A Review of Sales Forecasting Research Based on the Fusion of Online Consumer Sentiment and Time-Series Data

Hanyi Wang
International Business School Suzhou, Xi'an Jiaotong-Liverpool University, Suzhou, China
why399117@gmail.com

ABSTRACT

Sales forecasting in digital markets is no longer driven only by historical sales, price and promotion records. Online reviews, social media posts and other user-generated texts often appear before sales changes become visible, so they may provide early clues about attention, satisfaction and purchase intention. In this paper, online consumer sentiment is used as an umbrella term for affective signals extracted from these different forms of user-generated content. The paper reviews sales forecasting research based on the fusion of online consumer sentiment and time-series data. The discussion is organized around three connected layers: sentiment modeling, time-series forecasting and multimodal fusion. The reviewed studies suggest that sentiment features are most useful when demand is affected by reputation, product experience or online discussion, whereas time-series models remain essential for trend, seasonality and autocorrelation. Feature-level fusion is easy to implement, decision-level fusion is robust to noisy inputs, and deep or attention-based fusion can capture richer interactions when enough data are available. The paper also points out several practical limits, including noisy sentiment extraction, temporal misalignment, shallow fusion design, weak interpretability and deployment difficulty. Future work should pay more attention to aspect-level sentiment, lag structure, explainable forecasting and lightweight systems that can be transferred across product categories.

Keywords: Sales Forecasting; Online Consumer Sentiment; Time-Series Forecasting; Multimodal Fusion; Demand Prediction

1. INTRODUCTION

Sales forecasting supports inventory, pricing, promotion and capacity decisions. Overprediction can create excess stock, while underprediction can produce stockouts and lost sales. Conventional models rely on structured variables such as historical demand, prices, holidays and campaigns. Digital commerce adds another information source: consumers post reviews, ratings and social-media comments before, during and after purchase. These texts may reveal changes in attention or satisfaction before they are visible in the sales series. In this review, online consumer sentiment refers to affective information extracted from online reviews, social-media posts and other user-generated content. Online reviews are therefore treated as one important source of consumer sentiment, not as a synonym for social media.

Early evidence came largely from entertainment and e-commerce. Liu linked movie word of mouth to box-office revenue [1], and Dellarocas, Zhang and Awad showed that review information improved motion-picture sales forecasts [2]. Duan, Gu and Whinston identified reciprocal dynamics between sales and online word of mouth [3], while Tirunillai and Tellis connected broader user-generated content with market performance [4]. Chevalier and Mayzlin found sales effects from online book reviews [18], and Chintagunta et al. documented review effects across local movie markets [19]. Rui, Liu and Whinston showed that both the source and content of tweets mattered for movie sales [20]. Asur and Huberman used social-media activity to predict market outcomes [21]. Dijkman, Ipeirotis and Aertsen provided a Twitter-based sales case [22]. In product forecasting, Fan, Che and Chen combined reviews with sales history [5], whereas Zhang et al. added macroeconomic indicators and sentiment [6]. Together, these studies position sentiment as a time-sensitive external signal rather than a substitute for sales history.

Time-series research provides the complementary modeling foundation. LSTM introduced gated recurrence for long-range dependence [9], and the Transformer showed how self-attention can represent sequence relationships without recurrence [10]. Prophet separates trend, seasonality and holiday effects for interpretable business forecasting [11]. DeepAR learns probabilistic forecasts across related series [12], N-BEATS provides a strong neural baseline for univariate forecasting [15], and Temporal Fusion Transformer combines attention with variable selection for multi-horizon prediction [13]. These models make it possible to connect textual sentiment with temporal sales patterns, but they do not determine how the connection should be designed.

The literature remains fragmented. Sentiment studies often stop at classification, forecasting studies often ignore consumer text, and fusion studies sometimes add a single sentiment score without testing alignment or lag effects. This review therefore asks three questions: which sentiment features are useful, which temporal patterns should be modeled, and which fusion strategy fits different data and business conditions. Its contribution is a structured synthesis that links these choices rather than a new predictive model.

