Paradigm

Sentiment Analysis and Large Language Models in Foreign Exchange Forecasting: A Bibliometric Perspective

Constantin – Laurențiu STAMA*
National University of Science and Technology POLITEHNICA Bucharest, Bucharest, Romania
**Corresponding author, constantin.stama@stud.faima.upb.ro*

Cătălin – George ALEXE
National University of Science and Technology POLITEHNICA Bucharest, Bucharest, Romania
catalin.alexe@upb.ro

Gheorghe MILITARU
National University of Science and Technology POLITEHNICA Bucharest, Bucharest, Romania
gheorghe.militaru@upb.ro

Cătălina – Monica ALEXE
National University of Science and Technology POLITEHNICA Bucharest, Bucharest, Romania
catalina.alexe@upb.ro

Abstract. *Foreign exchange forecasting remains a demanding research area because prices respond quickly to news, policy communication, and shifts in market expectations. The scientific literature has shifted from macroeconomic and econometric models to machine learning and deep learning, with increasing attention to textual information and sentiment as complementary inputs. Recent studies confirm that sentiment features improve short-horizon forecasting, though traditional methods struggle with context, negation, and domain-specific language.*

This paper applies a bibliometric analysis to map how research at the intersection of foreign exchange prediction, sentiment analysis, and advanced language modeling has evolved. Bibliographic records were retrieved from the Web of Science database, based on 48 identified publications, with 34 studies published from 2019 onward selected for analysis. The study uses VOSviewer to identify co-authorship networks, co-citation structures, keyword co-occurrence clusters, and emerging thematic trends. The research asks how the literature structures the role of sentiment in exchange rate forecasting, how context-aware language representations are positioned relative to traditional sentiment pipelines, and which integration patterns are most common when combining textual signals with numerical time series.

The results show distinct research clusters linking forecasting architectures, financial text processing, and sentiment-based modeling, as well as a growing focus on context-aware language representations. The analysis also highlights persistent gaps related to end-to-end integration of text semantics, timing alignment between text and market data, and consistent evaluation framing. The paper maps the field and proposes research directions for more coherent sentiment-augmented forecasting designs.

Keywords: large language models, foreign exchange forecasting, sentiment analysis, bibliometric analysis, financial text processing, VOSviewer, Web of Science.

Introduction

Exchange rate forecasting remains a challenging research problem because FX prices react to heterogeneous information sources and because market conditions can change rapidly. Traditional forecasting approaches have often relied on macroeconomic fundamentals and econometric

DOI: 10.2478/picbe-2026-0455
© 2026 C. – C. – L. Stama; G. Alexe; G.e Militaru; C. – M. Alexe, published by De Gruyter.
This work is licensed under the Creative Commons Attribution 4.0 License.

specifications, while more recent work has expanded toward machine learning and deep learning models. Even when forecasting architectures improve, a persistent limitation is that many price moves are associated with how market participants interpret news, events, and policy communication, which are primarily expressed in text. This observation has motivated a growing body of research that treats textual information and sentiment as relevant complementary inputs for improving FX forecasting accuracy (Cheng et al., 2019; Farimani et al., 2022).

A second limitation concerns the measurement of sentiment. Lexicon-based methods and shallow NLP pipelines can miss meaning that depends on context, qualifiers, and domain-specific phrasing. For example, a statement like "The economy will grow, but only slightly" contains a positive token but communicates a constrained outlook. Context-aware language models are designed to represent that meaning more faithfully than word-list scoring, which is why recent studies discuss Large Language Models as a potential improvement for extracting market-relevant signals from financial text (Cheng et al., 2019; Yang et al., 2025).

Given the rapid expansion of this literature, the present paper adopts a bibliometric approach to clarify how research on sentiment-augmented FX forecasting is structured and how LLM-based methods are positioned within it. Following the data collection procedure documented in the research protocol, bibliographic records were retrieved from Clarivate Web of Science using Advanced Search queries that target FX forecasting, sentiment analysis, and LLM-related terms, with filtering by document type, research areas, and a publication window prioritizing recent years. The protocol reports that, for the core query on FX forecasting and sentiment, 48 records were identified and 35 studies from 2019 onward were selected for analysis. VOSviewer is used to map publication clusters, collaboration networks, and thematic trends.

The study is guided by three research questions. RQ1 asks how the literature structures the relationship between FX forecasting and sentiment-based textual signals, including dominant themes and recurring methodological choices. RQ2 asks how the literature describes the transition from traditional NLP sentiment extraction toward context-aware representations and LLM-based approaches in finance and FX-related settings (Cheng et al., 2019; Farimani et al., 2022; Yang et al., 2025). RQ3 asks what integration patterns are most common when combining text-derived signals with numerical FX time series and which integration gaps are repeatedly identified.

