International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/IJSRCSEIT

Micro-Financial Analysis And A Schematic View of Ai, Machine Learning and Big Data Analytics On Financial Markets

Kola Vasista
Financial Student Business Consultant, Temple University Small Business Development Center, Philadelphia, PA, USA

Article Info
Volume 7, Issue 3
Page Number: 503-508

Publication Issue :
May-June-2021

Article History
Accepted : 01 June 2021
Published : 15 June 2021

ABSTRACT

Forecast of financial time series is of the most critical issues in making financial decisions. In this regard, the Tehran Stock Exchange is of great importance for domestic and international financial markets. Based on the economic events and data of the past, it provides a profitable method for the future. This research is related to the usefulness of different machine learning methods in forecasting time series on financial markets. The main issue in this field is that economic managers and scientific society are still longing for more accurate forecasting algorithms. Fulfilling this request leads to an increase in forecasting quality and, therefore, more profitability and efficiency. In this paper, while we introduce the most efficient features, we will show how valuable results could be achieved by the use of a financial time series technical variables that exist on the Tehran stock market. This paper provides an analysis and schematic view of AI, ML and BDA on financial markets.

Index Terms : Machine Learning, Artificial Intelligence, Financial Markets, Big Data Analytics

I. INTRODUCTION

Governments, shareholders and organizations managers need to forecast exports and imports, recognizing the market situation, and the organization's future, respectively [4]. Managers take personal and professional decisions based on forecasting situations in the future. In many cases, forecasting is based on present and past. They, in fact, try to link two or more valuables so that they can be used for forecasting. On the subject of forecasting at financial markets, it is a priority to analyze data [5]. Two standard methods are used for analyzing financial series: technical analysis, basic analysis.

Basic analysis is an analyzing method that works on key figures and information on companies' financial statements, the country's macro-economy figures, and influencing agents on different economies. Therefore, selling or buying shares or any kind of asset is done after the information about the aforementioned items has been assessed [6]. Technical analysis is a method for forecasting markets based on assessing and studying price history and turnover on markets. Technical analysis is of mathematical formulas related to prices and turnover data that are used for modelling some aspects of shares prices for indices [7]. Technical analysis includes studying and assessing different indices, charts, and patterns that illustrate market trends and different shares status for investors. Machine learning methods, inspired by patternology and computational learning theory, investigate the studying and constructing algorithms that are able to learn and forecast based on data. Such algorithms do not follow a program's instructions yet, they forecast or make decisions based on modelling and sample input data. Machine learning methods are used on computational works where designing and programming exclusive algorithms with suitable function are hard or impossible.

Machine learning methods have a close relationship with computational statistics, and they often overlap with each other. This branch is about forecasting by computers, and has a secure connection with mathematical optimisation that it, in turn, introduces to the system method, theories and functions. Machine learning sometimes merges with data analysing. This subbranch focuses on exploratory analysis of data, and it is known as non-supervised learning. A machine learning method, in the data analysis field, is a method for designing complex models and algorithms used for forecasting. In industry, it is known as predictive analysis. Data and past samples are the first steps to learn the machine. In the field of computer science, theories are put forward in that basics or extracted to learning data, so that software can be developed meeting users' needs. Therefore, software development will be made possible on learning machines so that parts of software system basics that are not extractable through analysing and designing by analyses and designers can be achieved from past data. In learning the scope, this body of extracted basics is called a model. So, the goal is that for the extracted model to be as common as possible so that it can be used with a high degree of precision for data that is not created. Financial markets play an essential role in organizing modern society, socially and economically. Most assets are undoubtedly exchanged through the stock market worldwide, nowadays. National economies are profoundly affected by the stocks' value forecasting function on the stocks market. One of the most critical pieces of information for investors on stock markets is shares price information. Therefore, not only is shares price information challenging but also it is investors' favorite.

