International Journal of Artificial Intelligence, Data Science, and Machine Learning
Grace Horizon Publication | Volume 5, Issue 1, 167-172, 2024
ISSN: 3050-9262 | https://doi.org/10.63282/3050-9262.IJAIDSML-V5I1P116

Original Article

LLMs for Financial Document Processing

Venkata Sai Nageen Kanikanti
Director, Software Engineering.

***Abstract* - *The conventional methods of financial document analysis have been based on official numerical pointers, without considering the huge narrative data in company reports, regulatory submissions, morale reports, or sector statements. According to the recent progress in the sphere of -informed large language models (LLMs), machines have become capable of processing and encoding financial documents with better accuracy than at any previous point in history. This paper presents a multimodal LLM architecture, which makes use of financial reports to synthesize a narrative and lock it in with a time-series predictor in external downstream trainers that do not use the narrative, namely forecasting and control. As the energy and insurance sectors view of case studies evidence, document processing through LLM does increase the accuracy of prediction in addition, that it makes reading or comprehension superior, contributing to risk identification, harmonization of the rules, and decision-making. The findings highlight the value of LLMs as a methodological roadmap for financial document processing beyond the banking sector.*

***Keywords* - *Large Language Models (LLMs), Financial Document Processing, Natural Language Processing (NLP), Multimodal Learning, Financial Text Analytics, Time-Series Forecasting, Risk Assessment; Regulatory Compliance, Decision Support Systems, AI in Finance.*

1. Introduction

Financial decision-making would rely on processing and interpretation of financial documents being one of the main sources. Other industries like energy, insurance, retail and real estate also depend considerably on the streams of documents that are far rich in documents besides their numerical indicators like sustainability reports, regulatory filings, policy disclosures and litigation reports. LLMs enable reading, summarizing, and inferring such texts and transforming unstructured texts into structured knowledge that can be brought to time-sequences signals to make more accurate predictions and evaluate compliance, respectively. According to the One of the Researcher, among the aspects taken into account by insurance policies are the nature of natural disasters, health regulations, and litigation reports, as well as tracking the claims and premiums trends. Consumer sentiment, seasonal promotions, and news about the global supply-chain has had a devastating impact on retailing sales volume. One can quite safely assert that, regardless of the perspective that one is looking at when considering the financial reporting type, financial measures are already built upon both numerical measures (quantitative indicators) and narrative measures One of the Researcher.

In the past, models of forecasting have attracted concentration on figures. Statistics temporal analysis In application to statistical measures, regression, and autoregressive, and more recently neural network (LSTMs and transformers), they have all been more than capable to study trend over a time frame One of the Researcher. Their approaches have failed to take the narrative context (i.e. corporate storytelling, policy documentation or ESG disclosures). Text-based methods have also adopted sentiment, thematic or event-driven extraction of reports and news, but failed to integrate quantitative measures (narrative context) (One of the Researcher). This is a critical difference in approach that is left to a gap in the financial modelling where multimodal systems / methodologies would offer a more comprehensive view.

One of the investigators stated that, large language models have become a fast-changing element in NLP. These models propose embeddings of not only the sentiment but of the fine-tuned logic of financial text as well. These, along with the model of the numerical indicators in order, can compose one predictive system. It may be applicable particularly in non-banking contexts where both structured and unstructured information are highly influential in finance related decision making. The question which is being analyzed in the paper is the impact of multimodal learning in non-banking industries with regard to the accuracy of forecasts (Ang & Lim, 2022).

The work pursues three objectives. First, it produces a multimodal platform, which in turn tries to examine corporate text and market signals. Second, it compares this framework with two case studies in the energy and insurance industry, and, therefore, it is restricted to the sphere of application to the banking industry. Third, it focuses on the interpretability by means of related textual signals with the numerical process. In such a way, the specified study gives the future of financial forecasting that sheds the specific streams of data and that matures into a unified decision-making platform.

2. Literature Review

2.1. Foundations of Quantitative Forecasting in Finance

Conventional quantitative models, assumed to be working with non-metrical time-series information, have been the prevalent models in financial forecasting (Ang & Lim, 2022). The traditional market-pattern analysis variants were reflected in the conventional statistical analysis tools, which were moving averages, autoregressive integrated moving average (ARIMA), and a generalized autoregressive conditional heteroskedasticity (GARCH) (Lee & Yoo, 2020). These techniques have been satisfying in short term volatility forecasting and trend discovery however have failed in the description of non-linear and complex dynamics that are involved in the nature of financial systems.

