Artificial Intelligence Review in Stock Markets

Ujjwal Aggarwal[1], Anmol Saxena[2], Sarvjeet Herald[3]

[1]Research Student, Department of AI Systems, RoboGenius, Gurgaon, India
[2]Research Assistant, Department of AI Systems, RoboGenius, Gurgaon, India
[3]Professor, Department of AI Systems, RoboGenius, Gurgaon, India

Abstract: **Practice of Investing in stocks and shares of a company is widely popular for achieving higher returns. But success required high skills and expert domain knowledge. Digital technologies, especially automation and artificial intelligence, are challenging this notion to allow novice investors to make informed investment decisions. In this article, we review the adoption and accuracy of these modern approaches to help predict future prices with greater accuracy in the highly fluctuation market and in the presence of social media. Since digital data plays an important role in the prediction, we also discuss the types of data and how they are handled. In the context of more real world applications, we understand that the transitions in skillset and roles of the stakeholders in the stock market needs to be embraced and leveraged for success.**

Keywords: **Artificial Intelligence, Machine Learning, Neural Networks, Prediction & Forecasting, Stock Market**

1. Introduction

Investing in stock markets is gaining popularity among individuals for achieving higher return on their investments. But many are aspiring investors who are novice and not sure when and where to invest, how long to retain or even sell stocks for maximum gains. In this avenue, Artificial Intelligence (AI) applications are gaining popularity among individual investors for making informed decision. Based on the historical analysis of performance, AI applications uncover hidden patterns in the data, evaluate risks of investing and provide predictions on future performances to the investors. More recently, these techniques are also incorporating sentiments of people towards a company on social media. In this article, we review some of the recent AI algorithms and approaches used for making informed investment decision along with popular commercial applications leveraging it. We also discuss the transitions in the skillset of the stakeholders in trading.

2. Background

Dutch East India Company was the first company to issue its shares to general public, the process was manual. As sale and purchase of shares became popular among the people, there emerged a growing need to understand the future performances of the investment. Early practices involved studying the growth of the company by professionals who would advise people to invest. However, this process was not always reliable. Trend shifted towards data mining techniques, evolved into building intelligent systems which could automate laborious task of humans and more recently is directed towards transferring human expertise (i.e., skills and experience) to deliver knowledge, as an advisor would do and building upon by also considering real time factors across the regions.

The first report on AI in the Finance sector was Théorie de la Spéculation (i.e., the Theory of Speculation) written by Louis Bachelier. This work immensely contributed to the automation of stock market and trading. AI has been pulled to change in service modes such as software providing facility to take decision by analysing various aspects that can help to predict things with high accuracy [1]. It is extended further to provide output without any specific training i.e. there is no need to give data to teach systems as it will be able to learn itself. This approach which is popularly known as unsupervised learning tends to be more efficient in prediction, than the supervised learning approaches in the stock market. AI systems, such as deep mind, can manage, store and access the required information.

There are broadly two ways to find the future stock price probability i.e. daily prediction and monthly prediction [1]. In daily prediction, both historical data and trending news of company are used for prediction. In monthly predictions only the historical data is used. Usually the two are combined and the result of monthly prediction model is used to test the result of the daily prediction model. Factors that would affect stock market movement are, for example, good understanding about company, business, competitive landscape, company's dependencies and client and supplier type [6]. Investors, who want to invest in some company, check all these criteria's that can affect stock price. If we talk about any particular company, factors usually considered are, for example, History which repeats sometimes, randomization of market and people's rational behaviour that follows the market. Other features that play an important role in stock market prediction are:

- Correlated Assets
- Technical Indicators
- Fundamental Analysis
- Autoregressive Integrated Moving Average (ARIMA)
- Stacked Autoencoders
- Deep Unsupervised Learning

International Journal of Research in Engineering, Science and Management
Volume-2, Issue-11, November-2019
www.ijresm.com | ISSN (Online): 2581-5792

93

3. Data analysis approaches

Data that can be used to make predictions could be historical time series data or more recently emerging is the sentiments of the people towards a company or its performance. This market perception among the public can also pay critical role in prediction accuracy. In this section, we discuss the time series and the sentiment data analysis approaches.

