European Journal of
Technology
(EJT)

Robustness and Interpretability of Machine Learning Models in Financial Forecasting

Samir Vinayak Bayani, Ikram Ahamed Mohamed & Selvakumar Venkatasubbu

European Journal of Technology
ISSN 2520-0712 (online)
Vol.8, Issue 2, pp 54 - 66, 2024
www.ajpojournals.org

Robustness and Interpretability of Machine Learning Models in Financial Forecasting

Samir Vinayak Bayani[1*], Ikram Ahamed Mohamed[2] & Selvakumar Venkatasubbu[3]
[1]Broadcom Inc, USA
[2]Salesforce, USA
[3]New York Technology Partners, USA

Article history
Submitted 14.02.2024 Revised Version Received 20.03.2024 Accepted 06.05.2024

Abstract

Purpose: The practice of financial forecasting is an essential component of contemporary financial management and the process of making decisions on investments. As a result of the complexity and volatility of financial markets, traditional financial forecasting techniques often fail to adequately capture the situation.

Methodology/ Findings: One of the most effective methods for improving the precision and effectiveness of financial forecasting is the use of machine learning techniques. Through an examination of its potential, limitations, and future possibilities. They take linear and nonlinear possibilities into account in their study. Our focus is on linear methods, namely penalized regressions, and ensemble of models. The study considers a range of nonlinear techniques, including tree-based methods like boosted trees and random forests, and deep and shallow neural networks in both feed-forward and recurrent forms. They also think about ensemble and hybrid models, which combine features from several kinds of alternatives.

Implications to Theory, Practice and Policy: A brief overview of the tests used to measure outstanding predictive ability is provided. In the last part of this article, they focus on the possible applications of machine learning in the fields of economics and finance, and we provide an example that makes use of high-frequency financial data (Benti, Chaka, and Semie, 2023).

Keywords: *Boosting, Bagging, Forecasting, Neural Networks, Deep Learning, Penalized Regressions, Regularization, Nonlinear Models, Sieve Approximation, And Statistical Learning Theory.*

1.0 INTRODUCTION

This article provides an overview of the most recent advancements in the use of Machine Learning(ML) techniques in the forecasting of economic and financial time series. Estimation, model selection, and forecasting are all areas in which machine learning approaches have emerged as significant tools for practical researchers in the fields of economics and finance. In this day and age of Big Data, when seemingly endless datasets are readily available, it is of utmost significance to provide predictions that are both credible and resilient. In any case, what exactly is machine learning? Arthur L. Samuel, an early trailblazer in AI research, provided one of the most popular definitions. In his description, Samuel offers "the field of study that gives computers the ability to learn without being explicitly programmed." Machine learning, as we see it, is best described as the practice of learning (discovering) latent patterns in large datasets by combining effective statistical approaches with automated computer algorithms. For this reason, Statistical Learning Theory provides the statistical foundation for machine learning. Since improving statistical learning is our primary goal, this article will focus on statistical models specifically rather than machine learning as a whole. Machine learning approaches may be broadly grouped into three main types: Reinforcement learning, supervised learning, and unsupervised learning (Carvalho, Pereira, and Cardoso, 2019).

Supervised learning is the focus of this review. It is a kind of learning where the goal is to train a function that uses input-output pairs to map variables (explanatory and dependent) to one another.

Regression models are an example of a class that fits this description. Conversely, unsupervised learning is a subset of machine learning that seeks to unearth hidden patterns in unlabeled data sets. Such methods include data compression algorithms and cluster analysis, for instance. In conclusion, reinforcement learning is a technique that teaches an agent to do actions in a way that maximizes its reward. To achieve this aim, it explores and uses the knowledge it receives from repeated tries to maximize the reward. Several AI game players, like AlfaGo, and sequential therapies, like Bandit problems, rely on this basic idea (Li et al., 2022).

