Vol. 16 (2026): Proceedings of the 2026 International Conference on AI and Technological Innovation Applications (AITIA 2026)
E-ISSN: 3105-0522 P-ISSN: 3105-0514
DOI: https://doi.org/10.70267/aitia.2026316325

A Survey of Automated Time Series Forecasting: From Deep Learning to Foundation Models with Financial Applications

Zhuoran Yang*

Department of Information Technology, Hubei University of Police, Wuhan, China

**Corresponding author: Zhuoran Yang.*

Abstract

Time-series forecasting is a fundamental research topic in artificial intelligence and machine learning, with broad applications in finance, healthcare, energy management, transportation, and industrial systems. As the scale and complexity of temporal data continue to increase, forecasting methodologies have evolved significantly from traditional statistical models to modern intelligent forecasting systems. Understanding this technological evolution is essential for both researchers and practitioners seeking to develop more accurate, scalable, and automated forecasting solutions. This survey reviews the development of automated time-series forecasting methods from a historical and technological perspective. The evolution of forecasting paradigms is examined across multiple stages, including statistical learning methods, machine learning approaches, deep learning architectures, Transformer-based forecasting models, AutoML frameworks, foundation models, and large-language-model-based intelligent agents. Particular attention is given to how these paradigms differ in terms of automation capability, long-term forecasting performance, interpretability, and application adaptability. The survey highlights the key motivations driving major technological transitions, including limitations in statistical assumptions, challenges in feature engineering, difficulties in long-sequence modeling, and the growing demand for generalization and autonomous decision-making capabilities. In addition, representative forecasting applications in the financial domain are reviewed to illustrate the practical impact of modern forecasting technologies. Finally, current challenges and future research directions are discussed, including multimodal forecasting, foundation model development, intelligent agent systems, explainable forecasting, and computational efficiency. The findings suggest that time-series forecasting is gradually evolving from task-specific prediction models toward intelligent forecasting ecosystems capable of learning, reasoning, and autonomous decision-making.

Keywords

time series forecasting, machine learning, deep learning, transformer, foundation models, intelligent agents

1. Introduction

Time-series forecasting is a fundamental task in artificial intelligence and data science, enabling the prediction of future observations from historical time-series data [1, 2]. Accurate forecasting is essential for informed decision-making, planning, and resource allocation across diverse domains, including finance, healthcare, energy management, transportation, and industrial systems [3, 4]. The rapid growth of digital technologies and the increasing scale, complexity, and uncertainty of temporal data have further emphasized the need for automated, scalable, and intelligent forecasting methods [1, 4, 5].

Despite decades of research, traditional statistical methods and early machine learning approaches face limitations in capturing nonlinear relationships, long-term dependencies, and adapting to heterogeneous datasets. Statistical models such as ARIMA and GARCH provide strong interpretability but struggle with complex patterns [2, 4, 6]. Machine learning approaches improve predictive power but often rely on manual feature engineering, whereas deep learning architectures can automatically extract hierarchical features but still struggle with very long sequences [4, 5]. These challenges highlight the need for more sophisticated forecasting paradigms capable of generalization, automation, and adaptive learning [1, 4].

Over the past decade, forecasting methodologies have gradually evolved from statistical and machine learning models to deep learning, Transformer-based architectures, foundation models, and intelligent agents. Each transition was driven by the limitations of preceding paradigms, such as insufficient nonlinear modeling, limited automation, or challenges in handling long-term dependencies [2, 4, 5]. Understanding this progression provides context for evaluating the relative advantages and trade-offs of different forecasting paradigms without detailing every individual model, which is covered in Chapter 2.

Accordingly, this survey addresses three key questions: how forecasting methodologies have evolved from statistical models to foundation models and intelligent agents; how different forecasting paradigms compare in terms of automation capability, long-term forecasting performance, interpretability, and application adaptability; and what challenges and future research directions are likely to shape next-generation forecasting systems [1, 4, 5]. This concise structure ensures the Introduction conveys the importance, challenges, and research objectives while avoiding unnecessary repetition in later chapters.