A. Time Series Data Analysis Approach

Time Series Data widely used in various domains. Data that contains date and time along with other parameters is called Time Series Data. This data is broadly divided into two parts i.e., Univariate Time Series and Multivariate Time Series Data. ARIMA model is commonly used for the former. Multivariate Time Series Data requires verification of some significance to fit in the requirements of model. The main thing which necessarily follows is that the data should not contain any kind of trend or seasonality as these can affect prediction accuracy. Data preprocessing stage uses various techniques to exclude these from the data. It is widely popular with Indian Stock Markets for predicting the future value of the stock. Still, one can expect a deviation of ±5% mean percentage error for Nifty and other stock exchanges [8].

Daily generated data is recorded and based on this, prediction of stock price is done by visualizing patterns in the dataset and making histograms. In order to measure accuracy and future trend analysis, comparisons with the results from historical data also plays a major role [10]. Big data concepts are used to store, retrieve and process huge amounts of data on a single platform including translation from unstructured to a structured dataset [18]. Data Deep Learning techniques are gaining popularity in this field, especially to work with unstructured data for providing good results with unsupervised learning approach.

B. Sentiment Data Analysis Approach

In the Indian Stock Market, there are two main indicators which work for the price prediction i.e. BSE and NSE [6]. Sensex is used in Bombay Stock Exchange and Nifty is used in National Stock Exchange companies. The growing popularity of Social Media where public freely post their views about companies and business leaders can also affect future stock performances. News from Channels, blogs, twitter, facebook, etc. are also being included as socio-economic parameters aiding users in making informed decisions [18].

Still at a young age, techniques are required to differentiate between a fake rumour and an evidence backed perception. Opinions on web and social media by experts and general public can act as reviews on a product. Deriving context of a sentence, positive and negative, from a tweet of few words for example is very challenging for intelligent algorithms at this moment [7]. News from trustworthy publishers on the other hand contain facts that are much preferred for Sentiment Data Analysis. Also, debate frequently revolves around, to what extent the history of stock behavior is to be used for making future price predictions.

In stock market, risk is very important factor to measure. All the things go around risk factor for every stock price in market. Analyzing risk factors allow better decisions when investing money based on price prediction algorithms. Geometric Dispersion Theory (GDT) for example, is used to develop models which works on evaluating risk. GDT model satisfy convexity, non-negativity and asymmetry whereas the other models normally ignore these. This generally gives higher accuracy when measuring risks and helps in improving the accuracy whilst decreasing the risk factor.

Popular data sources for building models are, for example, S&P Financial data, Yahoo Finance data and News sentiment data [13]. Combination of Time Series Data with the sentiments could yield better results as fluctuations may occur when news arrives about a company such as replacement of CEO, product launches, mergers and acquisitions, etc. Algorithms which can find the correlation between these peaks, for example, news coupled with developed perception, could further improve accuracy of prediction of future share prices. Relying only on the historical data is useful when no new news of the company is published [17].

C. Artificial Intelligence Algorithms

Artificial Intelligence Algorithms are provided with wealth information as discussed earlier to provide accurate prediction results. These are also used to continuously track stock prices, analyze and operate on data to find useful information, check whether data is complete or incomplete by drawing patterns from the Real Time Data, store the data, measure to find the correlation between previous stock price and find conditions that make effect on the valuation of companies. More importantly, the objective is to transfer expert experiences and skills to make AI systems achieve a better result. Techniques which are currently used for making futures predictions are discussed in this section.

SVM, a classification and regression algorithm is good for extracting meaningful knowledge from unstructured and semi-structured large datasets. It minimizes the risk of over-fitting. This supervised learning model, works with the concept of hyperplanes for making predictions with high accuracy in the case of stock markets. All the training samples are plotted in a two-dimensional plane and these are assigned class levels with the hyperplane that classify these sample points based on their characteristics. This hyperplane includes some properties such as regularization parameter, gamma value and kernel to help classify data samples. Regularization parameter is used to decide the extent of data misclassification, gamma is for data influence limit.