Machine learning is a subset of artificial intelligence that's defined as the process of teaching a computer to learn from data. It does this by identifying patterns and relationships in training data so that the computer can make predictions about future values and events. While traditional methods use a set of predefined rules to make predictions, machine learning is able to learn and adapt from any amount of data. Machine learning can be used for a variety of purposes, such as predicting consumer behavior, understanding market trends, forecasting sales, or even predicting when a server might crash. In fact, it can be used for any problem where there is time-series data and a goal to predict the future.

Body

Based on historical data and a variety of economic indicators, financial forecasting is the process of making predictions about future financial events. For firms, investors, and regulators to develop well-informed judgments, accurate forecasting is very necessary. Time series analysis and regression models are two examples of traditional forecasting methodologies that have been 0the instruments of choice for a considerable amount of time. These methodologies, on the other hand, often run into difficulties when attempting to reflect the complexity and non-linearity of financial markets, particularly in the data-driven and volatile environment that exists today. In recent years, machine learning (ML) approaches have become more popular as a result of their capacity to analyze massive datasets, identify patterns, and generate predictions without human intervention.

The supervised machine learning techniques that are discussed in this article may be generally classified into two categories. Our primary concentration is on the estimation of specifications by regularization, which is sometimes referred to as shrinkage. The development of such techniques dates back at least to Tikhonov (1943). The bias-variance tradeoff in statistical estimating was first popularized by Willard James and Charles Stein's seminal papers (Stein, 1956; James and Stein, 1961). With the release of these works, regularized estimators rose to popularity in the statistical and econometric communities. An estimator proposed by Hoerl and Kennard (1970) called the Ridge Regression is considered first. We next proceed to present the LASSO estimator, a tool developed by Tibshirani in 1996, along with its several variations. Many other types of punishment also cross their minds. We also go back to the drawing of theoretical conclusions and dependent data (Linardatos, Papastefanopoulos, and Kotsiantis, 2020).

They believe that the achievement may be somewhat explained by comparing it to the functioning of the human brain. Regardless of what has been stated in the early literature, the empirical success of NN models is based on the mathematical fact that a linear combination of enough simple basis functions can approximate very complex functions arbitrarily well in any given metric. The practical success of NN models may be explained by this. The discovery of methods to decrease the volatility of estimated models sparked a rise in the popularity of regression trees. Nowadays, applied economists have access to a wider variety of algorithms, such as Boosted Trees and Random Forests. They also go over certain ensemble-based methods, including the Complete Subset Regression (Elliott et al., 2013; Elliott et al., 2015) and Bagging Breiman (1996). Along with the above-mentioned models, these approaches are also available. In addition, they provide a brief overview of what we call "hybrid methods," which are novel approaches to machine learning forecasting that combine ideas from linear and nonlinear models. They will begin by outlining tests with improved predictive power within the framework of machine learning techniques, and then we will provide empirical proof of the methodology (Masini, Medeiros, and Mendes, 2021).

General Framework

In contrast to x, which denotes a deterministic (non-random) integer, X, when raised to the uppercase, represents a random quantity. My remark about notation is brief. Bold characters, such as X and x, indicate multivariate items. These objects include vectors and matrices. For values of q greater than or equal to 1, the sign k · kq represents the 'q norm of a vector. We say that a set S has a cardinality of |S|.

For horizons h = 1 and H, we want to forecast YT +h by sampling T realizations and utilizing a random vector (Yt, Z0 t) = 0. The following assumption is taken into account throughout the work:

There is an initial assumption that (DGP). This covariance-stationary stochastic process, denoted by the equation {(Yt, Z0 t) 0} t=1, takes values on R d+1.

