

Machine Learning in Accounting & Finance: Architecture, Scope & Challenges

Ishan S. Kapoor¹, Sunint Bindra² & Monika Bhatia³

¹ Student, Del Norte HS, San Diego, USA

² Student, Dartmouth College, USA

³ Associate Professor, Department of Management, Gurugram University, Gurugram, USA

Correspondence: Monika Bhatia, Associate Professor, Department of Management, Gurugram University, Gurugram, USA. E-mail: monika.bh1@gmail.com

Received: January 21, 2022

Accepted: February 20, 2022

Online Published: April 5, 2022

doi:10.5539/ijbm.v17n5p13

URL: <https://doi.org/10.5539/ijbm.v17n5p13>

Abstract

Purpose: This paper discusses and presents the importance, scope, and limitations of machine learning in the area of financial decision-making. The purpose of the study is to find out the areas of application of machine/deep learning in the accounting and finance domain and also to identify challenges in adoption.

Design/methodology/approach: The current study is qualitative review-based research, where the systematic approach to reviewing the existing body of literature has been used. This article employs a thoughtful literature review of selected articles in identified journals that were subsequently evaluated through desktop analysis. All papers were selected based on the search in Google scholar. To enhance the quality of research, a scholarly filtration technique was employed. Only papers listed and accepted by the academia were shortlisted. The second criteria were to identify the keywords in the area of interest. The final step included only papers listed in established databases like Google Scholar, SCOPUS & ABDC.

Findings: The findings of the study indicate the importance of machine learning in financial decision-making and prediction. Advanced mathematical models such as unsupervised machine learning techniques have become the need of the hour to model complex non-linear relationships in financial systems, where complex business situations are resulting in the generation of 'Big Data' and 'Alternate Data'. However, there are many challenges in applying ML/DL models in these prediction models especially when the modeling in finance involves behavioral aspects of extremely dynamic customers and markets. The findings further indicate major research trends associated with machine learning in accounting and finance.

Originality/value: This is a novel study in the area of accounting and financial research, which requires considerable attention for interdisciplinary research.

Keywords: machine learning, deep learning, artificial intelligence, accounting research, systematic literature review

1. Introduction

Data is the new oil for the new world economy. With every single passing day, data is becoming more accessible due to technological advancements. Storage and processing of data have never been as easy and economical as is today. This has led to the development of the concept 'Big Data'. Data with high volume is known as big data, which can include both structured and unstructured data. This development has come with new opportunities and challenges. Decision-making in any area becomes more efficient if it is supported by the analysis of big data. Accounting and Finance are no exception to this premise, especially when it comes to predicting the outcome of any business decision. Although decision-making always involves future outcomes, however, the effect of any financial decision could be exceptionally phenomenal on the future of any business. Financial decisions involve day-to-day business operations on the one hand and long-term strategy on the other. A business needs to decide about the pricing of the product on the one hand and the pricing of securities on the other. A business needs to review the business risk as well as the financial risk. All these situations involve financial predictions involving a complex interaction between various datasets. Unfortunately, the existing financial theories are not able to handle such complex decision situations, though they may give some idea. Heaton J. B. et al (2018) commented

in their article "financial prediction problems such as those presented in designing and pricing securities, constructing portfolios, and risk management often involves large data sets with complex data interactions that currently are difficult or impossible to specify in a full economic model". They further say "applying deep learning methods to these problems can produce more useful results than standard methods in finance. In particular, deep learning can detect and exploit interactions in the data that are, at least currently, invisible to any existing financial economic theory". As they say that every new challenge brings a new opportunity as well, thereby, Machine learning techniques came to the rescue of business decision-making involving voluminous and complex data. Dixon M.& Halperin I. (2019) claim that even though machine learning has been prevalent in financial services for over four decades now, only in the past few years its impact has been felt in investment management and trading. Both computational and theoretical developments in machine learning have resulted in the increased practice of machine learning in the finance area. According to Dixon M.& HalperinI (2019) "Initial endorsers have been a variety of trading and hedge funds such as D.E. Shaw and World Qantas they have accepted new machine learning techniques. However, the extent to which Machine Learning can affect quantitative trading is still a subject of debate.

