Decision Analytics Journal 15 (2025) 100584

Contents lists available at ScienceDirect

Decision Analytics Journal

journal homepage: www.elsevier.com/locate/dajour

A systematic review of sentiment analytics in banking headlines

Muhunthan Jayanthakumaran [a], Nagesh Shukla [b,*], Biswajeet Pradhan [a], Ghassan Beydoun [a]

[a] Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and IT, University of Technology Sydney, 2007, Australia
[b] Department of Management, Griffith Business School, Griffith University, Brisbane South Campus, 4111, Australia

ARTICLE INFO

Keywords:
Sentiment analysis
Predictive modelling
Text mining
Banking headlines
Market trends

ABSTRACT

This systematic review investigates sentiment analysis of news headlines in the banking sector, a field susceptible to public sentiment, as demonstrated by phenomena like bank runs leading to rapid deposit withdrawals. We trace the evolution of analytic methods from traditional machine learning to advanced deep learning models, notably Bidirectional Encoder Representations from Transformer (BERT) and Generative Pre-trained Transformer (GPT). Our study highlights their applications including headline generation, sentiment measurement, fake news detection, and analysis of political bias. Despite significant advancements, we uncover research gaps, such as the ineffective use of these methodologies in banking analysis, the underuse of GPT, and a focus on performance rather than practical application. Looking ahead, we note the increasing significance of Large Language Model (LLM), the untapped potential of headline analysis in banking, and the growing interest in this area spurred by rapid technological advancements. Our findings emphasise the pivotal role of sentiment analysis in deciphering market trends and improving decision making in finance, underscoring its strategic importance in the banking industry.

1. Introduction

Sentiment analysis, the process of gauging and interpreting sentiment, emotions, and opinions from text data, plays a pivotal role in numerous domains, offering valuable insights and aiding decision-making processes. For example, Zuniga-Morales et al. [1] created a machine learning framework to use tweets and news headlines to identify how Mexico is perceived based on a country image analysis using Bidirectional Encoder Representations from Transformer (BERT), noting that the model performed better without data enhancements. Within the field of banking and finance, its importance is paramount. Banking institutions and financial markets are intricately tied to public perception and sentiment. This is evidenced by multiple research papers utilising the news to inform pricing of investments. Evidence for this, Das et al. [2] looked at how Graph Neural Network (GNN) can be used to predict stock market trends using financial news sentiment data.

News headlines play a crucial role as indicators of prevailing collective sentiment. They mirror the perceptions of individuals and investors about financial institutions, economic environments, and emerging market trends. The ability to effectively capture and analyse the sentiment embedded in these headlines is instrumental in guiding crucial decisions, reducing potential risks, and offering preliminary insights into future market fluctuations. Current events can both indirectly and directly impact banking decisions. For example, a bank might be indirectly affected by governmental policies, such as alterations in tariffs or changes in the central bank's cash rate. Karzanov [3], for example, examines headlines as a potential lever for low-risk stock prediction. A bank could be directly impacted as in the example of Australia and the Royal Commission into Misconduct in the Banking, Superannuation and Financial Services Industry. This event led to headlines about the poor performance of particular banks in which findings were identified. Thus, sentiment analysis within the banking industry is not just an academic exercise, but a practical necessity, with far-reaching implications for macroeconomic performance, investment strategies, and regulatory compliance.

This systematic review aims to provide a structured assessment of the existing literature on sentiment analysis in the context of banking and finance. The financial sector is progressively relying on nontraditional data sources to make informed decisions, and news headlines play a pivotal role in this information ecosystem. An example of this interrelationship is shown by Agarwal [4] who examine financial sentiment analysis and how it might be used to detect trends of a company's stock price, noticing that domain-specific words improve performance based on Twitter and news headlines datasets. By systematically examining the methodologies, applications, and outcomes of sentiment

* Corresponding author.
E-mail addresses: muhunthan.jayanthakumaran@student.uts.edu.au (M. Jayanthakumaran), n.shukla@griffith.edu.au (N. Shukla), biswajeet.pradhan@uts.edu.au (B. Pradhan), ghassan.beydoun@uts.edu.au (G. Beydoun).

https://doi.org/10.1016/j.dajour.2025.100584
Received 3 June 2024; Received in revised form 15 April 2025; Accepted 3 May 2025
Available online 7 May 2025