This ignores crucial non-stationary processes that manifest in time series applications. The first assumption is that long-memory procedures, including unit-root, are not allowed. Recasting economic forecasting as an issue of decision-making is central to this strategy. To minimize the projected prediction loss or risk, it is necessary to choose fh from a set of plausible models using a loss function. Calculating the risk that a forecasting technique presents is the next step after calculating fbh for fh in evaluating it. Absolute error and squared error are the most prevalent types of losses for L1 and L2 risk functions, respectively. More recent evaluations by Elliott and Timmermann (2008), Granger and Machina (2006), and Elliott and Timmermann (2016) all cite extensive debates on the subject (Nourah Alangari et al., 2023).

Penalized Regressions

When the regression parameter is not uniquely determined, penalized linear regression occurs. Typically, this occurs when the number of observations (n) is big, maybe even more than T, and/or when the correlation between the variables is strong. Solving the ordinary least squares problem inside a ball centered at the origin is the overarching goal. You can see that the constrained solution has a lesser mean squared error than the unrestricted OLS, even if it's biased. The goal of penalized regressions is to minimize the Lagrangian form of the unknown parameter vector β0 using the estimator βb (Shah and Konda, 2021).

$$
\begin{aligned}
Q(\boldsymbol{\beta}) &= \sum_{t=1}^{T-h} \left(Y_{t+h} - \boldsymbol{\beta}'\boldsymbol{X}_t\right)^2 + p(\boldsymbol{\beta}), \\
&= \|\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}\|_2^2 + p(\boldsymbol{\beta}),
\end{aligned}
$$

Ridge Regression

To deal with regressors that are highly linked, Hoerl and Kennard (1970) suggested the linear regression approach. The selected approach was ridging regression. While reducing the estimator's variance, a small amount of bias was meant to be added. If the scale matrix is diagonal and contains identical entries, then ridge regression occurs, which is an example of Tikhonov Regularization. If the parameter vector is subjected to the squared'2 norm, as in the ridge regression penalty, the regression will be penalized by

$$
p(\boldsymbol{\beta}) = \lambda \sum_{i=1}^{n} \beta_i^2 = \lambda \|\boldsymbol{\beta}\|_2^2.
$$

Using ridge regression, which involves reducing the coefficients linked to the least significant predictors to zero though they are never really zero yields an analytical answer that is simply calculable. Thus, without a truncation approach, it is useless for predictor selection (Benti, Chaka, and Semie, 2023).

The LASSO Operator is Short for "Least Absolute Shrinkage and Selection."

Tibshirani (1996) and Chen et al. (2001) implemented the LASSO to simultaneously regularize and select variables. In settings with an abundance of data and a greater number of characteristics (n) than observations (n), the widely used regularization method LASSO finds extensive use. In LASSO, the regression penalty is determined by the '1 norm of the parameter vector; that is, it is based on,

$$p(\boldsymbol{\beta}) = \lambda \sum_{i=1}^{n} |\beta_i| = \lambda \|\boldsymbol{\beta}\|_1.$$

Coordinate descent algorithms quickly calculate the LASSO solution. The maximum penalty is the sparse solution norm with the lowest convexity. Sparseness is defined as the presence of exactly one non-zero coefficient in the subset k < n of the solution. In other words, the algorithm picks just a tiny fraction of the variables. In addition to the impracticality of testing different model combinations, LASSO works best with a sum of regressors equal to or less than T. Despite its attractive features, the LASSO does have certain limitations. Several alternatives to its limitations that maintain its good features have been proposed as penalties (Carvalho, Pereira, and Cardoso, 2019).