Finance as a discipline has grown significantly and new theories and models keep on evolving. Apparently, these developments do not seem to be sufficient to meet the challenge posed by big data and also that of alternative data. Gone are the days when only the fundamentals of a firm along with macroeconomic indicators were considered adequate for the pricing of financial assets. Now, the information analysts need to consider is not the usual securities pricing like social media, new announcements, unstructured documents like earning announcements, etc. This is known as the use of 'Alternative Data' in the pricing and management of financial assets. Trading firms, asset management companies, fund managers, and other businesses involved directly or indirectly with the pricing of securities are now engaging experts from the area of machine learning for better decision-making about financial assets. These expert uses Natural Language Processing (NLP) to process alternate data and help financial analysts make more accurate and timely decisions.

The research in the area of use and application of machine learning is at a very nascent stage and is scattered as well, this propelled authors to write on this subject. This paper is organized into five sections. The first section talks about the introduction, followed by objectives and research method in sections two and three respectively. The architecture of machine learning models has been discussed in section four. The fifth section discusses how Machine Learning and in particular deep learning can be applied in Finance and accounting whereas, section six talks about the challenges faced while doing so. The next section addresses how to make the application of ML/DL more effective in the area of accounting and finance and the last section, presents the conclusion of the study.

1.1 Objectives of the Study

The overall objective of this study is to carry out a systematic review of the existing literature to find out the areas of application of machine/deep learning in the accounting and finance domain with the following particular objectives:

- To understand and present the architecture of machine learning/deep learning.
- To explore the areas of application of machine/deep learning in the accounting and finance disciplines.
- To identify limitations in the application of machine/deep learning in financial decision-making.

1.2 Research Method

A systematic literature review is an explicit and rigorous method of evaluating and interpreting available studies relevant to a particular topic area or phenomenon of interest by using possible keywords (Bennett et al., 2004). Adopting the standard guidelines firstly, an exploratory search was done by using the expression "machine/deep learning" and financial decisions of all the journals for the field of research (FOR) relating to Accounting and Finance. Thereafter desktop analysis method was followed to draw inferences about the objectives of the study.

1.3 Architecture of Machine Learning Modeling

The world is witnessing tremendous developments in providing solutions to problems through fully realistic modeling for many sectors including accounting & finance. Due to these developments, high-dimensional calibration problems can be solved in real-time. Wagner W. P. (2017) says that original expert systems, which were used to solve complex problems in a specified field, were rule-based programs in languages. However, the developments in the area of artificial intelligence have superseded those classic expert systems. Liu W. et al (2017) concluded that techniques like cognitive computing and augmented human intelligence techniques were used for a short time and now the techniques like deep learning neural networks or reinforcement learning, have

found significant success across industry and applications.

Gogas, P. & Papadimitriou T. (2021) commented in their article that the term Machine Learning (ML) was introduced by Arthur Samuel while working for IBM in 1959, mainly to describe how Artificial Intelligence systems can learn during pattern recognition tasks. P. Domingos (2012) says that "Machine Learning (ML) is a subfield of Artificial Intelligence (AI) that uses statistical techniques that provide computer models with the ability to learn from a dataset, allowing the models to perform specific tasks without explicit programming". Machine learning results in developing a model by training the complex data, which in turn helps in predictions for any complex business situation. Ha, Youngmin (2017) quoted (Alpaydin, 2014) saying that "Machine learning is computer programming to optimize a performance criterion using example data or past experience by using the theory of statistics". Heaton J. B. et al (2018) say that in ML, input data is used for the initial training of a model and this trained model can be used to make predictions. Whereas, Deep Learning (DL) is a subset of Machine Learning, where several layers of abstractions are used to make the same predictions. Dixon et al., 2016 commented in their article that deep learning is a kind of supervised learning method, which uses layers of neural networks. "Gogas, P. & Papadimitriou T. (2021) argued that "in the beginning, ML systems were considered indistinguishable from typical AI systems. But, Independent ML systems outnumber the ML components in AI systems due to the wide of practical applications of Machine Learning which aren't confined to the box of AI Systems. It is interesting to mention here that ML and AI are often used as synonyms due to several reasons, which may create confusion for non-experts.

According to Emerson Sophie et al. (2019) "There are many well-known machine learning networks like SVMs and KNN other than the neural networks (NN). Undoubtedly, artificial neural networks (ANN) have become a key technology in the development of ML". They further say "Bayesian networks are built from probability distributions and use probability laws for prediction and anomaly detection. KNN selects the most similar data points in the training data; this allows the algorithm to classify future data inputs in the same way. Some techniques are better suited to particular tasks than others". Learning is the specification of a neural network which approximates a certain non-linear function on some input space. According to Boyarshinov Victor (2005) "One generally should consider three different facets of designing an algorithm for an artificial intelligence application:

- (i) Constructing a training dataset
- (ii) Developing a training algorithm for discovering patterns
- (iii) Using side information for avoiding overfitting the training data".

