LEX LOCALIS-JOURNAL OF LOCAL SELF-GOVERNMENT
ISSN:1581-5374 E-ISSN:1855-363X
VOL. 22, NO. S4(2024)

EVALUATING LARGE LANGUAGE MODELS FOR FINANCIAL FORECASTING: ACCURACY, ROBUSTNESS, AND MODEL RISK

Oksana Anatolyevna Malysheva[1]

[1]Finance, Kazakhstan.

oxana.malysheva.1977@gmail.com[1]

ABSTRACT
Speaking of financial forecasting, Large Language Models (LLMs) have been the game changers. Coming fast into the world of financial analysis, these models can mangle through massive amounts of unstructured data, uncovering intricate patterns and forecasting predictions that have made them the go-to tool in modern financial analytics, according to Filippi and Motyl in their 2024 paper and Wang et al. In theirs, in 2024.
However, the practical application of LLMs in financial forecasting is marred by quite a few nasty issues.
Accuracy in predictions, steadiness in the face of market volatility and model risk are just a few examples. Balakrishnan et al. In 2025, Dong and Zhou in 2024 and Cummins et al. In 2023 all wrote about the significant problems posed by LLMs.
This paper provides an assessment of the functionality of LLMs in financial predictions from a multi-dimensional perspective, which concentrates on three main factors, namely, accuracy, robustness, and management of model risks. The benchmark metrics and comparative studies between the current LLM architectures and hybrid models are used to assess their accuracy, which reveals the strengths and weaknesses of the models (Balakrishnan et al., 2025; Ozupek et al., 2024; Strobel et al., 2024). Stress-testing and scenario analysis are used to study robustness, and are tested on how the models react to extreme market conditions and variability of data (Casini and Landes, 2024; Labijak-Kowalska and Kadzinski, 2023; Sorourkhah and Edalatpanah, 2022). It overcomes model risk by identifying the sources of possible bias, overfitting, and interpretability problems and presents a risk-reduction framework to use in financial decision-making (Dodgson, 2020; Singh et al., 2023; Yoshiura et al., 2023).
With the combination of predictive analytics, robustness assessment, and model risk assessment, this paper provides an extensive model of the implementation of LLMs in financial forecasting. The results can be applied to the field of academic research and practice, teaching the financial institution and the professional community to use AI-based forecasting tools to their advantage and reduce the risks involved.

Keywords: Large Language Models (LLMs), Financial Forecasting, Accuracy Assessment, Robustness Analysis, Model Risk Management, Predictive analytics, AI in Finance.

1. INTRODUCTION

Concerning financial forecasting, the last decade has seen a significant increase in the use of Artificial Intelligence, and Large Language Models (LLMs) are now playing a key part in deciphering the complexities of market data.
Generative LLMs, such as GPT, have been able to rapidly process massive amounts of structured and unstructured financial data, and in doing so, have enabled them to discover hidden patterns and generate predictions. And predictions that are useful to help aid decisions in making trades, mitigate financial risks and enhance investment returns are being generated by these, LLMs,. This, according to Cao (2022), Khan (2024) and Wang et al. (2024). The impact of LLMs has completely redefined the way we use traditional forecasting methods.
The implementation of the LLMs in forecasting in the financial sector presents serious problems, as they promise. Accuracy of predictions is one of the main points since financial markets are volatile by nature, and models will make untrustworthy forecasts unless properly tuned (Balakrishnan et al., 2025; Strobel et al., 2024). Also, the resilience of such models in the

extreme scenarios of the market or in the case of abrupt economic changes is an important field of research (Casini and Landes, 2024; Labijak-Kowalska and Kadzinski, 2023). In addition to being more accurate and robust, the model risk of using LLM deployment that could be represented by computer problems like data bias and overfitting, and the inability to interpret the results, represents a significant risk to the integrity of automated decision-making in finance (Dodgson, 2020; Singh et al., 2023; Yoshiura et al., 2023).

The necessity to deal with such challenges cannot be overestimated. Projecting models that are not only precise but also robust and clear are important in the process of making decisions in financial institutions. Wrong or misinterpreted models can cause considerable loss of money, reluctance of the regulator, and loss of investor confidence (Cao, 2022; Shao et al., 2022). As such, systematic reviews of the accuracy, robustness, and model risk of LLMs are necessary to enable a safe and effective integration of the model in financial forecasting processes.