It was when the machine-learning techniques were used that the nonlinear predictive models of the support-vector machine and random forests were adopted. As experimental experience indicates, these models are superior to a pure statistical approach because they are capable of detecting interactions inadequately captured in a linear context (Sardelich & Manandhar, 2018; Emami et al., 2023). The other one, most recent one is the application of deep learning as the next step, and the neural networks of the recurrent neural network (RNN) and long short-term memory (LSTM) come in the picture as sources of paramount significance. Their generalization of long-range dependencies enables them to more accurately predict sequence one of the Researcher. Moreover, time-series forecasting has been scaled to attention architecture, e.g. transformers, and offers the ability to rationalize contextual relationship at longer horizons and is resistant to vanishing-gradient issues.

2.2. Financial Document Processing with LLMs

As Lv et al. (2023) asserted that, as more sophisticated ways of forecasting are developed, financial document processing has become a focus. These reports and filings made by sustainability sources, regulatory filings, annual reports, and litigation reports are vital towards capturing investor perception, compliance risk, and sector valuation. LLMs provide a very powerful approach to automating key entity extraction, risk disclosure summary, and matching narrative indicators and documented indicators to facilitate vacuoles instead of mere sentiment analysis to understand the document better.

Previously Guo et al., (2019) stated that, handling textual information has relied on lexicon-based sentiment analysis in which the word dictionaries were tagged on the basis of their positive or negative polarity. Though such techniques might have been effective in deriving an overarching mood, they were ill suited to domain-specific language, irony or any subtle context that in financial markets is typical (Liu et al., 2020; Bellotti et al., 2021). The simple machine-learning classifiers such as Naive Bayes or logistic regression could only achieve small improvements and not much into semantics.

With the introduction of deep learning, and, most recently, large language models (LLMs) a new fundamentally new approach to financial text has emerged. The encoding potential of transformer-based models to encode the linguistic and semantic relationships breaks down complex sentences and derives salient features. The foregoing systems can extract sentence-level sentiment due to the dense embeddings generated by LLM, which are emotionally charged and replete with risk cues and context, thereby also capable of performing effective reasoning at higher semantics (Bellotti et al., 2021). These skills are particularly effective when a single phrase in a corporate disclosure filing or a regulatory filing pre-reveals a key future implication.

2.3. Unimodal Limitations in Financial Forecasting

Although the developments in numeric and text based forecasting are substantial, there are few unimodal forecasting pipelines (Liu et al., 2020). Only numerical-only approaches based on statistical or deep-learning can generate the discontinuities in market dynamics, not the qualitative causes of such a discontinuity; whether it is the result of a regulatory impact, an environmental risk, or a shift in consumer behavior. Text-based approaches, however, also prove useful when it comes to drawing conclusions regarding the tone, perspective, and background of the story, yet lack certain quantitative problems that affect the financial performance. Market mood, in its turn, can be determined by sentence-level sentiment indicators on the basis of textual analysis, yet such models are unable to consider price history, or the macroeconomic variables underlying objective market behavior (Chan & Qu, 2016).

It is the incompleteness of the data of financial forecasting, thus the unimodal gap. The lack of modality integration results in models that either omit contextual data--as in purely numerical systems--or disregard quantitative data, as in purely text-based systems. This weakness leads to the integration of models that are integrative in nature and can manage numerous data streams at once not only to accommodate the dimensions of the qualitative but also the quantitative (Chan & Qu, 2016).

2.4. Emergence of Multimodal Forecasting Approaches

To address these weaknesses, scientists were beginning to explore multimodal solutions when the data is expressed as a text and numbers. Two significant strategies have been mentioned i.e.; early fusion and late fusion (Loughran & McDonald, 2020). To connect textual embeddings to numbers at the input phase, early fusion is applied to generate an individual representation to the forecasting model. This can also be used to conceal the peculiarities of modalities in as far as it aids in cross-modal learning. In contrast, late fusion processes data models individually, e.g. textual data with natural language models and numerical sequences with, e.g. time-series models, then bringing together the results at the decision layer (Sawhney et al., 2020; One of the Researcher). This will not compromise modality-specific advantages, but needs to be well integrated.

The potential of multimodal schemes has already manifested itself in banking operations like credit score, fraud recognition and risk evaluation of a loan (Sawhney et al., 2020). Non-financial adoption is not however widespread. The available literature still has a part that remains in the field of stock exchange and banking and therefore creates a vacuum in areas to which multimodal signals would as well be applicable and even to greater measures.