Computers can generate highly convoluted and multi-dimensional data, iterate through multiple model architectures and determine the effectiveness of these models (Dhar, 2013). Constructing a training dataset is the first technical step in developing any model followed by developing a training algorithm. We need to follow several steps to construct deep learners. The first step in this direction is to divide the data into three subsets namely testing, training, and validation sets. The training sets determine the weight configuration of the model. The validation set prevents the model from being confined to suitable use for just the training data i.e. it prevents overfitting. Cortes and Vapnik (1995) talked about 'Support Vector Machine' an ML method to classify data. They further say "First, a set of labeled data (training data) is provided to the SVM. During this step, the algorithms perform a mathematical optimization based on the labeled data. The train examples that limit the maximum margin defined by the SVM during the training are called support vectors".

The algorithms used in the application of artificial intelligence belong to three broad categories which are supervised, unsupervised, and reinforcement learning. Windmann Alexander (2020) says that in supervised learning, the model is given labeled data in order to find patterns and make predictions about new, unseen data, whereas unsupervised learning detects patterns without labels.

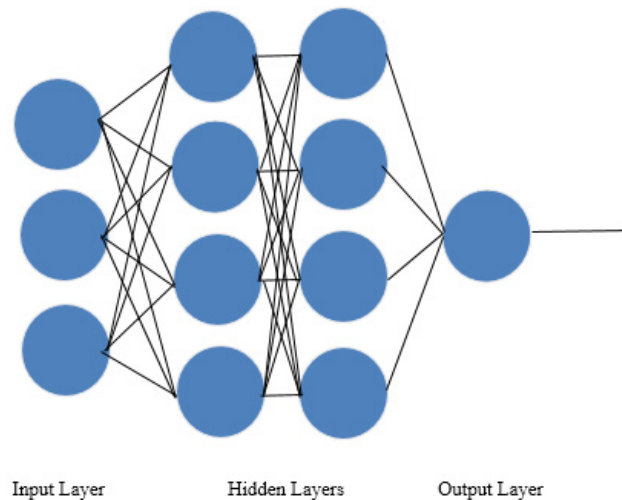


Figure 1. Depicting Architecture of a Deep Learning Model

In reinforcement learning, a model interacts with many factors from the environment to achieve its goals. This interaction can result in rewards or punishment depending on its effect. The ultimate goal of the modeling in reinforcement learning is to maximize the rewards of the program over time by including some constraints through its 'agent'. These machine learning models, thus developed can be successfully used in the accounting and financial sector

2. Machine Learning in Accounting and Finance

As discussed in the first section the emergence of 'Big Data' and 'Alternate Data' has led to new opportunities, where machine learning can be used in the accounting and finance area very effectively. Lindsay, R. K. et al. (1993) mentioned in their article that the very first use of artificial intelligence 'expert systems' can be traced at Stanford way back in the year 1960, though it became popular only in the 1990s. Dixon and Halperin (2019) say "finance has all the ingredients needed to use ML: Big data, the right resources, requirements to predict, and a suitable environment where everyone is trying to gain an edge. But it is also increasingly under scrutiny by regulators and some of the core research activities are resistant to "black-box" methodologies that are typical with "empirical" or "engineering" approaches". Financial modeling has never been as powerful and reliable as is today. This has become more evident after the financial crisis of 2007-08, when the regulatory bodies have re-oriented themselves toward 'data-driven' regulation and monitoring. Aziz, Saqib, and Dowling et al. (2019) commented in their article that M.L. is seemingly the next frontier for someone involved in finance, as it is highly dependent on practical understanding. According to Dhar (2013), several Machine Learning algorithms can be used for the financial modeling of data. Aziz et al. (2019) used the topography approach and studied 5,204 articles published from 1990 to 2018 and provided a comprehensive structure of machine learning application in finance.