2. REVIEW METHOD AND ANALYTICAL FRAMEWORK

This study uses a structured narrative review rather than a new empirical experiment. The final set contains 35 sources published between 1997 and 2023. Literature was identified through Google Scholar, Web of Science, Scopus and CNKI using combinations of sales forecasting, online consumer sentiment, online reviews, social media posts, time-series forecasting, multimodal fusion and demand prediction. Sources were retained if they examined user-generated content and demand, provided a methodological foundation for sentiment or temporal modeling, or addressed fusion and forecast interpretation. Peer-reviewed journals and major conference papers were prioritized; duplicates, non-scholarly commentary and sources without a clear connection to the review questions were not retained. Because the aim is conceptual integration rather than estimation of a common effect size, no meta-analysis was conducted.

The selected papers were grouped into three analytical layers. The first layer is sentiment modeling, which turns unstructured comments into polarity, emotion, topic or aspect-level variables. The second layer is time-series forecasting, which models sales history and structured covariates. The third layer is fusion, where textual and temporal information are combined before training, during representation learning or after separate predictions are produced. Figure 1 presents this three-layer view and provides the organizing logic for the following sections.

In practice, the framework also works as a checklist for building a forecasting system. Textual data must be cleaned and transformed into usable variables; sales and sentiment must be matched by product, market and time window; and the final model must be evaluated against sales-only baselines. This last point is important because a fusion model is only valuable if it improves a decision-relevant forecast rather than merely adding another data source.

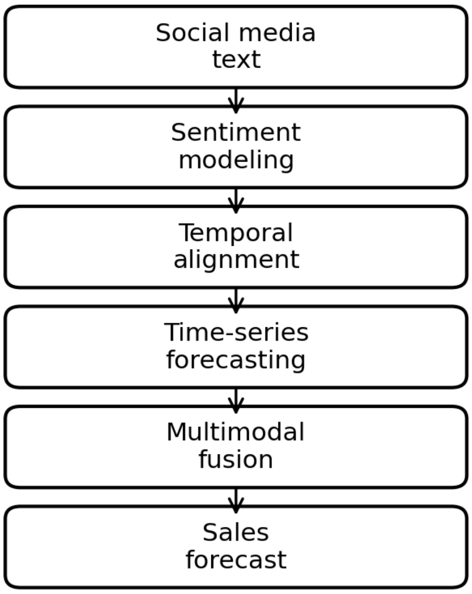

Figure 1. Three-layer framework for sentiment-time-series fusion in sales forecasting.

3. SENTIMENT MODELING FOR SALES FORECASTING

Sentiment analysis converts consumer text into variables that can be used in forecasting. Pang and Lee provide the broad foundation for opinion mining and sentiment analysis [7]. Hu and Liu show a more product-oriented approach by mining and summarizing customer review features [8]. Lexicon-based methods are transparent and inexpensive, but they often miss sarcasm, mixed opinions and domain-specific words. Review-mining research also shows that product features can be linked to perceived value and pricing power [23]. Deep learning methods improve contextual representation, and Zhang, Wang and Liu summarize this shift in their survey of deep learning for sentiment analysis [24]. In sales forecasting, the goal is not only to classify whether a review is positive or negative. The more useful question is whether the sentiment feature carries information that is not already contained in historical sales.

Aspect-level sentiment is particularly relevant for this topic. A product review may praise design but criticize delivery, or praise quality but complain about price. If such a review is compressed into one polarity score, useful information may be lost. Zhang et al. summarize aspect-based sentiment analysis as a set of tasks that identify both opinion targets and sentiment orientation [25]. Sun, Huang and Qiu show how BERT can be adapted to aspect-based sentiment analysis by constructing auxiliary sentences [26]. BERT itself is important because it represents words according to context rather than fixed dictionary meanings [14]. For forecasting, this means that sentiment variables can be designed more closely around actual purchase concerns.

The main weakness of sentiment data is noise. Online reviews may be duplicated, manipulated, emotionally extreme or written by users who are not representative of the broader market. Review volume can measure attention as much as satisfaction, and platform algorithms may amplify certain opinions. Therefore, sentiment features should be filtered, aggregated and tested with lag structures before being used in a sales model. A sentiment spike before a promotion, for example, may not have the same meaning as negative reviews after delivery problems. The timing of sentiment is often as important as the sentiment score itself.