2. Evolution of Time Series Forecasting

Over the past several decades, time-series forecasting has undergone a substantial evolution, driven by increasing data complexity, the need for automation, and the demand for robust long-term predictions. Early research relied primarily on statistical methods, including ARIMA, GARCH, and VAR, which offered strong interpretability and sound theoretical foundations [6–9]. However, these approaches were limited in handling nonlinear patterns, long-range dependencies, and high-dimensional datasets, creating the impetus for more flexible modeling techniques.

Machine learning approaches emerged to address these limitations, enabling models such as support vector machines and ensemble methods to capture nonlinear relationships and improve predictive performance. While these techniques reduced reliance on manual assumptions, they often required extensive feature engineering and still lacked robust handling of long sequences modeling [10–12]. Subsequently, deep learning architectures introduced hierarchical feature learning, enabling models to automatically extract complex temporal patterns. Despite this advancement, long-sequence modeling and generalization across different datasets remained challenging [13–15].

Transformer-based models represented the next step, leveraging attention mechanisms to capture long-range dependencies more effectively and enabling scalable parallel computation. These models improved the ability to forecast complex temporal dynamics but still relied on task-specific data and substantial computational resources. Building on this, foundation models pre-trained on massive temporal datasets offer enhanced generalization, transfer learning, and the potential for few-shot adaptation, addressing many limitations of previous paradigms [15–20]. Finally, intelligent forecasting agents integrate reasoning, planning, and tool usage to support decision-making autonomously, marking a shift from purely predictive models to adaptive, autonomous forecasting systems [21–23].

This progression highlights the primary motivations and trade-offs that have guided technological transitions in time-series forecasting, providing a concise framework to contextualize subsequent chapters without detailing individual model implementations. Figure 1 summarizes the major technological stages, key limitations, core breakthroughs, and developmental trends that have driven the progression toward increasingly automated and intelligent forecasting systems.

Figure 1: Evolution of automated time series forecasting from statistical models to intelligent forecasting agents

