Open Access Research Journal of
Science and Technology
Journals home page: https://oarjst.com/
ISSN: 2782-9960 (Online)

OARJ OPEN ACCESS RESEARCH JOURNALS

(Review Article)

Check for updates

Time series analysis and forecasting in finance: A data mining approach

Afiz Adewale Lawal [1, *], Omogbolahan Alli [2], Aishat Oluwatoyin Olatunji [3], Enuma Ezeife [4] and Ehizele Dean Okoduwa [5]

[1] *Department of Agricultural Economics, Ladoke Akintola University of Technology, Ogbomoso, Nigeria.*
[2] *Hult International Business School, Boston, Massachusetts, USA.*
[3] *Department of Computer Sciences, East Tennessee State University, Tennessee, USA.*
[4] *Ernst & Young LLC, McLean Virginia, USA.*
[5] *Yale School of Management, New Haven Connecticut, USA.*

Open Access Research Journal of Science and Technology, 2023, 09(01), 075-086

Publication history: Received on 02 July 2023; revised on 25 September 2023; accepted on 28 September 2023

Article DOI: https://doi.org/10.53022/oarjst.2023.9.1.0045

Abstract

Time series analysis and forecasting are essential methodologies in finance, playing a pivotal role in predicting market trends, evaluating economic conditions, and supporting decision-making. These methods rely on analyzing sequential data to uncover patterns, trends, and seasonal variations that drive financial phenomena. Traditional statistical models, such as ARIMA and GARCH, have long been utilized; however, their effectiveness is often constrained by assumptions like linearity and stationarity. Recent advancements in data mining techniques, including machine learning and artificial intelligence, have transformed the landscape of time series forecasting. These innovative approaches excel at handling non-linear relationships, high-dimensional data, and noise inherent in financial markets, making them indispensable for modern financial analytics. This paper scours the concept of time series analysis and data mining, examining their integration to improve forecasting accuracy. Additionally, it evaluates challenges such as data quality and computational requirements, while highlighting emerging opportunities, such as real-time forecasting and big data applications.

Keywords: Data Mining; Forecasting Model; Linearity; Neural Network; Stationarity; Time Series

1. Introduction

Forecasting in finance has always been a cornerstone of strategic decision-making, guiding investment strategies, risk management, and policy formulation. The analysis of time series data—sequentially recorded observations—allows researchers to identify patterns and relationships that provide insights into future trends. Time series analysis traditionally employs models like the Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH), which assume data stationarity and specific error distributions [1]. While these models are effective in many contexts, they often struggle with the inherent complexity of financial markets, where non-linear relationships, abrupt shifts, and high noise levels are common [2].

Data mining has emerged as a powerful complement to traditional time series methods, leveraging computational algorithms to uncover hidden patterns in large datasets [3]. Techniques such as Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and Recurrent Neural Networks (RNNs) allow for more sophisticated modelling of financial phenomena, particularly non-linear and dynamic behaviours [4]. These approaches can adapt to the evolving nature of financial markets, offering improved forecasting accuracy and resilience in the face of noisy and high-dimensional data [4].

Beyond supervised learning, unsupervised techniques like clustering and dimensionality reduction are gaining traction [5]. These methods enable efficient preprocessing of high-dimensional datasets, reducing computational complexity

* Corresponding author: Afiz Adewale Lawal

and improving model interpretability [6]. For instance, clustering can group similar time series data, enhancing pattern detection, while dimensionality reduction facilitates feature extraction by isolating the most relevant components [6, 7].

Despite their promise, the application of data mining to financial time series presents unique challenges. Issues such as non-stationarity, the presence of outliers, and the computational intensity of advanced algorithms remain significant hurdles [8]. The high-frequency nature of financial data necessitates the development of real-time analytics solutions, which add another layer of complexity [9].

2. Time series analysis in finance

2.1. Definition

Time series analysis is a statistical and computational method for analyzing sequences of data points collected at successive intervals over time [10]. Its primary objective is to understand the underlying structures and dynamics of the data, identify patterns, and forecast future values. In the context of finance, time series data includes variables such as stock prices, trading volumes, interest rates, and exchange rates, which are inherently dynamic and influenced by a multitude of factors [11].

2.2. Characteristics of Financial Time Series

Financial time series data are distinct from other types of time series due to several unique characteristics:

- **Volatility Clustering**: High-volatility periods are often followed by similar periods, and low-volatility periods exhibit the same pattern. This behaviour is a hallmark of financial time series and is modelled using techniques such as GARCH [12].
- **Non-Linearity**: Relationships between financial variables often deviate from linearity, requiring advanced models such as neural networks to capture these dynamics effectively [13].
- **Noise and Outliers**: Financial data are often noisy and subject to abrupt changes due to unforeseen events, such as political instability or economic crises [14].

2.3. Types of Time Series

Financial time series can be broadly categorized based on their statistical properties and the nature of the data.