4. TIME-SERIES MODELS AND TEMPORAL SALES DYNAMICS

Sales data are temporal by nature. Demand may follow seasonal cycles, product life-cycle patterns, holiday effects and promotion-driven shocks. Classical forecasting models are still useful when the series is relatively stable and the purpose is explanation or routine planning. Prophet is a good example of a business-oriented model because it decomposes time series into trend, seasonality and holiday effects, which makes the forecast easier for managers to inspect [11]. In many retail settings, this interpretability is not a minor benefit; it affects whether the forecast can actually be trusted and used.

Deep models become attractive when sales patterns are nonlinear or many related series are available. LSTM models sequential dependence [9], DeepAR produces probabilistic forecasts across related series [12], N-BEATS is a strong univariate baseline [15], and Temporal Fusion Transformer combines attention with variable selection for multi-horizon prediction [13]. The M4 competition benchmarked methods across 100,000 series [27], while M5 emphasized retail hierarchies and inventory decisions [28]. Informer reduces attention cost for long sequences [29]; Autoformer adds decomposition and autocorrelation [30]; N-HiTS uses hierarchical interpolation [31]; PatchTST uses patch-based representations [32]; and TimesNet models multiple periodic patterns [33]. Their value in sales forecasting still depends on data scale, forecast horizon and deployment cost.

For sentiment-time-series fusion, model selection should begin with the business task. Mature products with stable sales may gain little from sentiment because historical patterns already explain much of the variation. New products, entertainment goods and experience products may benefit more because online discussion forms before a long sales history exists [6]. Short-term inventory forecasting needs fast updating, while long-term planning needs stable assumptions. For this reason, sentiment fusion should be treated as a conditional tool rather than a universal upgrade.

Sections 3 and 4 describe sentiment and temporal modeling separately, but fusion requires an explicit interface between them. In a typical pipeline, sentiment is aggregated for each product and time window, then converted into lagged variables to represent delayed effects on demand. These signals may enter a forecasting model directly as exogenous or multi-source covariates, receive dynamic weights through attention or gating mechanisms, or be encoded jointly with the sales sequence as a learned cross-modal representation. The choice determines whether sentiment functions as a fixed auxiliary feature, a time-varying explanatory signal or a fully integrated representation. The next section compares these alternatives.

5. FUSION STRATEGIES FOR SENTIMENT AND TIME-SERIES DATA

Fusion turns separate textual and temporal sources into one forecasting problem. Feature-level, or early, fusion is the simplest design. Review volume, average polarity, negative-share ratio, topic share and aspect sentiment are aggregated by product and time window, shifted into lagged variables and concatenated with sales, price and promotion covariates. Fan, Che and Chen provide a clear example of combining online review information with a sales model [5]. This design is transparent and works well when the sentiment variables are stable and reasonably complete. Its limitation is that fixed concatenation assumes the same relevance across products and periods, and it may miss nonlinear interactions between a reputation shock and the current sales state.

Decision-level, or late, fusion trains separate models for sales history and consumer sentiment and then combines their forecasts through weighted averaging, stacking or another ensemble rule. It is useful when the two sources have different sampling frequencies, when sentiment is occasionally missing, or when organizations want to preserve an established sales model. Separate predictions also make source-specific errors easier to diagnose. The trade-off is that late fusion uses limited cross-modal interaction: the sentiment model cannot directly condition its representation on the current shape of the sales sequence.

Representation-level fusion addresses this limitation by learning the interaction inside the forecasting architecture. Sentiment embeddings and temporal states may be projected into a shared space, linked through cross-attention, or controlled by gates that increase the weight of text when a sudden discussion signal is informative. The multimodal learning literature emphasizes representation, alignment and co-learning rather than simple variable addition [16]. Tensor Fusion Network was designed for multimodal sentiment analysis, but its explicit modeling of cross-modal interactions illustrates one possible direction for sales forecasting [17]. Temporal Fusion Transformer also offers a practical mechanism because variable selection and attention can distinguish periods in which sentiment is useful from periods in which historical sales should dominate [13]. Deep fusion is therefore potentially richer, but it requires more observations, stronger regularization and clearer interpretation tools.

The three strategies should be compared against the same sales-only baseline. A practical model must align comments and sales by product, market and time window, test plausible lag structures, tolerate sparse sentiment and explain forecast changes. Feature-level fusion is a sensible baseline, decision-level fusion favors modularity and robustness, and representation-level fusion is justified only when the data support additional complexity. Table 1 summarizes these method families. Table 2 compares representative empirical studies, while Table 3 extends evaluation beyond average error to robustness, interpretation and deployment feasibility.

