International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Using Sentiment Analysis of Natural Language Processing to Analyse and Perform Risk Assessment in Finance

[1]Harpreet Singh Anand, [2]Ashritha R Murthy

[1]Student, B.E. (Hons.) Computer Science, Birla Institute of Technology and Science, Pilani Campus, Pilani

[2]Assistant Professor, Department of Computer Science and Engineering, JSS Science and Technology University, Mysuru, Karnataka, India

ABSTRACT

According to Forbes, unstructured data is growing at 55-65% each year and almost 90% of it has been generated in the recent two years. More data demands the need for more brains to process it. Today, companies use Artificial intelligence (AI) approaches to spend less time on data discovery and more time on deriving insights from the data. Combine sentiment analysis with the expertise of financial analysts to make more informed risk assessments. Human judgment can provide valuable context and validation. Finance and banking industry uses NLP for a variety of purposes like improved decision making, automation, data enrichment, etc. NLP in finance automates the manual processes of turning unstructured data into a more usable form. For example, information extraction on financial annual reports, Sentiment Analysis on financial news, ESG and asset management, Sentiment Analysis on tweets about companies, the capture of earning calls, and acquisition announcements. It adds context to the unstructured data and makes it more searchable and actionable. It even automates tedious/boring tasks reducing human interaction.

Keywords : Sentiment Analysis, NLP, AI, data, finance.

Introduction

Sentiment analysis in natural language processing (NLP) can be a valuable tool in the finance industry for risk assessment and decision-making. Here's how you can use sentiment analysis to analyze and perform risk assessment in finance:

Data Collection:

Gather relevant textual data from various sources such as news articles, social media, financial reports, and analyst opinions. These sources can provide insights into the market sentiment.

Preprocessing:

Clean and preprocess the text data by removing noise, stop words, and special characters. Tokenize the text into words or phrases.

Sentiment Analysis:

Use pre-trained sentiment analysis models or develop a custom model to classify the sentiment of the text as positive, negative, or neutral. There are various NLP libraries and tools available (e.g., NLTK, spaCy, TextBlob) that can help with sentiment analysis.

Feature Engineering:

Extract relevant features from the sentiment analysis results, such as sentiment scores or sentiment trends over time. These features can be used for risk assessment.

Market Impact Analysis:

Analyze how sentiment affects financial markets. For example, positive sentiment might lead to increased stock prices, while negative sentiment can lead to declines.

Risk Indicators:

Develop risk indicators based on sentiment analysis results. For instance, high levels of negative sentiment in news articles or social media posts related to a specific company or industry can be a risk indicator.

Risk Modeling:

Integrate sentiment-based risk indicators into financial models. These models can include risk assessment tools like Value at Risk (VaR), stress testing, or Monte Carlo simulations. By incorporating sentiment data, you can make these models more dynamic and responsive to changing market sentiment.

Real-time Monitoring:

Implement a real-time monitoring system that continuously analyzes sentiment data. This allows for timely risk assessment and quick decision-making in response to changing market sentiment.

Alerts and Triggers:

Set up alerts and triggers based on predefined thresholds for sentiment-based risk indicators. When these thresholds are crossed, it can trigger actions or alerts for risk management and mitigation.

Machine Learning and AI:

Utilize machine learning and artificial intelligence algorithms to predict future market movements and assess the impact of sentiment on various financial instruments.

Backtesting and Validation:

Continuously backtest the performance of sentiment-based risk assessment models to ensure their effectiveness. Adjust and fine-tune the models as needed based on historical data and performance.

Regulatory Compliance:

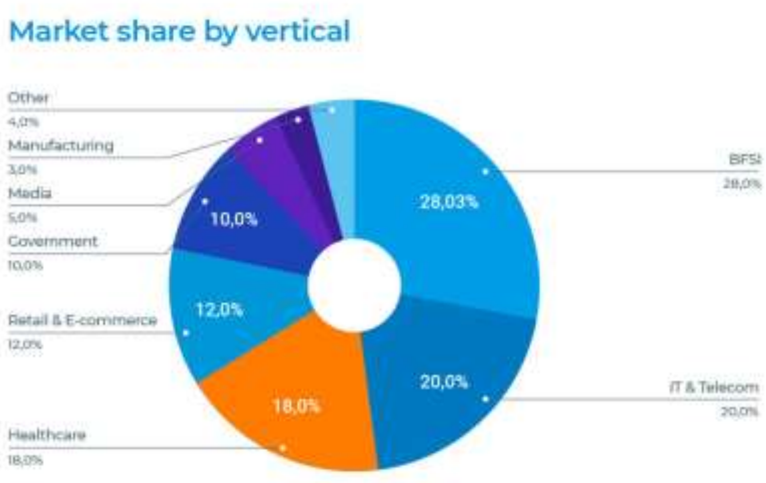
Ensure that your sentiment analysis and risk assessment processes comply with relevant financial regulations and reporting requirements.