Consistent with a bibliometric design, the hypotheses concern patterns observable in the literature rather than claims about model performance. H1 is that recent publications increasingly frame LLMs as a response to documented limitations of traditional sentiment pipelines in financial text processing. H2 is that multimodal FX forecasting studies frequently rely on externally computed, shallow text features, suggesting limited semantic integration. H3 is that the convergence of FX forecasting, sentiment analysis, and LLM-based representations forms an emerging and distinct research stream, with identifiable gaps around integration design and evaluation framing.

Literature review

Foreign Exchange forecasting: from fundamentals to data-driven models

Research on exchange rate forecasting has evolved from theory-driven econometric frameworks toward machine learning and hybrid approaches. Early models anchored in macroeconomic fundamentals, including purchasing power parity and interest rate parity, remain important for explaining longer-run movements, yet their empirical performance at short horizons is often weak when markets undergo regime shifts, policy surprises, or structural breaks. This limitation

motivates a broader line of work that combines market data with nonlinear models and text-derived signals instead of relying only on macro variables (Farimani et al., 2022).

More recent studies focus on architectures designed to learn nonlinear dependencies and adapt to fast-changing dynamics. Modular or composite neural designs have been explored as alternatives to single end-to-end models, with the aim of improving robustness and interpretability in forecasting setups (Bormpotsis et al., 2023). Attention-based mechanisms are also emphasized as a way to focus the model on the most relevant parts of the input when historical prices and external signals are both present (Dakalbab et al., 2025). In parallel, multivariate recurrent models continue to be tested for their capacity to handle multiple inputs and non-stationary patterns that appear in FX series (Chi et al., 2025).

Working definitions used in this paper

To maintain consistency in the paper's concepts, we use the following working definitions. By "FX forecasting" we refer to predicting an exchange rate outcome over a specified horizon, typically short to medium horizons, using information available up to a decision time. Depending on the experimental design, the target can be a direction of movement, a return, or a level, but the common focus is the incremental predictive value of additional signals under realistic timing constraints.

By "sentiment" we refer to a text-derived representation of market-relevant tone or stance expressed in financial discourse, such as news narratives or public commentary, which may encode expectations, perceived risk, and interpretation of events. By "sentiment feature" we mean a quantitative variable constructed from text and aligned in time with market data so it can be used in a forecasting model. By "event-based approach" we mean a method that organizes information around identifiable events, often distinguishing scheduled macro events from unscheduled shocks, and then models how the event context and related text are associated with subsequent FX movements (Semiromi et al., 2020; Yasir et al., 2019).

By "traditional NLP sentiment" we refer to lexicon-based scoring or standard supervised classifiers that rely on limited context windows or word-level features. By "transformer-based representation" we refer to contextual embeddings produced by transformer architectures, where the meaning of a token depends on surrounding text, enabling a richer representation of semantics than dictionary scoring (Cheng et al., 2019; Yang et al., 2025).

Why sentiment matters in Foreign Exchange and how it is measured

A practical limitation of purely fundamentals-based approaches is that short-horizon FX movement is often shaped by expectations and positioning, both of which are influenced by narratives, headlines, and the interpretation of events. For this reason, sentiment and related textual signals are usually treated as complementary information rather than a replacement for economic structure. Much of the recent literature therefore studies how to extract sentiment from financial text and integrate it into forecasting pipelines (Farimani et al., 2022).

Within sentiment extraction, an important distinction is between coarse polarity labels and context-aware interpretations. Fine-grained sentiment work highlights that financial language carries meanings that do not transfer cleanly from general-purpose sentiment dictionaries, particularly when domain-specific terms are involved (Cheng et al., 2019). Related work emphasizes that ambiguity, including polysemy and domain terminology, can distort sentiment measures when models do not resolve meaning in context, which directly affects downstream forecasting signals (Seifollahi and Shajari, 2019).

As a result, many forecasting pipelines combine market variables with sentiment features and use machine learning models that can capture nonlinear interactions. Studies that integrate sentiment into predictive models typically frame sentiment as an additive signal that can improve short-horizon performance when information flow is dense (Naeem et al., 2021). Other work combines sentiment with deep architectures such as GRU/CNN hybrids, treating sentiment as one input alongside historical price-based features (Mabrouk et al., 2025).