2.4. Stationary Time Series:

A stationary time series exhibits statistical properties that remain constant over time, including its mean, variance, and autocorrelation structure [15]. Stationarity is a desirable property because many time series forecasting models, such as ARIMA, are built on this assumption. Examples include log-transformed returns of stock prices, which often approximate stationarity after preprocessing [16].

2.5. Non-Stationary Time Series

Non-stationary series show variability in statistical properties over time. These series often contain trends or seasonality, requiring transformations like differencing or detrending to stabilize their statistical properties [17]. Financial data such as raw stock prices or GDP growth rates are common examples of non-stationary series [17, 18].

2.6. Components of Time Series

Time series data are typically decomposed into several components, each representing a different type of variation in the data.

2.6.1. Trend

The trend component reflects the long-term movement of the series, capturing persistent upward or downward directions. In financial markets, trends are indicative of macroeconomic factors, such as inflation or technological advancements, that influence long-term price movements [19]. For instance, the steady rise in the Nasdaq Composite Index over the past decade represents a trend driven by the growth of technology companies [20].

2.6.2. Seasonality

Seasonal variations refer to repetitive patterns or cycles occurring at regular intervals, such as daily, monthly, or yearly [21]. These patterns are often driven by predictable events, such as increased retail activity during holiday seasons, which can affect stock market behaviours. Seasonality analysis is particularly useful in identifying investment opportunities linked to recurring patterns [21].

2.6.3. Cycles

Cyclical components are irregular fluctuations that occur over longer periods, often tied to economic or business cycles [22]. Unlike seasonality, these variations do not follow a fixed period. Financial crises, such as the 2008 global financial meltdown, demonstrate cyclical behaviour driven by broader macroeconomic factors [23].

2.6.4. Residuals

Residuals capture the random noise or unexplained variability in the time series. While they are often considered irrelevant, residuals can sometimes reveal structural breaks or anomalies in financial data. Advanced techniques such as anomaly detection are employed to analyze these residuals in depth [24].

2.7. Importance of Time Series Analysis in Finance

- **Forecasting**: Enables predictions of future values, such as stock prices or interest rates, which are critical for investment strategies and risk assessment [25].
- **Modelling Dependencies**: Identifies temporal relationships between variables, aiding in understanding market dynamics and causative factors [25].
- **Detecting Structural Breaks**: Helps identify shifts in market conditions, such as regime changes or economic shocks, which are essential for adaptive strategies [25].

3. Data mining

Data mining refers to the computational process of discovering meaningful patterns, correlations, and structures from large datasets [26]. In finance, it has emerged as an essential tool for uncovering hidden relationships in time series data, such as stock prices, trading volumes, and interest rates. Financial markets generate vast amounts of data with complex temporal dependencies, non-linearities, and noise, making traditional statistical methods often insufficient for accurate forecasting [27]. Data mining, through its advanced algorithms, provides a framework to address these challenges and improve predictive accuracy [28].

By combining statistical techniques with machine learning, data mining enables analysts to handle high-dimensional data, extract key features, and model serpentine relationships [29]. These capabilities are especially valuable in finance, where understanding market dynamics, predicting trends, and managing risks are critical.

4. Applications of data mining in financial time series

In the context of financial time series analysis, data mining can be applied to tasks such as forecasting, anomaly detection, clustering, and feature selection [30]. Each application leverages different algorithms tailored to the characteristics of financial data. For instance, data mining techniques are widely used to identify patterns in stock prices, forecast exchange rates, and assess credit risks, all of which play a significant role in decision-making [30, 31]. The application can either be via supervised learning or unsupervised.

4.1. Supervised Learning in Finance

Supervised learning techniques have become a cornerstone of predictive modelling in finance. Artificial Neural Networks (ANNs), for example, mimic the structure of the human brain and are adept at modelling non-linear relationships in financial time series [32]. These models are particularly effective for tasks such as stock price prediction and fraud detection, where the data exhibits complex, non-linear patterns [33]. Unlike traditional models that rely heavily on assumptions like stationarity, ANNs can adapt to data with dynamic characteristics, making them versatile for financial applications [34].

Support Vector Machines (SVMs) also stand out in financial data mining due to their robustness in handling high-dimensional data and non-linear classification problems [35]. SVMs map input data into a higher-dimensional space where relationships between variables become clearer [36]. This capability is useful for predicting credit risk, bankruptcy, and other financial outcomes where the decision boundaries are not straightforward.

Gradient Boosting Machines (GBMs), including algorithms such as XGBoost and LightGBM, are another powerful supervised learning approach widely applied in finance [37]. These ensemble models combine the strengths of multiple weak learners to form a strong predictive model [37]. Their ability to handle missing data, identify feature importance, and minimize overfitting has made them popular in applications like portfolio optimization and credit scoring [38].

4.2. Unsupervised Learning in Finance

Unlike supervised learning, unsupervised learning methods analyze datasets without predefined labels. These techniques are particularly valuable for discovering underlying structures in data, making them essential for preprocessing, exploratory analysis, and feature engineering [39].

