with standard strategic indicators: RMSE and MAPE. The low values of these two variables indicate that the models are cost-effective.

Keywords: LSTM, LR, Machine Learning.

1. Introduction

A stock market forecast relates to determining a stock of a firm or future values for other financial entities which tend to be traded through exchanges [1]. Effective predictions of stock market rates are essential because they can provide substantial profit [2][1]. The concept of efficient markets shows that already accessible information reflects inventory prices and does not rely on new information and is therefore unexpected [2]. In this study, we proposed a stock market prediction model using the Neural Network [3]–[6]. The technique uses seven different features as training input parameters and provides the inventory as its "closing price" output. Neural networks are widely used for prediction, as they can learn from known instances and detect nonlinear and hidden dependency, even when noise is high in training. The currency is a part tool utilised as a money exchange device in the national budgetary exchange. It has a real impact on and promotes global financial perspectives. Of general trade, the consideration in fiscal principles can be unique, since a basic rule of monetary structures are needed that can be used throughout [4]. It is called Forex outside exchange or generally. Forex was intended as a market for exchanges [3], [7], [8].
An updated plot of the Artificial Bees Colony (ABC) to streamline ANN piles to anticipate numerous arrangements [9]. The ABC variation is the Artificial Bee Colony Differential Evolution, a combination of the unique ABC with two different change approaches used to estimate the currency exchange rate. [7], the study will examine the precision of stochastic ANN, LR models in anticipation of daily exchange rates. [8]. The direct pattern model, the autonomous moving average (MA) model, the artificial neural network and hereditary computations have been used for forecasting the currency rate[10]. The unique normal examination procedures and neural framework approach to foresee the currency debate also offers the consequences of different topological factors on neural frameworks’ accuracy and planning times. Reference [11]. A half-model of neural networks based back spread (Chang Jui-Fang, Kuan Chi-Ming, Lin Yu-Wen) [12] has been constructed using MSE, MAE (RMSE) and Base Back Propagation (GABPN).

These studies apply to LR, machine learning using the Artificial Neural Network (ANN), Short-Term Memory Networks (STM), and the Neural Network (NN), using one and two layer neurons to predict the exchange rate of multidevises [9]–[12]. Reference [15] presents a test of exchange rate dependency. Reference [16] utilizes an ANN model with moving normal exchange rules for checking exchange rate return consistency. Zhang and Hutchinson [17] provide the tick-by-tick experience of CHF/USD. Kuan and Liu use GBP, CAD, DEM, JPY, CHF to estimate both feedforward and intermittent neural networks versus USD. Wu [16] contrasts Taiwan/USD exchange rates with neural network and ARIMA models. Hann and Steurer [18] examine the neural network with a straight USD/DEM model. Episcopos and Davis [17] investigate the question of day-to-day forecasted returns depending on 5 Canadian exchange rates using ANNs and EGARCH. Reference [19] offers the use in the form of exchange rates of repeating neural networks. Additional models used in exchange rates by ANN incorporate [20] Neural networks are more suitable for problems with optimization. It is therefore utilised to optimise the neural network parameters for more accurate predictions. A programming language must be linked with special tools that enable the work to be carried out, whether it be data modelling or picture analysis.

Typical ML algorithms work well for numerous tasks, but show a decrease in the ability to simplify when problems are being measured. For more composite tasks, such as object identification, standard techniques tend to flop by taking the sophisticated functions embedded in the large space. In that situation, it is disappointing that the advanced stages of abstraction shortens the prediction process rather to relying on raw data or pixel values. This problem has long been avoided by separating the task into two main procedures. The initial procedure involves the construction of an intelligent extractor feature. This extractor works really hard for a certain task and requires skilled understanding. Then the informal training component of the classifier on the mined features can be achieved. ML is able to determine the predictive time series data.

As the rapid improvement of the economy and progress, currency markets have become a key part of our existence. The standard example amazingly anticipates the major aims of the open subject [3]. The offer exchange association gauge also shows the attempt to take a look later and an illustration of an exchange of protections. From late, several explorers are implementing machine learning calculations for expectations supplied by the exchange system. It contains many techniques to the Artificial Neural Network (ANN). The main objectives of this study are:

- To extract important features in KSE dataset
- To apply machine learning forecasting Algorithms for the prediction of stock price
- To apply comparative analysis between different algorithms

## 2. Related Work

Khan et al[13] purposed another strategy for stock market prediction utilizing machine learning. They use calculations to investigate the impact of web-based media and monetary news information on stock market figure precision more than ten days. The arbitrary woods more tasteful are viewed as predictable, and its gathering arrives at the most elevated exactness of 83.22 percent. Shunron et al [14] purposed another technique for Stock Market Using Machine Learning. they proposes another prediction calculation that utilizes LR to figure the following day value pattern by abusing the worldly relationship between worldwide stock markets and diverse monetary things. The NASDAQ has a prediction precision of 74.4 percent, the S&amp;P500 has a prediction exactness of 76%, and the D&H has a prediction exactness of 77.6 percent.