Event-based, semantic-context, and timing-aware approaches

Beyond general sentiment polarity, several studies argue that the market impact of text depends on the type of event and the economic domain it relates to. A standard distinction is between scheduled macroeconomic events, where timing is known, and unscheduled events, where surprise and interpretation dominate. Research linking news, FX behavior, and the economic calendar supports the view that timing and event context matter for both signal quality and evaluation design (Semiromi et al., 2020). Related approaches organize text around events and then connect event-labeled narratives to subsequent exchange-rate movements (Yasir et al., 2019).

Another direction shifts the question from whether text is positive or negative to what the text is about and which concepts are being emphasized. This motivates methods that extract salient points, topics, or conceptual signals that can serve as structured inputs to prediction models (Yono et al., 2019). In higher-frequency settings, this emphasis on context becomes more important because news arrives continuously and the market response can be uneven and sensitive to microstructure conditions (Xing et al., 2021).

Alternative data sources and behavioral indicators

Text signals are not limited to curated financial news. A broader literature examines online attention and behavioral indicators as proxies for market expectations. Social media, search activity, and other user-generated content can provide faster signals, but they also introduce higher noise and greater vulnerability to manipulation. Work using real-time Twitter-based indicators illustrates the potential predictive value of public discourse, while also highlighting the need for careful filtering, time alignment, and validation (Albahli et al., 2022). In FX applications, these alternative streams are usually treated as complementary sources whose usefulness depends on signal quality, temporal granularity, and how well they map onto tradable horizons.

From traditional NLP pipelines to Transformer-based models

Traditional NLP sentiment approaches commonly rely on staged pipelines such as tokenization, feature extraction, and classification. These methods can work in constrained settings, but they tend to be sensitive to domain vocabulary and context-dependent meaning. The literature on disambiguation in FX-related text highlights that resolving meaning in context is not an optimization detail, because it affects the validity of sentiment features used in forecasting (Seifollahi and Shajari, 2019). Fine-grained sentiment work similarly shows that word-list methods often miss important structure in financial language (Cheng et al., 2019).

Transformer-based language models changed this landscape by producing contextual embeddings whose meaning depends on surrounding text. In finance, this matters because the same expression can carry different implications depending on the narrative and economic setting. Recent work discusses the use of modern large language models for exchange-rate related tasks, including sentiment extraction and text-driven prediction pipelines (Yang et al., 2025). The practical motivation across these studies is consistent: richer representations can reduce

information loss relative to dictionary methods, especially when forecasting requires sensitivity to subtle shifts in tone, event framing, or policy communication.

Synthesis of viewpoints and the research gap

Taken together, the literature converges on two broad viewpoints. One viewpoint emphasizes model capacity and nonlinear learning, focusing on deep architectures and design choices such as modularity, attention, and multivariate recurrence to handle non-stationary FX dynamics (Bormpotsis et al., 2023; Dakalbab et al., 2025; Chi et al., 2025). The other viewpoint emphasizes information content, arguing that market-relevant text captures expectations and interpretation, and that forecasting improves when sentiment and event context are measured and aligned correctly in time (Farimani et al., 2022; Semiromi et al., 2020; Yasir et al., 2019). A practical tension between these lines is that many systems either rely on simplified sentiment signals that discard context, or they extract sentiment in a separate NLP stage that is not optimized for the forecasting objective and can introduce mismatch between text representation and prediction task (Cheng et al., 2019; Seifollahi and Shajari, 2019).

This leads to a concrete gap that motivates the present study. While prior work supports the usefulness of sentiment and event context, many approaches still depend on sentiment measures that are either too coarse or weakly contextualized, and the integration of richer semantic representations into forecasting architectures is often not treated as an end-to-end modeling problem. The perspective adopted in this paper is that contextual representations derived from modern transformer-based models can provide a more faithful encoding of market-relevant narratives, and that these representations should be integrated into FX forecasting models under careful timing constraints and evaluation design consistent with real information flow (Yang et al., 2025; Xing et al., 2021).

Methodology

This study applies bibliometric methods to map and summarize research at the intersection of foreign exchange (FX) forecasting, sentiment analysis, and recent developments in language-based artificial intelligence. The research design combines quantitative analysis of publication metadata with a qualitative screening step, so that the resulting dataset remains focused on papers that directly address the scope of the study.

The methodological framing is aligned with prior research that connects textual signals and market prediction, including approaches that leverage latent economic concepts and sentiment in news for financial forecasting (Farimani et al., 2021). In addition, evidence that foreign exchange risk can be relevant for forecasting activities in financial contexts supports the broader motivation of examining FX-related information effects (Ho et al., 2020).