Clustering algorithms, such as k-means and hierarchical clustering, group data points based on similarities. In finance, these methods are often used to segment time series into clusters with similar patterns, such as identifying stocks with comparable risk-return profiles [40]. This segmentation allows investors to create diversified portfolios by selecting assets from different clusters, thereby reducing overall risk [41].

Dimensionality reduction techniques, such as Principal Component Analysis (PCA) and t-SNE, address the challenge of high-dimensionality in financial data [42]. For example, a portfolio comprising multiple assets generates numerous time series, making it computationally expensive to analyze [42]. PCA reduces the dimensionality by identifying the most significant features, thus simplifying the analysis while preserving the data's essential structure [43]. This technique has been extensively used in identifying latent factors that drive asset prices, such as market, sectoral, or economic influences [44].

4.3. Hybrid Approaches

The integration of supervised and unsupervised methods has given rise to hybrid approaches, which combine the strengths of both paradigms [45]. For example, clustering can be employed to segment financial time series into similar groups, followed by supervised learning algorithms like ANNs or SVMs for predictive modelling within each cluster [46]. This layered approach enhances the model's accuracy by focusing on the unique characteristics of each segment, making it particularly suitable for complex financial problems like multi-asset risk assessment [45, 47].

5. Benefits and limitations of data mining in finance

The benefit of data mining approaches in finance heterogenous. They offer enhanced predictive accuracy by capturing non-linear relationships and temporal dependencies often overlooked by traditional methods [48]. Furthermore, they are scalable, making them suitable for analyzing high-frequency trading data and large-scale financial datasets. Their adaptability to changing market conditions allows them to remain relevant even in dynamic environments [48].

However, the application of data mining techniques is not without challenges. Financial datasets are often noisy, with missing values and outliers that can adversely affect model performance [49]. Overfitting is another common issue, particularly in models like ANNs, where excessive complexity can lead to poor generalization [50]. The computational requirements of advanced algorithms can be prohibitive, especially when applied to high-frequency or large-volume financial data [51]. Despite these limitations, the advancements in computational power and algorithmic efficiency continue to expand the scope and applicability of data mining in finance [51, 52].

6. Time series forecasting models

Forecasting financial time series is a critical aspect of financial decision-making, encompassing tasks such as predicting stock prices, estimating risk measures, and modelling macroeconomic indicators [53]. Over the years, various models have been developed to analyze and forecast time series data [54]. These models broadly fall into two categories: traditional statistical models and modern machine learning-based data mining models [55]. Each approach has distinct characteristics, applications, and limitations.

6.1. Traditional Statistical Models

Statistical models have historically been the cornerstone of time series forecasting, offering robust frameworks grounded in mathematical principles.

6.1.1. Autoregressive Integrated Moving Average (ARIMA)

The ARIMA model, one of the most popular traditional methods, forecasts future values based on past observations [56]. It combines three key components: autoregression (AR), differencing to achieve stationarity (I), and moving average (MA) [57]. ARIMA is particularly useful for time series that exhibit linear trends and do not have seasonal variations [58]. However, its reliance on linear assumptions makes it less effective for non-linear or highly complex data, which are common in financial markets.

6.1.2. Seasonal ARIMA (SARIMA)

extends ARIMA to handle seasonal patterns by incorporating seasonal differencing and additional seasonal AR and MA components [58, 59]. This makes SARIMA suitable for forecasting time series with periodic fluctuations, such as quarterly economic growth rates or seasonal sales trends [59].

6.1.3. Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

The GARCH model is widely used in finance to model and forecast volatility [60]. Unlike ARIMA, which focuses on mean values, GARCH captures time-dependent changes in variance, making it invaluable for risk management applications such as option pricing and portfolio optimization [61]. Financial time series often exhibit heteroskedasticity (changing variance), which GARCH effectively models by combining past error terms and variances [62]. While these statistical models are robust and interpretable, they have limitations in handling large datasets and capturing non-linear relationships, which are increasingly important in complex financial environments [61, 62].

6.2. Data Mining Models

Data mining approaches by leveraging machine learning have revolutionized time series forecasting by addressing the limitations of traditional models. These models excel in capturing non-linear dependencies, handling high-dimensional datasets, and adapting to changing patterns in real-time [63].

6.2.1. Recurrent Neural Networks (RNNs)

RNNs are a class of deep learning models specifically designed for sequential data. Unlike traditional models, RNNs leverage their internal memory to learn from the sequential nature of time series data [64]. This capability allows them to capture temporal dependencies effectively. In finance, RNNs are used for tasks such as stock price prediction, sentiment analysis of financial news, and portfolio management [65]. However, vanilla RNNs often suffer from issues like vanishing gradients, which limit their ability to learn long-term dependencies [66].

6.2.2. Long Short-Term Memory (LSTM)

LSTM networks, a specialized type of RNN, address the limitations of traditional RNNs by introducing memory cells and gating mechanisms [67]. These features allow LSTMs to retain relevant information over extended periods, making them highly effective for forecasting long financial time series with intricate patterns [68]. LSTMs have been successfully applied to tasks such as predicting currency exchange rates, detecting market anomalies, and modelling credit risk [69].