Table 1. Main Method Families in Sentiment-Time-Series Sales Forecasting.

Method family	Typical design	Use and limitation
Sentiment modeling	Lexicon, machine learning, BERT, aspect-level sentiment	Generates external demand signals; limited by noise and platform bias.
Time-series forecasting	Prophet, LSTM, DeepAR, N-BEATS, Temporal Fusion Transformer	Captures temporal structure; may lag behind sudden online attention.
Feature-level fusion	Concatenate sentiment scores, review volume and sales covariates	Practical and interpretable; weak at cross-modal interaction.
Deep/attention fusion	Joint representation learning with attention, gates or tensor fusion	Models complex interaction; requires more data and interpretation tools.
Decision-level fusion	Separate sentiment and sales models combined by averaging or stacking	Robust and modular; limited direct cross-modal interaction.

Table 2. Representative Studies on Sentiment-Enhanced Sales Forecasting.

Study	Context and data	Approach	Main implication
Dellarocas et al. [2]	Motion pictures; online reviews	Review metrics added to sales forecasting	Reviews provide predictive value beyond basic market information.
Fan et al. [5]	Product sales; reviews and sales history	Bass model with sentiment analysis	Combines diffusion dynamics with review-based demand signals.
Zhang et al. [6]	Product sales; reviews and macro indicators	Prospect theory with sentiment analysis	Shows that economic context can complement review sentiment.
Chintagunta et al. [19]	Movie box office; local review markets	Sequential rollout and market aggregation	Temporal and geographic aggregation affect estimated review impact.
Rui et al. [20]	Movie sales; Twitter messages	User- and content-level tweet features	The source and type of chatter matter, not only message volume.
Dijkman et al. [22]	Sales case; Twitter activity	Social-signal forecasting case study	Demonstrates feasibility but leaves transferability to other products open.

Table 3. Evaluation Criteria for Fusion Forecasting.

Criterion	Indicator	Purpose
Accuracy	MAE, RMSE, MAPE	Measure numerical forecast error
Direction	Increase/decrease hit rate	Check demand movement prediction
Robustness	Rolling-origin tests; missing-sentiment stress test	Assess stability across periods and data gaps
Interpretability	Feature attribution; lag/attention analysis	Explain why sentiment changes the forecast
Practicality	Latency, update cost, data availability	Assess deployment feasibility

6. APPLICATION SCENARIOS AND EVALUATION

E-commerce is the clearest application because reviews, ratings, browsing behavior and sales records can often be linked on the same platform. Fusion models may support daily or weekly forecasts for categories, stock keeping units and campaigns. In fast-moving consumer goods, online sentiment can signal reputation changes before retail sales fully respond. Tourism and cultural consumption are also relevant because destination image, service experience and event popularity are shaped by online discussion. Across these settings, sentiment is most useful when consumer perception changes quickly and can be aligned with a specific product, market and time window.

Evaluation should compare fusion models with strong sales-only baselines and report whether sentiment improves forecasts beyond price, promotion, seasonality and historical demand. MAE, RMSE and MAPE measure average error, while directional accuracy checks whether the model captures increases or decreases. Rolling-origin tests and missing-sentiment stress tests reveal whether gains survive across periods and data gaps. Interpretation should also show which lagged sentiment variables, attention weights or feature contributions changed the forecast. A small accuracy gain may be unattractive if it is unstable, opaque or costly to maintain; Table III therefore combines predictive and operational criteria.

The reviewed evidence supports the usefulness of online sentiment, but not uniformly. Historical sales often remain strongest for established products with regular purchase cycles. Sentiment and review volume may add more value for new products, experience goods or products exposed to public discussion. Firms should therefore add sentiment only when the product context and expected forecasting gain justify the collection, alignment and maintenance cost.

7. CHALLENGES AND FUTURE DIRECTIONS

Several challenges still limit the field. First, sentiment is noisy: posts may contain sarcasm, spam, bots or promotional manipulation. Domain-adapted language models, credibility weighting and aspect-level extraction can reduce but not remove this problem. Second, temporal alignment is difficult because comments arrive continuously while sales are often measured daily or weekly. Aggregation windows and lag choices can change the apparent relationship, so studies should report how these settings were selected.