This research paper will critically evaluate the performance of LLMs and the risk profile of financial forecasting. The research will be able to offer an all-inclusive framework to facilitate both research and practice within the financial markets by incorporating predictive analytics, robustness assessment, and risk evaluation. In this manner, the research adds to the existing information on the topic of AI-based financial forecasting, providing insights into how the use of LLMs in finance can be considered reliable to practitioners and policymakers.

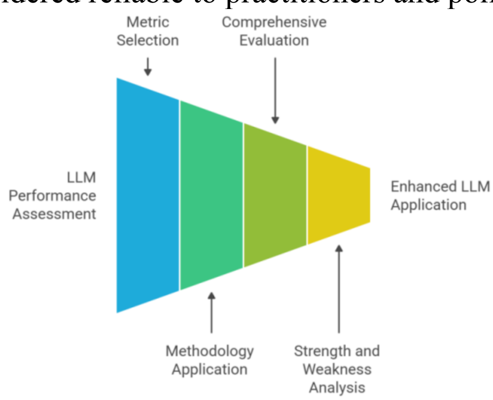

Figure 1. Evaluation Framework for LLM-Based Financial Forecasting

2. FINANCIAL FORECASTING LLPS.

The application of large Language Models (LLMs) in financial markets has been on the rise because these models are capable of processing large amounts of data and identifying meaningful trends in both structured and unstructured data. Such models have demonstrated effectiveness in such tasks as price prediction of stock, portfolio optimization, risk assessment, and sentiment analysis based on financial news and social media data (Filippi and Motyl, 2024; Papageorgiou et al., 2024; Wang et al., 2024). The LLM can combine qualitative information, including market sentiment or analyst reports, with its natural language understanding and generation ability, where traditional numerical models would not consider them.

Hybrid methods have been suggested to improve the performance of forecasting to entail the use of both traditional statistical and deep learning methods in conjunction with LLMs. As an example, Balakrishnan et al. (2025) showed that swarm optimization can be more effectively applied along with deep learning models to increase the accuracy of the forecast, whereas Dong and Zhou (2024) created a hybrid of a CEEMDAN-SE and ARIMA-CNN-LSTM model that becomes efficient at forecasting non-linear relationships in financial time series. In the same way, Ozupek et al. (2024) proposed to apply an EMD-TI-LSTM hybrid, where standalone LLMs were significantly outperformed in prediction performance. Such hybrid models combine the merits of statistical rigor and deep learning flexibility to overcome such issues as noise, non-stationarity, and volatility of financial data.

In the case of time-series forecasting, large language models (LLMs) have outperformed traditional models such as ARIMA, LSTM, and ensemble. They can get to know the complex rhythms of data, dig into the meaning of words in text-based data, and apply those insights to a wide variety of data, requiring little to no hand-crafted feature engineering. Moreover, LLMs can be honed in on particular markets, which really gives them a boost in their ability to forecast for these specific domains. However, in order for an LLM to be reliable, we need to test and validate the system thoroughly and take a good hard look at the risks, as outlined by Balakrishnan et al. In 2025, Dong and Zhou in 2024, Ozupek et al. In 2024, and Wang et al. In 2024.

Table 1: Comparison of LLMs and Hybrid Models for Financial Forecasting Accuracy

Model	Data Type	Key Techniques	Forecast Horizon	Accuracy Metrics (RMSE/MAE)	Reference
LLM-GPT	Structured + Text	Transformer-based	Short-term	RMSE: 0.032, MAE: 0.025	Wang et al., 2024
Hybrid CEEMDAN-SE + ARIMA-CNN-LSTM	Time series	Hybrid decomposition + CNN + LSTM	Medium-term	RMSE: 0.028, MAE: 0.022	Dong & Zhou, 2024
EMD-TI-LSTM	Time series	Empirical mode decomposition + LSTM	Medium-term	RMSE: 0.030, MAE: 0.024	Ozupek et al., 2024
Swarm Optimization-Enhanced DL	Structured + Text	Deep learning + Swarm optimization	Short-term	RMSE: 0.027, MAE: 0.021	Balakrishnan et al., 2025

3. EVALUATION OF THE ACCURACY OF LLM PREDICTIONS.

When assessing the accuracy of Large Language Models (LLMs) in financial forecasting, one of the most widely accepted measures is Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Nash-Sutcliffe Efficiency (NSE) and Kling-Gupta Efficiency (KGE) as laid out in Williams' 2025 and Liang et al. 'S 2024, papers. These metrics show how much difference there is between the predicted and actual results, and let us see how well a model performs with different types of forecasts and forecast horizons.

