ITJ | The International Tax Journal

Machine learning models in financial econometrics: A critical assessment

Dr. Padmalosani Dayalan [1], **Dr. Anju Lata Gajpal** [2], **Dr Soumya R** [3], **Dr. P. Kannaiah** [4], **Renukhadevi M** [5], and **Abhishek Goswami** [6]

[1] Faculty, College of Economics and Business Administration, University of Technology and Applied Sciences, Ibra, Sultanate of Oman, Email: d.padmalosani@gmail.com
[2] Kushabhau Thakre, Patrakarita Avam, Janshanchar Vishwavdyalaya, Raipur, Chhattisgarh – 492013, Email: anjugajpal@gmail.com
[3] Associate Professor, Department of Commerce and Management, Surana College Autonomous, South End Campus, Bangalore 560004, Email: soumyasrinath2011@gmail.com
[4] Faculty Member, Institute of Cooperative Management, Hyderabad, Telangana, India, Email: Dr.p.kannaiah@gmail.com
[5] Assistant Professor, First year Engineering, Dr. D. Y. Patil Institute of Technology, Pimpri, Maharashtra – 411018, Email: renukadevi.m@dypvp.edu.in
[6] Faculty, National Institute of Rural Development & Panchayati Raj, Hyderabad, Telangana, India, Email: abhishek.nird@gmail.com

Abstract---The present paper is a critical assessment of the application of machine learning (ML) models in financial econometrics. Even though machine learning models, including random forests, gradient boosting, support vectors machines, and deep learning, have demonstrated to exhibit high predictive power in various domains, such as asset pricing, risk management, credit scoring and volatility forecasting, their application results in serious methodological and practical problems. We compare the benefits of ML methods with old econometric methods, which primarily include the flexibility of the method, nonlinear modelling, and high dimension data processing. Meanwhile, we also present issues of interpretability, overfitting, data-snooping bias, and the conflict between prediction accuracy and economic theory. In developing the argument based on the available empirical evidence and theoretical standpoints, we contend that ML is not a replacement to econometrics, but rather a supplementary tool to econometrics in combination with structural modelling and economic intuition. At the end of the paper, the future directions are discussed, which are explainable AI, hybrid modelling frameworks, and incorporation of domain knowledge to make financial applications more reliable and policy relevant.

How to Cite:
Dayalan, P., Gajpal, A. L., Soumya, R., Kannaiah, P., Renukhadevi, M., & Goswami, A. (2025). Machine learning models in financial econometrics: A critical assessment. *The International Tax Journal*, *52*(5), 1688–1696. Retrieved from https://internationaltaxjournal.online/index.php/itj/article/view/181

The International tax journal ISSN: 0097-7314 E-ISSN: 3066-2370 © 2025

Submitted: 27 July 2025 | Revised: 18 August 2025 | Accepted: 09 September 2025

Keywords---Machine Learning; Financial Econometrics; Predictive Modelling; Asset Pricing; Risk Management; Volatility Forecasting; Model Interpretability; Explainable AI; Hybrid Models.

1. Introduction

Statistical and econometric models have long been used in financial econometrics to analyse asset pricing, volatility and risk. Linear regression, autoregressive conditional heteroskedasticity (ARCH/GARCH) models and vector autoregressions (VARs) are examples of traditional methods that offer interpretable forms with economic theory underpinnings [1]. The rapid pace of financial market evolution, together with the availability of high frequency and high dimensional data has, however, exposed the limitations of conventional econometric techniques, and in their capacity to address nonlinearities, complex interactions, and big data.

Machine learning (ML) is a potential alternative, and it offers flexible methods of pattern recognition, predictive modelling and data-driven decision making. Techniques that have been actively used in the assessment of credit risks, allocation in portfolio, prediction of volatility, and algorithmic trading include random forests, gradient boosting machines, support vector machines and deep neural networks [2]. They are robust on predictive accuracy and response to a wide variety of forms of financial data, both structured numeric time series and unstructured text and sentiment.