6.3. Transformer Models

Recently, transformer-based architectures have gained traction in time series forecasting. Originally developed for natural language processing, transformers use self-attention mechanisms to focus on relevant parts of the input sequence, making them highly effective for handling long-range dependencies [70]. In finance, transformers are being explored for multi-asset portfolio forecasting and high-frequency trading applications [71].

6.4. Hybrid Models

Hybrid models combine the strengths of traditional statistical methods and data mining approaches to achieve improved forecasting accuracy. For instance, ARIMA-LSTM models integrate the linear forecasting capabilities of ARIMA with the non-linear learning power of LSTMs [72]. These hybrid models are particularly effective in capturing both short-term trends and long-term dynamics, making them suitable for complex financial tasks such as multi-asset portfolio optimization and real-time risk assessment [72, 73].

7. Evaluation metrics for forecasting models

Evaluating the performance of forecasting models is crucial to ensure their effectiveness. Commonly used metrics include:

- **Mean Absolute Error (MAE):** Measures the average magnitude of errors without considering their direction, making it intuitive and easy to interpret [74].
- **Root Mean Squared Error (RMSE):** Highlights larger errors by squaring them before averaging, making it more sensitive to outliers. This metric is particularly useful in financial forecasting, where large errors can have significant implications [75].
- **Mean Absolute Percentage Error (MAPE):** Expresses forecasting errors as a percentage of actual values, facilitating comparisons across datasets with different scales [76].

In addition to these general-purpose metrics, specific evaluation methods are often employed to assess models in finance such as **volatility metrics** for models like GARCH that focus on volatility prediction, metrics such as realized volatility and variance ratios are used to measure accuracy [61, 62, 77], and **directional accuracy** that evaluates whether the model correctly predicts the direction of change in the financial series, which is crucial for trading strategies and investment decision-making [78].

8. Importance of evaluation metrics in finance

Financial time series are often characterized by volatility, non-stationarity, and noise, making them inherently difficult to forecast. A robust evaluation framework enables researchers and practitioners to quantify the accuracy of their models and compare different methodologies objectively [79]. Additionally, these metrics provide insights into the model's limitations, guiding further refinement and adaptation [79, 80].

9. Practical considerations

The selection of evaluation metrics should align with the financial context and the specific objectives of the forecasting task [81]. In high-frequency trading, minimizing RMSE might be less critical than achieving high directional accuracy, while in portfolio risk assessment, accurately modelling tail risks and volatility becomes paramount [81, 82].

It is also important to recognize that financial time series often exhibit structural breaks, regime shifts, and other complexities that can distort metric calculations [83]. Combining multiple evaluation metrics can provide a more comprehensive assessment, ensuring that models are both accurate and robust across varying market conditions [84].

10. Challenges and opportunities

The application of time series analysis and data mining in finance offers remarkable potential but also presents significant challenges. Understanding these hurdles and leveraging emerging opportunities is crucial for advancing the field and ensuring practical utility in real-world financial systems.

10.1. Noise and Outliers

Financial time series often contain noise and outliers caused by unexpected market events, economic shocks, or data recording errors. These irregularities can distort forecasts and degrade model performance [85]. Traditional statistical models like ARIMA are particularly vulnerable to such anomalies, whereas advanced data mining approaches such as neural networks may overfit the noise, reducing their generalization capabilities [86]. Robust preprocessing techniques, including outlier detection and noise reduction, are essential to mitigate this issue [87].

10.2. Non-Stationarity

Non-stationarity is a hallmark of financial time series, where mean, variance, or autocorrelation structures change over time [88]. Structural breaks, such as economic crises or regulatory changes, exacerbate this issue, making it difficult for models to adapt. While differencing techniques and advanced statistical methods like GARCH address some aspects of non-stationarity, data mining models require additional mechanisms, such as dynamic parameter tuning or transfer learning, to remain effective [89].

10.3. High Dimensionality

The increasing availability of financial data has led to high-dimensional datasets, which include multiple features such as macroeconomic indicators, social media sentiment, and trading volumes [90]. High dimensionality complicates model training and increases the risk of overfitting [91]. Dimensionality reduction techniques, such as Principal Component Analysis (PCA) or feature selection algorithms, are often employed to address this challenge [92].

10.4. Real-Time Forecasting Demands

In applications such as high-frequency trading, real-time forecasting is critical. Models must process vast amounts of streaming data with minimal latency [93]. Traditional methods struggle with the computational demands of real-time analysis, whereas data mining models require substantial hardware and software resources to maintain performance [94].

10.5. Interpretability

While data mining models, such as neural networks, provide high accuracy, they often operate as "black boxes," offering limited interpretability [95]. In finance, where transparency is essential for regulatory compliance and investor trust, this lack of interpretability is a significant drawback [95, 96]. Efforts to develop explainable AI (XAI) techniques are ongoing to bridge this gap [96].