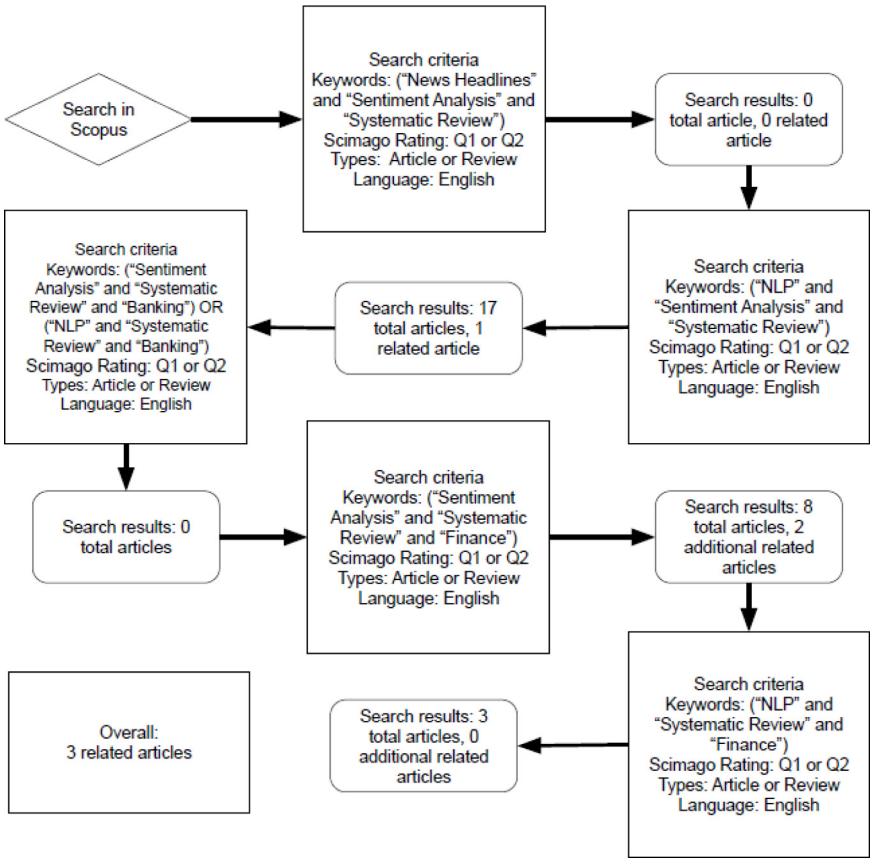

Fig. 1. Flowchart representing related works approach.

analysis in financial news, this review seeks to provide a comprehensive overview of the state of research in this domain.

The financial landscape is characterised by its complexity, volatility, and sensitivity to news and events. Understanding the role of sentiment in shaping financial markets has far-reaching implications, ranging from predicting stock market movements to assessing the impact of news sentiment on financial stability. However, rapid growth in this field has led to a diverse array of research approaches, making it essential to discern emerging trends and identify avenues for further investigation.

In subsequent sections, this article embarks on a systematic exploration of sentiment analysis within banking. By critically evaluating the methodologies used, the applications explored, and the outcomes observed, our goal is to provide a clear and nuanced understanding of the landscape. Our objective is not only to summarise existing research, but also to identify gaps in the literature, potential areas for refinement, and the broader implications of sentiment analysis in financial decision making.

2. Previous works

Previous systematic reviews in sentiment analysis have focused primarily on topics not related to news headlines and the banking sector. To ensure that this research is novel, the research process highlighted in Fig. 1 was applied. In particular, the process leads to three articles deemed relevant. To begin with a comprehensive Scopus search with the query [“News Headlines” AND “Sentiment Analysis” AND “Systematic Review”] yielded no results.

To broaden our scope, we conducted an additional search using the query [“NLP” AND “Sentiment Analysis” AND “Systematic Review”], resulting in the identification of seventeen articles and reviews per Fig. 1. These works encompassed diverse applications, including sentiment analysis in the context of application of Natural Language Processing (NLP) to health fields [5], virtual consumerism [6], and housing safety [7]. Among these, Sharma et al. [8] systematic review was most relevant in that it considered big data in financial institutions. Their study looks generally at the concept of NLP, big data and finance. They suggest that researchers tend to use social media for research purposes as it is more available than confidential financial data. Large Language Model (LLM) is touched as a possibility to improve thematic analysis. However, Sharma et al. [8] did not specifically consider headlines or banking strategy beyond data. Rather, it focused on papers that discuss the usage of NLP techniques within finance with regard to big data.

Furthermore, our searches using the terms [“Sentiment Analysis” AND “Systematic Review” AND “Banking”] and [“NLP” AND “Systematic Review” AND “Banking”] yielded zero articles. Expanding the search term to [“Sentiment Analysis” AND “Systematic Review” AND “Finance”] produced eight articles. Among these, Warin and Stojkov [9] examined the application of machine learning in finance, focussing on the technical analysis of financial market trends, indicating the potential value of such techniques for policymakers and the scientific community. However, this paper primarily emphasised metadata analysis rather than article content. The second relevant article, Ferreira et al. [10], explored the application of artificial intelligence to stock market trading and indirectly referenced financial sentiment analysis. The search term [“NLP” AND “Systematic Review” AND “Finance”] yielded three articles; two were within prior searches and the additional article was not directly related.

Table 1 summarises the three articles referenced in Fig. 1 in the context of why it was considered relevant, what the reviews themselves focused on, and why it does not cover the area of study proposed by this paper.

As seen in Table 1, these papers highlight the interest in the use of techniques such as NLP, machine learning, and sentiment analysis in