Irrespective of these developments, the use of ML in financial econometrics poses serious problems. The problem with most ML models is that they are black box and this limits the interpretability and theory underpinning. The use of critical analysis is contributed by the possibility of overfitting, data-snooping bias and trade-off between prediction and inference. Furthermore, statistical regularities are not the only factors that determine the behaviour of financial markets, but economic behaviour, institutional structure and regulatory environment that could not be factored into purely data-driven models [3].

It is three things that brought this paper to the critical assessment of ML models in financial econometrics. We will start by looking at the methodology developments and the empirical successes of machine learning in finance. Second, we indicate the limitations and risks of their use, particularly in such cases, where interpretability and theoretical consistency are the requirements [4]. Third, we present our conclusions about the future research directions with the promise of having hybrid models that will combine machine learning and econometric theory with the domain knowledge. Thus, we wish to position machine learning as a complement, and not a substitute, to the analysis of financial markets which what econometrics can offer.

2. Literature reviews

The literature may be grouped into four general themes: asset pricing, volatility prediction, risk management and credit grade, and methodological criticism. Recent research shows that ML methods are more predictive of asset returns as compared to the older econometric models. Ghosh, I., Sanyal, M. K., & Jana, R. K. (2017)[5] trained deep learning and regularization on many firm characteristics and found significant out-of-sample predictive improvements over linear models. In a similar tone, Jia, F., & Yang, B. (2021)[6] highlight that machine learning techniques could reveal nonlinear and interaction-based asset pricing at the expense of understanding.

For volatility modeling, random forests, long short-term memory (LSTM) networks and support vector regression were some of the widely studied ML methods to enhance forecasting accuracy. Research by Kou, G., Chao, X., Peng, Y., Alsaadi, F. E., & Herrera-Viedma, E. (2019)[7] recommends that the deep neural networks would be able to replicate intricate market dynamics that GARCH-type models would not have been able to replicate. Nonetheless, there is still conflicting evidence as certain studies show that the marginal gains in forecasting are possibly unlikely to appear in economically substantial benefits.

ML techniques have become very popular in credit risk modeling in academia and industry. Addo, Li, X., & Tang, P. (2020)[8] demonstrate that ensemble techniques and neural networks have a superior ability in classifying the occurrence of defaults than logistic regression models. Trends and Challenges of Management Education in India has been used in value-at-risk (VaR) estimation to estimate tail risk [9]. However, the issue of regulatory risk in terms of transparency and model validation remains high-ranking.

Although machine learning has such benefits as predictive, an expanding literature has criticized its shortcomings in financial econometrics. The authors warn against the notion of using predictive performance as an alternative to causal inference. Economic processes influence financial markets and models that lack theoretical basis are prone to data-snooping and random findings. To address this issue, the hybrid approaches that combine the econometric theory with ML approaches have been proposed. To illustrate, Müller, O., Fay, M., & Brocke, J. V. (2018a)[10] suggested structural machine learning models in which economic constraints are considered in predictive models that would strike a compromise between interpretability and flexibility.

The potential of machine learning in financial econometrics is not the only thing represented in the literature. Even though the empirical literature indicates an enhancement in the predictive accuracy of assets pricing, volatility forecasting, and risk management, concerns on interpretability, robustness and economic relevance have been on the frontline [11]. The new pact is that ML cannot replace econometric approaches particularly where transparency and theoretical insight is required.

3. Methodology

The research methodology will be created to critically analyze the role of machine learning (ML) in financial econometrics by synthesizing the organized literature review and assessing it based on the conceptual framework in fig.1. Instead of undertaking empirical modelling, this paper focuses more on comparative analysis of methodological strengths, limitations and application in financial world.