Financial Forecasting using Machine Learning Algorithms

Regression Models

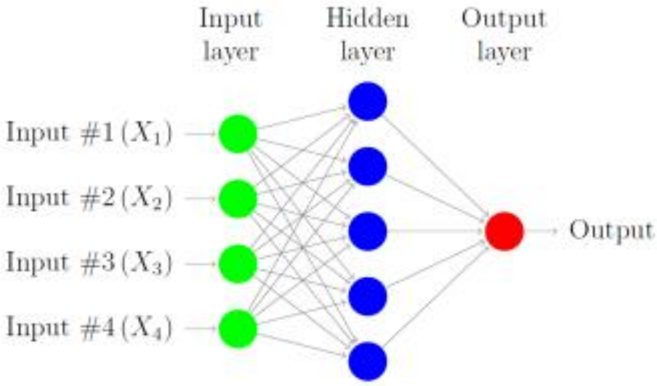
Within the field of machine learning, regression models provide a powerful tool for financial forecasting. The ability to model relationships between dependent and independent variables is made possible by them. Many algorithms have been used for financial result prediction, including linear, ridge, and LASSO regression. When the data shows linear connections, these models work well (Li et al., 2022).

Analysis of Time Series

One of the most important tools for financial forecasting is time series analysis. When applied to this field, ML approaches may capture intricate temporal connections. Machine learning has been used to improve the forecasting accuracy of methods such as GARCH (Generalized Autoregressive Conditional Heteroskedasticity) and autoregressive integrated moving average (ARIMA) (Linardatos, Papastefanopoulos and Kotsiantis, 2020).

Neural Networks

Because of their remarkable ability to simulate complex, non-linear patterns, artificial neural networks (ANNs) have become quite popular in the field of financial forecasting. The subfield of machine learning known as deep learning has made significant strides in this domain. The analysis of financial data has been enhanced using convolutional neural networks (CNNs) and recurrent neural networks (RNNs), with recurrent networks excelling in sequential financial data.



A kind of nonlinear sieves, NN models have found several applications in economic forecasting. The parameters are represented by weights, whereas activation functions often stand for the fundamental functions S(·). Attempting to make a dismal analogy to the human brain, the parts that make up the whole are called hidden neurons. Graphical representations of requirements, like the one in Figure, are commonly created using a single hidden layer NN model. In the picture, the green circles stand for the input layer components that make up the model's covariates (Xt). Four input variables are shown in the example. The output layer is represented by the red circles, while the hidden layer is symbolized by the blue ones. The example's hidden layer consists of five neurons. As an example of a linear combination of inputs, consider the arrows that go from green to blue circles: $0,jXt + γ0,j, where j ranges from 1 to 5. Finally, the arrows that link the red and blue circles show the linear combination of the hidden layer's outputs: β0 + P5 j=1 βjS(γ0 jXt + γ0,j) (Nourah Alangari et al., 2023).

Ensemble Methods

A strong strategy for merging the prediction abilities of several models has arisen from ensemble approaches. This includes techniques like gradient boosting and random forests. By combining the results of many base models, ensembles improve financial forecasting by reducing the likelihood of overfitting and increasing the precision of predictions.

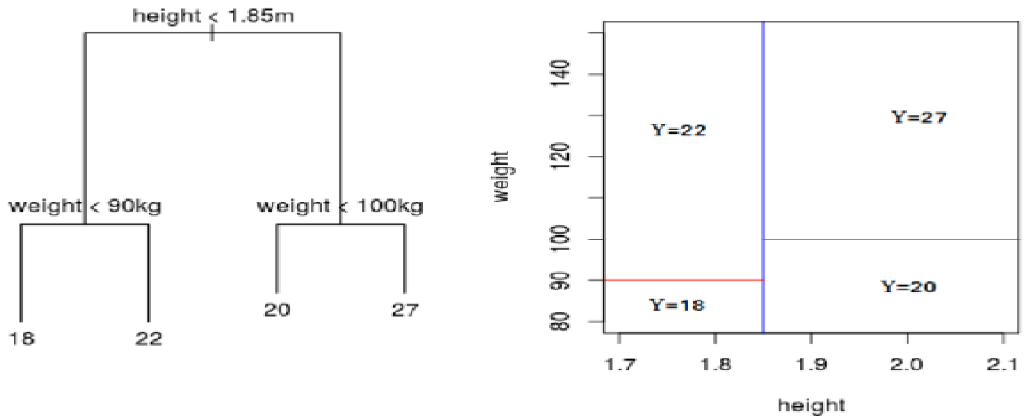
Learning using Reinforcement

The use of ML's reinforcement learning subfield, algorithmic trading, in particular, is on the rise in the financial sector. In different dynamic settings, this method enables models to make consecutive judgments. By analyzing past results and incorporating lessons learned, reinforcement learning algorithms may improve their approach and ultimately achieve better financial results (Shah and Konda, 2021).