Mustansar et al[15] purposed another strategy for the prediction of stock trade utilizing machine learning. They present the primary exhaustive investigation of information from the Karachi Stock Exchange (KSE) and the Saudi Stock Exchange (SSE) to gauge stock market volume for ten unique organizations. They utilized and contrasted key machine learning calculations with foresee stock trade volume in this investigation. The exactness measurements on the dataset, gathered more than a half year by creping the KSE and SSE sites, were utilized to look at the exhibition of these calculations.

Osman et al [16] purposed another strategy for the stock market a machine learning prediction. Particle swarm advancement (PSO) and least square help vector machine are converged in the proposed calculation (LS-LR). To enhance LS-LR for normal market value prediction, the PSO calculation is utilized. The proposed model depends on the survey of recorded stock information just as mechanical measurements. To forestall overfitting and nearby minima issues and increment prediction exactness,
the PSO calculation chooses the best free boundary blend for LS-LR. The proposed model was approved and contrasted with a fake neural organization with the Liebenberg-Marguardt (LM) calculation utilizing thirteen benchmark monetary datasets. The acquired outcomes uncovered that the proposed model has higher prediction precision and that the PSO calculation can streamline LS-LR.

Sidra et al [17] purposed another technique for stock value prediction utilizing machine learning. The entirety of the relapse models give far reaching impacts on various boundaries. The discoveries show that the LSTM-based univariate model, which utilizes one-week earlier information as contribution to gauge the open estimation of the NIFTY 50 time arrangement for the next week, is the most dependable.

Stock et al [18] purposed another technique for stock market prediction utilizing machine learning. In the wake of applying LR, effectively estimate the stock market pattern for the following month and determined exactness as per estimations.

Zexin et al [19] purposed another strategy for stock value prediction utilizing machine learning. Each article's dataset, segment, model, and results are talked about in this archive. The discoveries of the investigation are introduced utilizing the most generally utilized exhibition measurements: Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Mean Square Error (MSE), exactness, Sharpe proportion, and return rate are largely instances of root mean square mistake.

Iqbal et al [20] purposed another strategy for stock market proficient machine learning. To address the issue of stock market prediction, software engineering is fundamental. To develop an anticipating technique, a few methodologies are utilized. This postulation investigates various cutting edge systems for stock anticipating. The neural organization was discovered to be the best methodology for gauging stock markets at that point. The strategies were tried on PSO (Pakistan State Oil) and S&P500 informational indexes.

Mehak et al [21] purposed another technique for stock market prediction utilizing machine learning. The Single Layer Perception (SLP), Multi-Layer Perception (MLP), Radial Basis Function (RBF), and Linear Regression (LR) machine learning procedures are looked at. A significant number of these highlights are researched separately too. When contrasted with different strategies, the MLP calculation worked the best. The oil rate property was found to be the most basic in estimating market results. The discoveries demonstrate that machine learning techniques can be utilized to estimate the productivity of the KSE-100 file.

Kranthi et al [22] purposed another strategy for stock market prediction utilizing machine learning. Their investigation utilizes a machine learning procedure known as Linear Regression (LR) to estimate value esteems for huge and little capitalizations, just as in three separate markets, utilizing standard and regularly updated costs.

Nargis et al [23] purposed another technique for stock market prediction utilizing machine learning. This paper expects to incorporate an investigation of machine learning's utilization of stock market prediction to choose what can be accomplished later on.

Deshmukh et al [24] purposed another strategy for stock market prediction utilizing machine learning. In this task, they utilized machine learning to conjecture different parts of a portfolio or file, like expected opening and shutting costs, file esteem, etc. Accordingly, financial backers and dealers will actually want to settle on better and quicker choices.

Rohit et al [25] purposed another strategy for stock market prediction utilizing machine learning. They examine how Machine Learning can be utilized to anticipate a stock. Machine Learning approaches utilized for prediction incorporate Naive Bay Random Forests, Perceptions, and Linear Regressions. Straight help vector machines are usually perceived as perhaps the most remarkable machine learning strategies for text grouping.

Strader et al [26] purposed another technique for Machine Learning Stock Market. Every class' investigations are examined for run of the mill discoveries, uncommon discoveries, deficiencies, and regions that require further examination. The last segment incorporates an audit of the outcomes just as ideas for additional examination.