A third challenge is shallow fusion. Many studies still insert a single sentiment score into a regression or neural model and then describe the result as multimodal forecasting. This is useful as a baseline, but it does not fully explain how consumer emotion, attention volume and sales dynamics interact. The fourth challenge is interpretability. Black-box fusion is difficult to use in operational settings if managers cannot see which variable caused a forecast shift. LIME is one possible tool because it explains individual predictions through local interpretable approximations [34]. SHAP offers another route by connecting feature attribution with Shapley values [35]. Both methods are useful starting points, although they need to be adapted carefully for sequential forecasting.

Deployment is also underdeveloped. A paper may train a model once, but enterprise use requires pipelines, monitoring, retraining and error diagnosis. Sentiment features can become stale as platform vocabulary or product life-cycle stage changes. Future work should therefore consider lightweight architectures, multilingual sentiment, cross-category transfer and human-in-the-loop decisions alongside forecast accuracy.

8. CONCLUSION

This paper reviewed sales forecasting research based on the fusion of online consumer sentiment and time-series data. The two sources play different roles: sentiment may provide early information about attention, satisfaction and purchase intention, while time-series data remain necessary for trend, seasonality and autocorrelation. The three-layer framework is an organizing synthesis rather than a new forecasting model. Its value lies in connecting sentiment construction, temporal modeling and fusion choice, and in showing why the same design may not work equally well across products and forecast horizons.

Overall, sentiment-time-series fusion should be used selectively. It is most promising when online discussion is active, when historical sales are limited or when consumer perception changes before sales records show the change. It is less necessary when demand is stable and historical sales already explain most variation. Future studies should move from simple sentiment-score addition toward aspect-level features, explicit lag modeling, explainable fusion and deployable forecasting systems. For firms, the framework can serve as a method-selection guide rather than a one-size-fits-all solution.

ACKNOWLEDGMENT

The author thanks the supervisor and academic guidance team for helpful suggestions during the proposal stage.

REFERENCES

[1] Y. Liu, “Word of mouth for movies: Its dynamics and impact on box office revenue,” Journal of Marketing, vol. 70, no. 3, pp. 74-89, 2006.
[2] C. Dellarocas, X. Zhang, and N. F. Awad, “Exploring the value of online product reviews in forecasting sales: The case of motion pictures,” Journal of Interactive Marketing, vol. 21, no. 4, pp. 23-45, 2007.
[3] W. Duan, B. Gu, and A. B. Whinston, “The dynamics of online word-of-mouth and product sales: An empirical investigation of the movie industry,” Journal of Retailing, vol. 84, no. 2, pp. 233-242, 2008.
[4] S. Tirunillai and G. J. Tellis, “Does chatter really matter? Dynamics of user-generated content and stock performance,” Marketing Science, vol. 31, no. 2, pp. 198-215, 2012.
[5] Z.-P. Fan, Y.-J. Che, and Z.-Y. Chen, “Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis,” Journal of Business Research, vol. 74, pp. 90-100, 2017.
[6] C. Zhang, Y.-X. Tian, Z.-P. Fan, Y. Liu, and L.-W. Fan, “Product sales forecasting using macroeconomic indicators and online reviews: A method combining prospect theory and sentiment analysis,” Soft Computing, vol. 24, no. 9, pp. 6213-6226, 2020.
[7] B. Pang and L. Lee, “Opinion mining and sentiment analysis,” Foundations and Trends in Information Retrieval, vol. 2, no. 1-2, pp. 1-135, 2008.
[8] M. Hu and B. Liu, “Mining and summarizing customer reviews,” in Proc. 10th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, Seattle, WA, USA, 2004, pp. 168-177.