Research questions and hypotheses

The analysis is guided by three research questions. The first question examines which themes dominate the literature that links FX forecasting with sentiment-based signals. The working hypothesis is that the literature organizes into connected thematic areas, typically separating sentiment and text-processing topics from forecasting models and evaluation practices, while still showing strong links between them. The second question evaluates how these themes have evolved in recent years within the selected corpus. The corresponding hypothesis is that more recent work places greater emphasis on machine learning and deep learning approaches and on sentiment signals derived from digital information sources. The third question investigates collaboration

patterns among authors within the filtered dataset. The hypothesis is that co-authorship links are limited, suggesting a fragmented collaboration structure in this research niche.

Data source and database selection

The bibliometric dataset was retrieved from the Clarivate Web of Science database. This database was selected because it offers standardized bibliographic metadata suitable for systematic quantitative analysis and because it supports reproducible search queries that can be rerun and audited. Its coverage across economics, business finance, and computer science makes it appropriate for research positioned at the boundary between finance and artificial intelligence.

Search strategy and query formulation

The literature search was implemented in three sequential steps, each targeting a different component of the overall research scope. The first step focused on FX forecasting studies that explicitly include sentiment analysis. The query used was:

TS=("foreign exchange" OR forex) AND TS=(forecast* OR prediction) AND TS=("sentiment analysis" OR sentiment)

This query returned 48 publications. After applying the selection criteria described below, 34 publications were retained and treated as the core dataset for bibliometric mapping.

The second step expanded the search toward recent work on Large Language Models in financial contexts, with the goal of identifying emerging directions relevant to FX and financial forecasting. The query used was:

TS=("large language model*" OR LLM OR GPT) AND TS=(finance OR "foreign exchange")

The third step targeted papers that provide clear methodological structures that can serve as reference models for reporting research design and empirical workflows. The query used was:

TS=("exchange rate forecasting" AND "sentiment analysis")

The second and third steps were used to support positioning and gap identification, while the first step remained the basis for the main bibliometric analysis.

Selection criteria and screening logic

All retrieved records were screened using predefined rules. Only Articles and Proceedings Papers were included, while editorials and book chapters were excluded. The screening prioritized records indexed in research domains aligned with the topic, namely Economics, Business Finance, Computer Science, and Artificial Intelligence. The publication period was set to prioritize 2019 onward in order to reflect contemporary methodological practice, while allowing the inclusion of earlier publications when needed to support theoretical framing. Relevance was assessed through inspection of the title, abstract, and keywords. To be included, a study needed to address FX forecasting, sentiment extraction from text, or model designs that incorporate textual information into forecasting pipelines in a way that is directly applicable to the present scope.

Analysis procedure and software

The retained publications were analyzed along methodological and structural dimensions. The review tracked how sentiment or textual information is obtained and represented, what predictive model families are used, and what evaluation choices are reported. The analysis also examined how text-derived signals integrate with numerical market data or forecasting architectures, since integration choices often drive performance differences across studies.

For the bibliometric mapping, VOSviewer (version 1.6.19) was used to generate networks from the Stage 1 core dataset. The outputs include a keyword co-occurrence map highlighting thematic clusters, an overlay visualization showing temporal distribution, and a co-authorship network summarizing collaboration links. The supplementary sets contextualized findings and supported gap discussion without changing the core dataset definition.