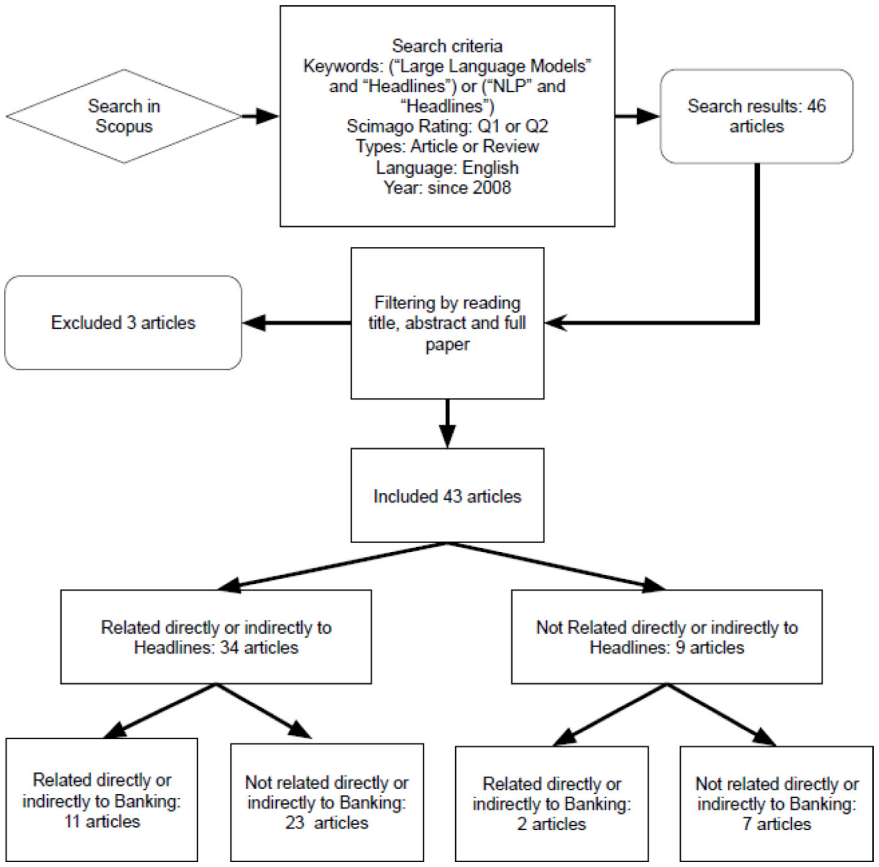

Fig. 2. Flowchart representing systematic review approach.

Table 1
Overview of Relevant Previous Works.

Paper	Relevance	Focus	Limitations relative to the area of study
Sharma et al. [8]	Financial institutions	NLP, Big Data, Finance	Not headline-specific; primarily focused on big data. Limited relevance to banking strategy beyond data analysis.
Warin and Stojkov [9]	Financial markets	Machine Learning, Technical Analysis	Emphasis on meta data analysis rather than article content; not directly related to sentiment analysis or headlines.
Ferreira et al. [10]	Stock market trading	Artificial Intelligence Sentiment Analysis	Indirectly references financial sentiment analysis; lacks a direct focus on banking headlines or practical applications within banking.

the finance field. However, the general exploration of these articles, combined with the limited search results, underscores the scarcity of research dedicated to applying sentiment analysis methodologies to the banking industry and leveraging headline data. Additionally, it becomes evident that recent advancements in this field, such as Bidirectional Encoder Representations from Transformer (BERT) and Generative Pre-trained Transformer (GPT), have not been comprehensively integrated into existing systematic reviews within the scope of papers considered. Hence, a detailed and focused review in this area is essential to provide insight into the methodologies currently employed and their relevance to the banking sector and to address the research gap made evident in Fig. 1 and Table 1.

This paper aims to bridge this research gap by assessing the applicability of existing sentiment analysis techniques to the banking industry. The significance of this inquiry lies in the role of confidence within the financial sector in macroeconomic performance. Headlines are often considered a reflection of the sentiments prevailing among financial sector participants, rendering them potential indicators of confidence. For example, the collapse of a bank would likely result in negative headlines, influencing perceptions of both the affected institution and the banking industry as a whole. In contrast, responsible rate management or initiatives promoting Environmental, Social, and Corporate Governance (ESG) may generate more positive sentiments.

In summary, this review intends to evaluate whether the sentiment analysis techniques investigated in previous research are relevant and effective within the context of the banking industry, using the unique characteristics of headline data.

3. Review methodology

To conduct a comprehensive systematic review of sentiment analysis in the banking industry, with a focus on the analysis of headlines, we implemented a structured methodology. This methodology enabled us to identify relevant research articles, categorise them by their subject matter and methodology, and draw meaningful insights from the literature. A summary of the approach can be seen in Fig. 2. A breakdown of the articles considered is given in Section 4.

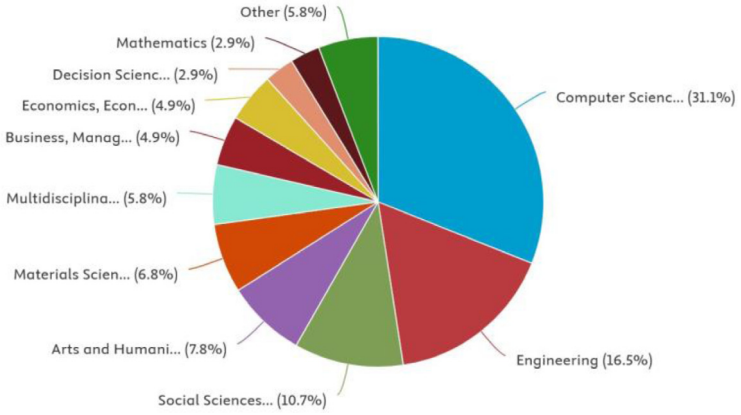

Fig. 3. Frequency of subject area by articles considered.

3.1. Search strategy and exclusion criteria

As seen in Fig. 2, our search strategy primarily involved accessing the Scopus database, a reliable source of academic publications. We designed a specific search query to capture articles related to sentiment analysis and banking, with an emphasis on headlines. Our search query considered keywords and phrases such as "sentiment analysis", "banking", "headlines", and their synonyms. To ensure a broad foundation, the final search criteria considered ["Large Language Models" and "Headlines"] or ["NLP" and "Headlines"].

In particular, it is not possible to limit the search to banking itself, given the limited research in the area. The applicability to banking is determined on the basis of the nature of the topic. Implicit in the search is the importance of capturing headlines and methods to explain the content of the headlines in a quantitative and automated fashion.