Regression Trees

Nonparametric regression trees estimate a nonlinear function fh(Xt) from unknown data using local predictions and recursive variable partitioning. In this two-dimensional case, the graph of a tree on the left side of Figure 5 corresponds to the partitioning on the right side. As an example, let's pretend we wish to predict a basketball player's performance based on their past statistics. At the top of the tree, we can see that the participants with a height of more than 1.85 meters are separated from those with a lower stature.

Two nodes, one on each side of the graph, use weights to create distinct groups of players based on their height. The terminal nodes provide the category predictions after averaging the scores in each group. To find the optimal moment to prune branches, each tree node requires an ideal variable and an ideal observation (Benti, Chaka, and Semie, 2023).



The various areas are referred to as R1,..., RJ. At zero, there is the root node. Two new nodes, 2j + 1 and 2j + 2, are offshoots of the parent node at j.

Xsj t is the threshold variable associated with each parent node, where sj is an element of the set S = {1, 2,..., p}.

J is the set of all parent nodes, while T is the set of all terminal nodes. Figure 6 illustrates this point.

J = {0, 2, 5} represents the parent nodes in the example, whereas T = {1, 6, 11, 12} represents the terminal nodes.

This allows us to express the approximation model as

$$h_D(\boldsymbol{X}_t) = \sum_{i \in \mathbb{T}} \beta_i B_{\mathbb{J}i}\left(\boldsymbol{X}_t; \boldsymbol{\theta}_i\right),$$

where

$$B_{\mathbb{J}i}\left(\boldsymbol{X}_t; \boldsymbol{\theta}_i\right) = \prod_{j \in \mathbb{J}} I(X_{s_j,t}; c_j)^{\frac{n_{i,j}(1+n_{i,j})}{2}} \times \left[1 - I(X_{s_j,t}; c_j)\right]^{(1-n_{i,j})(1+n_{i,j})},$$

$$I(X_{s_j,t}; c_j) = \begin{cases} 1 & \text{if } X_{s_j,t} \le c_j \\ 0 & \text{otherwise,} \end{cases}$$

$$n_{i,j} = \begin{cases} -1 & \text{if the path to leaf } i \text{ does not include parent node } j; \\ 0 & \text{if the path to leaf } i \text{ include the } \textbf{right-hand} \text{ child of parent node } j; \\ 1 & \text{if the path to leaf } i \text{ include the } \textbf{left-hand} \text{ child of parent node } j. \end{cases}$$

Comparison of Forecasts

There is a rapid proliferation of predicting models and techniques in the growing body of ML research. This is why statistical methods for comparing models are so important. You may also utilize the many forecasting-related tests released since Diebold and Mariano's (1995) original work to evaluate these ML models. Diebold and Mariano (1995) proposed a straightforward t-test as a means to determine if two competing approaches share the same unconditional expected loss under the null hypothesis. Harvey et al. (1997) were the first to suggest the idea of a small sample adjustment.

This was recently also discussed by Diebold (2015). A limitation of the test proposed by Diebold and Mariano (1995) is that the statistics diverge under null when the competing models are nested. However, if the prediction models include a rolling window structure, as shown by Giacomini and White (2006), the test holds water. Diebold and Mariano's (1995) statistics might diverge under the null, as recently shown by McCracken (2020) using a fixed estimate window. This illustrates how critical it is to use a rolling window approach when making predictions. But as Hansen (2005) shows, when contrasted to less-than-stellar alternatives, White's (2000) test could be too cautious. A significant advancement in the field of forecasting, the model confidence set (MCS) was introduced by Hansen et al. (2011). Incorporating the model with the best performance within the specified loss function and confidence level is crucial when building a multi-class structure (MCS). Because the MCS considers the dataset's possible limits, an abundance of models with insufficiently informative data will produce an MCS with few models, and vice versa for data with sufficiently informative data. To begin, MCS does not presuppose the correctness of any given model.