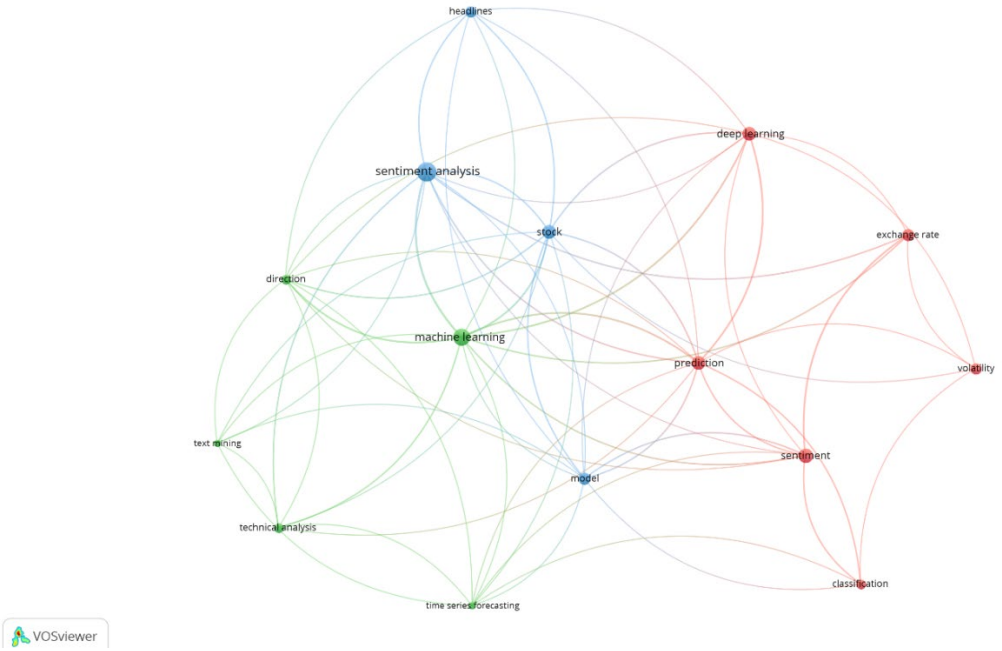


Figure 1. Keywords co-occurrence network identifying sentiment extraction, deep learning, and multimodal clusters.

Source: Authors' processing in VOSviewer, using WoS indexed articles.

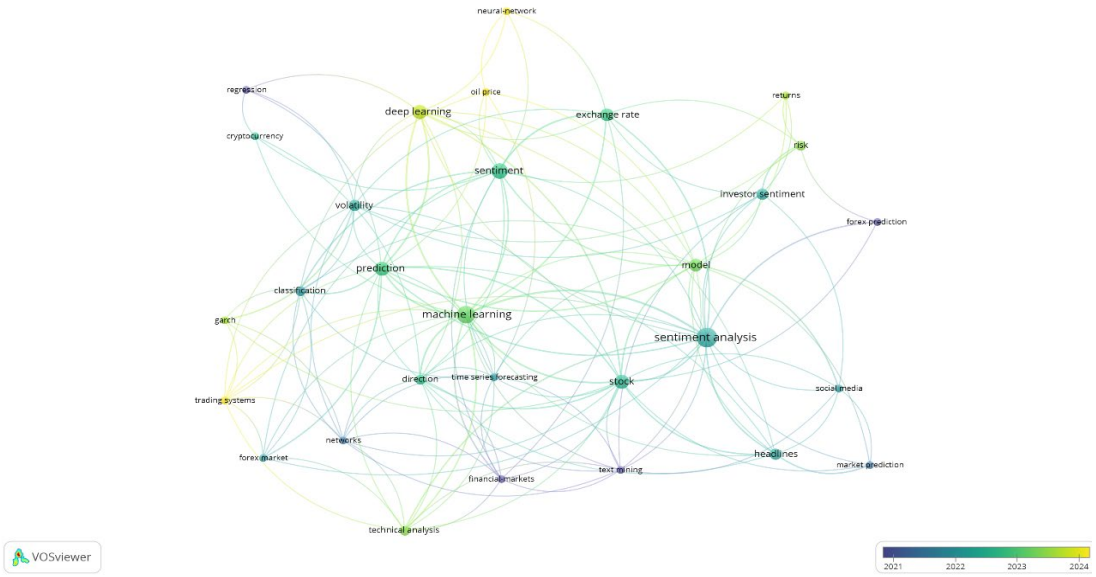


Figure 2. Overlay visualization map showing temporal evolution (blue=older, yellow=recent: LLM/ChatGPT emergence).

Source: Authors' processing in VOSviewer, using WoS indexed articles.

To analyze collaboration patterns within the filtered corpus, a co-authorship network was constructed in VOSviewer based on the core dataset (Stage 1) and colored by average publication year. As shown in Figure 3, the network is dominated by small groups with relatively limited links between them, indicating a fragmented collaboration structure in this research niche. The most visible clusters are centered around Cheng (2019), Xing (2021), and Yang (2025), with low connectivity across clusters.