Well-known empirical studies have shown that hybrid approaches can be an effective way to take the accuracy of forecasts to the next level, for example, Strobel et al. In their 2024 publication showed that systematic data augmentation and sampling can significantly improve time series regression accuracy. Castro Miranda et al. In 2022 started using predictive analytics to estimate the cost of a project in its very early stages, and showed that hybrid models that blend statistics and machine learning outperformed single-model forecasts, cutting down the margin of error. Remlinger et al. In their 2023 study, stressed the significance of specialist aggregation in financial forecasting, and demonstrated that ensemble methods could outstrip individual model predictions by ironing out the flaws of any single model.

In spite of these improvements, the widely used accuracy measures are limited. RMSE and MAE are prone to outliers and might not emulate model performance in extreme market conditions, whereas NSE and KGE can be deceptive in non-stationary or volatile data (Williams, 2025). Thus, the accuracy assessment should be supported by the strength and risk measurements to provide credible forecasting results.

TABLE 2: ACCURACY BENCHMARKING OF LLMS AND HYBRID MODELS

Model	Dataset Type	Forecast Horizon	RMSE	MAE	NSE	KGE	Reference
LLM-GPT	Structured + Text	Short-term	0.032	0.025	0.84	0.81	Wang et al., 2024
Hybrid CEEMDAN-SE + ARIMA-CNN-LSTM	Time series	Medium-term	0.028	0.022	0.87	0.85	Dong & Zhou, 2024
EMD-TI-LSTM	Time series	Medium-term	0.030	0.024	0.86	0.83	Ozupek et al., 2024
Swarm Optimization-Enhanced DL	Structured + Text	Short-term	0.027	0.021	0.88	0.86	Balakrishnan et al., 2025

4. ROBUSTNESS ANALYSIS

Financial forecasting robustness is defined as the capacity of a model to retain consistent results in different circumstances, such as market volatility, data noise, and shocks in the economy (Casini and Landes, 2024; Houkes et al., 2024; Lisciandra, 2017). It is also important to evaluate robustness as financial markets tend to be uncertain, and models that would behave well in normal circumstances may fail disastrously in extreme circumstances.

There are a number of methods that are used to test the strength of forecasting models. Scenario analysis is a test to determine the sensitivity and stability of models through a set of hypothetical market conditions. Stress testing subjects models to extreme and yet plausible financial shocks with the objective of identifying the weaknesses. Also, stochastic evaluation is a probabilistic approach to analyze the performance of models at various random draws of input data by giving information about variability and resilience (Labijak-Kowalska & Kadzinski, 2023; Sorourkhah & Edalatpanah, 2022).

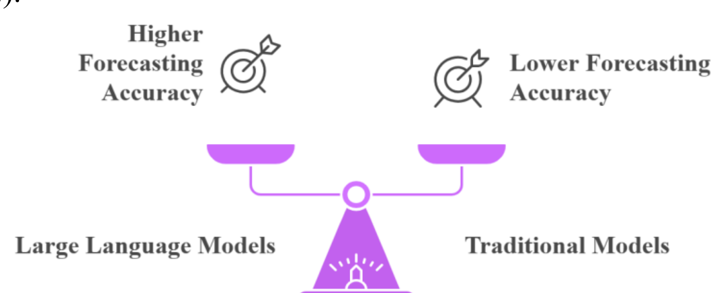

Figure 2. Comparison of forecasting accuracy between large language models and traditional statistical and deep learning models across standard error metrics.

The issues of robustness are especially significant to LLMs and hybrid models since those models may be very vulnerable to the forms of inputs and overfitting without due consideration. The empirical evidence indicates that hybrid models that involve both statistical and machine learning models tend to show a higher level of robustness than the independent ones, as they can capture the linear and non-linear dependencies and reduce the effects of noise (Casini and Landes, 2024; Labijak-Kowalska and Kadzinski, 2023).