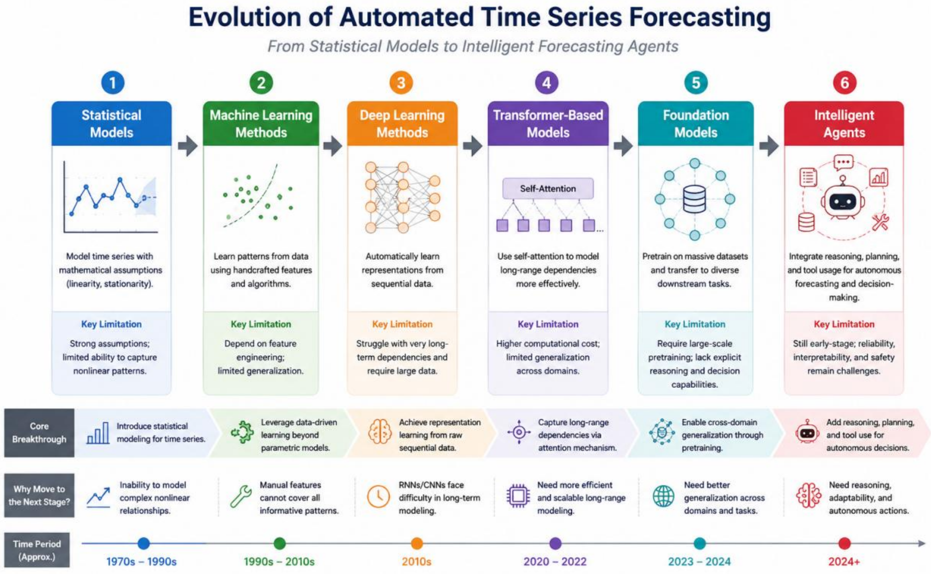

Table 1: Comparison Framework of Major Time Series Forecasting Paradigms

Paradigm	Representative Models	Automation Capability	Long-Term Forecasting	Interpretability	Application Adaptability
Statistical Learning	ARIMA, SARIMA, GARCH, VAR	Low	Low	High	Low
Machine Learning	SVM, Random Forest, XGBoost, LightGBM	Medium	Medium	Medium	Medium
Deep Learning	RNN, LSTM, GRU, CNN	Medium-High	Medium	Low	High
Transformer-Based Forecasting	Informer, Autoformer, FEDformer, PatchTST	High	High	Low	High
AutoML Forecasting	AutoTS, HPO, NAS-based Systems	Very High	Medium-High	Medium	High
Foundation Models	Chronos, TimeGPT, TimesFM, Lag-Llama	Very High	High	Low	Very High
Intelligent Agents	LLM-based Agents, Multi-Agent Systems	Highest	High	Medium	Very High

3. Major Forecasting Paradigms: From Statistical Models to Transformers

3.1 Statistical and Machine Learning Methods

Statistical and machine learning methods represent the foundational stages of time-series forecasting. Early forecasting research relied primarily on statistical approaches that model temporal patterns using mathematical assumptions and historical observations [1, 6]. These methods are highly interpretable, computationally efficient, and supported by well-established theoretical foundations. However, their effectiveness often depends on assumptions such as linearity and stationarity, limiting their ability to capture complex nonlinear relationships frequently observed in real-world data [6, 24].

To overcome these limitations, machine learning approaches introduced more flexible data-driven modeling strategies. Compared with statistical methods, machine learning models can better capture nonlinear patterns and interactions among variables, often resulting in improved forecasting performance across diverse applications [5, 24]. In addition, they reduce reliance on predefined statistical assumptions and enable more adaptive forecasting systems.

Despite these advantages, machine learning methods still require substantial feature engineering and domain expertise. The quality of forecasting performance often depends on manually designed input features, while long-term temporal dependencies remain difficult to model effectively. Consequently, although machine learning approaches improve predictive capability, their level of automation remains limited [5, 24].

Overall, the transition from statistical methods to machine learning reflects a broader shift from assumption-driven forecasting toward data-driven forecasting. Statistical approaches emphasize interpretability and theoretical rigor, whereas machine learning methods prioritize predictive performance and flexibility. Nevertheless, both paradigms face limitations in handling increasingly complex temporal patterns and large-scale datasets. These challenges motivated the emergence of deep learning approaches capable of automatic representation learning and more advanced sequence modeling [5, 25].

3.2 Deep Learning Methods

The increasing complexity of forecasting tasks exposed several limitations of traditional machine learning approaches. Although methods such as Support Vector Machines, Random Forests, and XGBoost improved the ability to model nonlinear relationships, their performance often depended on handcrafted features and extensive domain expertise [24]. As forecasting datasets became larger and more dynamic, researchers increasingly sought approaches capable of automatically learning informative representations from raw temporal data [5].

Deep learning addressed this challenge through representation learning, enabling models to extract hierarchical features automatically without extensive feature engineering [25]. Compared with traditional machine learning methods, deep learning approaches significantly improve nonlinear modeling capabilities and reduce reliance on manually designed features. This development marked an important step toward more automated forecasting systems [5, 25].

Several neural network architectures were widely adopted in forecasting research, including Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, Gated Recurrent Units (GRUs), and Convolutional Neural Networks (CNNs) [14, 15, 25]. Although these architectures differ in design, they share the common objective of learning temporal patterns directly from sequential data. Their success expanded the application of forecasting models across finance, energy management, transportation, and industrial analytics [5].

Despite these advantages, deep learning methods still face important limitations. Most models require substantial training data and computational resources. More importantly, recurrent architectures process information sequentially, making it difficult to capture very long-range dependencies while limiting efficient parallel computation. As forecasting horizons and dataset scales continued to increase, these challenges became increasingly significant [5, 14].