Ishita et al [27] purposed another technique for stock market prediction utilizing machine learning. Machine learning, which makes conjectures dependent on the estimations of current stock market lists via preparing on their past qualities, is another improvement in stock market prediction innovation. Machine learning utilizes a few models to make right predictions. They center on market valuation assessment utilizing relapse and LSTM-based machine learning. Open, close, medium, high, and volume are for the most part factors to recall.

3. Methodology

For this research we have used Karachi Stock Exchange Dataset for the last 10 years. Historical data has been collected from Kaggle, an open source website which contains 8 features. Below table shows the complete description of features of dataset:

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timestamp</td>
<td>The time at which each instance or entry has been collected from stock market</td>
<td>It can be in the form of dd/mm/yyyy</td>
</tr>
<tr>
<td>Open Price</td>
<td>Open price for each day according to each</td>
<td>Million USD</td>
</tr>
</tbody>
</table>

Table 1. Features Description
timestamp

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>High Price</td>
<td>Highest price on that day in which data has been collected.</td>
<td>Million USD</td>
</tr>
<tr>
<td>Low Price</td>
<td>Lowest price on that day in which data has been collected.</td>
<td>Million USD</td>
</tr>
<tr>
<td>Close Price</td>
<td>Final price of stock on that day in which data has been collected.</td>
<td>Million USD</td>
</tr>
<tr>
<td>Volume of Stock</td>
<td>Turnovers in the price of stock</td>
<td>Million USD</td>
</tr>
<tr>
<td>Volume of Currency</td>
<td>Turnovers in the price of exchange rates of currencies</td>
<td>Million USD</td>
</tr>
<tr>
<td>Weighted Price</td>
<td>The average of the shared prices of all companies invested in stock</td>
<td>Million USD</td>
</tr>
</tbody>
</table>

The removal of unwanted data from a dataset is called Preprocessing. For this purpose, we have normalized some features, removed outliers and significant attributed analysis has been also conducted.

**Linear Regression Prediction**

Linear regression is the investigation of two different variables in order to define a single relationship and is a valuable measure of technical and quantitative market analysis. Traders can trace stock movements on an ordinary distribution – a bell curve – to detect if a stock is overbought or over-sold.

**Time Series Prediction**

Time series prediction is a strategy for predicting occurrences over a period of time. The strategies anticipate future events by evaluating the tendencies of the past, assuming that future trends resemble previous trends

**Long-Short Term Memory**

Using the Long Stock Predicting Stock Prices (LSTM) model. In sequence prediction issues, LSTMs are highly powerful because they can store past knowledge. This is relevant in our scenario, because the preceding stock price is essential for forecasting its future price.
Karachi Stock Exchange Price Prediction using Machine Learning Regression Techniques

Figure 3. Flowchart for LSTM Prediction model for KSE Market

Features
The project uses the following characteristics:
1. Opening price: this refers to the price used for initial exchange security on the opening day.
2. High prices: the highest stock exchange price achieved for a specific day. It frequently exceeds the final price or opening price.
3. Low Price: The lowest stock exchange price achieved for a certain day. The current low is usually less than the starting or ending price.
4. Simple moving average of 10 days: this is the average closing price of each stock for the last 10 days.
5. Simple average movement over the last 50 days: this is the average closing price of every stock for the last 50 days.
6. Exponential 10-day moving average: this is the exponential average over the last ten days of each stock closing price
7. Exponential average movement over 50 days: This is the exponential average price of any stock for the last 50 days. The exponential moving average (EMA) refers to a type which gives more weight to recent prices to make them more responsive as new information enters.

K-Fold CV is where a given data set is split into a K number of sections/folds where each fold is used as a testing set at some point. Let’s take the scenario of 3-Fold cross validation (K=3). Here, the data set is split into 3 folds. In the first iteration, the first fold is used to test the model and the rest are used to train the model. In the second iteration, 2nd fold is used as the testing set while the rest serve as the training set. This process is repeated until each fold of the 3 folds have been used as the testing set.

$$Cv^k = \frac{1}{k} \sum_{i=1}^{k} \text{MSE}_i \tag{3.11}$$

where $k = 3$

4. Results
In this section we have presented the output of the artificial neural network, long and short term memory, linear regression and convolutional neural network for predicting the future values for training for different number of years.

Linear Regression
From Sklearn library we have imported the linear regression model to train on the training data and to evaluate the model on testing values. After training the model on KSE Dataset we have applied linear regression algorithm for prediction of the KSE up low close and open values of each instance.