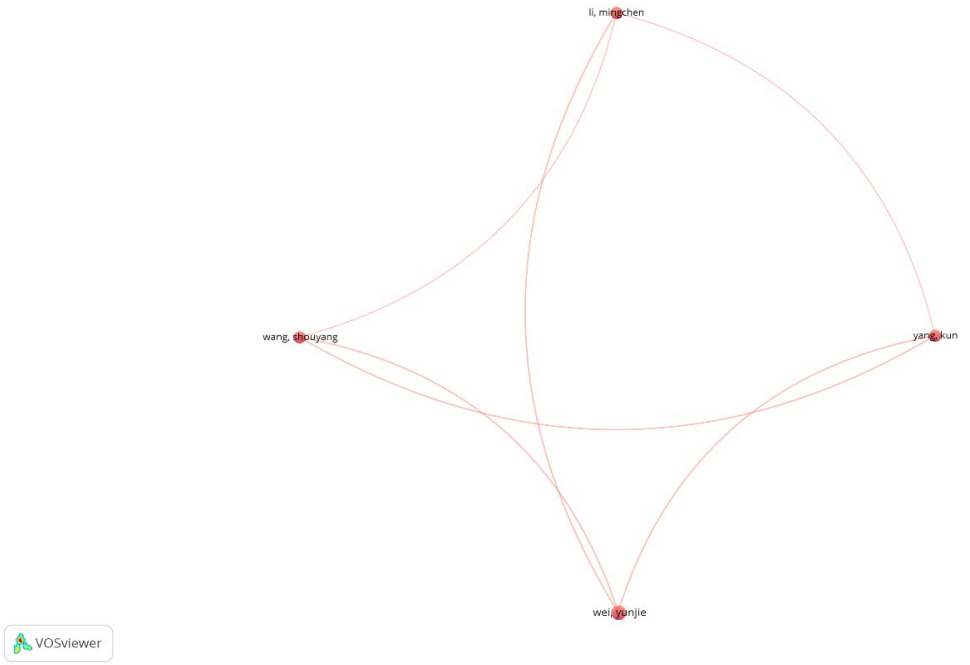


Figure 3. Authors co-authorship network (colored by average publication year) revealing clusters around Cheng (2019), Xing (2021), Yang (2025).

Source: Authors' processing in VOSviewer, using WoS indexed articles.

Results and discussions

The VOSviewer maps built from the Stage 1 dataset describe how research connecting sentiment analysis, machine learning, and exchange rate prediction is organized and how it has shifted over time. In Figure 1, the keyword network forms three main groupings. One grouping is centered on sentiment-related terms such as "sentiment analysis" and "investor sentiment". A second grouping is dominated by modeling terms including "machine learning", "deep learning", and "prediction". A third grouping connects exchange-rate terms, such as "exchange rate" and "forex market," with application-oriented terms, such as "trading systems". The structure suggests that sentiment extraction, model development, and forecasting applications have developed as distinct lines of work, while remaining closely linked through shared concepts and frequent co-occurrence.

Figure 2 adds a time perspective using the overlay visualization. Older terms in the network are concentrated around broad themes such as sentiment analysis and machine learning, while newer terms move toward deep learning and neural-network methods and toward exchange-rate-specific prediction tasks. Within the more recent part of the map, the appearance of terms associated with ChatGPT or Large Language Models is consistent with a broader shift in financial text research toward models that capture richer linguistic context. This pattern supports the interpretation that LLM-based approaches are emerging as an extension of prior deep learning work in financial sentiment and market prediction, rather than as an isolated research direction.

Figure 3 presents the co-authorship network for the core set of papers most closely related to LLM–Forex links in the filtered corpus. The network shows a small and relatively compact collaboration structure, with visible connections around authors such as Yang, Wang, Li, and Wei. Given the limited size and density of this network in the selected dataset, the collaboration structure appears concentrated, consistent with an emerging field in which only a small number of teams repeatedly publish on closely related LLM-based FX forecasting questions.

These bibliometric patterns provide context for interpreting the substantive contributions of specific studies within each thematic area. Cheng et al. (2019) introduced a fine-grained approach to sentiment analysis of foreign exchange news, arguing that more granular classification aligns sentiment signals more closely with market movements than coarse polarity schemes.

Seifollahi and Shajari (2019) examined word sense disambiguation for sentiment analysis of news headlines in the Forex context and reported improvements when semantic ambiguity is explicitly addressed. Their findings are consistent with the broader methodological point that semantic precision matters in financial text processing, particularly when market language includes polysemy and domain-specific usage.

Xing et al. (2021) extended sentiment-based forecasting to higher-frequency settings and provided evidence that sentiment signals extracted from news streams are relevant at intraday horizons, where microstructure and behavioral factors may play a larger role in price formation than in longer-horizon forecasts. More recently, Yang et al. (2025) evaluated the use of ChatGPT for processing financial text and predicting exchange rate movements, reporting that modern LLM-based text processing can outperform traditional NLP pipelines in extracting market-relevant information. Within the scope of this paper, that study serves as direct motivation for examining the connection between LLM semantic representations and FX forecasting workflows.