Human Expertise:

One of the approaches is Natural Language Processing (NLP) which helps companies make sense of unstructured data. NLP is a subfield of AI that helps computers understand human language. We use NLP every day when our phone autocorrects the spelling or recommends the next word for our message. The typical applications of NLP in Finance are:

- Classifying Financial Documents
- Recognizing Financial Entities
- Understanding Entities in Context
- Extracting Financial Relationships
- Normalization and Data Augmentation
- Financial Deidentification
- Financial Document Splitting

By 2025, almost 30% of the applications of Natural Language Processing will be carried out inside Banking, Financial Services, and Insurance. Banking has historically been the natural promoter of the application of AI, more specifically, NLP for finance, in data automation.



NLP in Finance (BFSI sector) is already used for:

- Automatic loan, credit applications
- Automatic calculations of fees
- Customer onboarding
- Risk Management
- Asset Management, ESG
- Compliance
- Content Enrichment, etc.

The Role of Natural Language Processing in Financial Services

NLP Applications in Finance and Banking Sector

Below are the applications of Natural Language Processing in the finance industry.

Classification of Financial Documents

In today's fast and complex ecosystem, it is difficult to manage financial information. It is because privacy is important as the data is highly confidential and sensitive. We can use various NLP techniques to classify financial documents.

For instance, the finance industry uses text classification to predict various financial outcomes. it can automatically classify different types of agreements (loan, service, consulting agreements, etc).

We can also use **Sentiment Analysis** to analyze large volumes of textual data and understand various entities in it. Sentiment Analysis is an NLP technique that companies use for various things like analyzing reports and customer feedback, gauging market sentiment, etc.

Recognizing Financial Entities

NLP helps us identify and classify named entities in text, such as people, locations, dates, numbers, etc. to make recommendations or predictions. **Named Entity Recognition (NER)** is an NLP approach that finds and extracts entities from unstructured textual documents. For instance, it can recommend solutions based on news articles about a particular organization.

We can also use it to extract investment signals from news headlines. Banks and NBFCs (Non-Banking Financial Companies) use **NER** to extract **key** information from customer data.

Normalization & Data Augmentation

The text data is preprocessed to a suitable form before it is used in training NLP models. Normalization reduces variations in word forms and improves the model's performance. When we normalize text, we reduce its randomness and bring it closer to a predefined standard. We also reduce the variability of, for example, Company names, to disambiguate and be able to match/ link with other databases, such as SEC Edgar.

Data Augmentation in our libraries is the ability to use extracted information, such as Company Names, to query data sources and obtain more information, like Company's SIC code, Trading Symbol, Address, Financial Period, etc.

Financial De-identification

De-identification is a general term for any process of removing the association between a set of identifying data and the data subject. It consists of algorithms and processes that can be applied to documents, records, and data to remove any information, which can lead to the identification of the person the document is concerned with. It protects the privacy of the individuals when addressed by people who should not know the person's identity.

Financial Document Splitting

In NLP, splitting is the process of dividing text into smaller pieces, like sections, paragraphs, sentences, phrases. For instance, we can use NER to detect headers and subheaders in Financial Documents.

Text Summarization

This approach generates a concise summary of a text document. We can use it to extract insights and useful relationships between entities from financial reports and news articles.

In Finance, Text Summarization helps us extract headers from financial news and summarize financial news.

Limitations of NLP in Financial Services

Natural Language Processing In Finance, automates processes, reduces errors, provides customer support 24/7, and boosts revenue. But here are some **challenges** that the finance industry faces when using NLP.

Data Privacy

One of the biggest challenges faced by the banking and financial institutions is **data privacy**. Banks hold sensitive customer data that must be protected. When they use AI and NLP, they have to share data with third-party providers to train the machine learning algorithms. Here, the security concerns raise like:

1. Who will have access to the data?
2. How will the data be used?

Further, there are regulatory concerns when using NLP in the banking and finance sector. For instance, there can be fears about biased decision-making when a bank uses NLP to make lending decisions. Therefore, transparency needs to be ensured around how NLP systems make decisions.