Long Short-Term Memory (LSTM), part of Recurrent Neural Network (RNN), is used to generate output at every iteration. This output works as input for next iteration, learning from its errors, which are generated in each epoch and modify the output by adjusting the weights of network. This approach efficiently works with Time Series Data and provides good results [2]. Generally, all Neural Network based approaches give good

International Journal of Research in Engineering, Science and Management
Volume-2, Issue-11, November-2019
www.ijresm.com | ISSN (Online): 2581-5792

94

results with high accuracy. For example, data is collected from Shanghai Stock Exchange Composite (SSEC) index analyzed algorithms such as Neural Network, Single-factorization Machine based Neural Network, Multi-Factorization Machine based Neural Network, Empirical Mode Decomposition and Factorization Machine based Neural Network (EMD2FNN), Single-EMD2FNN and Multi-EMD2FNN. In all these, multi-EMD2FNN showed highest accuracy when predicting closing prices on daily basis [3].

Convolutional Neural Network (CNN), for example, are popular, to capture motifs such as fluctuation and measure the situation of the company for every ups and downs from the graph of stock time series dataset. Nine stocks (combination of volatility and capitalizations) and CNX NIFTY50 are, for example, considered for future value prediction in stock market. Hare back propagation multilayered Neural Network (NN) is used, trained by Levenberg Marquardt Algorithm (LMA) [4]. For back propagation multilayered NN, 10 neurons or nodes are used in hidden layer. LMA was used to solve curve fitting problem in data to efficiently find solution for all networks as large cap, mid cap and small cap. This study included stock data before and after demonetization which had impacted stock markets. This model also gives high accuracy for closing price in any day. However, most of the platforms of stock market use Logistic Regression to predict future price rather than Linear Regression but Artificial Neural Network (ANN) is by far more suitable in prediction and classification of stock into three main categories like buy, hold and sell price [5]. Table below summarizes the predicting accuracy of the algorithms discussed in this section.

Table 1
Summary of Accuracy of Algorithms

Algorithms	Accuracy
Linear Regression	81%
SVM-RBF	98%
Random Forest	99%
Gradient Boosting	99%
Single-EMD2FNN	84%
Multi-EMD2FNN	99%
Motif Extraction and CNN	92%
RNN and LSTM	97.5%

4. Discussion & Conclusion

There are various AI-powered digital platforms that provide informed decision-making tools to novice and expert investors, for example, samco, moneycontrol, valuenotes, economics times, livemint, cofagr, icicdirect, zerodha, allianceresearch, stockmarkettips, stockwarepro, geojit, etc. These platforms provide recommendations by recognizing activities such as what type of shares are visited most, how much time anyone spend on which share, how frequently any stock visited in a day. They also provide information about previous trend and send alert to users about high returning stock at the end of the day.

Money control, for example, is the most visited site in India. It provides reports related to capital gain, performance, profit & loss, cash flow, tax, etc. along with tools to analyse investor's

investment, provide filtered news and corporate actions and recommendations from experts of stock market. Valuenotes on the other hand mainly provides strong leads from the editorial content along with reports of company researches that can aid investors to choose right stock to invest.

Investing, for example, provides all the information at one place to allow fundamental and technical stock analysis. One of the best tools of this site is stock screener by which stocks are shortlisted by applying some indicators. Economics Times and Livemint Market provide latest news throughout the day with an option to receive updates and briefs. Each is using AI to power their systems as discussed in this article and is changing the roles and skillsets of different stakeholders in the stock market trading.

Stakeholders in the stock market such as companies, investors, companies, brokers, depositories, government, trader and stock exchange are widely adopting AI for transparency and higher growth. While government is responsible for ensuring that fair practices are adopted by the stakeholders and regulating the market, it should also consider curating and making more data open, especially in the structured forms within the digital repositories. Options for Single Sign-On should also be considered by the government. Depositories consider blockchain to preserve privacy whilst ensuring auditability in the chain of transactions. Companies should be at the forefront of maintaining active data pipelines to introduce real time audits. Brokers are already implementing AI platforms, but the future holds delegating entire lifecycle to the AI and minimize human advisor involvement. All to some degree will require knowledge of stages of AI project lifecycle, especially data curation and cleaning so ensure that the algorithms learn from high quality data. These algorithms also require becoming explainable, so that thinking process adopted by the AI is understandable by traders and individual investors. Having a skillset in AI is essential for success of the future investors.

