SentimentPulse: AI-Driven Sentiment Analysis for Financial News

Sandhya Reddy Kallem, Dhruv Maheshbhai Ranpariya, Yashwanth Thalla,
Hema Durga Prasad Chittala, Sravan Kumar Ganji

Seidenberg School of Computer Science and Information Systems
Pace University, New York, NY, USA

sk44000n@pace.edu
dr46704n@pace.edu
yt73563n@pace.edu
hc63409n@pace.edu
sg66926n@pace.edu

***Abstract*— News events, within this modern information era, make profound effects within the financial markets. However, news items arrive in huge chunks, which makes the identification of actionable insights from them difficult for investors. SentimentPulse solves that with an AI-powered sentiment analytics financial news platform. It runs sorted categorizations of news into either positive, negative, or neutral sentiments based on high-level NLP methodologies, more so transformer-based models like BERT. With features such as enterprise mention detection, historical sentiment trends, and user-filterable results, SentimentPulse enables users to make data-driven investment decisions. This paper outlines the design, development, and deployment of the platform, emphasizing real-time processing, scalability, and impact on financial decision-making.**

***Keywords*— Sentiment Analysis, Financial News, Natural Language Processing (NLP), BERT (Bidirectional Encoder Representations from Transformers), Machine Learning, Transformer Models.**

I. Introduction

Financial markets work well if timely and precise information is provided. News is used by investors for checking on market sentiment, but through digital growth, identifying relevant and material insights has become impossible to find. Traditional tools, such as news aggregators, still lack proper sentiment analysis, and investors themselves have to go through large data sets of information.

SentimentPulse fills this gap by offering an AI-driven, real-time solution for sentiment analysis in financial news articles. This platform leverages highly trained machine learning models on financial datasets to output the sentiment scores to users, highlighting the relevant company mentions. It intuitively helps investors make decisions faster and confidently.

II. Literature

Sentiment financial analysis has grown into a key tool in financial markets, with many studies evidencing its predictive power. The work of Bollen et al. (2011) identified a striking correspondence between Twitter sentiment and stock market performance, showing just how easily financial dynamics can be swayed by popular sentiment. However, traditional models often struggle to make sense of the tortuous language that is common in financial news, which seriously limits their usefulness. This gap has therefore led to calls for domain-specific systems such as SentimentPulse, which uses sophisticated machine learning in trying to surmount these challenges.

Recent advancements in NLP, especially the transformer-based models like BERT, have really upgraded the perspective towards SA. Unlike typical systems, BERT learns context bidirectionally, hence giving a more in-depth understanding of complex financial texts. Despite such developments, what is available at the moment with common news aggregators or sentiment analysis programs is deceptively simple, normally lacking real-time capabilities and the level of customization that would be needed to support multiple investment approaches. This is a need to which SentimentPulse responds with real-time, accurate sentiment analysis, tailored for the financial industry, hence offering investors practical insights.

III. Current Solutions

Traditional ways to analyze financial news involve news aggregators, stand-alone sentiment analyzers, and combined approaches. The many news aggregators such as Google News give access to a wide range of data, but news aggregation software offers no facility for analyzing sentiment; interpretation of the mood beneath the articles is left to the user. Standalone solutions of sentiment analysis provide basic sentiment classification; they remain very generic in order to deal with such levels of complexity in financial news, hence making the interpretations incomplete or inaccurate. Hybrid models try to close this gap by putting news aggregation and sentiment assessment together. These systems hardly ever offer true real-time processing with rich insights for a specific end-user tailored need.

SentimentPulse overcomes these limitations by incorporating sophisticated natural language processing technology along with specific knowledge of the financial industry. It offers real-time sentiment analysis, providing immediate information about news relevant to the market. Users can personalize their output using filters for stock symbols, sectors, and sentiment types for maximum relevance of the delivered results. Historical sentiment graphs are also provided by SentimentPulse, enabling users to view changes in sentiment over time and helping them optimize their investment decision-making.

IV. Project Requirements

A. Functional Requirements

1) *User Authentication:* Secure login using email or OAuth.
2) *News Aggregation:* Fetch articles from global financial sources.
3) *Sentiment Analysis:* Classify articles as positive, neutral, or negative.
4) *Company Mention Detection:* Highlight companies referenced in articles.
5) *Historical Sentiment Trends:* Provide graphical representations for sentiment data.
6) *Filtering Options:* Allow filtering by sentiment type, stock symbols, and date ranges..

B. Non-Functional Requirements

1) *Performance:* Fast and scalable to handle high data volumes.

2) *Reliability:* Minimal downtime with robust error handling.

3) *Security:* Adherence to data privacy regulations.

4) *Compatibility:* Accessible on web and mobile platforms.