[9] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997.
[10] A. Vaswani et al., “Attention is all you need,” in Advances in Neural Information Processing Systems 30, 2017.
[11] S. J. Taylor and B. Letham, “Forecasting at scale,” The American Statistician, vol. 72, no. 1, pp. 37-45, 2018.
[12] D. Salinas, V. Flunkert, J. Gasthaus, and T. Januschowski, “DeepAR: Probabilistic forecasting with autoregressive recurrent networks,” International Journal of Forecasting, vol. 36, no. 3, pp. 1181-1191, 2020.
[13] B. Lim, S. O. Arik, N. Loeff, and T. Pfister, “Temporal Fusion Transformers for interpretable multi-horizon time series forecasting,” International Journal of Forecasting, vol. 37, no. 4, pp. 1748-1764, 2021.
[14] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” in Proc. NAACL-HLT, Minneapolis, MN, USA, 2019, pp. 4171-4186.
[15] B. N. Oreshkin, D. Carpov, N. Chapados, and Y. Bengio, “N-BEATS: Neural basis expansion analysis for interpretable time series forecasting,” in Proc. Int. Conf. Learning Representations, 2020.
[16] T. Baltrusaitis, 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.
[17] A. Zadeh, M. Chen, S. Poria, E. Cambria, and L.-P. Morency, “Tensor Fusion Network for multimodal sentiment analysis,” in Proc. EMNLP, Copenhagen, Denmark, 2017, pp. 1103-1114.
[18] J. A. Chevalier and D. Mayzlin, “The effect of word of mouth on sales: Online book reviews,” Journal of Marketing Research, vol. 43, no. 3, pp. 345-354, 2006.
[19] P. K. Chintagunta, S. Gopinath, and S. Venkataraman, “The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets,” Marketing Science, vol. 29, no. 5, pp. 944-957, 2010.
[20] H. Rui, Y. Liu, and A. B. Whinston, “Whose and what chatter matters? The effect of tweets on movie sales,” Decision Support Systems, vol. 55, no. 4, pp. 863-870, 2013.
[21] S. Asur and B. A. Huberman, “Predicting the future with social media,” in Proc. IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2010, pp. 492-499.
[22] R. M. Dijkman, B. Ipeirotis, and F. Aertsen, “Using Twitter to predict sales: A case study,” arXiv preprint arXiv:1503.04599, 2015.
[23] N. Archak, A. Ghose, and P. G. Ipeirotis, “Deriving the pricing power of product features by mining consumer reviews,” Management Science, vol. 57, no. 8, pp. 1485-1509, 2011.
[24] L. Zhang, S. Wang, and B. Liu, “Deep learning for sentiment analysis: A survey,” arXiv preprint arXiv:1801.07883, 2018.
[25] W. Zhang, X. Li, Y. Deng, L. Bing, and W. Lam, “A survey on aspect-based sentiment analysis: Tasks, methods, and challenges,” arXiv preprint arXiv:2203.01054, 2022.
[26] C. Sun, L. Huang, and X. Qiu, “Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence,” in Proc. NAACL-HLT, Minneapolis, MN, USA, 2019, pp. 380-385.
[27] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, “The M4 Competition: 100,000 time series and 61 forecasting methods,” International Journal of Forecasting, vol. 36, no. 1, pp. 54-74, 2020.
[28] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, “M5 accuracy competition: Results, findings, and conclusions,” International Journal of Forecasting, vol. 38, no. 4, pp. 1346-1364, 2022.
[29] H. Zhou et al., “Informer: Beyond efficient Transformer for long sequence time-series forecasting,” in Proc. AAAI Conference on Artificial Intelligence, vol. 35, no. 12, pp. 11106-11115, 2021.
[30] H. Wu et al., “Autoformer: Decomposition Transformers with Auto-Correlation for Long-Term Series Forecasting,” in Advances in Neural Information Processing Systems, vol. 34, pp. 22419-22430, 2021.
[31] C. Challu, K. G. Olivares, B. N. Oreshkin, F. G. Ramirez, M. M. Canseco, and A. Dubrawski, “N-HiTS: Neural hierarchical interpolation for time series forecasting,” in Proc. AAAI Conference on Artificial Intelligence, vol. 37, no. 6, pp. 6989-6997, 2023.
[32] Y. Nie, N. H. Nguyen, P. Sinthong, and J. Kalagnanam, “A Time Series is Worth 64 Words: Long-term Forecasting with Transformers,” in Proc. International Conference on Learning Representations, 2023.
[33] H. Wu et al., “TimesNet: Temporal 2D-variation modeling for general time series analysis,” in Proc. International Conference on Learning Representations, 2023.
[34] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why should I trust you?: Explaining the predictions of any classifier,” in Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1135-1144.
[35] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in Advances in Neural Information Processing Systems, vol. 30, 2017.