Therefore, while deep learning substantially improved forecasting performance and automation capability, it did not fully address the challenges of long-sequence modeling and scalability. These limitations motivated the development of Transformer-based forecasting models, which introduced attention mechanisms to model global temporal dependencies more effectively and ultimately became the dominant paradigm for long-term forecasting tasks [16].

3.3 Transformer-Based Forecasting

The limitations of deep learning models in long-sequence forecasting motivated the adoption of Transformer-based architectures. Unlike recurrent neural networks, which process information sequentially, Transformers use self-attention mechanisms to model relationships between time steps directly. This design significantly improves the ability to capture long-range temporal dependencies while enabling highly parallel computation [16].

The introduction of Transformers marked an important advancement in time-series forecasting. Compared with traditional recurrent architectures, Transformer-based models offer stronger scalability and improved performance for long-horizon forecasting tasks. Representative models such as Informer, Autoformer, FEDformer, and PatchTST further adapted the Transformer framework to forecasting applications and demonstrated strong results across a variety of benchmark datasets. Their success established Transformer-based forecasting as one of the dominant paradigms in modern time-series prediction research [18, 19, 26, 27].

Despite these advantages, Transformer-based forecasting models still face several Slimitations. Their training process often requires large amounts of data and substantial computational resources. In addition, most Transformer models remain task-specific, meaning that they are typically optimized for particular datasets or forecasting scenarios. As a result, their ability to generalize across diverse domains remains limited [16, 19].

These challenges motivated the emergence of foundation models for time-series forecasting. By leveraging large-scale pretraining and transfer learning, foundation models aim to learn general temporal representations that can be reused across multiple tasks and domains. Therefore, the development of Transformer-based forecasting not only improved long-term forecasting performance but also laid the foundation for the next generation of automated forecasting systems [20, 28–31].

4. Automated and General-Purpose Forecasting: AutoML, Foundation Models, and Agents

4.1 AutoML Forecasting

As forecasting systems become increasingly complex, selecting appropriate models, features, and hyperparameters has become a major challenge. AutoML aims to automate key stages of the forecasting pipeline, including feature selection, model selection, hyperparameter optimization, and performance evaluation [32, 33]. By reducing manual intervention, AutoML improves efficiency, reproducibility, and accessibility for non-expert users.

The primary contribution of AutoML is forecasting automation rather than improvements in model generalization. Through automated search and optimization strategies, AutoML can identify suitable forecasting configurations with reduced human effort. This capability is particularly valuable in large-scale forecasting applications where manual experimentation is costly and time-consuming [32, 34].

Despite these advantages, AutoML remains dependent on predefined search spaces and optimization objectives. Its focus is primarily on improving workflow efficiency rather than fundamentally enhancing forecasting intelligence. These limitations motivated the development of foundation models that seek to learn transferable forecasting knowledge through large-scale pretraining [20, 32].

4.2 Foundation Models for Time Series

SFoundation models represent a fundamental paradigm shift in time-series forecasting. Traditional forecasting systems are typically trained for specific datasets and tasks, requiring substantial retraining when applied to new domains. In contrast, foundation models are pretrained on large-scale collections of heterogeneous time-series data and subsequently adapted to downstream forecasting tasks through transfer learning [20].

This shift from task-specific learning to general-purpose forecasting enables stronger scalability, transferability, and cross-domain generalization. Representative models such as Chronos, TimeGPT, TimesFM, and Lag-Llama demonstrate the potential of large-scale pretraining to learn reusable temporal representations that can be applied across diverse forecasting scenarios [28–31]. Rather than focusing on individual model architectures, the key innovation lies in introducing a unified forecasting paradigm that leverages pretrained knowledge [20, 29].

Compared with conventional forecasting approaches, foundation models reduce dependence on handcrafted model design and enable more flexible adaptation to new tasks. However, challenges related to interpretability, computational cost, robustness, and domain adaptation remain active research topics [20, 30]. Nevertheless, foundation models are widely regarded as a major milestone in the evolution of automated forecasting systems and provide the foundation for the next generation of intelligent forecasting technologies.