References

[1] Y. Pan, "Heading toward Artificial Intelligence 2.0", Engineering, vol. 2, no. 4, pp. 409-413, 2016.

[2] D. Karmiani, R. Kazi, A. Nambisan, A. Shah and V. Kamble, "Comparison of Predictive Algorithms: Backpropagation, SVM, LSTM and Kalman Filter for Stock Market", in *Amity International Conference on Artificial Intelligence (AICAI)*, Dubai, United Arab Emirates, 2019, pp. 288-234.

[3] F. Zhou, H. Zhou, Z. Yang and L. Yang, "EMD2FNN: A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction". *Expert Systems with Applications*, 115, pp.136-151, 2019

[4] S. Chopra, D. Yadav and A. Chopra, "Artificial Neural Networks Based Indian Stock Market Price Prediction: Before and After Demonetization", *International Journal of Swarm Intelligence and Evolutionary Computation*, vol. 8, no. 1, 2019.

[5] S. Agrawal, D. Thakkar, D. Soni, K. Bhimani and C. Patel, "Stock Market Prediction using Machine Learning Techniques", *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 5, no. 2, pp. 1099-1103, 2019.

[6] A. Bhardwaj, Y. Narayan, Vanraj, Pawan and M. Dutta, "Sentiment Analysis for Indian Stock Market Prediction Using Sensex and Nifty", *Procedia Computer Science*, vol. 70, pp. 85-91, 2015.

[7] S. Idrees, M. Alam and P. Agarwal, "A Prediction Approach for Stock Market Volatility Based on Time Series Data". *IEEE Access*, 7, pp.17287-17298, 2015.

[8] W. Souma, I. Vodenska and H. Aoyama, "Enhanced news sentiment analysis using deep learning methods". *Journal of Computational Social Science*, 2(1), pp.33-46, 2019.

[9] P. Kohli, S. Zargar, S. Arora and P. Gupta, "Stock Prediction Using Machine Learning Algorithms", in *Applications of Artificial Intelligence Techniques in Engineering*, SIGMA, 2018

[10] M. Komaki and B. Malakooti, "Geometric Dispersion Theory for Portfolio Accurate Out-of-Sample Predictions of Stock Market". *SSRN Electronic Journal*, 2019.

[11] R. Deaves, J. Lei and M. Schrrder, "Forecaster Overconfidence and Market Survey Performance". *SSRN Electronic Journal*, 2015.

[12] M. Asghar, F. Rahman, F. Kundi and S. Ahmad, "Development of stock market trend prediction system using multiple regression", *Computational and Mathematical Organization Theory*, vol. 25, no. 3, pp. 271-301, 2019.

[13] M. Asghar, F. Rahman, F. Kundi and S. Ahmad, *Computational and Mathematical Organization Theory*. Springer US, pp. 271-301, 2019.

[14] S. Aziz and M. Dowling, "Machine Learning and AI for Risk Management", in *Disrupting Finance*, 2018.

[15] S. Herald, S. Clarke and B. Christianson, "A Non-Transitive Trust Model for Key Distribution", *Journal of Information Assurance and Security*, 2010.

[16] E. Fama, "Random Walks in Stock Market Prices", *Financial Analysts Journal*, vol. 51, no. 1, pp. 75-80, 1995.

[17] J. Bollen, H. Mao and X. Zeng, "Twitter mood predicts the stock market", *Journal of Computational Science*, vol. 2, no. 1, pp. 1-8, 2011.

[18] N. Arora and P. M, "Financial Analysis: Stock Market Prediction Using Deep Learning Algorithms", *SSRN Electronic Journal*, 2019.