We limited the publication date to articles published from 2008 onwards to ensure the relevance of the research in contemporary contexts. Furthermore, we focus on articles published in English, as it is a widely accessible language for academic literature.

In the process of our search, we established the additional exclusion criteria of articles that are not related to text analysis, sentiment analysis, or banking.

3.2. Data collection and screening

Following the execution of our search query, we obtained a list of potentially relevant articles from the Scopus database. We then performed a preliminary screening of these articles to assess their relevance based on the criteria and research goals. Articles that met the criteria were retained for further analysis, while those that did not meet the criteria were excluded.

3.3. Categorisation of articles and analysis and insights

To provide a comprehensive overview of the literature, we categorised the included articles into several groups based on their topic matter and methodologies. For example, we identified articles related to headline analysis in banking, articles applying sentiment analysis to banking without direct focus on headlines, and articles that focused on NLP techniques in the broader context of finance or economics.

We systematically reviewed the selected articles and extracted relevant insights and findings. These insights were categorised into the following areas:

- Purpose and objectives of the studies.
- Methodologies and techniques used, including specific NLP models and algorithms.
- Outcomes and results, with a focus on the performance and applicability of the methods.
- Trends in the field of sentiment analysis in banking, including popular models and emerging approaches.

To facilitate a clear understanding of trends and patterns in the reviewed literature, we used visual representations, including charts and graphs. These visuals helped highlight the frequency of certain methodologies, the evolution of research over time, and the distribution of articles across subject areas.

Our systematic review methodology allowed us to provide a comprehensive overview of sentiment analysis in the banking industry, with a specific emphasis on headline analysis. It allowed us to identify gaps in the literature, emerging trends, and areas with untapped research potential.

3.4. Insights from high-level overview

In this section, we present high-level insights derived from our analysis of the reviewed articles. These insights cover various aspects, including the distribution of the topics covered, the growth in the number of articles and citations over the years, and the geographical origin of the reviewed literature. These findings provide a comprehensive view of sentiment analysis in the banking and finance sector, setting the stage for a more detailed exploration of methodologies, applications, and outcomes in subsequent sections.

- Fig. 3 presents a breakdown of the topic areas covered in the reviewed articles. In particular, most papers fall within Computer Science (31.1%) and Engineering (16.5%), suggesting a potential focus on methodological advancements rather than practical applications.
- Fig. 4 illustrates the number of articles and citations related to those examined. Notably, the articles considered are all post 2021 highlighting the novelty of this area. Dale [11] is the most cited paper in 2021 with a total count of 119. Of the sample considered, it is one of the earliest papers to talk to trust associated with Generative Pre-trained Transformer (GPT)-3.
- Fig. 5 displays the geographical distribution of the origin of the reviewed articles. Most of the papers originate from India, the United States, and China, which may reflect the population sizes of these countries. However, there is potential to foster global interest in this research area, especially given its universal relevance in the banking sector.
- Fig. 6 uses the author keywords from Scopus to segment the articles reviewed into clusters. In particular, there are four clusters. The yellow cluster directly relates to headlines and financial sentiment analysis linking to this article's purpose directly. In

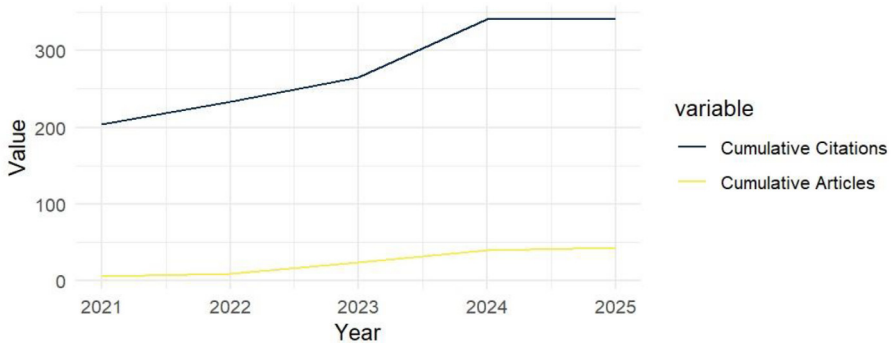

Fig. 4. Citation frequency of articles considered.

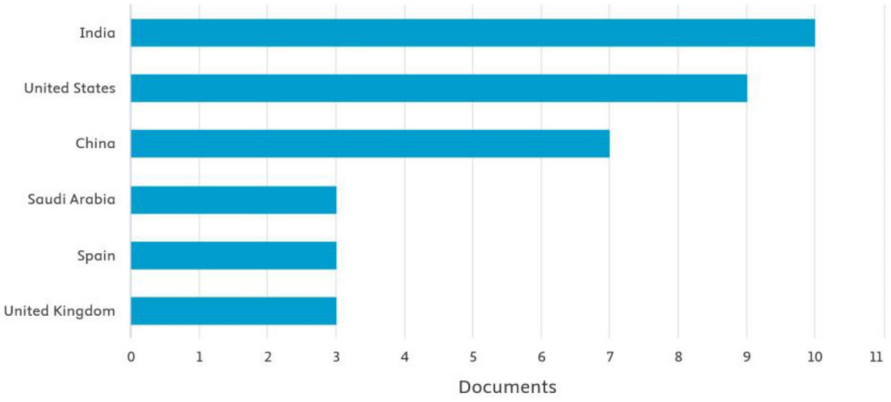

Fig. 5. Country of origin of articles considered for countries with at least 3 documents.

blue, it implies a use of machine learning and deep learning for the purpose of clickbait analysis. Purple indicates the link between Large Language Model (LLM) and AI. For Green, the techniques involving the techniques of Bidirectional Encoder Representations from Transformer (BERT), word2Vec and Glove to sarcasm detection. Generative AI and ChatGPT are directly linked as expected for red, but interestingly it connects to the concept of bias, suggesting either a measure for bias through this or an inherent bias based on calibration.