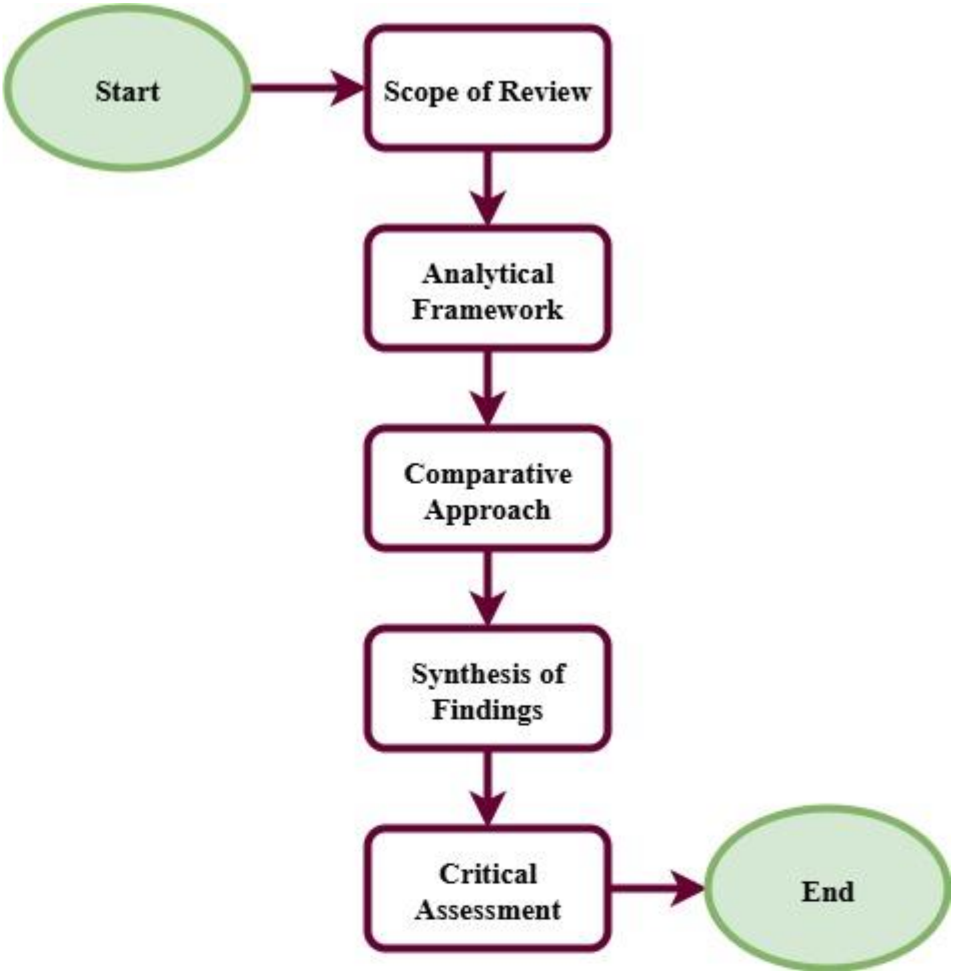

Fig. 1 Methodology Flowchart

Scope of Review

We pinpoint and screen a sample of studies that use machine learning to predict assets prices, volatility, credit risk modeling, and risk management. Peer-reviewed journal articles, working papers and official industry reports published in the period between 2010 and 2025 were selected as it was necessary to cover both the seminal contributions and the latest developments [12].

Analytical Framework

The evaluation can be informed by three criteria:

Predictive Performance: We compare evidence on the out of-sample predictive power of ML to that of traditional econometric models, including high-dimensional setting and nonlinear relationships.

Interpretability and Theoretical Consistency: We explore the consistency with or inconsistency of ML models with existing economic and econometric theory, with the trade-off between predictive models and explanatory power.

Practical and Regulatory Issues: We examine the viability of the adoption of ML in financial practice with respect to concerns of robustness, overfitting, transparency, and regulatory compliance.

Comparative Approach

To provide systematic evaluation, the paper contrasts classes of ML models, including tree-based techniques, support vector machine, neural networks and ensemble-based methods, to existing econometric tools, including OLS, VAR and GARCH. It is a qualitative comparison, which focuses on the differences in methods and implications on inference and decision-making.