However, investing in the stock market is an essential part of the economy. Therefore, forecasting, especially in developing countries like Iran, is very important, to manage the stock market for achieving stable development. That eases decision-making for the stock market executives under the current uncertainties. Investors are also able to forecast shares price or overall index and make logical decisions, accordingly. Considering the importance of the subject, machine-learning algorithms have already presented remarkable functions based on studies. Thus, most papers in shares price field center around intelligent methods, nowadays [4]. Therefore, doing more research in this field seems necessary. Since machine learning methods are empowered to model complicated engineering problems and nonlinear systems, they are known as a suitable method for forecasting share prices. The main objective of this paper is to increase the accuracy of forecasting Iran's stock market share prices as a financial time series using learning algorithms. To achieve that goal, it is vital to select the most suitable features, main parameters' settings, and effective ones in chosen algorithms.

II. BACKGROUND AND DEFINITIONS

Researchers in computer science and statistics have developed advanced techniques to obtain insights from large disparate data sets. Data may be of different types, from different sources, and of different quality (structured and unstructured data). These techniques can leverage the ability of computers to perform tasks, such as recognising images and processing natural languages, by learning from experience. The application of computational tools to address tasks traditionally requiring human sophistication is broadly termed 'artificial intelligence' (AI). As a field, AI has existed for many years. However, recent increases in computing power coupled with increases in the availability and quantity of data have resulted in a resurgence of interest in potential applications of artificial intelligence.5 These applications are already being used to diagnose diseases, translate languages, and drive cars; and they are increasingly being used in the financial sector as well.

There are many terms that are used in describing this field, so some definitions are needed before proceeding. 'Big data' is a term for which there is no single, consistent definition, but the term is used broadly to describe the storage and analysis of large and/or complicated data sets using a variety of techniques including AI. Analysis of such large and complicated datasets is often called 'big data analytics.' A key feature of the complexity relevant in big data sets analytics often relates to the amount of unstructured or semi-structured data contained in the datasets.

This report defines AI as the theory and development of computer systems able to perform tasks that traditionally have required human intelligence. AI is a broad field, of which 'machine learning' is a sub-category. Machine learning may be defined as a method of designing a sequence of actions to solve a problem, known as algorithms, which optimise automatically through experience and with limited or no human intervention. These techniques can be used to find patterns in large amounts of data (big data analytics) from increasingly diverse and innovative sources. Figure 1 gives an overview.

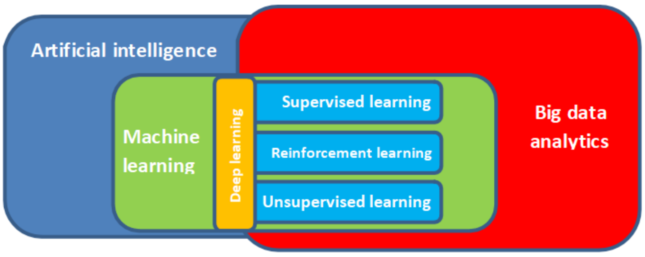

Figure 1 : A schematic view of AI, machine learning and big data analytics

Many machine learning tools build on statistical methods that are familiar to most researchers. These include extending linear regression models to deal with potentially millions of inputs, or using statistical techniques to summarise a large dataset for easy visualisation. Yet machine learning frameworks are inherently more flexible; patterns detected by machine learning algorithms are not constrained to the linear relationships that tend to dominate economic and financial analysis. In general, machine learning deals with (automated) optimisation, prediction, and categorisation, not with causal inference.10 In other words, classifying whether the debt of a company will be investment grade or high yield one year from now could be done with machine learning. However, determining what factors have driven the level of bond yields would likely not be done using machine learning.

There are several categories of machine learning algorithms. These categories vary according to the level of human intervention required in labelling the data:

- In 'supervised learning', the algorithm is fed a set of 'training' data that contains labels on some portion of the observations. For instance, a data set of transactions may contain labels on some data points identifying those that are fraudulent and those that are not fraudulent. The algorithm will 'learn' a general rule of classification that it will use to predict the labels for the remaining observations in the data set.