Table 2: Representative Foundation Models for Time Series Forecasting

Model	Organization	Year	Core Idea	Strength
Chronos	Amazon	2024	Pretrained forecasting model	Zero-shot forecasting
TimeGPT	Nixtla	2024	General-purpose foundation model	Transfer learning
TimesFM	Google	2024	Large-scale pretrained model	Cross-domain generalization
Lag-Llama	Open Source	2024	Probabilistic forecasting	Uncertainty estimation
MOMENT	Carnegie Mellon & Collaborators	2024	Universal time-series representation	Multiple downstream tasks

4.3 LLM-Based Intelligent Agents

The emergence of foundation models has enabled forecasting systems to move beyond prediction toward intelligent decision support. Unlike traditional forecasting models that focus solely on estimating future values, large language models and intelligent agents can integrate reasoning, planning, tool usage, and external knowledge into forecasting workflows [21, 35].

In forecasting applications, intelligent agents do not replace forecasting models. Instead, they serve as coordination frameworks that combine forecasting models, foundation models, external tools, and domain knowledge to support complex decision-making [21, 23]. Such systems can automatically collect data, select forecasting methods, evaluate prediction quality, generate explanations, and assist users in interpreting forecasting outcomes.

This capability extends forecasting from a prediction-oriented task to a broader decision-support process. Although challenges related to reliability, explainability, safety, and computational cost remain unresolved, intelligent agents represent an important step toward autonomous forecasting systems [23, 35]. Their integration with foundation models is increasingly viewed as a promising direction for the future development of automated forecasting [20, 23].

5. Financial Applications

Financial forecasting is one of the most important application domains of time-series forecasting technologies. Financial markets generate large volumes of sequential data, including stock prices, exchange rates, trading volumes, and cryptocurrency prices. Accurate forecasting plays a critical role in investment decision-making, risk management, portfolio optimization, and market monitoring. Consequently, financial forecasting has become a key benchmark for evaluating the effectiveness of forecasting methodologies [1, 6, 24].

Despite its practical importance, financial forecasting remains highly challenging. Financial time series often exhibit strong nonlinearity, volatility, noise, structural breaks, and sensitivity to external events such as economic policies, geopolitical developments, and investor sentiment. These characteristics make accurate prediction significantly more difficult than in many other forecasting tasks [1, 5, 6].

Modern forecasting technologies have demonstrated substantial value in addressing these challenges. Statistical methods provide interpretable baselines for trend and volatility analysis, while machine learning and deep learning approaches improve the ability to capture nonlinear market dynamics [5, 6, 25]. More recently, Transformer-based models, foundation models, and intelligent forecasting agents have enabled stronger long-term forecasting capability, cross-domain knowledge transfer, and enhanced decision-support functions [16, 20]. These advances have expanded the role of forecasting systems from pure prediction toward integrated financial intelligence.

Nevertheless, forecasting uncertainty remains unavoidable in financial markets. Even the most advanced models cannot fully anticipate unexpected events or rapidly changing market conditions. Therefore, future research is increasingly focused on improving model robustness, explainability, adaptability, and real-time decision support. As automated forecasting technologies continue to evolve, financial applications are expected to remain one of the most important environments for evaluating and advancing intelligent forecasting systems [5, 20].

6. Challenges, Future Directions, and Conclusion

Despite the significant progress achieved in automated time-series forecasting, several important challenges remain unresolved. As forecasting systems evolve from traditional statistical models toward foundation models and intelligent agents, addressing these issues is essential for improving robustness, scalability, and real-world applicability [20, 36].

One major challenge is data quality and distribution shift. Real-world time-series data are often affected by missing values, noise, outliers, and non-stationary distributions. In many practical scenarios, models trained on historical data may experience significant performance degradation when deployed in dynamic environments where data distributions continuously evolve. Developing models capable of adapting to such shifts remains a critical research direction [36].