Synthesis of Findings

The synthesis of findings across areas by the methodology points to both cross-domain (e.g., prediction vs. inference trade-offs, challenges of interpretability) and domain-specific (e.g., credit scoring vs. asset pricing) results [13]. The following systematic synthesis offers the foundation of the critical discussion and research agenda. In table 1, illustrates, Machine Learning vs. Econometric Models in Financial Applications with all the aspects.

Table 1. Machine Learning vs. Econometric Models in Financial Applications

Aspect	Machine Learning Models	Econometric Models
Objective	Focus on prediction and pattern recognition	Focus on inference, causality, and theory testing
Data Handling	Handle high-dimensional, nonlinear, and unstructured data	Suited for structured, low-to-medium dimensional data
Model Types	Random Forests, Gradient Boosting, SVM, Neural Networks	OLS, VAR, ARCH/GARCH, Cointegration models
Interpretability	Often "black box"; limited transparency	High interpretability; grounded in economic theory
Flexibility	Very flexible; captures complex relationships	Rigid functional forms; assumes linearity or stationarity
Overfitting Risk	High (requires regularization, cross-validation)	Lower if models are correctly specified
Regulatory Acceptance	Limited (due to opacity and validation challenges)	Widely accepted and trusted in policy/regulatory use
Forecasting Performance	Often superior in high-frequency and nonlinear environments	Performs well in structured and stable settings
Computational Demand	High (requires advanced algorithms and hardware)	Relatively low; computationally efficient
Best Use Cases	Asset pricing, credit risk, portfolio optimization	Policy analysis, risk modeling, structural inference

4. Results and discussion

A survey of the literature shows a less obvious perspective of how machine learning (ML) can be applied in the field of financial econometrics. Despite its excellent predictive power in certain cases, the weaknesses of ML emphasise the necessity of supplementing it and not replacing the conventional econometric methodologies [14].

The predictive power of the ML models has been found to be observed as the same trend among the literature, the only difference being that it is more predictive of the nonlinear and high-dimensional data. The neural networks and the regularization methods have proved that they can isolate useful signals in extremely large collections of firm characteristics in pricing of assets. Similarly, in credit risk modeling, ensemble methods outperform logistic regression in default classification tasks. These gains clarify why ML is clearly more useful in prediction-based applications.

Despite these improvements, the magnitude of predictive gains varies. ML architectures such as LSTMs and support vector regression have been found to offer small improvements in volatility prediction compared to GARCH models but with excessive added complexity to pay off these small gains. Predictive ability in back tests is not, moreover, economically significant advantage, and this raises questions of overfitting and resilience in live markets.

Low interpretability of ML models is one of the frequent issues. Despite the fact that the tree-based methods may be partially transparent, the deep learning methods are predominantly black-box. The opportunity to compromise between flexibility and interpretability lies in promising hybrid methods where economic constraints are included in the ML models. In practice, the implementation of ML illustrates the issue of data quality, the expense of computation, and validation of the models. Regulators and financial institution are still terrified to adopt opaque models in decision making particularly in risk sensitive areas, which are credit scoring and capital adequacy testing. Due to this reason, a trend of increasing popularity of explorable AI (XAI) tools is growing and the idea is that such tools should be capable of providing interpretable model outputs and at the same time be predictive.

There is a growing conclusion that ML should be used as a supplement to the econometric models, which is being pointed out in the discussion. ML is optimal in prediction-based tasks such as pricing assets or building portfolios, as it discloses very nonlinear associations. Tasks that require inferencing, where causality and theory matter, cannot yet be done without conventional econometrics. The future of financial econometrics is likely to be the hybrid framework that can have the predictive ability of ML and explainability and theoretical rigor of econometrics. In fig.2 to 4 representations such as Numerical Data for Scale: 1 = Low, 5 = High, Performance Metrics on Scale 1-5, Application for Suitability (Scale 1-5).