Dixon and Halperin (2019) state that both the growing amounts of suitable data and the growth in computation powers of machines point in favor of a boom for Machine Learning. They further say that "Data vendors such as Bloomberg, Thomson Reuters, and Raven Pack are providing processed news sentiment data tailored for systematic trading models. See Guide (2019) and de Prado (2018) for an extensive discussion of machine learning and big data in quantitative investing and trading". Hang et al (2020) say that due to the proliferation of Fin-Tech in the recent past, the application of DL has become very widespread even in the finance and banking sector. Heaton J. B. et al. (2018) argue that deep learning can determine several hidden patterns and trends in data that can't be seen using traditional economic theory.

The Financial Services sector of any economy is considered to be very dynamic due to the behavioral implications of its players especially that of customers and markets. Thereby, the job of analysts becomes very difficult when it comes to modeling the same. Mullainathan, Sendhil, and Jann Spiess (2017) said in their article that the crucial role of ML in finance services is simulating the various elements of the markets such as the customers and products. P. Domingos (2012) says that ML is being used to improve several key areas of the finance industry such as payment transactions, detecting frauds, investment process including return forecasting,

portfolio construction, and risk modeling. Emerson Sophie et al. (2019) say that "recently, there has been a proliferation of ML techniques and growing interest in their applications in finance, where they have been applied to sentiment analysis of news, trend analysis, portfolio optimization, risk modeling among many use cases supporting investment management". Liu et al. (2017) say that many easy to use programming languages such as R, Python and ML focused framework such as Tensor Flow (TF) has found wide commercial application across multiple industries from automated trading systems in the finance industry to the health sector where ML algorithms assist decision making in the fertility treatments. So, there are many ways to apply M.L. techniques in the financial sector which have been discussed in the following section:

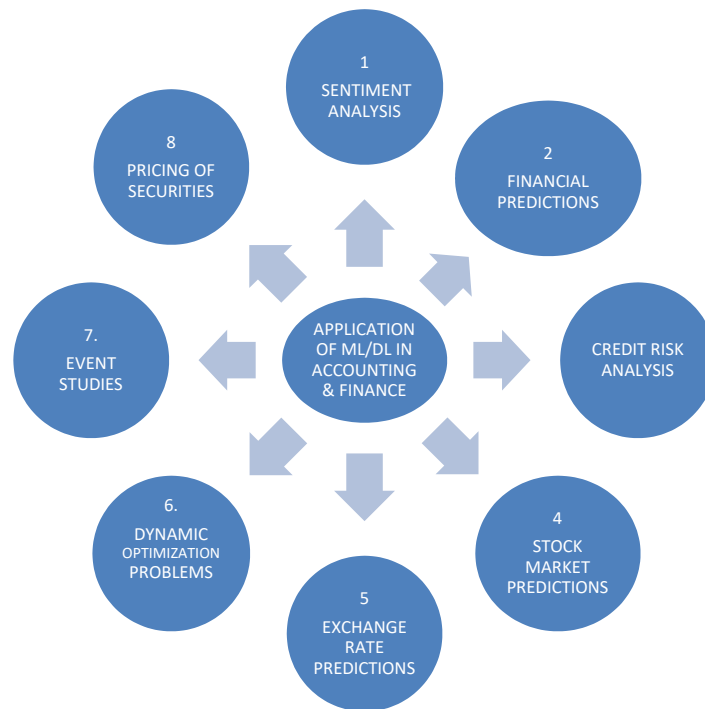


Figure 2. Depicting Application of Machine/Deep Learning in Accounting & Finance

ML in Sentiment Analysis for Prediction: Sentiment analysis is an established method of studying the future price and returns of any security. This method can be used very efficiently to predict the future price of financial securities as well as the returns on securities. Latest studies show that the process of sentiment analysis can be improved by using machine learning. Renault (2020) in his study analyzed a million messages sent on one of the micro-blogging platforms for evaluating the effectiveness of several pre-processing methods and ML approaches for performing sentiment analysis. He concluded that "the pre-processing method and the size of the dataset have a strong impact on the correlation between investor sentiment and stock returns". Thus, the investors' opinion about any stock, security, or financial asset can be modeled, which can give a very fair idea about the future price of the security than just getting a hint through traditional sentiment analysis. This development can indeed change the whole landscape of the trading psychology of the market. Becker and Reinganum (2018) provided a thorough overview of the growth of big data and sentiment analysis research over the last 30 years, highlighting the use of techniques such as NLP, SVMs, and ANNs for the analysis of news, conference calls, reports, and social media activity.