Another key challenge is model interpretability and explainability. Although deep learning models, Transformer-based architectures, and foundation models have demonstrated strong predictive performance, their internal decision-making processes are often difficult to interpret. In high-stakes domains such as finance, healthcare, and energy systems, transparency and trustworthiness are essential requirements. Therefore, improving model interpretability remains an important research objective for future forecasting systems [37].

In addition, uncertainty quantification remains a fundamental limitation in modern forecasting systems. Even advanced deep learning models may produce overconfident predictions when faced with noisy or out-of-distribution data. Probabilistic modeling approaches, such as Bayesian approximations, provide one direction for improving robustness under uncertainty [38].

Computational efficiency and scalability also represent significant concerns, particularly for foundation models. The increasing scale of modern forecasting systems leads to substantial computational cost during both training and deployment. Balancing performance with efficiency is therefore a key consideration for practical adoption in real-world environments [20].

Looking forward, multimodal forecasting, foundation models, and intelligent agents are expected to become major research directions. Future systems may integrate heterogeneous data sources, including numerical time-series data, text, images, sensor signals, and external knowledge, to improve forecasting accuracy and robustness [39].

In addition, intelligent agents based on large language models are expected to transform forecasting systems from pure prediction tools into comprehensive decision-support frameworks. Such systems may not only generate forecasts but also reason, select appropriate models, explain results, and assist with decision-making under uncertainty [21, 35].

Overall, automated time-series forecasting is evolving toward more adaptive, explainable, and intelligent systems. Despite remaining challenges in robustness, interpretability, and efficiency, continued advancements in foundation models, multimodal learning, and agent-based systems are expected to significantly enhance future forecasting capabilities [20, 23].

References

[1] R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 3rd ed. Melbourne, Australia: OTexts, 2021. Available: https://otexts.com/fpp3/

[2] J. D. Hamilton, Time Series Analysis. Princeton, NJ, USA: Princeton University Press, 1994.

[3] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019," Applied Soft Computing, vol. 90, Art. no. 106181, 2020, doi: 10.1016/j.asoc.2020.106181.

[4] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and machine learning forecasting methods: Concerns and ways forward," PLOS ONE, vol. 13, no. 3, Art. no. e0194889, 2018, doi: 10.1371/journal.pone.0194889.

[5] B. Lim and S. Zohren, "Time-series forecasting with deep learning: A survey," Philosophical Transactions of the Royal Society A, vol. 379, no. 2194, Art. no. 20200209, 2021, doi: 10.1098/rsta.2020.0209.

[6] G. E. P. Box and G. M. Jenkins, Time Series Analysis: Forecasting and Control, 2nd ed. San Francisco, CA, USA: Holden-Day, 1976.

[7] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," Econometrica, vol. 50, no. 4, pp. 987–1007, 1982, doi: 10.2307/1912773.

[8] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," Journal of Econometrics, vol. 31, no. 3, pp. 307–327, 1986, doi: 10.1016/0304-4076(86)90063-1.

[9] C. A. Sims, "Macroeconomics and reality," Econometrica, vol. 48, no. 1, pp. 1–48, 1980, doi: 10.2307/1912017.

[10] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995, doi: 10.1007/BF00994018.

[11] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[12] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[13] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[14] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

[15] K. Cho, B. van Merriënboer, Ç. Gülçehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder–decoder for statistical machine translation," in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 2014, pp. 1724–1734.

[16] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in Advances in Neural Information Processing Systems, vol. 30, 2017.

[17] Q. Wen, T. Zhou, C. Zhang, W. Chen, Z. Ma, J. Yan, and L. Sun, "Transformers in time series: A survey," in Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, 2023, pp. 6778–6786, doi: 10.24963/ijcai.2023/759.

[18] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, "Informer: Beyond efficient transformer for long sequence time-series forecasting," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, no. 12, 2021, pp. 11106–11115, doi: 10.1609/aaai.v35i12.17325.

[19] H. Wu, J. Xu, J. Wang, and M. Long, "Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting," in Advances in Neural Information Processing Systems, vol. 34, 2021.