Giacomini and White (2006) developed Conditional Equal Predictive Ability (CEPA) as an expansion of the Diebold and Mariano (1995) evaluation. Knowing how different models compare to one another and when one model is better than the others is crucial for practical application. Recently, a widely applicable approach to administering conditional predicting ability assessments was suggested by Li et al. (2020). There are several tests available in the literature that may be used to compare the predictions generated by various ML approaches, which is a key generalization (Carvalho, Pereira, and Cardoso, 2019).

Data Acquisition and Preprocessing

Data Sources

The reliability and authenticity of the data used for financial forecasting are of the utmost importance. Market data from the past, economic indicators, public opinion in the press, and even social media trends are common sources. Machine learning models rely heavily on financial data sources and APIs like Quandl and Bloomberg to give them up-to-date and correct information.

Data Quality and Cleaning

The accuracy of predictions is dependent on the precision of the data used. Outliers, inconsistent data, and missing values are commonplace in financial datasets. To make sure the input data is accurate, data cleaning processes are essential. These include filling in missing numbers, finding outliers, and reducing noise (Linardatos, Papastefanopoulos, and Kotsiantis, 2020).

Enhancement of Features

By identifying valuable characteristics in raw data, feature engineering aims to enhance machine learning models' predicting ability. Things like creating lag characteristics, developing technical indicators, or merging data from many periods could fall under this category when it comes to financial forecasting. Expertise in the relevant field is often necessary to identify crucial traits at this point.

Machine Learning Techniques for Use in Economic and Financial Projection

Linear Methods

There is a large body of work exploring the potential applications of penalized regressions to economics and financial forecasting, and practical economists increasingly use these methods in their toolboxes.

One of the best uses of penalized regressions is in macroeconomic forecasting. By comparing the ada LASSO to linear autoregressive and factor models, Medeiros, and Mendes (2016) show that the technique performs better when it comes to US inflation predictions. When a wide collection of predictors is considered, high-dimensional linear models tend to yield fewer forecasting errors for macroeconomic variables (Medeiros and Vasconcelos, 2016). Additionally, their findings show that ada LASSO hyperparameters may be fine-tuned to improve prediction accuracy (Nourah Alangari et al., 2023).

Nonlinear Methods

The use of nonlinear ML algorithms for economic and financial forecasting has been covered in several papers. The focus of the earlier research is mostly on NN methods. The most common models used in the first papers were autoregressive models with nonlinear parameters. It was seen to be an enhancement to add even a few additional variables. Think about the work of Terasvirta et al. (2005) and the sources cited within it. Nonlinear models have recently made a triumphant return because of the availability of large datasets. Consider the work of Medeiros et al. (2021).

Inflation forecasting literature has been debunked by their demonstration that multi-variable ML models routinely beat benchmarks across all forecasting horizons. Random Forests are determined to be the most effective model. Possible nonlinearities between inflation and prior significant macroeconomic variables and the Random Forest's distinctive variable selection approach both contribute to the model's impressive performance. All the hallmarks of high-quality research are present in the work of Gu et al. (2020). Using ML predictions of future stock returns based on a plethora of factors, the authors show that investors may reap substantial economic benefits. Tree and neural network-based models perform well. Coulombe et al. (2020) demonstrate substantial improvements in forecasting macroeconomic time series using nonlinear ML algorithms. Borup and Schutte (2020) use penalized regressions, ensemble techniques, and random forest to predict the increase in US employment from 2004 to 2019, using data from Google searches. Google search data is highly predictive, according to their analysis. Borup et al. (2020) use a wealth of data from Google Trends search volumes and machine learning algorithms to calculate the first claims for unemployment insurance in the US every week. Their data includes both historical and current claims.