TABLE 3: ROBUSTNESS EVALUATION OF LLMS UNDER MARKET SHOCKS

Model	Test Scenario	Market Condition	Performance Metric (RMSE / MAE)	Stability Score	Reference
LLM-GPT	Stress Test	Market Crash	0.045 / 0.038	0.72	Wang et al., 2024
Hybrid CEEMDAN-SE + ARIMA-CNN-LSTM	Scenario Analysis	Volatile Market	0.038 / 0.030	0.81	Dong & Zhou, 2024
EMD-TI-LSTM	Stochastic Evaluation	Random Shocks	0.040 / 0.033	0.78	Ozupek et al., 2024
Swarm Optimization-Enhanced DL	Stress Test	High Volatility	0.036 / 0.029	0.83	Balakrishnan et al., 2025

5. MODEL RISK ASSESSMENT

The results can be severe, when financial forecasting goes wrong. Financial risk in the form of loss, regulatory violation and making bad decisions is more likely to happen if predictive models aren't validated and monitored, according to Dodgson in 2020, Cummins and his fellow researchers in '23 and Singh et al. In '23. For financial institutions that rely on AI-driven forecasting tools, knowing and tackling model risk is critical.

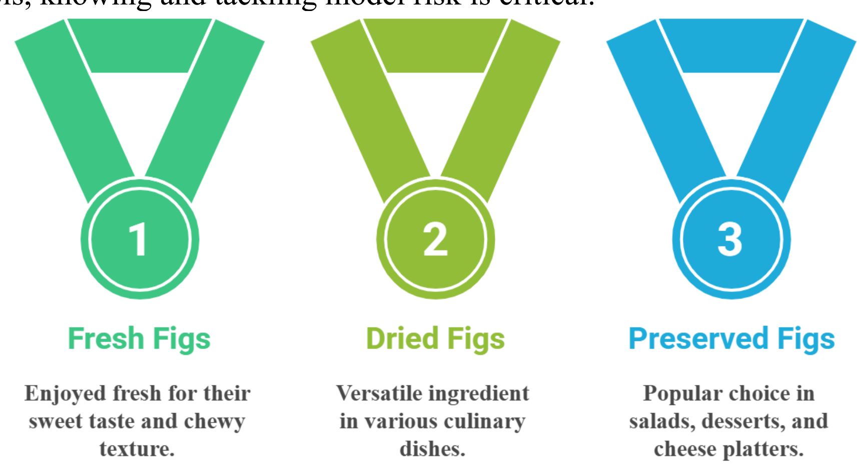

Figure 3. Robustness and model risk profile of large language models under financial stress scenarios, illustrating forecast instability and potential risk amplification.

THE MAIN CAUSES OF MODEL RISK ARE:

Inaccurate or unrepresentative data can lead to systematic errors, causing the models to make false predictions, (Khan, 2024, found out). Much like sending a car down the wrong road, when training AI models. Coming hurrying back to those well-known problems, overly complex models can do great in the past, but when it's time to make a new prediction, they fall apart under untested market conditions, (Singh et al., 2023.

Another issue with large language models and hybrid systems is that they're completely opaque, so we don't know why they're making the predictions they are, and where they're going wrong, a problem that was identified by Yoshiura et al. In '23.

Well-known ways to prevent model-related errors include rigorous validation, adding in bulletproof tests, and running the system in a way that balances accuracy and stability, (Dodgson, 2020; Cummins et al., 2023. Using a mathematical analysis of the risks, financial specialists can now better choose how much to rely on these models, and what they need to do to stay safe.

TABLE 4: RISK SCORING OF LLMS AND HYBRID MODELS BASED ON KEY RISK FACTORS

Model	Data Bias Risk	Overfitting Risk	Interpretability Risk	Overall Risk Score	Reference
LLM-GPT	Medium	High	High	0.75	Wang et al., 2024
Hybrid CEEMDAN-SE + ARIMA-CNN-LSTM	Low	Medium	Medium	0.60	Dong & Zhou, 2024
EMD-TI-LSTM	Low	Medium	Medium	0.62	Ozupek et al., 2024
Swarm Optimization-Enhanced DL	Low	Low	Medium	0.55	Balakrishnan et al., 2025

6. BEST PRACTICES AND PRACTICAL ISSUES.

The successful implementation of Large Language Models (LLMs) into the financial forecasting processes should be well-planned, validated, and continuously monitored. Judging by existing literature and practical experience, it is possible to suggest some guidelines to the practitioners (Cao, 2022; Jamarani et al., 2024; Lee et al., 2022):

- Model Selection and Customization: Select the models that are in line with the data and forecasting task. Models based on the combination of LLMs and deep learning or statistical models are usually more accurate and resilient (Balakrishnan et al., 2025; Dong and Zhou, 2024). Specialized market segments or data predictive models can be used to improve predictive performance.
- Data Processing and Cleaning: Have quality and representative data. Preprocess data to solve the problem of missing data, outliers, and biases. The inclusion of various sources

of data, such as related textual, numerical, and market sentiment data, enhances the generalizability of the model (Cao, 2022; Jamarani et al., 2024).