Across the reviewed literature, the methodological progression follows a recurring pattern in which each step reduces specific weaknesses of earlier approaches while introducing new constraints. Dictionary-based sentiment methods provided an initial demonstration that text contains signals correlated with market dynamics, but they are limited by context insensitivity and a fixed mapping between words and sentiment. Machine learning classifiers improved this by learning task-specific decision boundaries, but they remain dependent on labeled datasets and can degrade under distribution shifts. Deep learning approaches, including recurrent models applied to sentiment classification, improved representation learning and the ability to capture sequential patterns, but many implementations still treat sentiment extraction as a separate pipeline that feeds predefined sentiment variables into downstream forecasting models.

A consistent implication across these strands is that sentiment quality can materially affect forecasting quality, yet sentiment is often operationalized as an intermediate feature rather than a jointly optimized component of the forecasting model. This separation can create bottlenecks, because relevant information may be compressed or lost when text is reduced to a small set of sentiment scores before prediction. LLM-based approaches address this limitation by enabling richer text representations and, in some implementations, a closer coupling between text understanding and prediction. In this sense, LLM-based forecasting can be framed as a shift from two-stage pipelines toward more integrated semantic modeling, in which representation learning and prediction are more closely connected.

The bibliometric synthesis also highlights several areas that remain underexplored. One gap is the limited number of studies that systematically integrate LLM-based text representations into Forex forecasting with clear evaluation protocols and comparable baselines. A second gap concerns hybrid designs that combine LLM-driven semantic information with econometric structure and

established technical indicators. While deep learning methods are increasingly common, the question of how to blend distributed semantic representations with explicit economic reasoning remains open in the literature covered by the dataset. A third gap relates to interpretability. Even when LLM-based methods show strong predictive performance, many studies provide limited transparency on which semantic cues drive predictions, which can be a constraint in high-stakes financial settings. A fourth gap is domain adaptation. Because many LLMs are trained on general-purpose corpora, their performance and robustness in FX contexts may depend on adaptation to financial language, on the stability of terminology over time, and on the ability to remain reliable as market narratives change.

Taken together, the results suggest that LLM-based approaches can be understood as a convergence point for multiple research trajectories. Earlier work established that sentiment and news carry information relevant to exchange-rate dynamics. Subsequent advances in machine learning and deep learning improved the ability to extract and use that information. Recent work on LLMs extends these developments by offering stronger contextual language understanding and more flexible semantic representations. Within this framing, LLM-augmented Forex forecasting appears theoretically grounded and increasingly feasible, while still requiring clearer methodological standards for evaluation, integration strategies that connect semantics with economic structure, and more substantial evidence of interpretability and robustness.

Conclusion

This paper used bibliometric methods to map research on foreign exchange forecasting that incorporates sentiment analysis and, more recently, Large Language Models. The dataset was collected from Clarivate Web of Science using a reproducible query strategy. The initial search returned 48 records, and 34 publications from 2019 onward were retained as the core corpus for detailed mapping. Using VOSviewer, the study visualized keyword structures, topic evolution over time, and co-authorship relations, which helped clarify how this research area is currently organized and where it is expanding.

A central finding is that sentiment-augmented forecasting has moved steadily from early lexicon-based approaches toward machine learning and deep learning pipelines. Across the literature reviewed, sentiment signals are repeatedly treated as useful inputs relative to models that rely solely on price-based indicators or purely econometric specifications, although the magnitude of the improvement depends on how sentiment is extracted and integrated into forecasting. At the same time, the literature highlights recurring limitations of earlier NLP methods and of many two-stage pipelines, especially when financial text requires contextual interpretation, disambiguation, and domain-specific understanding. These constraints help explain why performance gains are often inconsistent across market regimes, sources, and time horizons.

Recent work suggests that LLM-based text processing can reduce some of these limitations by producing richer contextual representations of financial language. Studies that evaluate LLMs in financial text settings report stronger performance than traditional NLP pipelines in extracting market-relevant information and, in some cases, in supporting exchange-rate prediction tasks (Yang et al., 2025). In the context of the present synthesis, this positions LLMs as a plausible next step in the broader trend toward deeper representation learning for financial forecasting, rather than a separate line of work.

Even with this momentum, the bibliometric evidence indicates that systematic research on end-to-end or tightly integrated LLM-augmented FX forecasting architectures remains limited. Much of the existing work either focuses on sentiment extraction as a standalone task or evaluates

language models in broader financial contexts without fully specifying how LLM-derived signals should be combined with econometric structure, technical indicators, or market microstructure features in a controlled forecasting design. This gap is important because the practical value of LLMs for FX forecasting depends less on isolated text performance and more on how language-based representations interact with numeric inputs, evaluation protocols, and market conditions.