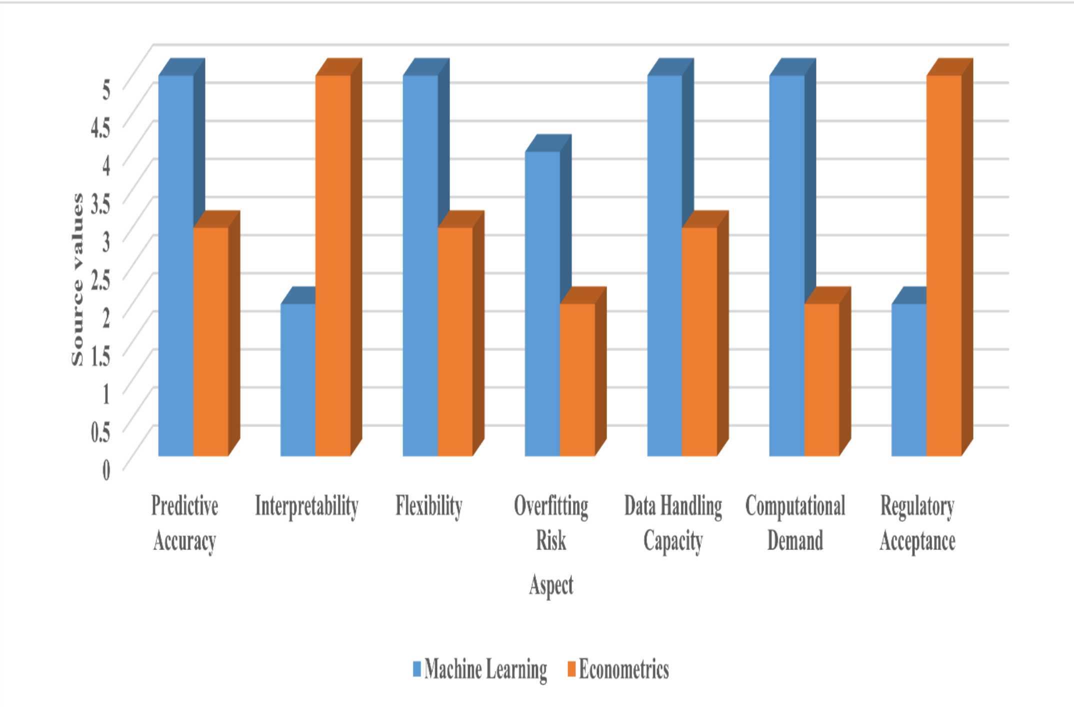

Fig. 2 Numerical Data (Scale: 1 = Low, 5 = High)