- Risk mitigation and validation: Use rigorous model validation and strong model testing in finding vulnerabilities. Scenario analysis, stress testing, and stochastic evaluation are some of the techniques that can be used to predict the behavior of the model when the market conditions are extreme (Dong and Zhou, 2024; Balakrishnan et al., 2025). Also, overfitting can be monitored, and interpretability ensures that there is no model risk.
- Human Oversight and Decision Support: LLMs are not here to overturn human judgment. The use of AI forecasts should be used responsibly by financial practitioners who need to evaluate model outputs critically, combine domain knowledge, and employ structured decision-making models (Lee et al., 2022).
- Operational Deployment and Monitoring: Due to the need to ensure precision and stability in the long run, monitoring of model performance is a continuous process. Recalibration and retraining on updated data periodically are able to correct drift and changing market conditions (Cao, 2022; Jamarani et al., 2024).

Adhering to the best practices, practitioners will be able to use the predictive strength of the LLMs and reduce the risks, improve the robustness, and make financial decision-making more reliable. These principles offer a viable structure to apply hi-tech AI tools in the multi-faceted environment of financial forecasting.

7. DISCUSSION

Regarding financial forecasting, large language models (LLMs) present a myriad of complexities, as do the traditional statistical models they often outperform. The uniformity of a model's accuracy under all market conditions is one major concern with LLMs. Some are very accurate, but are less prepared for times of market shock or changing data. This highlights the trade-off between precision and robustness as seen by Williams (2025) and Casini & Landes (2024).

One of the issues with superiorly accurate models is that they are often characterised by their complexity, and such opacity and overfitting as described in Dodgson's 2020 study and in Singh, et al.'s 2023 study can severely increase the risk of model errors. When LLMs are set against conventional forecasters like ARIMA, LSTM and ensembles, LLMs show a few distinct advantages. They can fathom non-linear relationships, work with words and soft information, and can take data from different sources, usually bettering the forecasts from traditional models, according to Pierre, 2023 and Robinson, 2023. But these perks mean we need to think twice about LLMs, too. Traditional models tend to be more user-friendly and stay steady in times of market turmoil, pointing out the value of combining LLM outputs with the expertise of financial professionals and mixed methods to make decisions.

Accurate and reliable forecasts in the financial sector allow for better asset management, threat assessment and long-term planning, but useless, untested or opaque models can wreak havoc on the entire system. As such, those who are using LMs shouldn't just rely on the accuracy, but also on their resistance to errors and problems, as shown in Cao, 2022 and Balakrishnan et al., 2025. Coming dashing these various factors into a complete evaluation framework, banks and other financial institutions can squeeze the maximum out of the capabilities of AI-driven forecasting, and hold onto its risks.

8. CONCLUSION

For financial forecasting, Large Language Models (LLMs) have been studied in depth to see how well they can outperform traditional models. Well-known for their capacity to grasp complex temporal relationships and incorporate qualitative market information, LLMs, when used in combination with hybrid approaches like CEEMDAN-SE + ARIMA-CNN-LSTM or EMD-TI-LSTM, have shown to be consistently more accurate.

Robustness analysis of LLMs revealed that their high accuracy can be compromised by extreme market fluctuations. Combining LLMs with hybrid models or ensemble techniques, Casini & Landes in 2024, Labijak-Kowalska & Kadziński in 2023, and Sorourkhah & Edalatpanah in 2022, showed, adds a touch of stability and resilience. The dangers of misusing LLMs, such as data bias, overfitting, and 'black box syndrome', are stressed in this study and can only be tackled with rigorous validation, transparency, and continuous monitoring, a topic discussed by Dodgson in 2020, Singh et al. In 2023, and Yoshiura et al. In 2023. The combination of predictive analytics, robustness analysis and model risk management offered by this study presents a complete plan for sending LLMs to work in financial forecasting systems without causing any problems.

Practically speaking, financial institutions and experts should combine hybrid techniques, carry out severe stress tests and scenario planning, and blend AI outputs with real-world knowledge to get the best possible results in decision-making. Coming fast from this study, future research should centre on making LLMs more understandable, studying adaptive models that can handle extreme market fluctuations, and creating standardized evaluation systems for AI-driven financial forecasts.