2.5. Gaps and Opportunities in Non-Banking Financial Domains

Multimodal forecasting has many opportunities in non-banking financial sectors-energy, insurance, retail, and real-estate sectors. These industries have huge generation of numerical and textual data. The price of commodities and sustainability reports reported by the energy firms shows long-term commitment promises of becoming carbon neutral. Insurance companies publish regulatory reports and claims statements that can influence the premium flows and risk in the market (Guo et al., 2019). The retailers issue the sales data and the strategic news of consumer promotions or alterations in the supply chain. Real estate developers would provide their demand estimates and also provide reviewed policies on zoning and building One of the Researcher.

Despite this richness, there are not many powerful multimodal models, which have been applied to these fields. The existing research is either inclined to single out the numerical indicators or restrict the quantitative analysis to the dichotomy of sentiment, and the complete integrative potential is not exhausted (Bellotti et al., 2021). The absence of systematic multimodal strategies creates unpredictable holes of forecasting reliability particularly in cases where external influences or policy discourses that are not expected are the cause of the sector movements. There is, therefore, a possible opportunity as suggested in the literature: multimodal architectures that integrate textual embeddings using LLM with time-series analysis of numerical data (Bellotti et al., 2021; One of the Researcher). Such systems will not only ensure a higher degree of forecast accuracy but will also ensure greater interpretability. Multimodal strategies can also help to make decisions in uncertain situations where quantitative signs are presented, revealing their relationship with stories One of the investigators.

3. Methodology

3.1. Research Design

The hybrid computational design, i.e. the natural language processing and deep learning time-series analysis will be used in this study. The design is based on the understanding that there is no textual and numeric data that can be applied to explain the financial behaviors within non-banking sectors (One of the Researcher). The integration of the advantages of the two modalities is then aimed to give more accurate and context specific predictions. The research design has three components, all of which are connected with each other: text processing, which generates meaning and sentiment out of financial texts; numerical time-series modeling, which generative historical and cyclical patterns out of structured data; and multimodal fusion, a pooling of these free representations into one forecasting model (Lee & Yoo, 2020). A combination of these elements forms an end-to-end pipeline capable of receiving a wide range of financial inputs and making predictions that also include quantitative and qualitative stories.

3.2. Data Collection

To ensure that it is not limited to the banking field, this study focuses on two significant sectors, i.e. energy and insurance. These industries have been selected because they generate a large amount of both structured and unstructured data, and each of them has a direct impact on the outcomes of the forecasts.

- Energy sector datasets include annual reports, sustainability disclosures, and climate policy statements from listed energy firms. These textual records are paired with commodity price histories, demand indices, and production statistics to provide a balanced view of qualitative and quantitative signals (Sardelich & Manandhar, 2018).
- Insurance sector datasets incorporate regulatory filings, claims outlooks, and litigation disclosures as the textual corpus. Numerical indicators include premium revenues, loss ratios, and payout histories, which capture sector-specific performance.

The models will be trained on data spanning the course of several years in order to learn both the cyclic behavior (e.g., seasonal energy demand or regular insurance claim cycles) and structural behavior (e.g., the changes in long-term policy or risks related to climate). Longitudinal nature of the data sets enables one to perform the forecast in which the short term variation and the long term direction are considered.

3.3. Financial Document Processing with LLMs

LLM pipeline, once pipelines are targeted at domain-specific, long-form content, financial documents (regulatory filings, ESG disclosures, or insurance claim reports) are handled with a LLM pipeline. The pipeline is having a preprocessing (cleaning, tokenization, segmentation), and consequently, generating dense embeddings through LLMs. These embeddings are sentiment on one hand and entities (e.g., risk factors, compliance requirements, and financial metrics) or contextual laws between points on the other. Numerical indicators are then adjusted against the processed outputs in time so as to allow the forecasting of the integrated results and analysis of the compliance.

The creation of dense embeddings is done using large language models (LLMs) after the preprocessing. Such embeddings differ with the conventional bag-of-words sentiment or lexico-based sentiment measures which encode one-dimensional sentiment, semantics and context measures (Bellotti et al., 2021). The documents are mapped into high dimensional form of numerical vectors that give the underlying meaning of those documents, their tone and their implications in future.

The textual embeddings are averaged and dated by the year of publication as a mechanism to make the data align along the time dimension of the numeric data. This will enable the process of correlating corporate disclosures, or regulatory filings or sustainability reports with the time-windows (Sardelich & Manandhar, 2018). The output is a rational sequence of textualizations which may be used with the numerical models.