11. Conclusion

Time series analysis and forecasting are indispensable tools for financial decision-making, investment strategies, and risk management. As financial markets grow increasingly complex, the need for sophisticated predictive models has never been more critical. While traditional statistical methods, such as ARIMA and GARCH have laid the foundation for time series forecasting, the incorporation of data mining approaches has revolutionized this field, offering the potential for greater accuracy and adaptability in an era of big data and rapid market changes.

Data mining techniques, particularly supervised and unsupervised learning models such as artificial neural networks (ANNs), support vector machines (SVMs), and clustering algorithms, are pushing the boundaries of forecasting capabilities. However, the application of these models in financial forecasting is not without challenges. Non-stationarity, noise, outliers, and high-dimensionality are just some of the issues that can impact the reliability of forecasts. The ever-changing nature of financial markets, characterized by volatility and unexpected events, further complicates the task of building robust predictive models. Additionally, while machine learning models offer high accuracy, they often operate as "black boxes," creating concerns over interpretability and transparency in decision-making.

Despite these challenges, significant opportunities lie ahead. The integration of machine learning with traditional financial models, the use of big data and cloud computing, and the rise of real-time analytics and automation are poised to transform financial forecasting. These advancements will enable faster, more accurate, and more adaptive models that can respond to market changes in real-time. Furthermore, the ongoing development of explainable AI (XAI) techniques will address concerns about the transparency of machine learning models, ensuring that financial practitioners can trust and understand the predictions being made.

Ultimately, time series forecasting powered by data mining approaches will continue to evolve, and ass the field matures, it holds the promise of enabling more informed decision-making, reducing risks, and driving innovations in financial markets.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Mustapa FH, Ismail MT. Modelling and forecasting S&P 500 stock prices using hybrid Arima-Garch Model. InJournal of Physics: Conference Series 2019 Nov 1 (Vol. 1366, No. 1, p. 012130). IOP Publishing.

[2] Kobeissi YH. Multifractal financial markets: An alternative approach to asset and risk management. New York: springer; 2013.

[3] Fulcher BD, Little MA, Jones NS. Highly comparative time-series analysis: the empirical structure of time series and their methods. Journal of the Royal Society Interface. 2013 Jun 6;10(83):20130048

[4] Abiodun OI, Jantan A, Omolara AE, Dada KV, Umar AM, Linus OU, Arshad H, Kazaure AA, Gana U, Kiru MU. Comprehensive review of artificial neural network applications to pattern recognition. IEEE access. 2019 Oct 4;7:158820-46.

[5] Abukmeil M, Ferrari S, Genovese A, Piuri V, Scotti F. A survey of unsupervised generative models for exploratory data analysis and representation learning. Acm computing surveys (csur). 2021 Jul 9;54(5):1-40.

[6] Boppiniti ST. Big Data Meets Machine Learning: Strategies for Efficient Data Processing and Analysis in Large Datasets. International Journal of Creative Research In Computer Technology and Design. 2020;2(2).

[7] Jain AK, Murty MN, Flynn PJ. Data clustering: a review. ACM computing surveys (CSUR). 1999 Sep 1;31(3):264-323.

[8] Benndorf J. Closed loop management in mineral resource extraction: turning online geo-data into mining intelligence. Springer Nature; 2020 Feb 13.

[9] Milosevic Z, Chen W, Berry A, Rabhi FA, Buyya R, Calheiros RN, Dastjerdi AV. Real-time analytics. Big Data: Principles and Paradigms. 2016;2016:39-61.

[10] Anderson TW. The statistical analysis of time series. John Wiley & Sons; 2011 Sep 28.

[11] Preis T. Econophysics—complex correlations and trend switchings in financial time series. The European Physical Journal Special Topics. 2011 Mar;194(1):5-86.

[12] Mohamed IA. Using GARCH Model in the Analysis of Trade Liberalization and Poverty in Developing Countries. International and Comparative Law. 2011 Mar 14;6(32).

[13] Inglada-Perez L. A comprehensive framework for uncovering non-linearity and Chaos in financial markets: Empirical evidence for four major stock market indices. Entropy. 2020 Dec 18;22(12):1435.

[14] Makridakis S, Bakas N. Forecasting and uncertainty: A survey. Risk and Decision Analysis. 2016 Jan 1;6(1):37-64.

[15] Weiß CH. Stationary count time series models. Wiley Interdisciplinary Reviews: Computational Statistics. 2021 Jan;13(1):e1502.

[16] Kitagawa G. Introduction to time series modeling with applications in R. Chapman and Hall/CRC; 2020 Aug 10.

[17] Castle JL, Doornik JA, Hendry DF. Modelling non-stationary ‘big data’. International Journal of Forecasting. 2021 Oct 1;37(4):1556-75.

[18] Hunter J, Burke SP, Canepa A. Multivariate modelling of non-stationary economic time series. Basingstoke: Palgrave Macmillan; 2017 May 8.

[19] Waheeb W, Ghazali R. A novel error-output recurrent neural network model for time series forecasting. Neural Computing and Applications. 2020 Jul;32(13):9621-47.

[20] Arsov D. Top 10 Relatively Strongest American Sector Indices. Available at SSRN 2534103. 2014 Dec 4.