[20] R. Bommasani et al., "On the opportunities and risks of foundation models," arXiv preprint arXiv:2108.07258, 2021.

[21] S. Yao, J. Zhao, D. Yu, N. Du, I. Shafran, K. Narasimhan, and Y. Cao, "ReAct: Synergizing reasoning and acting in language models," in Proceedings of the 11th International Conference on Learning Representations, 2023.

[22]T. Schick, J. Dwivedi-Yu, R. Dessì, R. Raileanu, M. Lomeli, L. Zettlemoyer, N. Cancedda, and T. Scialom, "Toolformer: Language models can teach themselves to use tools," in Advances in Neural Information Processing Systems, vol. 36, 2023.

[23] L. Wang et al., "A survey on large language model based autonomous agents," Frontiers of Computer Science, vol. 18, Art. no. 186345, 2024, doi: 10.1007/s11704-024-40231-1.

[24] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 Competition: Results, findings, conclusion and way forward," International Journal of Forecasting, vol. 34, no. 4, pp. 802–808, 2018, doi: 10.1016/j.ijforecast.2018.06.001.

[25] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[26] T. Zhou, Z. Ma, Q. Wen, X. Wang, L. Sun, and R. Jin, "FEDformer: Frequency enhanced decomposed transformer for long-term series forecasting," in Proceedings of the 39th International Conference on Machine Learning, PMLR, vol. 162, 2022, pp. 27268–27286.

[27] Y. Nie, N. H. Nguyen, P. Sinthong, and J. Kalagnanam, "A time series is worth 64 words: Long-term forecasting with transformers," in Proceedings of the International Conference on Learning Representations, 2023.

[28] A. F. Ansari et al., "Chronos: Learning the language of time series," arXiv preprint arXiv:2403.07815, 2024.

[29] A. Garza, C. Challu, and M. Mergenthaler-Canseco, "TimeGPT-1," arXiv preprint arXiv:2310.03589, 2023.

[30] A. Das, W. Kong, R. Sen, and Y. Zhou, "A decoder-only foundation model for time-series forecasting," in Proceedings of the 41st International Conference on Machine Learning, PMLR, vol. 235, 2024, pp. 10148–10167.

[31] K. Rasul et al., "Lag-Llama: Towards foundation models for probabilistic time series forecasting," arXiv preprint arXiv:2310.08278, 2024.

[32] F. Hutter, L. Kotthoff, and J. Vanschoren, Eds., Automated Machine Learning: Methods, Systems, Challenges. Cham, Switzerland: Springer, 2019, doi: 10.1007/978-3-030-05318-5.

[33] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," Journal of Machine Learning Research, vol. 13, pp. 281–305, 2012.

[34] S. Y. Shah et al., "AutoAI-TS: AutoAI for time series forecasting," arXiv preprint arXiv:2102.12347, 2021.

[35] T. Kojima, S. S. Gu, M. Reid, Y. Matsuo, and Y. Iwasawa, "Large language models are zero-shot reasoners," in Advances in Neural Information Processing Systems, vol. 35, 2022, pp. 22199–22213.

[36] X. Jin, Y. Park, D. C. Maddix, H. Wang, and Y. Wang, "Domain adaptation for time series forecasting via attention sharing," in Proceedings of the 39th International Conference on Machine Learning, PMLR, vol. 162, 2022, pp. 10280–10297.

[37] Z. C. Lipton, "The mythos of model interpretability," Queue, vol. 16, no. 3, pp. 31–57, 2018, doi: 10.1145/3236386.3241340. Also available as arXiv:1606.03490.

[38] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," in Proceedings of the 33rd International Conference on Machine Learning, PMLR, vol. 48, 2016, pp. 1050–1059.

[39] T. Baltrušaitis, C. Ahuja, and L.-P. Morency, "Multimodal machine learning: A survey and taxonomy," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 41, no. 2, pp. 423–443, 2019, doi: 10.1109/TPAMI.2018.2798607.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Open Access