Several limitations of this study should be acknowledged. The search and metadata extraction were restricted to Web of Science, so relevant studies indexed primarily in other databases may not be captured. The core dataset was defined through queries that prioritize explicit links between FX forecasting and sentiment, which can exclude adjacent work in other asset classes that may still offer transferable methods. In addition, LLM research evolves quickly, and some relevant contributions may exist as preprints or technical reports that are not yet fully represented in peer-reviewed indexing.

Future work can build on this mapping in a few clear directions. One priority is the development and empirical evaluation of LLM-augmented FX forecasting architectures using comparable baselines and consistent testing protocols. Another is to study integration strategies that combine LLM-derived semantic representations with econometric assumptions or domain constraints, rather than treating text as an isolated feature stream. Research is also needed on domain adaptation, including how fine-tuning or specialized pretraining affects performance and stability in financial language. Finally, interpretability and robustness remain necessary for practical adoption, particularly through tests across currency pairs, time periods, and market regimes, and through methods that clarify which textual cues drive predictions and why.

References

Albahli, S., Irtaza, A., Chesneau, C., et al. (2022). A machine learning method for prediction of stock market using real-time Twitter data. *Electronics*, 11(20).

Bormpotsis, C., Sedky, M., & Patel, A. (2023). Predicting Forex currency fluctuations using a novel bio-inspired modular neural network. *Big Data and Cognitive Computing*, 7(3).

Cheng, Z., Qi, T.M., et al. (2019). Fine-grained sentiment analysis of foreign exchange news. In *Proceedings of the 5th International Conference on Information Management (ICIM)*, pp. 279-284.

Chi, D. T. K., Kien, H. N. T., & Nguyen, T. Q. (2025). Enhancing forex market forecasting with feature-augmented multivariate LSTM models using real-time data. *Knowledge-Based Systems, 330*(Part A), 114500.Dakalbab, F., Kumar, A., et al. (2025). Advancing Forex prediction through multimodal text-driven model and attention mechanisms. *Intelligent Systems with Applications*, 26.

Farimani, S. A., Vafaei Jahan, M., Milani Fard, A., & Haffari, R. (2021). Leveraging latent economic concepts and sentiments in the news for market prediction. In P. Larrañaga & A. Bifet (Eds.), *2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 171–180). IEEE.

Farimani, S.A., Jahan, M.V., & Fard, A.M. (2022). From text representation to financial market prediction: A literature review. *Information*, 13(10).

Ho, T., Nguyen, Y., Parikh, B., & Vo, D.-T. (2020). Does foreign exchange risk matter to equity research analysts when forecasting stock prices? Evidence from U.S. firms. *International Review of Financial Analysis, 72*, 101568.

Mabrouk, N., Chihab, M., & Chihab, Y. (2025). A FOREX trading system based on semi-supervised news classification, market sentiment analysis, and GRU-CNN deep learning

models. *International Journal of Advanced Computer Science and Applications*, 16(9), pp. 846-859.

Naeem, S., Mashwani, W.K., et al. (2021). Machine learning-based USD/PKR exchange rate forecasting using sentiment analysis of Twitter data. *CMC-Computers Materials & Continua*, 67(3), pp. 3451-3461.

Seifollahi, S., & Shajari, M. (2019). Word sense disambiguation application in sentiment analysis of news headlines: an applied approach to FOREX market prediction. *Journal of Intelligent Information Systems*, 52(1), pp. 57-83.

Semiromi, H.N., Lessmann, S., & Peters, W. (2020). News will tell: Forecasting foreign exchange rates based on news story events in the economy calendar. *North American Journal of Economics and Finance*, 52.

Xing, F.Z., Hoang, D.H., & Vo, D.V. (2021). High-frequency news sentiment and its application to Forex market prediction. In *Proceedings of the 54th Annual Hawaii International Conference on System Sciences*, pp. 1583-1592.

Yang, K., Deng, R.X., et al. (2025). The power of ChatGPT in processing text: Evidence from analysis and prediction in the exchange rate markets. *Financial Innovation*, 11(1).

Yasir, M., Durrani, M.Y., et al. (2019). An intelligent event-sentiment-based daily foreign exchange rate forecasting system. *Applied Sciences-Basel*, 9(15).

Yono, K., Izumi, K., & Shimada, T. (2019). Extraction of focused topic and sentiment of financial market by using supervised topic model for price movement prediction. In *Proceedings of the IEEE Conference on Computational Intelligence for Financial Engineering and Economics (CIFEr)*, pp. 149-155.