- 'Unsupervised learning' refers to situations where the data provided to the algorithm does not contain labels. The algorithm is asked to detect patterns in the data by identifying clusters of observations that depend on similar underlying characteristics. For example, an unsupervised machine learning algorithm could be set up to look for securities that have characteristics similar to an illiquid security that is hard to price. If it finds an appropriate cluster for the illiquid security, pricing of other securities in the cluster can be used to help price the illiquid security.
- 'Reinforcement learning' falls in between supervised and unsupervised learning. In this case, the algorithm is fed an unlabelled set of data, chooses an action for each data point, and receives feedback (perhaps from a human) that helps the algorithm learn. For instance, reinforcement learning can be used in robotics, game theory, and self-driving cars.
- 'Deep learning' is a form of machine learning that uses algorithms that work in 'layers' inspired by the structure and function of the brain. Deep learning algorithms, whose structure are called artificial neural networks, can be used for supervised, unsupervised, or reinforcement learning.

Recently, deep learning has led to remarkable results in diverse fields, such as image recognition and natural language processing (NLP). Deep learning algorithms are capable of discovering generalisable concepts, such as encoding the concept of a 'car' from a series of images. An investor might deploy an algorithm that recognises cars to count the number of cars in a retail parking lot from a satellite image in order to infer a likely store sales figure for a particular period. NLP allows computers to 'read' and produce written text or, when combined with voice recognition, to read and produce spoken language. This allows firms to automate financial service functions previously requiring manual intervention.

Machine learning can be applied to different types of problems, such as classification or regression analysis. Classification algorithms, which are far more frequently deployed in practice, group observations into a finite number of categories. Classification algorithms are probability-based, meaning that the outcome is the category for which it finds the highest probability that it belongs to. An example might be to automatically read a sell-side report and label it as 'bullish' or 'bearish' with some probability, or estimate an unrated company's initial credit rating. Regression algorithms, in contrast, estimate the outcome of problems that have an infinite number of solutions (continuous set of possible outcomes). This outcome can be accompanied with a confidence interval. Regression algorithms can be used for the pricing of options. Regression algorithms can also be used as one intermediate step of classification algorithm.

It is important to note what machine learning cannot do, such as determining causality. Generally speaking, machine learning algorithms are used to identify patterns that are correlated with other events or patterns. The patterns that machine learning identifies are merely correlations, some of which are unrecognisable to the human eye. However, AI and machine learning applications are being used increasingly by economists and others to help understand complex relationships, along with other tools and domain expertise.

Many machine learning techniques are hardly new. Indeed, neural networks, the base concept for deep learning, were first developed in the 1960s.

However after an initial burst of excitement, machine learning and AI failed to live up to their promises and funding dissipated for over a decade, in part because of the lack of sufficient computing power and data. There was renewed funding and interest in applications in the 1980's, during which many of the research concepts were developed for later breakthroughs.

III. MICRO-FINANCIAL ANALYSIS

From a micro-financial point of view, the application of AI and machine learning to financial services may have an important impact on financial markets, institutions and consumers. In this section, potential changes to incentives and behaviour and how they may affect financial stability, for better or worse, are considered.

Possible effects of AI and machine learning on financial markets

Since AI and machine learning have the potential to substantially enhance the efficiency of information processing, thereby reducing information asymmetries, applications of AI and machine learning have the potential to strengthen the information function of the financial system. The mechanisms whereby this improvement may occur include:

AI and machine learning may enable certain market participants to collect and analyse information on a greater scale. In particular, these tools may help market participants to understand the relationship between the formulation of market prices and various factors, such as in sentiment analysis. This could reduce information asymmetries and thus contribute to the efficiency and stability of markets.

a) AI and machine learning may lower market participants' trading costs. Moreover, AI and machine learning may enable them to adjust their trading and investment strategies in accordance with a changing environment in a swift manner, thus improving price discovery and reducing overall transaction costs in the system.

b) Nonetheless, if many market participants come to use similar AI and machine learning programmes in areas such as credit scoring or financial market activities, the consequent correlated risks may entail financial stability risks. If machine learning-based traders outperform others, this could in the future result in many more traders adopting similar machine learning strategies (even if this may also reduce the profitability of such strategies). While there is no evidence of this occurring to date, this could become relevant with greater adoption of such trading strategies. As with any herding behaviour in the market, this has the potential to amplify financial shocks. Moreover, advanced optimisation techniques and predictable patterns in the behaviour of automated trading strategies could be used by insiders or by cybercriminals to manipulate market prices.