<table>
<thead>
<tr>
<th>Actual</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>67511</td>
<td>173.600004</td>
</tr>
<tr>
<td>675608</td>
<td>171.919998</td>
</tr>
<tr>
<td>676105</td>
<td>172.000000</td>
</tr>
<tr>
<td>676602</td>
<td>187.789993</td>
</tr>
<tr>
<td>677099</td>
<td>187.029999</td>
</tr>
</tbody>
</table>

Figure above shows the predicted and actual values of KSE market of stock exchange. We have evaluated the linear regression model on the basis of accuracy score i.e. 0.993.
**Figure 5. Prediction of Close price using Linear Regression**

**Artificial Neural Networks**
From Sklearn library we have imported the ANN model to train on the training data and to evaluate the model on testing values.

After training the model on KSE Dataset we have applied ANN algorithm for prediction of the KSE up low close and open values of each instance.

<table>
<thead>
<tr>
<th>Actual</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>675111</td>
<td>173.660004</td>
</tr>
<tr>
<td>675608</td>
<td>171.919998</td>
</tr>
<tr>
<td>676105</td>
<td>172.000000</td>
</tr>
<tr>
<td>676602</td>
<td>187.789993</td>
</tr>
<tr>
<td>677099</td>
<td>187.029999</td>
</tr>
</tbody>
</table>

Figure above shows the predicted and actual values of KSE market of stock exchange. We have evaluated the ANN Model on the basis of accuracy score i.e. 0.987.

**Figure 6. Prediction of Close price using ANN**

**Convolutional Neural Networks**
From Sklearn library we have imported the CNN model to train on the training data and to evaluate the model on testing values.

After training the model on KSE Dataset we have applied CNN algorithm for prediction of the KSE up low close and open values of each instance.

<table>
<thead>
<tr>
<th>Actual</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>675111</td>
<td>173.660004</td>
</tr>
<tr>
<td>675608</td>
<td>171.919998</td>
</tr>
<tr>
<td>676105</td>
<td>172.000000</td>
</tr>
<tr>
<td>676602</td>
<td>187.789993</td>
</tr>
<tr>
<td>677099</td>
<td>187.029999</td>
</tr>
</tbody>
</table>

Figure above shows the predicted and actual values of KSE market of stock exchange. We have evaluated the CNN Model on the basis of accuracy score i.e. 0.990.
Karachi Stock Exchange Price Prediction using Machine Learning Regression Techniques

Figure 7. Prediction of Close price using CNN

Figure 8. Comparison of Accuracies between ML Models for the prediction of Close Price

Now we convert data into time series ans apply ML Models

Figure 9. Data Converted to Time-series

Data has been converted to time series using minmaxscalar function

Prediction using LSTM

We have evaluated the CNN Model on the basis of accuracy score i.e. 0.85.

Figure 10. Prediction of Close price using LSTM

Prediction using Artificial Neural Networks

From Sklearn library we have imported the ANN model to train on the training data and to evaluate the model on testing values.

After training the model on KSE Dataset we have applied ANN algorithm for prediction of the KSE up low close and open values of each instance.

Figure shows the predicted and actual values of KSE market of stock exchange

We have evaluated the ANN Model on the basis of accuracy score i.e. 0.777.
Prediction using Convolutional Neural Networks
From Sklearn library we have imported the CNN model to train on the training data and to evaluate the model on testing values. After training the model on KSE Dataset we have applied CNN algorithm for prediction of the KSE up low close and open values of each instance.

Prediction above shows the predicted and actual values of KSE market of stock exchange

We have evaluated the CNN Model on the basis of accuracy score i.e. 0.86

Figure 11. Prediction of Close price using ANN

Figure 12. Prediction of Close price using CNN

Figure 13. Comparison of Accuracies between ML Models for the prediction of Close Price

5. Conclusions

Predicting stock market returns is a challenging task since stock prices change quickly and depend on a range of factors that produce convoluted patterns. Nevertheless, the history data on the company's website is restricted to a number of characteristics, including high and low stock prices, stock price opening and closing, stock traded volumes and similar information, which is not enough. New variables were produced by combining data from the present variables to improve the accuracy of the expected value. The ANN model is utilised for the next day's closing price estimate of the stock and the RF model is also employed for a comparison study. The RMSE, MAPE, and MBE comparison studies reveal clearly that ANN exceeds LR in terms of stock price prediction. The
best values produced by the ANN model are RMSE (0.42), MAPE (0.77) and MBE, according to the findings (0.013). In order to obtain better financial sector results, in the future, deep learning models may be created that take into account financial news items and financial measurements such as a closing price, traded volume, profit and loss statements, etc.

**References**