DL Hierarchical Decision Models for Financial Prediction: These models can be used very effectively for financial predictions and classification. Deep learning predictor enables to include all data points of importance to an estimating problem, which was not possible in traditional models. This means, practically all non-linear and complex data interactions are also accounted for in the prediction model, making the model more robust and reliable. Lirong G. et al. (2020) commented in their article that deep learning is better in terms of speed and accuracy than traditional pricing models and solves the challenges of pricing popular arithmetic Asian options.

Credit Risk Analysis through Default Probabilities: Credit analysts perform an analysis of the prospective borrower's ability to meet the debt obligations and is a very important area of finance, especially in the banking and non-banking financing sector. Analysts quantify the risk of loss of non-payment of principal or interest dues by using many theories and models of quantitative finance. This is another very important area of finance where machine learning can be used very effectively. Deep learning models in the area of credit risk analysis use multi-dimensional input space. To give an example, in image processing, these dimensions can be looked at as entire objects, then part of the objects, then certain features of the object, and finally pixels of the object. Such a map of features can be created to determine the credit-deservedness of a company. Zhu et al. (2016) used a novel ML technique referred to as RS-RAB for determining the credit risk of several small and medium-sized enterprises in China. Through empirical analysis, they determined that the performance of this novel method was superior to the other three methods they discussed.

ML/DL and Stock Market & Exchange Rate Predictions: Many technological advancements including computing technologies have resulted in the production and availability of a large amount of data related to stock markets and exchange rates. The use of ML/DL models is increasing in this field due to the ability of these models to deal with non-linear data more efficiently than the traditional models. Huang et al. (2020) carried extensive literature review on selected 40 articles out of the 140 articles published during 2014-18 and concluded that the stock market domain was the one in which deep learning was most widely used. Zheng et al. (2017) examined the performance of DBN and concluded in their research that the DBN model outperformed the FNN model.

Reinforcement Deep Learning for Dynamic Optimization Problems: Most of the problems in the financial sector are considered dynamic as they involve variables or revolve around such variables, which are very dynamic, and their value changes with time. Dynamic optimization is a decision-making process where algebraic mathematical models are formulated based on predictions. Reinforcement deep learning models are proving very effective in this area as well. Kolm et al. (2019) say in their research that reinforcement learning allows us to solve these dynamic optimization problems. Financial decisions related to pricing, hedging, investment and portfolio allocation, asset-liability management are the best examples of dynamic optimization, and ML/DL models can result in very successful modeling. Reinforcement learning models enable agents to learn through a sequence of decisions through the 'trial and error' method while incorporating feedback from experience.

ML and Event Studies: Event studies measure how past events can impact the current return. When coupled with ML/DL models, the event study can result in more accurate predictions about returns especially when alternate data is included in the model. Mullainathan et al. (2017). Traditional Time series data can now be used alongside highly informational event streams where the data is measured at irregular intervals and these streams are multiplying. Deep learning approaches extend options to build models that can extract meaningful patterns to solve a variety of situations arising from a particular event.

ML and Evaluation of Investment Proposals: Machine Learning is emerging as a very important tool for the evaluation of investment proposals. Investment decisions are the most important decisions in the area of finance other than financing and dividend decisions. Investment decisions can make or break a business. Looking at the importance of these decisions, it becomes imperative to use such an evaluation tool, which provides the most accurate information for accepting or rejecting any investment proposal. Emerson Sophie et al. (2019) in their study also argued that machine learning has been finding its application in many areas of business, especially in the investment process including return forecasts, risk modeling, and construction of the portfolio. They further say that "Quantitative investing, traditionally a leading field in adopting new techniques is found to be the most common source of use cases in the emerging literature".

ML in the field of Computational Finance: The application of quantum computation in solving financial problems is growing very fast. Literature review suggests that the machine learning models can be used to maximize portfolio returns, find opportunities to make a profit and do accurate credit scoring. Román Orús, et al. (2019) used quantum amplitude estimation in their study and proved that it can speed-up Monte Carlo modeling which has a link to analyzing risks and deriving prices.

To sum up, this section will quote Emerson Sophie et al. (2019), who identified seven main themes where the ML/DL models can be applied successfully; which include: return forecasting; decision-making; portfolio construction; fraud detection; language processing, and sentiment analysis.