Our systematic review methodology enables us to offer a comprehensive overview of sentiment analysis in the banking industry, with a specific focus on headline analysis. It also facilitates the identification of research gaps, emerging trends, and areas ripe for further investigation.

4. Analysis of trends in reviewed articles

This section presents a comprehensive analysis of the trends observed in the literature pertaining to sentiment analysis in the banking sector, with a specific focus on news headlines. The objective of this review is to systematically examine the evolution of research methodologies, thematic concentrations, and the application of various techniques within this domain. By analysing trends in the accumulated literature, we aim to identify the prevailing research directions, highlight the methods and tools most frequently used, and uncover potential gaps in the current body of research. This analysis provides valuable information on the state of sentiment analysis in banking, particularly in how it has adapted and responded to the challenges and opportunities presented by advances in methodologies. In addition, we explore the extent to which these developments have been incorporated into practical applications within the banking industry. Through this evaluation, we seek to understand the implications of these trends for future research and the potential impact on banking practice.

We classify the articles reviewed in our systematic analysis into four distinct categories based on their relevance to headlines and the banking sector. A high-level reasoning is provided for why it is allocated to headlines or banking, or if it is not. This is covered in Fig. 7. The details of links to headlines and banking, such as what is meant to be headline generation, are discussed in greater detail in Section 4.2.

Per Fig. 7 each category is defined into four groups based on whether the link to the headlines or the link to banking is "Not related" or not:

1. **Headline but not banking related:** Articles in this category are related to headlines but not specifically focused on the banking domain. They may provide valuable information on the sentiment analysis methods used in short-form data. Szczepanski et al. [12] for example looked at how to measure fake news detection using Bidirectional Encoder Representations from Transformer (BERT)-based models. Although banking was not specifically referenced, the techniques used for fake news detection referenced are usable in a banking context. Similarly, Awais and Muhammad Adeel Nawab [13] considered how text summarisation could be used in the Urdu language. This form of summarisation involves looking at headlines and articles, which allows for applicability of the techniques to a wider context. Shao et al. [14] considered how energy vehicle sales could be forecasted using news headlines as an explanatory variable. This link allows for a better understanding of market demand through headlines, which is a concept of interest from a strategic perspective for business management. In particular, the three examples provided highlight valuable insights into headline data modelling, while not being specifically applied to the banking industry.

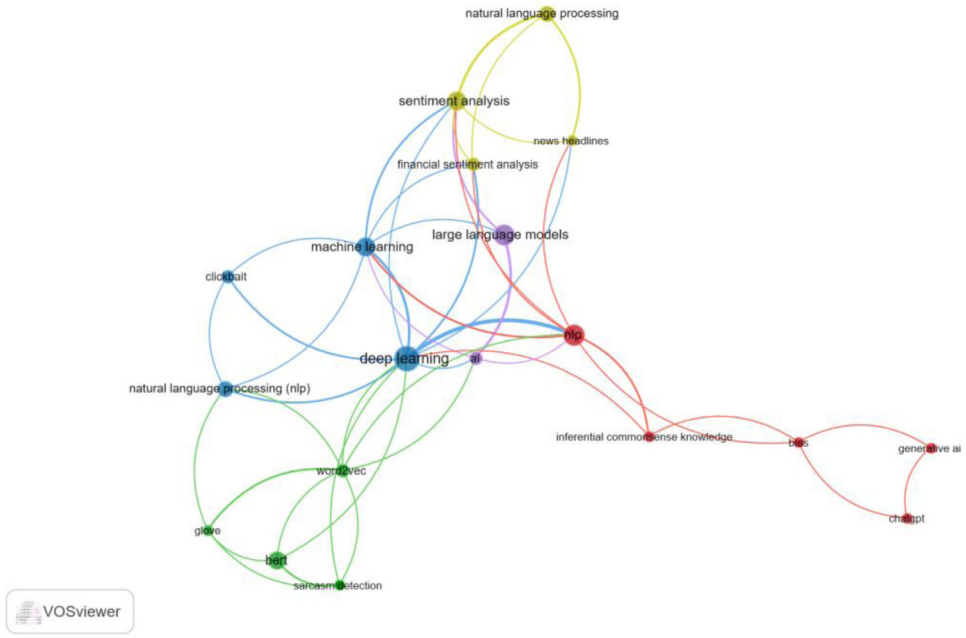

Fig. 6. Map of index keywords.

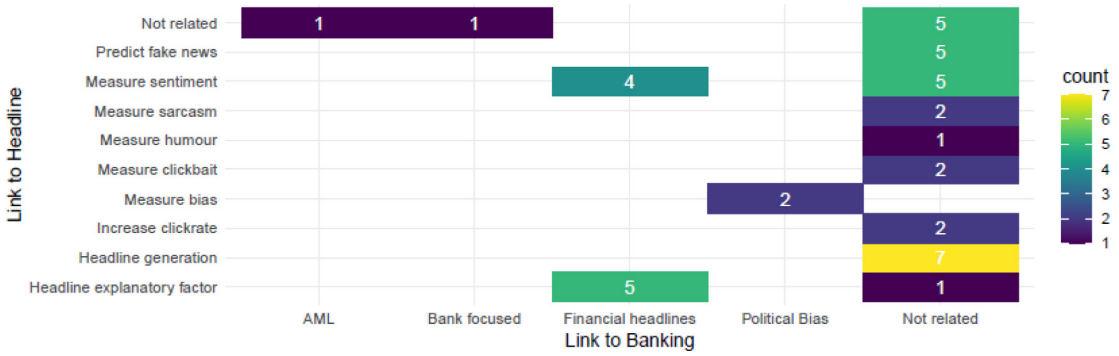

Fig. 7. Article relevance to headlines and banking split by categorisation.