3.4. Time-Series Modeling

Using the numerical pipeline, structured indicators are run using prices, volumes of demand, premium flows, and industry-specific ratios. The signs are also sequential in nature and deep learning architectures like Long Short-Term Memory (LSTM) networks and transformers are especially adept at that (Rodrigues et al., 2019). LSTM applications rely on the capacity to replicate long-term dependencies and periodicities, whereas transformer applications relies on the capacity to find attention-based correlations in a long sequence. Sufficient history is provided to the model and sufficient output layers to give projections of financial indicators in the future. The fact that it tested both the two architectures gives the study a wide range of the changes over time, the local variations to global sectoral behaviour.

3.5. Fusion Strategy

The multimodal framework centers on the fusion mechanism. A late fusion approach is chosen in this research, whereby modality-specific information is kept within each modality only to be assembled at the conclusion of the processing (Emami et al., 2023).

- The textual pipeline produces embeddings that encapsulate semantic and sentiment information.
- The numerical pipeline generates hidden states representing temporal dynamics.

In the fusion step, the results of these are added and run through a dense layer of neurons that learn to know cross-modal interactions. The prediction head which is the last one exports the forecast value. The plan will contribute to the fact that the two modalities supplement each other: qualitative data that is provided by LLMs and time series prediction capabilities of deep learning models. Baselines-text-only and numerical-only pipelines are compared with the model to see the value contributed by the multimodal integration.

4. Evaluation

Model performance is assessed using a range of forecasting metrics**:**

- Root Mean Squared Error (RMSE): measures the average magnitude of prediction errors.
- Mean Absolute Percentage Error (MAPE): captures percentage-based deviations, useful for comparing across industries.
- Directional Accuracy: evaluates the ability of the model to correctly predict the direction of change, a key measure for financial decision-making.

To make it strong, cross-validation over time is done in the multi-time period to prevent overfitting and to test the generalization capacity (Guo et al., 2019). Case study of insurance industry and case study of energy industry illustrate the results of applying the framework to different data environments. This is done to develop a fusion of both textual and numerical data on basis of a reasonable multimodal design. The time series of the financial indicators and the processing of the text can be processed and the deep learning models can be applied to retrieve the rich semantic indicators in the corporate documents respectively (Bellotti et al., 2021). Fusion when these insights are combined into one forecasting pipeline, the output is compared to unimodal baselines, and is tested on strong metrics. Similar treatment of other non-banking industries like energy and insurance suggests that the paper may offer a deeper and more realistic explanation to financial forecasting as far as multimodal techniques are concerned.

5. Experiments and Results

The data set used (emerging energy and insurance) are experimented on. LLMs in the energy industry handle sustainability disclosures and climate policy reports converting material risks and commitments and fusing them with commodity price history and demand indices. The regulatory filings and disclosures of litigation are under processing in the insurance industry, which is to recognize compliance issues and arising risks which is incorporated with the premium revenues and loss ratios. As the results reveal, the accuracy of the forecasts and the overall detection of sector-specific risks made when using LLMs improve in comparison with unimodal baselines.

Directional accuracy is also improved with multimodal integration implying that the models are better positioned to signal the upward or downward direction of the market (One of the Researcher). The comparison of graphs of the predicted and the actual curves reveals a more similar result than the feature attribution analysis shows, which demonstrates the relevance of both text and numbers. This is the case when negative sentiment in the sustainability reporting is correlated with reductions in energy demand expectations and positive regulatory attitudes in insurance filings are correlated with premium growth expectations (Chan & Qu, 2016).

6. Discussion

This paper illustrates that non-banking financial services improve the decision-making process considerably when using the financial document processing by LLM. LLMs can be used to give more valuable inputs to be integrated with numerical time-series models by deriving structured information about financial reports (ignoring restatements), including e.g. ESG commitments, regulatory risks, or litigation disclosures. With this connectivity, not only forecasting is enhanced but also it helps in monitoring compliance, detection of fraud and also evaluations of risks. It highlights interventions of the LLMs in the modification of unstructured finance documents into practical intelligence.

The suggested framework provides a viable basis on how to come up with new predictive knowledge in industries. As an example, energy firms can correlate the policy documents and sustainability reports with commodity demand and price dynamics, and thus predict how the change in regulatory stories or environmental pledges will shape market dynamics (Chan & Qu, 2016). According to the One of the Researcher. Insurance companies can combine regulatory reports, climate risk reports, and claims perspectives with financial ratios to have a more refined representation of risks that relate to extreme weather events, litigation, or health-related claims. Likewise, retailers are able to align the sales reports and financial statements with the consumer sentiment as determined by the announcements made by the management or through news and media reports, and therefore forecast their demand more quickly than the market (Bellotti et al., 2021). Real-estate businesses can use the combination of construction cost indexes, rental yield information and regulatory disclosures to predict the financial implications of zoning changes or changes to fiscal policy.