[21] Mills TC. Modelling trends and cycles in economic time series: historical perspective and future developments. Cliometrica. 2009 Oct;3(3):221-44.

[22] Zarnowitz V. What is a business cycle?. National Bureau of Economic Research; 1991 Oct 1.

[23] Dimsdale N, Thomas R, Dimsdale N, Thomas R. Business, Monetary and Credit Cycles in Theory. UK Business and Financial Cycles Since 1660: Volume I: A Narrative Overview. 2019:7-34.

[24] Buda TS, Caglayan B, Assem H. Deepad: A generic framework based on deep learning for time series anomaly detection. InPacific-Asia conference on knowledge discovery and data mining 2018 Jun 3 (pp. 577-588). Cham: Springer International Publishing.

[25] Hyndman RJ, Lee AJ, Wang E. Fast computation of reconciled forecasts for hierarchical and grouped time series. Computational statistics & data analysis. 2016 May 1;97:16-32.

[26] Mining WI. Data mining: Concepts and techniques. Morgan Kaufinann. 2006;10(559-569):4.

[27] Lommers K, Harzli OE, Kim J. Confronting machine learning with financial research. arXiv preprint arXiv:2103.00366. 2021 Feb 28.

[28] Talagala TS, Hyndman RJ, Athanasopoulos G. Meta-learning how to forecast time series. Monash Econometrics and Business Statistics Working Papers. 2018 May;6(18):16.

[29] Maitra R. A statistical perspective on data mining. Journal of the Indian Society of Probability and Statistics. 2002;6:28-77.

[30] Shaukat K, Alam TM, Luo S, Shabbir S, Hameed IA, Li J, Abbas SK, Javed U. A review of time-series anomaly detection techniques: A step to future perspectives. InAdvances in Information and Communication: Proceedings of the 2021 Future of Information and Communication Conference (FICC), Volume 1 2021 (pp. 865-877). Springer International Publishing.

[31] Choi K, Yi J, Park C, Yoon S. Deep learning for anomaly detection in time-series data: Review, analysis, and guidelines. IEEE access. 2021 Aug 26;9:120043-65.

[32] Huang W, Lai KK, Nakamori Y, Wang S. Forecasting foreign exchange rates with artificial neural networks: A review. International Journal of Information Technology & Decision Making. 2004 Mar;3(01):145-65.

[33] Raschka S, Julian D, Hearty J. Python: deeper insights into machine learning. Packt Publishing Ltd; 2016 Aug 31.

[34] Zhang C, Li R, Shi H, Li F. Deep learning for day-ahead electricity price forecasting. IET Smart Grid. 2020 Aug;3(4):462-9.

[35] Nair BB, Mohandas VP, Sakthivel NR. A genetic algorithm optimized decision tree-SVM based stock market trend prediction system. International journal on computer science and engineering. 2010 Jan 1;2(9):2981-8.

[36] Hao PY, Kung CF, Chang CY, Ou JB. Predicting stock price trends based on financial news articles and using a novel twin support vector machine with fuzzy hyperplane. Applied Soft Computing. 2021 Jan 1;98:106806.

[37] Jha SB, Babiceanu RF, Pandey V, Jha RK. Housing market prediction problem using different machine learning algorithms: A case study. arXiv preprint arXiv:2006.10092. 2020 Jun 17.

[38] Pham BT, Prakash I. Evaluation and comparison of LogitBoost Ensemble, Fisher's Linear Discriminant Analysis, logistic regression and support vector machines methods for landslide susceptibility mapping. Geocarto International. 2019 Feb 23;34(3):316-33.

[39] Abukmeil M, Ferrari S, Genovese A, Piuri V, Scotti F. A survey of unsupervised generative models for exploratory data analysis and representation learning. Acm computing surveys (csur). 2021 Jul 9;54(5):1-40.

[40] Lim T, Ong CS. Portfolio diversification using shape-based clustering. Journal of Financial Data Science. 2021;3(1):111.

[41] Botte A, Bao D. A machine learning approach to regime modeling. Two Sigma. 2021.

[42] Ray P, Reddy SS, Banerjee T. Various dimension reduction techniques for high dimensional data analysis: a review. Artificial Intelligence Review. 2021 Jun;54(5):3473-515.

[43] Hasan BM, Abdulazeez AM. A review of principal component analysis algorithm for dimensionality reduction. Journal of Soft Computing and Data Mining. 2021 Apr 15;2(1):20-30.

[44] Gu S, Kelly B, Xiu D. Autoencoder asset pricing models. Journal of Econometrics. 2021 May 1;222(1):429-50.

[45] Gururangan S, Marasović A, Swayamdipta S, Lo K, Beltagy I, Downey D, Smith NA. Don't stop pretraining: Adapt language models to domains and tasks. arXiv preprint arXiv:2004.10964. 2020 Apr 23.

[46] Otchere DA, Ganat TO, Gholami R, Ridha S. Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. Journal of Petroleum Science and Engineering. 2021 May 1;200:108182.

[47] Binu D. Cluster analysis using optimization algorithms with newly designed objective functions. Expert Systems with Applications. 2015 Aug 15;42(14):5848-59.