2. **Headline and banking related:** Articles in this category directly relate to both headlines and the banking sector. They are highly pertinent to the objective of the systematic review. Anamika and Subramaniam [15] examined determining the investor sentiment for cryptocurrency-specific news headlines using NLP techniques and then applying this to determine the influence it has on cryptocurrency returns. Zhang et al. [16] considered how to improve financial sentiment analysis to better understand the sentimental orientation of financial texts (ie bullish or bearish), highlighting the value of mining financial industry stock-related opinions. Ma et al. [17] observe that the use of ChatGPT provides better performance to predict Chinese equity premiums compared to more traditional Bag-of-Words (BoW) methods. Overall, these are directly related to headlines and indirectly related to banking. The indirect nature is demonstrated through the relationship with financial markets, and it is required to apply an indirect lens given limited research on specifically banking behaviour.
3. **Not headline or banking related:** Articles that were not related to headlines, banking, or sentiment analysis. These did not satisfy the exclusion criteria and were included in the analysis due to the methodologies applied considered relevant. Scola and Segura-Bedmar [18] consider how BERT can be used to detect sarcasm for social media-type data. While social media data does not directly relate to headlines, for example, it can be considered a short-form text data and therefore techniques can be broadly applicable. Wahlster [19] is a systematic review that considers the development of Large Language Model literature, concluding that there are benefits in human trainers combining with interactive machine learning. In general, these topics do not have any relation to banking or headlines, but there are techniques and considerations that are applicable to text data analysis and could be seen as useful in approaching headline studies.
   Cobb [20] suggest how Large Language Model (LLM) might be beneficial in an alternative area of archaeological work, concluding on the potential value it could have in the process highlighting the development of the space. Similarly, Rathje et al. [21] showcased the value of BERT in multilingual psychological text analysis, detecting psychological constructs such as sentiment, discrete emotions, offensiveness and moral foundation. This broad applicability suggests that it could be used for banking purposes.
4. **Banking but not related to headlines:** Articles that focus on banking but do not address headlines or sentiment analysis were included due to the subject matter. There are two articles within this category. One is a critical review by Kute et al. [22] that examines the implementation of techniques such

as Machine Learning and Deep Learning in the realm of Anti-Money Laundering (AML). Although AML is indirectly linked to banking, since banks must address money laundering issues, the article centres on the application to AML detection rather than focussing on headline data. The second article, authored by Mathebula et al. [23], investigates sentiment classification using real-time reviews, drawing on feedback about South African banks as a data source. This article addresses banking directly by analysing review data instead of using headlines.

4.1. Purpose

The analysis of headlines in the literature reveals several predominant themes and applications. These themes provide valuable insight into the multifaceted nature of headline analysis within the field of sentiment analysis and natural language processing.

Fig. 7 provides a visual breakdown of articles related to headlines and the banking sector. The links to the headlines are discussed in more detail in the following.