Although the advantages are self-evident, the methodology is not without challenges. To achieve time-series correlation of textual data, exact preprocessing must be used to guarantee time correlation. Misalignment that pairs off a different time window than the one represented in the documents may distort the results, and reduce the predictive validity (Sardelich & Manandhar, 2018). Computational needs are other hindrances. The large language models also need considerable resources, particularly on extended financial statements or regulatory submissions, or policy documents. This makes the framework resource intensive especially where small organisations are concerned and have very little computation resources (Emami et al., 2023).

There is an unanswered question of interpretability. Although fusing schemes are capable of enabling a robust integration, deep neural architecture often conceals the relative importance of each modality. Information on how far a prediction was fueled by text feelings or number relationships is an important factor to the decision-maker who would like a clear representation of the decision-making process to be transparent and responsible One of the Researcher. On the whole, this discussion demonstrates that the idea of multimodal integration is a promising trend in the financial forecasting of nonbanking industries, but it requires further adjustments to overcome the problems of data alignment, computation costs, and interpretation of the model.

7. Conclusion and Future Work

In conclusion, the findings presented within the current research, it is possible to suggest a LLM model of processing non-banking sector financial documents. The model derives its findings based on regulatory filings, sustainability reports and corporate disclosures and blends them with time-series indicators. This method does not only help in developing strong performance in predicting but also in increasing interpreting and oversight in terms of compliance. The results point out that the LLMs offer a methodological framework on the basics of turning document-intensive financial systems into structured and data-driven decision platforms.

The research provides a methodological blueprint and proves its efficiency with references to the empirical evidence. The research opens the door to more potent, more thorough predictive systems through the unification of financial narrative and quantitative study and has straightforward practical consequences. There are a number of directions that can be taken by future work. First, real-time multi modal forecasting would make use of real-time news and high-frequency prices as they are being streamed. Second, multimodal AI explainability would improve transparency and allow regulators and practitioners to have a better grasp of predictions. Third, it would be a challenge to extend the approach to other areas, like retail or real-estate, as it would be a test of its ability to be generalized. Lastly, reinforcement learning may allow proactive, dynamic decision-support systems.

References

[1] Ang, G., & Lim, E. P. (2022, May). Guided attention multimodal multitask financial forecasting with inter-company relationships and global and local news. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 6313-6326).

[2] Bellotti, A., Brigo, D., Gambetti, P., & Vrins, F. (2021). Forecasting recovery rates on non-performing loans with machine learning. *International Journal of Forecasting*, *37*(1), 428-444.

[3] Chan, Y. Y., & Qu, H. (2016, January). Finavistory: Using narrative visualization to explain social and economic relationships in financial news. In *2016 International Conference on Big Data and Smart Computing (BigComp)* (pp. 32-39). IEEE.

[4] Emami, H., Dang, X. H., Shah, Y., & Zerfos, P. (2023). Modality-aware Transformer for Financial Time series Forecasting. *arXiv preprint arXiv:2310.01232*.

[5] Guo, W., Wang, J., & Wang, S. (2019). Deep multimodal representation learning: A survey. *Ieee Access*, *7*, 63373-63394.

[6] Lee, S. I., & Yoo,S. J. (2020). Multimodal deep learning for finance: integrating and forecasting international stock markets. *The Journal of Supercomputing*, *76*(10), 8294-8312.

[7] Liu, Y., Li, Y., & DeGeronimo, B. A. D. (2020). FinBERT: A Pretrained Language Model for Financial Communications. *Journal of Computational Finance*, *24*(4), 15.

[8] Loughran, T., & McDonald, B. (2020). Textual analysis in finance. *Annual Review of Financial Economics*, *12*(1), 357-375.

[9] Lv, X., Xiong, X., & Geng, B. (2023). Increasing the prediction performance of temporal convolution network using multimodal combination input: Evidence from the study on exchange rates. *Frontiers in Physics*, *10*, 1008445.

[10] Rodrigues, F., Markou, I., & Pereira, F. C. (2019). Combining time-series and textual data for taxi demand prediction in event areas: A deep learning approach. *Information Fusion*, *49*, 120-129.

[11] Sardelich, M., & Manandhar, S. (2018). Multimodal deep learning for short-term stock volatility prediction. *arXiv preprint arXiv:1812.10479*.

[12] Sawhney, R., Mathur, P., Mangal, A., Khanna, P., Shah, R. R., & Zimmermann, R. (2020, October). Multimodal multi-task financial risk forecasting. In *Proceedings of the 28th ACM international conference on multimedia* (pp. 456-465).