Data Quality

NLP systems need voluminous amounts of data to work effectively. But the banks may not have sufficient data on certain products or customers. Further, machine learning models need clean and well-structured data as input and the data available to banks may not be of **high quality**. Here comes the need of data cleaning processes that are expensive and time-consuming.

No Justification for Rejections

NLP-based systems for financial services can significantly impact a person's life. For example, there can be a huge effect on a customer's future if the system does not approve his/her loan request.

If the system is not able to discern the **bias** and only analyzes information based on its design, how can financial institutions explain rejection to clients? Without proper justification, it is difficult for them to explain their decision.

High Cost of Investment

Minor and small-scale organizations can not afford advanced NLP-based systems as they are quite expensive. Apart from the software and additional hardware costs, regular updates need to be scheduled and implemented. Systems can be unavailable for an extended period of time if there's a problem with the update.

The Future of NLP in Finance

The top key benefits of utilizing NLP in Finance are:

- Accuracy
- Consistency
- Scaling
- Efficiency
- Process Automation

Natural Language Processing has transformed a number of industries like Healthcare, Education, Business, Data Science, Banking and Finance. Banks use NLP-powered chatbots to enhance communication with customers and better answer their queries. Financial institutions use NLP to manage risks, and automate routine tasks.

Conclusion

In the future, NLP will help the banks identify new revenue streams, make lending decisions, and provide personalized financial advice. We see as NLP evolves, it will have a profound impact on the financial industry.

Incorporating sentiment analysis into financial risk assessment can provide a more comprehensive view of market dynamics. However, it's important to remember that sentiment analysis has limitations, and it should be used in conjunction with other financial data and analysis methods to make well-informed decisions. But there are certain risks associated with using Natural Language Procassinf in the financial sector. NLP algorithms for decision-making are hard for humans to comprehend. Further, there is a risk to human employment as NLP can replace human workers in various roles. In the future, NLP-powered systems can have access to sensitive data such as health information and financial records that can be used to violate our privacy rights.

References

1. Ahir K, Govani K, Gajera R, Shah M (2020) Application on virtual reality for enhanced education learning, military training and sports. Augment Hum Res 5:7
2. Akaichi J, Dhouioui Z, López-Huertas Pérez MJ (2013) Text mining facebook status updates for sentiment classification. In: 2013 17th international conference on system theory, control and computing (ICSTCC), Sinaia, 2013, pp 640–645.
3. Al-Natour S, Turetken O (2020) A comparative assessment of sentiment analysis and star ratings for consumer reviews. Int J Inf Manage.
4. Audrino F, Sigrist F, Ballinari D (2018) The impact of sentiment and attention measures on stock market volatility. Available at SSRN: https://ssrn.com/abstract=3188941 or https://doi.org/10.2139/ssrn.3188941
5. Bach MP, Krsti Z, Seljan S, Turulja L (2019) Text mining for big data analysis in financial sector: a literature review. Sustainability 2019(11):1277
6. Bharti SK, Babu KS (2017) Automatic keyword extraction for text summarization: a survey. CoRR. abs/1704.03242.
7. Bidulya Y, Brunova E (2016) Sentiment analysis for bank service quality: a rule-based classifier. In: 2016 IEEE 10th international conference on application of information and communication technologies (AICT). https://doi.org/10.1109/icaict.2016.7991688
8. Brindha S, Prabha K, Sukumaran S (2016) A survey on classification techniques for text mining. In: 2016 3rd international conference on advanced computing and communication systems (ICACCS), Coimbatore, 2016, pp 1–5. https://doi.org/10.1109/ICACCS.2016.7586371
9. Cambria E (2016) Affective computing and sentiment analysis. IEEE Intell Syst 31(2):102–107. https://doi.org/10.1109/MIS.2016.31
10. Chen CC, Huang HH, Chen HH (2020) NLP in FinTech applications: past, present and future
11. Cook A, Herron B (2018) Harvesting unstructured data to reduce anti-money laundering (AML) compliance risk, pp 1–10
12. Dohaiha H, Prasad PWC, Maag A, Alsadoon A (2018) Deep learning for aspect-based sentiment analysis: a comparative review. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2018.10.003
13. Elagamy MN, Stanier C, Sharp B (2018) Stock market random forest-text mining system mining critical indicators of stock market movements. In: 2018 2nd international conference on natural language and speech processing (ICNLSP). https://doi.org/10.1109/icnlsp.2018.8374370
14. Gandhi M, Kamdar J, Shah M (2020) Preprocessing of non-symmetrical images for edge detection. Augment Hum Res 5:10. https://doi.org/10.1007/s41133-019-0030-5
15. Gupta A, Bhatia P, Dave K, Jain P (2019) Stock market prediction using data mining techniques. In: 2nd international conference on advances in science and technology, pp 1–5
16. Jani K, Chaudhuri M, Patel H, Shah M (2019) Machine learning in films: an approach towards automation in film censoring. J Data Inf Manag. https://doi.org/10.1007/s42488-019-00016-9
17. Jha K, Doshi A, Patel P, Shah M (2019) A comprehensive review on automation in agriculture using artificial intelligence. Artif Intell Agric 2:1–12
18. Kou G, Yang P, Peng Yi, Xiao F, Chen Y, Alsaadi F (2019) Evaluation of feature selection methods for text classification with small datasets using multiple criteria decision-making methods. Appl Soft Comput 86:105836. https://doi.org/10.1016/j.asoc.2019.105836
19. Kowsari K, Jafari Meimandi K, Heidarysafa M, Mendu S, Barnes L, Brown D (2019) Text classification algorithms: a survey. Information 10:150
20. Kumar BS, Ravi V (2016) A survey of the applications of text mining in financial domain. Knowl Based Syst 114:128–147
21. Kundalia K, Patel Y, Shah M (2020) Multi-label movie genre detection from a movie poster using knowledge transfer learning. Augment Hum Res 5:11. https://doi.org/10.1007/s41133-019-0029-y
22. Mudinas A, Zhang D, Levene M (2019) Market trend prediction using sentiment analysis: lessons learned and paths forward. arXiv:1903.05440
23. Panchiwala S, Shah MA (2020) Comprehensive study on critical security issues and challenges of the IoT world. J Data Inf Manag. https://doi.org/10.1007/s42488-020-00030-2
24. Pandya R, Nadiadwala S, Shah R, Shah M (2019) Buildout of methodology for meticulous diagnosis of K-complex in EEG for aiding the detection of Alzheimer's by artificial intelligence. Augment Hum Res. https://doi.org/10.1007/s41133-019-0021-6
25. Parekh V, Shah D, Shah M (2020) Fatigue detection using artificial intelligence framework. Augment Hum Res 5:5
26. Pathan M, Patel N, Yagnik H, Shah M (2020) Artificial cognition for applications in smart agriculture: a comprehensive review. Artif Intell Agric. https://doi.org/10.1016/j.aiia.2020.06.001