3. Challenges in Adopting Machine Learning in Accounting & Finance Area

The world is witnessing perpetual growth in the application of machine learning models in many diverse areas including accounting and finance. There are many success stories propelling us to seek confidence in the

adoption of ML/DL modeling in the accounting and finance area, but we have an equal number of anecdotal stories as well if not more. Many funds depended upon ML or DL models but failed to achieve their predicted performance and a few of them even failed completely. Not only this, there are many examples when predictions about stock price, exchange rates, valuation of securities, and financial health of institutions using ML/DL models failed; raising fingers at the very idea of adopting ML/DL in the area of accounting and finance. "Despite an impressive array of recent successful applications of supervised machine learning (Dixon et al., 2016; Heaton, Polson, and Witte, Heaton et al.; Bayer and Stemper, 2018; Feng et al., 2018) and reinforcement learning (Halperin, 2017; Buhler et al., 2018) in financial modeling, there is an equally if not larger crowd of anecdotal stories of machine learning leading to fund failure, poor performance on financial data, or even seed research projects failing to make the light of day". Israel, Ronen et al. (2020), also commented that because machine learning has done so many amazing things, it may seem an inescapable conclusion that it will dominate financial tasks like stock picking. In order to develop realistic expectations about the benefits of machine learning for asset management, we must understand what makes finance different. Thereby, it becomes imperative to discuss the peculiarity of financial modeling.

The accounting and financial modeling involve understanding the behavior of customers and markets. So, experts developing financial models need to have strong knowledge of behavioral finance other than financial concepts, theories, and of course coding experience. Mullainathan, Sendhil, and Jann Spiess (2017) concluded in their research article that "machine learning algorithms are now technically easy to use as one can download convenient packages in R or Python that can fit decision trees, random forests, or LASSO (Least Absolute Shrinkage and Selection Operator) regression coefficients. This also raises the risk that they are applied naively or their output is misinterpreted". This emphasis upon the knowledge and understanding of the user about Dixon et al., 2016 commented in their article that "while problem areas requiring natural language processing and text mining fall heavily into well-established engineering practices, financial modeling requires a blend of engineering and financial modeling experience which is still uncharted territory". They brought out four common problems which limit the adoption of ML in the financial sector, which they call four horsemen as 1. The Tinkerer; 2. The Historian; 3. The Historian; 4. The Miner. According to Dixon et al. (2016), the 'Tinkerers' are statistically illiterate, thus leading to their inability to integrate ML/DL modeling with that of financial time series analysis, financial modeling, and interpretation. The 'Historians' on the other hand assume the stationarity of co-variant data, failing to predict effectively non-stationary data through supervised modeling. It is imperative to mention that only closely related LSTM models based on change-point detection are successful in such cases. The 'Miners' on the other hand fail to visualize the wider picture or context and end up applying one successful classifier model to all areas of accounting and finance, which results in a big failure of models. Lastly, the 'Puritans' are those experts who strictly work on rationalistic terms and expect every minute detail to be documented meticulously. They fail to accept that ML/DL modeling is ever-evolving and can discover new useful patterns, thus end up ignoring the non-linear effects of the predictions, which is the main premise of ML/DL models.

The next challenge in applying machine learning in the accounting and finance field is 'over-fitting'. Huang et al. (2020) summarised three important aspects of deep learning applications in their study and presented unfavorable impacts of over-fitting and sustainability while applying DL models in providing solutions. Machine learning in accounting and finance can be computationally very expensive due to the dynamic and volatile nature of the sector.

As we know that ML/DL models are as good as is the data used to develop them. This emphasizes the importance of quality data for the successful application of ML/DL modeling in accounting and finance. If the data has high noise, adopting machine learning becomes very challenging, especially in the area of asset management and returns prediction. Israel, Ronen et al. (2020) commented that "asset management, and return prediction in particular, is a small data science with low signal-to-noise ratios, making it very different from disciplines where machine learning has thrived. As a result, adapting machine learning in finance is a more difficult proposition than many commentators appreciate". Arnott R. et al. (2019) also concluded that though machine learning techniques can offer relatively better performance, however, for times we will have false patterns and results due to the notorious nature of financial data having a low signal-to-noise ratio.

Another challenge in the successful adoption of ML/DL models in the accounting and finance sector is 'explanation' and 'validation', which is sometimes known as a problem of 'black-box'. Bracke et al. (2019) say "Machine learning based predictive techniques are seeing increased adoption in a number of domains, including finance. However, due to their complexity, their predictions are often difficult to explain and validate". Machine learning models developed in the financial sector especially for predicting mortgage default may consist of

thousands of large decision trees deployed parallel to each other, which makes it very difficult to explain how the model has worked intuitively.