1. **Headline Generation:** This theme includes articles focused on the conversion of full articles into concise headlines. These studies aim to effectively summarise the core content in a summarised form. Gorenz and Schwarz [24] tested the humour production abilities of ChatGPT compared to professional comedy writers in producing satirical headlines similar to *The Onion*, and based on a survey findings suggested the jokes were at least equally funny for ChatGPT regardless of comedic task and expertise of human comedy writer. Fatima et al. [25] utilised part-of-speech tagging to support headline text generation approach using Generative Pre-trained Transformer (GPT) suggested the generated headlines had a significant improvement in performance. Singh et al. [26] suggest an approach to summarisation and headline generation. Using deep learning they suggest this approach leads to effective headlines.
   Although the overarching theme is headline generation, this includes general summarisation techniques. For example, Benedetto et al. [27] uses a variety of language models, including LLM, to summarise Italian legal news documents. They concluded that the latest LLM type models outperform models like Bidirectional and Auto-Regressive Transformers (BART) in terms of informativeness. They do note, however, that nuances of legal reasoning were not captured, potentially driven by not having models specialised to the legal domain. In addition, they raise concerns about privacy concerns, high computational demands, and potential bias in output. This concept of bias in AI-generated output in relation to summarisation was also captured by Fang et al. [28]. They used seven representative LLM, such as ChatGPT and Large Language Model Meta AI (Llama), to generate headlines using news articles and conclude the potential for racial and gender bias in the output. A recurring theme in the literature review, this was also discussed by Breazu and Katsos [29] who questioned the perspective that ChatGPT is reproducing. They used the context of the UK referendum of 2016 and compared the generated headlines with actual headlines on eastern European Roma migrants and identified that ChatGPT was less biased and less sensationalised compared to the right-wing media. Most of these articles highlight methods in which headlines could be generated with higher quality.
2. **Explanatory Headlines:** Under this theme, headlines are used as explanatory factors in various models. These studies explore the use of headline sentiment to predict market movements or other outcomes. Chandola et al. [30] look at the prediction of the directional movement of stock prices using deep learning, leveraging headline data as an explanatory factor, and finding its inclusion beneficial. Liu and Huang [31] consider the prediction of crude oil price forecasting, noting that prior studies tend to focus on it as a time series or econometric variable prediction problem. Although some do consider raw news headlines or topic models, they do not consider event information in detail. They suggest a framework for extracting underlying related events and a text sentiment analysis algorithm to extract sentiment, leading to better performance compared to benchmark methods.
   An alternative perspective is suggested by Glasserman and Lin [32] who reference the look-ahead bias in stock price predictions where it is informed specifically by GPT sentiment analysis. There are look-ahead bias limits backtesting capabilities, as the LLM is trained on many years of data. In addition, there is distraction bias, since it has general knowledge on the companies that could influence the results. To overcome this, they anonymised the headlines, noting that surprisingly, this improved in-sample performance suggests that distraction bias might be larger than look-ahead bias. Overall, this showcases an application of headline data as input into other analysis and highlights that it has the potential to be a significant contributor, particularly to prices in existing literature.
3. **Sentiment Measurement:** This category comprises studies that focus on measuring the sentiment expressed in headlines. The goal is to gauge public perception, or the emotional tone conveyed. Zhou et al. [33] suggests a model based on BERT targeted at short text classification. They experiment with a number of datasets, including a headline news corpus, and determine that this proposed approach offers a better performance than standard BERT models by looking at accuracy and F1 value performance metrics. Jannani et al. [34] analyse perception and discourse by categorising headlines into themes such as politics, business, and education and obtaining the sentiment. They use techniques such as Latent Dirichlet Allocation (LDA) and BERT to derive an index representing well-being. Dangi et al. [35] proposes the value of using the artificial rabbits optimised robust random vector functional link network to improve sentiment analysis accuracy, suggesting a limitation of existing sentiment analysis work being a lack of consideration of randomisation-based neural networks.
   There are papers in this area focused on broader language applicability of sentiment measurement. Pan et al. [36] evaluated transformer models for financial-targeted sentiment analysis in Spanish. They compiled a corpus of financial tweets and news headlines in Spanish and identified which Spanish-specific LLM. Benitez et al. [37] considered if data augmentation techniques using machine learning and deep learning can improve classification for Spanish data, noting that Spanish large label data are scarce. They do cite concerns with augmentation techniques impacting semantic integrity, however, despite this, the augmentation techniques were suggested to have the capability to enhance sentiment and emotion analysis. Overall, this showcases application of sentiment analysis for predictions and the potential to help further refine analysis in this area.
4. **Fake News, Clickbait Detection and Click Rate:** This area of research deals with the detection of fake news through analysis of headlines, leveraging technology to discern the authenticity of news content. This category also considers clickbait headlines, which while might not necessarily be entirely false, are sensationalist. Supriya et al. [38] use the dissimilarity between headlines and articles to determine if clickbait. They measure this through a Sentence-BERT (SBERT) method suggesting this provides better performance than current state-of-the-art models. Farokhian et al. [39] uses two parallel BERT networks, one based on headline data and the other checking for important text in an article's newsbody to detect fake news, noting that this approach leads to a higher performance. Devarajan

et al. [40] proposes a framework for fake news detection using deep Natural Language Processing (NLP) model based on four layers: publisher layer, social media network layer, edge layer embedded and cloud layer. They suggest a better performance as a result of this. García-Ferrero and Altuna [41] propose a dataset called NoticIA which consists of Spanish news articles with clickbait headlines, each paired with a single-sentence summary written by humans. They suggest its value in task-specific models for clickbait summarisation.

As a different view, DeVerna et al. [42] suggests that ChatGPT can be used for fact checking to determine false headlines. However, using ChatGPT specifically, while accurate 90%, has the potential harm of decreasing belief in true headlines and increasing belief in false headlines. In a similar vein, another perspective was provided by Bajaj and Vishwakarma [43] which suggested that it is possible to bypass deep learning clickbait detection mechanisms by adversarial modifications of headline data. By highlighting this weakness, they suggest that improvements would be beneficial to overcome this weakness. Notably the majority of papers were focused on how to better distinguish fake news or clickbait using more complex methodologies.

An alternative angle are studies focused on how to increase click rate. This is something that is useful for a business trying to sell a product, for example, in this scenario, the headline could be the title of an email. Nguyen et al. [44] propose an approach to build an email open rate predictor for marketing purposes. The improvement in approach is based on using headline characteristics as predictors, similar to a computer science view, but also appending this with additional marketing-based variables. This was identified to provide better performance. Qiu and Golman [45] use NLP to identify headlines with a salient question, convey importance, and appear surprising, with negative valence, have higher click rates, but not necessarily long-term reader engagement.

5. **Measuring Specific Elements:** These studies investigate specific elements of headlines, such as political bias, humour, or sarcasm, in order to quantify and identify unique headline attributes. Shatnawi et al. [46] examine headlines, but in the context of detecting humour. They propose a variation of the BERT model which is an ensemble of different state-of-the-art pretrained models suggesting good performance as a result.

   Goel et al. [47] apply a number of different neural techniques, such as Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) and Convolutional Neural Network (CNN) in an ensemble model to detect sarcasm for a news headline and a Reddit dataset. They identified that using an ensemble model with word embedding outperformed other state-of-the-art models. Sharma et al. [48] consider sarcasm detection for headlines, Reddit and Twitter. They suggested the use of an ensemble approach that considers text embedding and also includes fuzzy evolutionary logic before making the final classification. This approach was suggested to be more accurate than earlier state-of-the-art models.