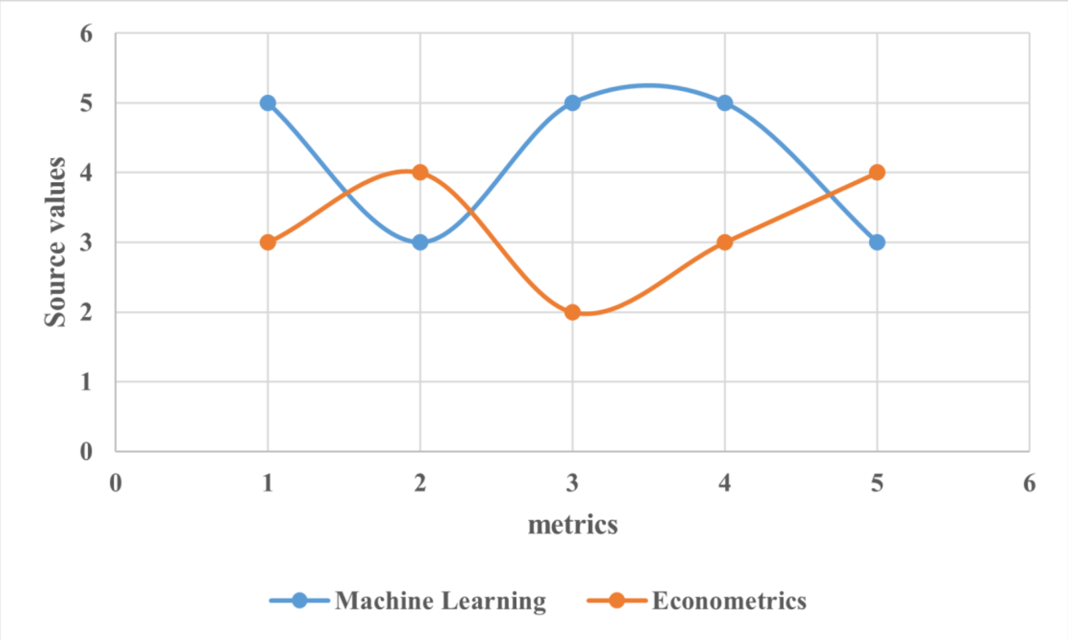

Fig. 3 Performance Metrics (Scale 1–5)

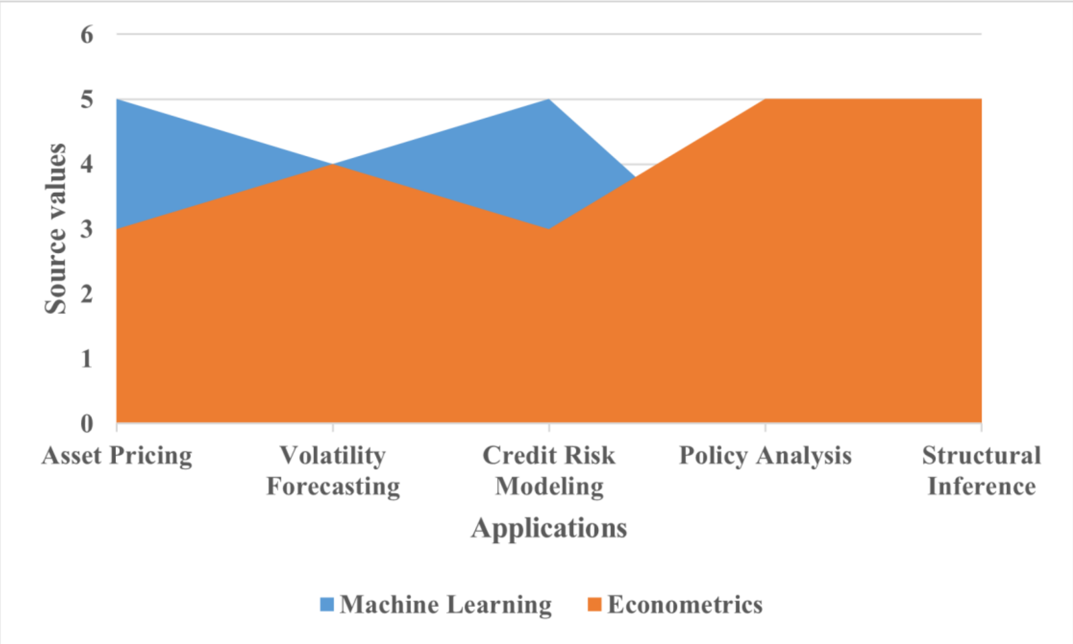

Fig. 4 Application Suitability (Scale 1–5)

5. Conclusion

The purpose of this paper was to evaluate the role of machine learning in financial econometrics, identify its strong and weak sides. It has been stated that the predictive advantages of the ML techniques in the application are significant in asset pricing, the prediction of volatility and credit risk modeling where big, intricate multi-dimensional data sets are considered. These capabilities position ML as a valuable addition to the financial econometric toolkit.