4. How to Use Machine Learning Effectively in Accounting and Finance?

In this section, we will discuss certain guidelines to enhance the successful application of machine learning in the accounting and finance area.

There are many players in developing models in the area of accounting and finance. Bracke et al. (2019) in their article identified at least six stakeholders, who are involved in creating and deploying a financial model. Out of those six players; four are directly involved in creating and deploying the architecture as developers; first and second line checkers and management responsible for the application. The other two stakeholders look after the regulatory part. Thereby, the success or failure of any ML/DL model developed for any problem in the accounting and finance sector will primarily depend on the knowledge, experience, understanding of financial theories, and behavior of customers and markets of developers as well as checkers.

To make DL/ML models work effectively in the financial sector, the quality of data used for modeling is of paramount importance. To deal with the high noise data, Arnott R. et al. (2019) suggested adopting back-testing protocols to tackle the low signal-to-noise ratio of financial data. Bracke et al. (2019) proposed a technique for solving the 'black box' problem part of several ML applications by making the use of the QII method of Datta *et al* (2016) in a real-world example: an ML model to predict mortgage defaults. Rundo, Francesco et al. (2019) argued in their article that ML/DL based financial models outperform traditional financial models as they identify significant information from presumably irrelevant information and can process and analyze a large amount of data including alternative data very effectively and efficiently.

In a nutshell, to ensure the successful application of ML/DL modeling in solving any problem in the accounting and finance sector is to define the problem itself. Once the problem is defined clearly, then the next step should be to decide whether the problem has a scheduling component requiring reinforcement learning or not. The next factor to be reviewed by the developers and checkers is to evaluate the computational feasibility to analyze the full set of data or reduced data through unsupervised learning. All assumptions of the modeling should be defined based on the type of data used. The developers would need to develop intuition by establishing a toy model. The experiment should be designed in such a way that the model output could be explained and defensible by reduction ad absurdum. Diagnostics, both from machine learning and statistics should be employed to characterize the significance of the model while using minimum algorithms and widgets. Further, the developers need to review the assumptions regarding the loss function used to train a model and its constraints ahead of analyzing the performance of specific algorithms. ML/DL team should ensure that the data sets used for training machine learning algorithms are not misbalanced to avoid the problem of the 'inspection paradox'. Lastly, the temptation to prematurely feature engineer should be avoided, without first assessing the signal in a back-testing or live simulation environment. The development of new back-testing strategies is considered to improve upon the reproducibility of ML/DL models in finance and is seeking the attention of researchers.

5. Conclusion

There is a growing body of literature suggesting a sharp increase in the use of machine learning techniques in the area of accounting and finance. Advanced mathematical models such as unsupervised machine learning techniques have become the need of the hour to model complex non-linear relationships in financial systems, where complex business situations are resulting in the generation of 'Big Data' and 'Alternate Data'. Heaton J. B. et al (2018) also concluded that deep learning models are likely to exert greater and greater influence in the practice of finance, particularly where prediction is of paramount importance. However, there are many challenges in applying ML/DL models in these prediction models especially when the modeling in finance involves behavioral aspects of extremely dynamic customers and markets. Thereby, the success of any ML/DL model developed for solving any problem in this sector will primarily depend on the knowledge, experience, understanding of financial theories as well behavior of customers and markets of developers and checkers, of course along with the modeling experience.

This article objected to presenting the scope and application of ML/DL in the financial sector in an easily understandable and lucid language without using a lot of technical jargon. Further, the challenges in the application of ML/DL in financial decision-making have also been addressed. This work can further be extended to draw classifications of ML/DL models used in the area of sentiment analysis, credit analysis, mortgage default predictions, asset pricing, investment analysis, portfolio construction, and return forecasting.