[48] Ndikum P. Machine learning algorithms for financial asset price forecasting. arXiv preprint arXiv:2004.01504. 2020 Mar 31.

[49] Ball RT, Ghysels E. Automated earnings forecasts: Beat analysts or combine and conquer?. Management Science. 2018 Oct;64(10):4936-52.

[50] Bin Rafiq R, Modave F, Guha S, Albert MV. Validation methods to promote real-world applicability of machine learning in medicine. InProceedings of the 2020 3rd International Conference on Digital Medicine and Image Processing 2020 Nov 6 (pp. 13-19).

[51] Yu JG, Selby B, Vlahos N, Yadav V, Lemp J. A feature-oriented vehicle trajectory data processing scheme for data mining: A case study for Statewide truck parking behaviors. Transportation research interdisciplinary perspectives. 2021 Sep 1;11:100401.

[52] Hu YJ, Lin SJ. Deep reinforcement learning for optimizing finance portfolio management. In2019 amity international conference on artificial intelligence (AICAI) 2019 Feb 4 (pp. 14-20). IEEE.

[53] Hendahewa DN, Perera AA, De Meraal TD, Indikadulle IK, Perera SD, Rathnayake HM. Comparative Analysis between Supervised Machine Learning and Time Series Models for stock price prediction. In2021 International Conference on Data Analytics for Business and Industry (ICDABI) 2021 Oct 25 (pp. 354-358). IEEE.

[54] Oswal N. Predicting rainfall using machine learning techniques. arXiv preprint arXiv:1910.13827. 2019 Oct 29..

[55] Dogan A, Birant D. Machine learning and data mining in manufacturing. Expert Systems with Applications. 2021 Mar 15;166:114060.

[56] Fan D, Sun H, Yao J, Zhang K, Yan X, Sun Z. Well production forecasting based on ARIMA-LSTM model considering manual operations. Energy. 2021 Apr 1;220:119708.

[57] Liu T, Liu S, Shi L, Liu T, Liu S, Shi L. ARIMA Modelling and Forecasting. Time Series Analysis Using SAS Enterprise Guide. 2020:61-85.

[58] Khashei M, Bijari M, Hejazi SR. Combining seasonal ARIMA models with computational intelligence techniques for time series forecasting. Soft computing. 2012 Jun;16:1091-105.

[59] Al-Gounmeein RS, Ismail MT. Forecasting the exchange rate of the Jordanian Dinar versus the US dollar using a Box-Jenkins seasonal ARIMA model. International Journal of Mathematics and Computer Science. 2020;15(1):27-40.

[60] Chong CW, Ahmad MI, Abdullah MY. Performance of GARCH models in forecasting stock market volatility. Journal of forecasting. 1999 Sep;18(5):333-43.

[61] Livnat J, Singh J. Machine Learning Algorithms to Classify Future Returns Using Structured and Unstructured Data. The Journal of Investing. 2021 Feb 12.

[62] Sharma P, _ V. Forecasting stock index volatility with GARCH models: international evidence. Studies in Economics and Finance. 2015 Oct 5;32(4):445-63.

[63] Hassanien AE, Darwish A, editors. Machine learning and big data analytics paradigms: Analysis, applications and challenges, 2020.

[64] DiPietro R, Hager GD. Deep learning: RNNs and LSTM. InHandbook of medical image computing and computer assisted intervention 2020 Jan 1 (pp. 503-519). Academic Press.

[65] Hewamalage H, Bergmeir C, Bandara K. Recurrent neural networks for time series forecasting: Current status and future directions. International Journal of Forecasting. 2021 Jan 1;37(1):388-427.

[66] Kag A, Saligrama V. Training recurrent neural networks via forward propagation through time. InInternational Conference on Machine Learning 2021 Jul 1 (pp. 5189-5200). PMLR..

[67] Tao L, Varshney V, Li Y. Benchmarking machine learning models for polymer informatics: an example of glass transition temperature. Journal of Chemical Information and Modeling. 2021 Oct 18;61(11):5395-413.

[68] Lara-Benítez P, Carranza-García M, Riquelme JC. An experimental review on deep learning architectures for time series forecasting. International journal of neural systems. 2021 Mar 16;31(03):2130001..

[69] Ozbayoglu AM, Gudelek MU, Sezer OB. Deep learning for financial applications: A survey. Applied soft computing. 2020 Aug 1;93:106384.

[70] Luo Q, Zeng W, Chen M, Peng G, Yuan X, Yin Q. Self-Attention and Transformers: Driving the Evolution of Large Language Models. In2023 IEEE 6th International Conference on Electronic Information and Communication Technology (ICEICT) 2023 Jul 21 (pp. 401-405). IEEE.

[71] Nunes M, Gerding E, McGroarty F, Niranjan M. Artificial neural networks in fixed income markets for yield curve forecasting. Available at SSRN 3144622. 2018 Mar 20.