It is critical to use many ML models and compare their results. Techniques like back-testing, out-of-sample testing, and cross-validation are used to evaluate how well models perform when presented with fresh data. It is possible to choose the best model for a given financial forecasting job by using model selection criteria, which include information criteria such as AIC and BIC (Shah and Konda, 2021).

Evaluation Metrics

To determine the efficacy of machine learning models, it is essential to evaluate their performance in financial forecasting. The accuracy and robustness of these models are evaluated using a variety of measures. Key assessment metrics are as follows:

To determine the efficacy of machine learning models, it is essential to evaluate their performance in financial forecasting. The accuracy and robustness of these models are evaluated using a variety of measures. Key assessment metrics are as follows (Benti, Chaka, and Semie, 2023):

The Mean Absolute Error (MAE) quantifies the typical magnitude of errors and is a metric for evaluating precision.

One such statistic is the Mean Squared Error (MSE), which gives more weight to larger mistakes since it squares them.

RMSE, or Root Mean Squared Error, is the square root of the Mean Squared Error; it gives the error in units that match those of the target variable (Carvalho, Pereira, and Cardoso, 2019).

Metrics for Risk

Value at Risk (VaR): Determines the most a portfolio of investments may lose over a certain period, given a certain degree of confidence.

If VaR is violated, the anticipated loss beyond VaR is measured by the Conditional Value at Risk (CVaR).

Sharpe Ratio: Evaluates a portfolio's performance after adjusting for volatility and return (Linardatos, Papastefanopoulos, and Kotsiantis, 2020).

Ethical Considerations

Bias and Fairness When it comes to banking, bias in machine learning algorithms might prevent people from getting the help they need. Unfair results, and more severe impacts on already-vulnerable populations, might result from biases in training data, such as past lending practices.

An urgent ethical challenge is the need to address prejudice and ensure fairness in model predictions (Nourah Alangari et al., 2023). Concerns about Personal Data Protection Sensitive personal information is often included in financial data. To preserve people's privacy, machine learning algorithms need to handle this data carefully. Protecting sensitive financial data requires strong data anonymization, encryption, and access restrictions (Shah and Konda, 2021).

Responsibility and Openness It is imperative that financial machine-learning models be open and responsible. To make sure that banks can explain their decision-making processes and take responsibility for the results of their models, industry standards and regulatory agencies are constantly changing. Customers and regulators are more likely to have faith in a model if it is transparent (Benti, Chaka, and Semie, 2023). Adherence to Regulations Worldwide, regulatory agencies are adjusting to the growing role of artificial intelligence in the financial sector. There is a heavy moral burden to adhere to rules and regulations, such as the General Data Protection Regulation (GDPR) in Europe and the ever-changing standards set forth by financial authorities. Banks and other financial organizations should check that their AI systems are ethical and follow all applicable laws (Carvalho, Pereira, and Cardoso, 2019).

2.0 CONCLUSION AND RECOMMENDATIONS

This paper provides a brief overview of recent advances in time-series modeling and forecasting using machine learning and high-dimensional statistics. We provided options that were both linear and nonlinear. Ensemble and hybrid models are also taken into consideration. We wrap off by quickly going over several tests for better-predicting abilities. To sum up, machine learning has changed the game for financial forecasters by becoming an indispensable tool in their decision-making process. Using a variety of algorithms, case studies, data gathering, assessment measures, difficulties, and new trends, this article has offered a thorough review of machine learning's usage in the financial sector. Investment banks, asset managers, risk managers, and fintech companies have all seen machine learning in action, which has shown its revolutionary potential. Ethical concerns and difficulties abound in this transition, however.