References

- Alcazar, et al. (2020). Classical Versus Quantum Models in Machine Learning: Insights from a Finance Application. *Machine Learning: Science and Technology*, 1(3), 035003 <https://doi.org/10.1088/2632-2153/ab9009>
- Arnott, R., Harvey, C. R., & Markowitz, H. (2019). A Backtesting Protocol in the Era of Machine Learning. *The Journal of Financial Data Science*, 1(1), 64-74.
- Aziz, S., Dowling, M. M., Hammami, H., & Piepenbrink, A. (2019). Machine Learning in Finance: A Topic Modeling Approach. <http://dx.doi.org/10.2139/ssrn.3327277>
- Becker, Y. L., & Reinganum, M. R. (2018). *The Current State of Quantitative Equity Investing*. CFA Institute Research Foundation.
- Belloni, et al. (2014). Inference on Treatment Effects after Selection among High-dimensional Controls. *The Review of Economic Studies*, 81(2), 608-650.
- Boyarshinov, V. (2005). Machine Learning in Computational Finance. Ph. D Thesis, Rensselaer Polytechnic Institute Library, Troy, New York.
- Bracke, P., & Datta, et al. (2019). Machine Learning Explainability in Finance: An Application to Default Risk Analysis. *Bank of England Working Paper No. 816*. <http://dx.doi.org/10.2139/ssrn.3435104>
- Cortes, C., & Vapnik, V. (1995). Support Vector Machine. *Machine Learning*, 20, 273-297.
- Dixon, M., & Halperin, I. (2019). The Four Horsemen of Machine Learning in Finance.
- Domingos, P. (2012). A Few Useful Things to know about Machine Learning. *Communications of the ACM*, 55(10), 78-87.
- Emerson, et al. (2019). Trends and Applications of Machine Learning in Quantitative Finance. *8th International Conference on Economics and Finance Research (ICEFR 2019)*. Retrieved from <https://ssrn.com/abstract=3397005>
- Gogas, P., & Papadimitriou, T. (2021). Machine Learning in Economics & Finance. *Computational Economics*, 57, 1-4. <https://doi.org/10.1007/s10614-021-10094-w>
- Ha, Y. (2017). *Machine Learning in Quantitative Finance*, Ph.D.Thesis, University of Glasgow.
- Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep Learning in Finance: Deep Portfolios. *Applied Stochastic Models in Business and Industry*, 3(1), 3-12. <https://doi.org/10.1002/asmb.2209>
- Huang, et al. (2020). Deep learning in finance and banking: A literature Review and classification. *Frontiers of Business Research in China*, 14, 2-24. <https://doi.org/10.1186/s11782-020-00082-6>
- Israel, et al. (2020). Can Machines 'Learn' Finance? *Journal of Investment Management*, Available at
- Kolm, P. N., & Ritter, G. (2019). Modern Perspectives on Reinforcement Learning in Finance. *The Journal of Machine Learning in Finance*, 1(1). <http://dx.doi.org/10.2139/ssrn.3449401>
- Lindsay, R. K. et al. (1993). DENDRAL: A Case Study of the First Expert System for Scientific Hypothesis Formation. *Artificial Intelligence*, 61(2), 209-261.
- Lirong, G., Huamao, W., & Zhaojun, Y. (2020). Machine Learning Solutions to Challenges in Finance: An Application to the Pricing of Financial Products. *Technological Forecasting and Social Change*, 153.
- Liu W. et al. (2017). A Survey of Deep Neural Network Architectures and their Applications. *Neuro Computing*, 234, 11-26.
- Lopez de Prado, M. (2016). Mathematics and economics: A Reality Check. *The Journal of Portfolio Management*, 43(1), 1-4. <https://doi.org/10.3905/jpm.2016.43.1.001>
- Mullainathan, S., & Jann, S. (2017). Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives*, 31(2), 87-106. <https://doi.org/10.1257/jep.31.2.87>
- Renault, T. (nd.). Sentiment Analysis and Machine Learning in Finance: A Comparison of Methods and Models on One Million Messages. *Digit Finance*, 2, 1-13.
- RománOrús, et al. (2019). Quantum computing for finance: Overview and prospects. *Reviews in Physics*, 4, 100028, <https://doi.org/10.1016/j.revip.2019.100028>
- Rundo, F. et al. (2019). "Machine Learning for Quantitative Finance Applications: A Survey. *Appl. Sci.*, 9(24),

5574. <https://doi.org/10.3390/app9245574>

Wagner, W. P. (2017). Trends in Expert System Development: A Longitudinal Content Analysis of over Thirty Years of Expert System Case Studies. *Expert Systems with Applications*, 76, 85-96.

Windmann, A. (2020). Machine Learning in Finance. University of RWTH Aachen.

Zheng, J., Fu, X., & Zhang, G. (2017). Research on Exchange Rate Forecasting based on a Deep Belief Network. *Neural Computing and Application*, 31, 573-582.

Zhu, Y. et al. (2016). Predicting China's SME Credit Risk in Supply Chain Finance Based on Machine Learning Methods. *Entropy*, 18(5), 195. <https://doi.org/10.3390/e18050195>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).