27. Renault T (2019) Sentiment analysis and machine learning in finance: a comparison of methods and models on one million messages. Digit Finance. https://doi.org/10.1007/s42521-019-00014-x

28. Shah G, Shah A, Shah M (2019) Panacea of challenges in real-world application of big data analytics in healthcare sector. Data Inf Manag. https://doi.org/10.1007/s42488-019-00010-1

29. Shah D, Dixit R, Shah A, Shah P, Shah M (2020) A Comprehensive analysis regarding several breakthroughs based on computer intelligence targeting various syndromes. Augment Hum Res 5:14. https://doi.org/10.1007/s41133-020-00033-z

30. Shah K, Patel H, Sanghvi D, Shah M (2020) A comparative analysis of logistic regression, random forest and KNN models for the text classification. Augment Hum Res 5:12. https://doi.org/10.1007/s41133-020-00032-0

31. Sohangir S, Wang D, Pomeranets A et al (2018) Big data: deep learning for financial sentiment analysis. J Big Data 5:3. https://doi.org/10.1186/s40537-017-0111-6

32. Xing FZ, Cambria E, Welsch RE (2018a) Natural language based financial forecasting: a survey. Artif Intell Rev 50:49–73. https://doi.org/10.1007/s10462-017-9588-9

33. Yang Li, Li Y, Wang J, Sherratt R (2020) Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. IEEE Access 8:1–1. https://doi.org/10.1109/ACCESS.2020.2969854

34. https://www.johnsnowlabs.com/examining-the-impact-of-nlp-in-financial-services/#:~:text=We%20can%20also%20use%20Sentiment,%2C%20gauging%20market%20sentiment%2C%20etc.

35. https://www.linkedin.com/pulse/using-natural-language-processing-understand-identify-huma-firdaus/

36. https://www.symanto.com/blog/how-to-use-nlp-for-more-accurate-business-analysis-and-risk-assessment/

37. https://www.refinitiv.com/perspectives/ai-digitalization/four-ways-to-apply-nlp-in-financial-services/