The field of financial forecasting is well-positioned for future innovations because of the ongoing development of machine learning and the emergence of new technologies such as quantum computing. The development of artificial intelligence in the financial sector will also be 00000000000heavily influenced by regulations and compliance requirements. It is critical to recognize the benefits and drawbacks of machine learning as we enter this new age of financial forecasting. Keep an eye out for any biases, assumptions in the models, and data quality issues. Also, because this area combines AI with finance, it is very multidisciplinary, thus specialists from both fields need to work together all the time. Finally, the use of machine learning in financial forecasting is an exciting new development in the financial industry that might lead to better choices, better risk management, and more accessible and efficient financial markets. Participants in the industry, regulators, and researchers must all keep ethical considerations in mind as they navigate this ever-changing landscape.

Machine learning (ML) has been defined in a variety of ways over the years; however, one of the most prominent comes from AI pioneer Arthur L. Samuel, who describes it as "the field of study that gives computers the ability to learn without being explicitly programmed." With a more precise definition in mind, we propose that ML be defined as the process of learning (discovering) hidden patterns in large datasets by combining automated computer algorithms with powerful statistical methods. Thus, Statistical Learning Theory provides ML with its statistical underpinnings. Because of our emphasis on statistical models, this article is less about ML generally and more on advances in Statistical Learning. Supervised, unsupervised, and reinforcement learning are the three main categories into which ML approaches fall. This study focuses on supervised learning, a kind of data-driven learning in which the goal is to train a function that, given a set of input-output pairs, transfers an input (explanatory variables) to an output (dependent variable). For instance, this category includes regression models. Contrarily, data compression techniques and cluster analysis are examples of unsupervised learning, a subset of ML that seeks for hidden patterns in data sets without accompanying labels. Finally, an agent may maximise its reward by learning to do certain behaviours in a given environment via reinforcement learning. In order to do this, it learns to maximise its reward via repeated trials and then uses this information to explore and exploit new areas. This is the meat and potatoes of many AI game players (like AlfaGo) and sequential treatments (like Bandit issues).

REFERENCES

Benti, N.E., Chaka, M.D. and Semie, A.G. (2023). Forecasting Renewable Energy Generation with Machine Learning and Deep Learning: Current Advances and Future Prospects. *Sustainability*, [online] 15(9), p.7087. doi:https://doi.org/10.3390/su15097087.

Carvalho, D.V., Pereira, E.M. and Cardoso, J.S. (2019). Machine Learning Interpretability: A Survey on Methods and Metrics. *Electronics*, 8(8), p.832. doi:https://doi.org/10.3390/electronics8080832.

Li, X., Xiong, H., Li, X., Wu, X., Zhang, X., Liu, J., Bian, J. and Dou, D. (2022). Interpretable deep learning: interpretation, interpretability, trustworthiness, and beyond. *Knowledge and Information Systems*, 64(12), pp.3197–3234. doi:https://doi.org/10.1007/s10115-022-01756-8.

Linardatos, P., Papastefanopoulos, V. and Kotsiantis, S. (2020). Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy*, [online] 23(1), p.18. doi:https://doi.org/10.3390/e23010018.

Masini, R.P., Medeiros, M.C. and Mendes, E.F. (2021). Machine learning advances for time series forecasting. *Journal of Economic Surveys*, 08(06). doi:https://doi.org/10.1111/joes.12429.

Nourah Alangari, El, M., Mathkour, H. and Ibrahim Almosallam (2023). Exploring Evaluation Methods for Interpretable Machine Learning: A Survey. *Information*, 14(8), pp.469–469. doi:https://doi.org/10.3390/info14080469.

Shah, V. and Konda, S.R. (2021). Neural Networks and Explainable AI: Bridging the Gap between Models and Interpretability. *INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY*, [online] 5(2), pp.163–176. Available at: https://ijcst.com.pk/index.php/IJCST/article/view/387 [Accessed 10 Mar. 2024].