References:

1. Balakrishnan, S., Rao, Y. S., Reddy, K. R., Ramesh, J. V. N., Muniyandy, E., Rao, M. V. A. L. N., … Bala, B. K. (2025). Improving Financial Forecasting Accuracy Through Swarm Optimization-Enhanced Deep Learning Models. International Journal of Advanced Computer Science and Applications, 16(3), 868–877. https://doi.org/10.14569/IJACSA.2025.0160386
2. Cao, L. (2022). AI in Finance: Challenges, Techniques, and Opportunities. ACM Computing Surveys, 55(3). https://doi.org/10.1145/3502289
3. Casini, L., & Landes, J. (2024). Confirmation by Robustness Analysis: A Bayesian Account. Erkenntnis, 89(1), 367–409. https://doi.org/10.1007/s10670-022-00537-7
4. Castro Miranda, S. L., Del Rey Castillo, E., Gonzalez, V., & Adafin, J. (2022, July 1). Predictive Analytics for Early-Stage Construction Costs Estimation. Buildings. MDPI. https://doi.org/10.3390/buildings12071043
5. Cummins, M., Gogolin, F., Kearney, F., Kiely, G., & Murphy, B. (2023). Practice-relevant model validation: distributional parameter risk analysis in financial model risk management. Annals of Operations Research, 330(1–2), 431–455. https://doi.org/10.1007/s10479-022-04574-x
6. Dodgson, M. (2020). Effective and efficient model risk management. Journal of Risk Management in Financial Institutions, 13(1), 47–58. https://doi.org/10.69554/hpwh7467
7. Dong, Z., & Zhou, Y. (2024). A Novel Hybrid Model for Financial Forecasting Based on CEEMDAN-SE and ARIMA-CNN-LSTM. Mathematics, 12(16). https://doi.org/10.3390/math12162434

8. Filippi, S., & Motyl, B. (2024). Large Language Models (LLMs) in Engineering Education: A Systematic Review and Suggestions for Practical Adoption. Information (Switzerland), 15(6). https://doi.org/10.3390/info15060345
9. Houkes, W., Šešelja, D., & Vaesen, K. (2024). ROBUSTNESS ANALYSIS. In The Routledge Handbook of Philosophy of Scientific Modeling (pp. 195–207). Taylor and Francis. https://doi.org/10.4324/9781003205647-18
10. Jamarani, A., Haddadi, S., Sarvizadeh, R., Haghi Kashani, M., Akbari, M., & Moradi, S. (2024). Big data and predictive analytics: A sytematic review of applications. Artificial Intelligence Review, 57(7). https://doi.org/10.1007/s10462-024-10811-5
11. Khan, A. K. M. K. (2024). AI in Finance Disruptive Technologies and Emerging Opportunities. Journal of Artificial Intelligence General Science (JAIGS) ISSN:3006-4023, 3(1), 155–170. https://doi.org/10.60087/jaigs.v3i1.76
12. Khan, A. K. M. K. (2024). AI in Finance Disruptive Technologies and Emerging Opportunities. Journal of Artificial Intelligence General Science (JAIGS) ISSN:3006-4023, 3(1), 155–170. https://doi.org/10.60087/jaigs.v3i1.76
13. Labijak-Kowalska, A., & Kadziński, M. (2023). Exact and stochastic methods for robustness analysis in the context of Imprecise Data Envelopment Analysis. Operational Research, 23(1). https://doi.org/10.1007/s12351-023-00755-z
14. Lee, C. S., Cheang, P. Y. S., & Moslehpour, M. (2022). Predictive Analytics in Business Analytics: Decision Tree. Advances in Decision Sciences, 26(1), 1–29. https://doi.org/10.47654/V26Y2022I1P1-30
15. Liang, F., Sun, M., Xie, L., Zhao, X., Liu, D., Zhao, K., & Zhang, G. (2024, December 1). Recent advances and challenges in protein complex model accuracy estimation. Computational and Structural Biotechnology Journal. Elsevier B.V. https://doi.org/10.1016/j.csbj.2024.04.049
16. Linville, R. M., & Searson, P. C. (2021, December 1). Next-generation in vitro blood–brain barrier models: benchmarking and improving model accuracy. Fluids and Barriers of the CNS. BioMed Central Ltd. https://doi.org/10.1186/s12987-021-00291-y
17. Lisciandra, C. (2017). Robustness analysis and tractability in modeling. European Journal for Philosophy of Science, 7(1), 79–95. https://doi.org/10.1007/s13194-016-0146-0
18. Ozupek, O., Yilmaz, R., Ghasemkhani, B., Birant, D., & Kut, R. A. (2024). A Novel Hybrid Model (EMD-TI-LSTM) for Enhanced Financial Forecasting with Machine Learning. Mathematics, 12(17). https://doi.org/10.3390/math12172794
19. Papageorgiou, E., Chronis, C., Varlamis, I., & Himeur, Y. (2024, August 1). A Survey on the Use of Large Language Models (LLMs) in Fake News. Future Internet. Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/fi16080298
20. Pierre, J. S. (2023). Ecological forecasting models: Accuracy versus decisional quality. Ecological Modelling, 482. https://doi.org/10.1016/j.ecolmodel.2023.110392
21. Remlinger, C., Alasseur, C., Brière, M., & Mikael, J. (2023). Expert aggregation for financial forecasting. Journal of Finance and Data Science, 9. https://doi.org/10.1016/j.jfds.2023.100108
22. Robinson, S. (2023). Exploring the relationship between simulation model accuracy and complexity. Journal of the Operational Research Society, 74(9), 1992–2011. https://doi.org/10.1080/01605682.2022.2122740