[72] Júnior DS, de Oliveira JF, de Mattos Neto PS. An intelligent hybridization of ARIMA with machine learning models for time series forecasting. Knowledge-Based Systems. 2019 Jul 1;175:72-86.

[73] Naqvi A. Artificial intelligence for asset management and investment: a strategic perspective. John Wiley & Sons; 2021 Feb 9.

[74] Klimberg RK, Sillup GP, Boyle KJ, Tavva V. Forecasting performance measures–what are their practical meaning?. InAdvances in business and management forecasting 2010 Nov 17 (pp. 137-147). Emerald Group Publishing Limited.

[75] Liemohn MW, Shane AD, Azari AR, Petersen AK, Swiger BM, Mukhopadhyay A. RMSE is not enough: Guidelines to robust data-model comparisons for magnetospheric physics. Journal of Atmospheric and Solar-Terrestrial Physics. 2021 Jul 1;218:105624.

[76] Mao W, Wang W, Jiao L, Zhao S, Liu A. Modeling air quality prediction using a deep learning approach: Method optimization and evaluation. Sustainable Cities and Society. 2021 Feb 1;65:102567.

[77] Kim HY, Won CH. Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. Expert Systems with Applications. 2018 Aug 1;103:25-37.

[78] Adegboye A, Kampouridis M. Machine learning classification and regression models for predicting directional changes trend reversal in FX markets. Expert Systems with Applications. 2021 Jul 1;173:114645.

[79] Reed MS, Ferré M, Martin-Ortega J, Blanche R, Lawford-Rolfe R, Dallimer M, Holden J. Evaluating impact from research: A methodological framework. Research Policy. 2021 May 1;50(4):104147.

[80] Zhou J, Gandomi AH, Chen F, Holzinger A. Evaluating the quality of machine learning explanations: A survey on methods and metrics. Electronics. 2021 Mar 4;10(5):593.

[81] Hewamalage H, Ackermann K, Bergmeir C. Forecast evaluation for data scientists: common pitfalls and best practices. Data Mining and Knowledge Discovery. 2023 Mar;37(2):788-832.

[82] Chen Y, Lin W, Wang JZ. A dual-attention-based stock price trend prediction model with dual features. IEEE Access. 2019 Oct 8;7:148047-58.

[83] Dufays A, Rombouts JV. Relevant parameter changes in structural break models. Journal of Econometrics. 2020 Jul 1;217(1):46-78.

[84] Hyndman RJ, Ahmed RA, Athanasopoulos G, Shang HL. Optimal combination forecasts for hierarchical time series. Computational statistics & data analysis. 2011 Sep 1;55(9):2579-89.

[85] Leung CK, Thulasiram RK, Bondarenko DA. An efficient system for detecting outliers from financial time series. InFlexible and Efficient Information Handling: 23rd British National Conference on Databases, BNCOD 23, Belfast, Northern Ireland, UK, July 18-20, 2006. Proceedings 23 2006 (pp. 190-198). Springer Berlin Heidelberg.

[86] Cuevas-Tello JC, González-Grimaldo RA, Rodríguez-González O, Pérez-González HG, Vital-Ochoa O. Parallel approach for time series analysis with general regression neural networks. Journal of applied research and technology. 2012 Apr;10(2):162-79.

[87] Ahuja A, Al-Zogbi L, Krieger A. Application of noise-reduction techniques to machine learning algorithms for breast cancer tumor identification. Computers in Biology and Medicine. 2021 Aug 1;135:104576.

[88] Rady EH, Fawzy H, Fattah AM. Time series forecasting using tree based methods. J. Stat. Appl. Probab. 2021;10(1):229-44.

[89] Patil PR, Parasar D, Charhate S. A literature review on machine learning techniques and strategies applied to stock market price prediction. Emerging Trends in Data Driven Computing and Communications: Proceedings of DDCIoT 2021. 2021:121-35.

[90] Souma W, Vodenska I, Aoyama H. Enhanced news sentiment analysis using deep learning methods. Journal of Computational Social Science. 2019 Jan 1;2(1):33-46.

[91] Ray P, Reddy SS, Banerjee T. Various dimension reduction techniques for high dimensional data analysis: a review. Artificial Intelligence Review. 2021 Jun;54(5):3473-515.

[92] Zebari R, Abdulazeez A, Zeebaree D, Zebari D, Saeed J. A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. journal of applied science and technology trends. 2020 May 15;1(1):56-70.

[93] Furht B, Escalante A, editors. Handbook of data intensive computing. Springer Science & Business Media; 2011 Dec 9.

[94] Simsek MU, Yildirim Okay F, Ozdemir S. A deep learning-based CEP rule extraction framework for IoT data. The Journal of Supercomputing. 2021 Aug;77(8):8563-92.

[95] Chromik M, Schuessler M. A Taxonomy for Human Subject Evaluation of Black-Box Explanations in XAI. Exss-atec@ iui. 2020 Mar 17;1:1-7.

[96] Shah V, Konda SR. Neural Networks and Explainable AI: Bridging the Gap between Models and Interpretability. INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY. 2021 Jun 30;5(2):163-76.