5) *Usability: Intuitive interface requiring minimal user training.*

V. System Design

A. System Architecture

The platform architecture comprises:

1) *Frontend:* Built with ReactJS for responsive user interfaces.
2) *Backend:* Flask and ExpressJS handle API integrations and data processing.
3) *Database:* MongoDB stores user preferences and sentiment data.
4) *Cloud Hosting:* Google Cloud Platform ensures scalability and availability.
5) *Machine Learning:* PyTorch and Huggingface models provide sentiment analysis

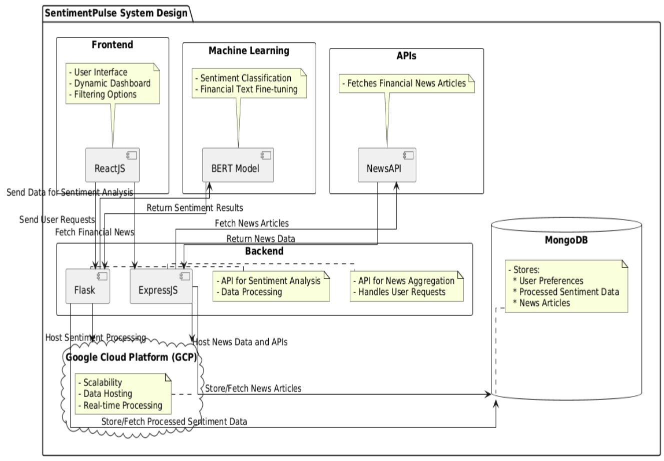

Fig. 1 System Design.

B. Frontend

The frontend offers:

1) *Dashboards:* Summarize key insights with sentiment scores and company mentions.
2) *Filters:* Enable users to customize news views based on preferences

C. Backend

The backend ensures secure and efficient processing, with features such as:

1) API integration with NewsAPI for real-time news fetching.
2) Connection to sentiment analysis models for classification

VI. Machine Learning

A. Introduction

SentimentPulse uses BERT for sentiment analysis, fully differentiating between the various types of financial news articles. Excellence in performance in BERT is achieved by the deep, contextualized understanding of language, which is a point of high relevance with financial texts where the sentiment can often depend on subtlety. Owing to its bidirectional nature, it captures the complete meaning of the sentences based on contextual analysis in both the forward and backward directions.

B. Model Selection

It chose BERT because it outperforms other models in terms of contextual understanding-particularly in the case of financial texts that rely on subtle language cues to denote sentiment. This is ensured by its high-end, transformer-based architecture, which makes it perfect for SentimentPulse.

C. Data Preparation

It all starts with proper data preprocessing in a machine learning pipeline; this involves the tokenization and cleaning of the news articles to remove irrelevant information and noise. Fine-tuning of BERT is conducted on financial-specific datasets, ensuring the model gets tuned for the domain. Data balancing was conducted to make equal numbers of the sentiment classes (positive, neutral, and negative) in order to reduce bias during training, hence improving the overall performance.

D. Model Training

The fine-tuned BERT model is iteratively trained until its accuracy in validation datasets reaches an initial 80%. This will be further improved with more rounds of fine-tuning and increasing the training dataset, including implementation of the feedback loop to improve its performance.

E. Model Evaluation

Standard evaluation metrics such as accuracy, precision, recall, and F1-score are then applied to measure performance by their means. This makes the model robust and reliable to prove its appropriateness for a real-world application.

F. Prediction and Deployment

The model is then deployed on the backend after successful training and verification. This deployment now enables the SentimentPulse to process current news articles and instant sentiment predictions. Investors, with this real-time capability, are warranted to get timely and actionable insights that can help power their decision-making processes. SentimentPulse ensures, through this architecture, that scalable, responsive, and accurate sentiment analysis is delivered in the dynamic financial environment.

VII. Conclusions

SentimentPulse really unlocks the true potential of AI-powered sentiment analysis in financial decision-making by fusing state-of-the-art machine learning models with user-centered design in proactive insights for smart investment decisions with confidence. The next development is going to include multilingual support, predictability analytics, and further data source addition.

References

[1] Lusardi, A., & Mitchell, O. S. (2011). Financial literacy around the world: An overview. Journal of Pension Economics & Finance, 10(4), 497-508.

[2] Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185(4157), 1124-1131.

[3] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

[4] Bollen, J., Mao, H., & Zeng, X. (2011). Twitter's mood predicts the stock market. Journal of Computational Science, 2(1), 1-8.

[5] Huang, E. H., & Lehmann, E. L. (1998). Nonparametric risk management with general risk measures. Finance and Stochastics, 2(4), 387-407.

[6] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

[7] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30, 5998-6008.

[8] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. NAACL-HLT.

[9] Schumaker, R. P., & Chen, H. (2009). Textual analysis of stock market prediction using breaking financial news. ACM Transactions on Information Systems, 27(2), 1-19.

[10] Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 25(2), 383-417.

[11] Mandelbrot, B. (1963). The variation of certain speculative prices. The Journal of Business, 36(4), 394-419.

[12] Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10‐Ks. The Journal of Finance, 66(1), 35-65.

[13] Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of Internet stock message boards. The Journal of Finance, 59(3), 1259-1294.

[14] Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. The Journal of Finance, 62(3), 1139-1168.

[15] Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? The American Economic Review, 71(3), 421-436.

[16] Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. The Journal of Economic Perspectives, 21(2), 129-151.

[17] Brown, G. W., & Cliff, M. T. (2004). Investor sentiment and the near-term stock market. Journal of Empirical Finance, 11(1), 1-27.

[18] Li, F. (2010). The information content of forward-looking statements in corporate filings—A naïve Bayesian machine learning approach. Journal of Accounting Research, 48(5), 1049-1102.

[19] Kamath, U. R., Liu, J., & Whitaker, J. (2019). Deep Learning for NLP and Speech Recognition. Springer.

[20] Cao, L., & Tay, F. E. H. (2001). Financial forecasting using support vector machines. Neural Computing & Applications, 10(2), 184-192..