However, the main impediments are the interpretability, overfitting and economic theory-congruency issues. ML can uncover trends, which are overlooked with conventional models, but it may also lack transparency and causality which is required to infer, analyze the policy, and adhere to regulations. In practice, the threat of data-snooping bias and the unfeasibility of testing black-box models restrict their broader application in financial decision-making.

The findings validate the notion that machine learning should not be considered as a replacement of the econometric procedures but as an augmentation. These are the most promising views of the future in the context of hybrid models of the predictive flexibility and the theoretical sketch and design of the exploratory artificial intelligence approaches that could enhance the trust and responsibility. As financial data continues to become multifaceted and scale-driven, machine learning coupled with econometrics may offer an opportunity to get even closer to more precise, solid, and theoretically well-informed information on financial markets.

References

1. Ampomah, E. K., Qin, Z., & Nyame, G. (2020). Evaluation of Tree-Based Ensemble Machine Learning Models in predicting stock Price direction of movement. Information, 11(6), 332. https://doi.org/10.3390/info11060332
2. Cerniglia, J. A., & Fabozzi, F. J. (2020). Selecting Computational Models for Asset Management: Financial Econometrics versus Machine Learning—Is There a Conflict? The Journal of Portfolio Management, 47(1), 107–118. https://doi.org/10.3905/jpm.2020.1.184
3. Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., . . . Williams, M. D. (2019). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002
4. Ghosh, I., Jana, R. K., & Sanyal, M. K. (2019). Analysis of temporal pattern, causal interaction and predictive modeling of financial markets using nonlinear dynamics, econometric models and machine learning algorithms. Applied Soft Computing, 82, 105553. https://doi.org/10.1016/j.asoc.2019.105553
5. Ghosh, I., Sanyal, M. K., & Jana, R. K. (2017). Analysis of causal interactions and predictive modelling of financial markets using econometric methods, Maximal overlap Discrete wavelet Transformation and Machine Learning: A Study in Asian context. In Lecture notes in computer science (pp. 664–672). https://doi.org/10.1007/978-3-319-69900-4_84
6. Jia, F., & Yang, B. (2021). Forecasting Volatility of Stock Index: Deep Learning Model with Likelihood-Based Loss Function. Complexity, 2021(1). https://doi.org/10.1155/2021/5511802
7. Kou, G., Chao, X., Peng, Y., Alsaadi, F. E., & Herrera-Viedma, E. (2019). Machine Learning Methods For Systemic Risk Analysis In Financial Sectors. Technological and Economic Development of Economy, 25(5), 716–742. https://doi.org/10.3846/tede.2019.8740
8. Li, X., & Tang, P. (2020). Stock index prediction based on wavelet transform and FCD-MLGRU. Journal of Forecasting, 39(8), 1229–1237. https://doi.org/10.1002/for.2682
9. Siddi, Someshwar, Trends and Challenges of Management Education in India (April 20, 2022). International Journal Of Special Education Vol.37, No.3, 2022, https://ssrn.com/abstract=4088573
10. Müller, O., Fay, M., & Brocke, J. V. (2018a). The Effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics. Journal of

Management Information Systems, 35(2), 488–509. https://doi.org/10.1080/07421222.2018.1451955

11. Müller, O., Fay, M., & Brocke, J. V. (2018b). The Effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics. Journal of Management Information Systems, 35(2), 488–509. https://doi.org/10.1080/07421222.2018.1451955
12. Petropoulos, A., & Siakoulis, V. (2021). Can central bank speeches predict financial market turbulence? Evidence from an adaptive NLP sentiment index analysis using XGBoost machine learning technique. Central Bank Review, 21(4), 141–153. https://doi.org/10.1016/j.cbrev.2021.12.002
13. Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2016). A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. Tourism Management, 58, 51–65. https://doi.org/10.1016/j.tourman.2016.10.001
14. Xie, M., Zhao, S., & Lv, K. (2024). The impact of green finance and financial technology on regional green energy technological innovation based on the dual machine learning and spatial econometric models. Energies, 17(11), 2521. https://doi.org/10.3390/en17112521