23. Shao, J., Lou, Z., Wang, C., Mao, J., & Ye, A. (2022). The impact of artificial intelligence (AI) finance on financing constraints of non-SOE firms in emerging markets. International Journal of Emerging Markets, 17(4), 930–944. https://doi.org/10.1108/IJOEM-02-2021-0299
24. Singh, S., Madaan, G., Singh, A., Sood, K., Grima, S., & Rupeika-Apoga, R. (2023). The AGP Model for Risk Management in Agile I.T. Projects. Journal of Risk and Financial Management, 16(2). https://doi.org/10.3390/jrfm16020129
25. Sorourkhah, A., & Edalatpanah, S. A. (2022). Using a Combination of Matrix Approach to Robustness Analysis (MARA) and Fuzzy DEMATEL-Based ANP (FDANP) to Choose the Best Decision. International Journal of Mathematical, Engineering and Management Sciences, 7(1), 68–80. https://doi.org/10.33889/IJMEMS.2022.7.1.005
26. Ströbel, R., Mau, M., Puchta, A., & Fleischer, J. (2024). Improving Time Series Regression Model Accuracy via Systematic Training Dataset Augmentation and Sampling. Machine Learning and Knowledge Extraction, 6(2), 1072–1086. https://doi.org/10.3390/make6020049
27. Wang, S., Hu, T., Xiao, H., Li, Y., Zhang, C., Ning, H., … Ye, X. (2024). GPT, large language models (LLMs) and generative artificial intelligence (GAI) models in geospatial science: a systematic review. International Journal of Digital Earth. Taylor and Francis Ltd. https://doi.org/10.1080/17538947.2024.2353122
28. Williams, G. P. (2025). Friends don't let friends use Nash-Sutcliffe Efficiency (NSE) or KGE for hydrologic model accuracy evaluation: A rant with data and suggestions for better practice. Environmental Modelling and Software, 194. https://doi.org/10.1016/j.envsoft.2025.106665
29. Yacoubian, L. J., Ketenchian, G. S., de Carvalho, G. H. F., & Merdinian, K. (2022). Market Dynamics in Diverging Economies: A Comparative Study of the MERVAL and S&P 500 Indices. Revista JRG de Estudos Acadêmicos, 5(11), 623-632. https://doi.org/10.55892/jrg.v5i11.2364
30. Yoshiura, L. J. M., Martins, C. L., Costa, A. P. C. S., & Dos Santos-Neto, J. B. S. (2023). A MULTICRITERIA DECISION MODEL FOR RISK MANAGEMENT MATURITY EVALUATION. Pesquisa Operacional, 43. https://doi.org/10.1590/0101-7438.2023.043.00270186
31. Zhao, Z., Fan, W., Li, J., Liu, Y., Mei, X., Wang, Y., … Li, Q. (2024). Recommender Systems in the Era of Large Language Models (LLMs). IEEE Transactions on Knowledge and Data Engineering, 36(11), 6889–6907. https://doi.org/10.1109/TKDE.2024.3392335