IV. CONCLUSION

Many current providers of AI and machine learning tools in financial services may fall outside the regulatory perimeter or may not be familiar with applicable law and regulation. Where financial institutions rely on third-party providers of AI and machine learning services for critical functions, and rules on outsourcing may not be in place or not be understood, these servicers and providers may not be subject to supervision and oversight. This paper provided an analysis and schematic view of AI, ML and BDA on financial markets.

V. REFERENCES

[1]. Chen Cang. Research on Quantitative Timing Strategies Based on Depth Learning D]. Lanzhou, University of Finance and Economics, 2017.

[2]. Li Xiang. Multifactor Quantitative Stock Selection Scheme Planning Based on XGBoost Algorithm D]. Shanghai. Shanghai Normal University, 2017.

[3]. Jiang Minqi, Liu Jiapeng and Zhang Lu. An improved Stacking framework for stock index prediction by leveraging tree-based ensemble models and deep learning algorithms J]. Physical A: Statistical Mechanics and its Applications, 2020, 541, 122272.

[4]. Kola Vasista, "Foreign Capital Issuance and Participants in the Securities Market", International Journal of Research and Analytical Reviews, VOLUME 2, ISSUE 4, OCT. – DEC. 2015

[5]. Kola Vasista, "A Research Study On Major International Stock Market", International Journal of Research and Analytical Reviews, VOLUME 4, ISSUE 3, JULY – SEPT. 2017

[6]. Kola Vasista, "A Review On The Various Options Available For Investment", International Journal Of Creative Research Thoughts - IJCRT (IJCRT.ORG), Volume 7, Issue 2, April 2019, ISSN: 2320-2882

[7]. Kola Vasista, "A Detailed Study On The Factors Influencing The Price Of a Stock", International Journal of Novel Research and Development, Volume 2, Issue 8, August 2017, ISSN: 2456-4184

[8]. Kola Vasista, "Objectives And Importance Of Capital Markets And The Role Of Financial Institutions", International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN:2349-5162, Vol.2, Issue 9, page no.475-478, September-2015, Available :http://www.jetir.org/papers/JETIR1701762.pdf

[9]. Kola Vasista, "An Overview On Provident Fund, Pension Funds, Pfrda, Insurance Companies And Irda", International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN:2349-5162, Vol.5, Issue 10, page no.284-287, October-2018, Available :http://www.jetir.org/papers/JETIR1810A93.pdf

[10]. Li Hanshen. Quantitative Trading Strategy of Stock Index Futures based on SVM D]. Zhejiang University, Hangzhou, 2016.

[11]. Guo Bin, Zhang Wei and Zhang Yongjie. The five-factor asset pricing model tests for the Chinese stock market J]. Pacific-Basin Finance Journal, 2017, 43: 84-106.

[12]. Dattatray P. Gandhmal, K. Kumar. Systematic analysis and review of stock market prediction techniques J]. Computer Science Review, 2019, 34, 100190.

[13]. Sun Hongquan. Research on quantitative trading D]. University of Technology, Beijing, 2016.

[14]. Chen Zheshi, Li Chunhong and Sun Wenjun. Bitcoin price prediction using machine learning: An approach to sample dimension engineering J]. Journal of Computational and Applied Mathematics, 2020, 365, 112395.

Cite this article as :

Kola Vasista, "Micro-Financial Analysis And A Schematic View of Ai, Machine Learning And Big Data Analytics On Financial Markets", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 3 Issue 7, pp. 503-508, September-October 2018.

Journal URL : https://ijsrcseit.com/CSEIT228316