   Swati et al. [49] create a framework for predicting political bias in headlines, suggesting that existing methodologies do not recognise the potential of common-sense reasoning. They overcome this, they introduce a neural network framework which leverages inferential commonsense knowledge. They concluded that this leads to better performance. Swati et al. [50] also look at political bias in headlines, but in how to determine it for low-resource languages. The framework leverages commonsense knowledge as well and notes that the proposed framework is the best performing. These articles focused on how extending existing techniques could lead to a better fit of performance for purpose.

Despite the diverse and valuable contributions of these various approaches, their direct application to the field of banking has been notably limited. Although none of the reviewed literature directly addresses banking-related sentiments, indirect connections are observed in the examination of stock prices, energy markets, and financial headlines. In particular, researchers commonly employ financial headlines to gain insight into potential price movements within financial markets. Furthermore, some studies consider the degree of political bias in headlines related to the banking sector and explore topics like antimoney laundering, which is associated with financial services.

However, there remains a significant research gap in comprehensive understanding how headline analysis can be leveraged to gain insight into banking behaviour and decision-making processes. This gap presents an intriguing research opportunity for future research in the field.

4.2. Methodologies employed in literature

Fig. 8 provides a summary of the prevalence of different methodologies utilised in articles considered within the systematic review. In instances where GPT and BERT are mentioned, they have been distinctly categorised apart from LLM, due to their respective widespread adoption.

A summary of the methodologies represented in our review is discussed in this section (see Fig. 8). Firstly, the **Neural Networks & Deep Learning** category encompasses approaches based on artificial neural networks designed to identify complex patterns through deep hierarchical structures. These methods excel in learning from large volumes of data and capture subtle relationships that are not easily identified by traditional methods. Specific methodologies include *Neural Networks*, computational models inspired by biological neurons effective in pattern recognition and predictive analytics; *Deep Learning*, involving advanced multi-layered neural networks designed to learn intricate data representations used in domains such as image and speech recognition; *GRU, LSTM, and Recurrent Neural Network (RNN)*, types of recurrent neural networks specialising in sequential data analysis ideal for language processing and time-series forecasting; *CNN*, neural networks tailored specifically for spatial data like images and videos; *Artificial Rabbits Optimization*, a biologically inspired algorithm used to solve complex optimisation tasks efficiently; and *Ensemble* methods that combine predictions from multiple neural networks to enhance accuracy and reliability.

The **Large Language Models** group comprises extensive neural architectures trained on vast textual datasets to produce coherent, contextually relevant language, uniquely addressing tasks related to language generation and comprehension. Key methodologies include *GPT*, known for generating humanlike text useful in automated content creation, summarisation, and conversational AI; *BERT*, notable for its deep contextual understanding essential for search optimisation and text classification; *Llama*, Meta's large language model optimised for efficient training and fine-tuning; and general *LLM* architectures instrumental in complex comprehension and generative tasks.

Traditional Machine Learning methods involve conventional statistical and algorithmic approaches for predictive modelling and classification, valued for their interpretability and computational efficiency compared to deep learning models. Included in this category are *Machine Learning* algorithms that learn patterns to predict outcomes effectively, *Support Vector Machine (SVM)* noted for powerful classification capabilities, especially in high-dimensional spaces, and *Regression* methods useful in forecasting continuous outcomes and exploring relationships within datasets.

Natural Language Processing (NLP) techniques focus specifically on analysing, interpreting, and generating human language, facilitating automated textual analysis and sentiment identification. Specific methodologies include general *NLP* approaches foundational for

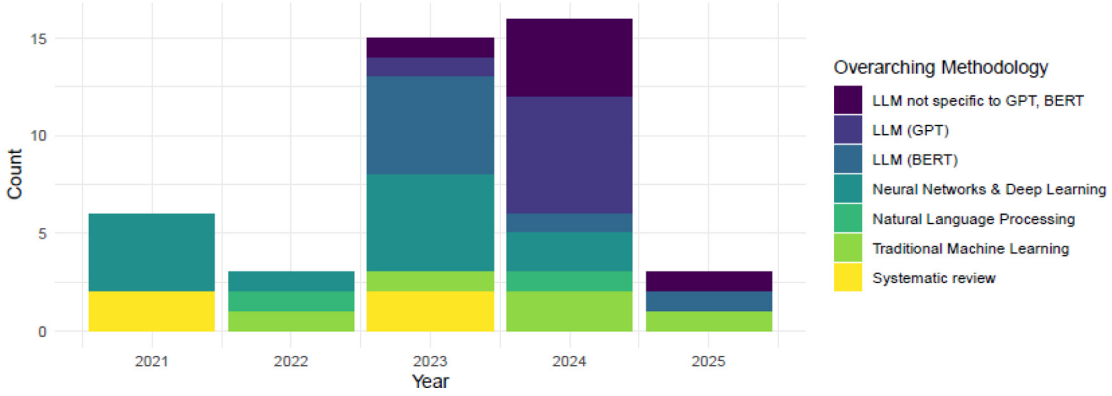

Fig. 8. Methodologies observed in articles reviewed by year.

language-based tasks; *Lexicon*based methods employing curated word lists for basic language processing; *Sentiment Analysis*, designed to systematically evaluate emotions and opinions prevalent in social networks and market research; and *Event Ontology*, which provides structured semantic frameworks to categorise textual descriptions of events.

Systematic Reviews represent a distinct category that does not focus on developing or applying machine learning models directly. Instead, these studies take a meta-analytical approach to synthesise findings in a body of literature, comparing methodologies, summarising results, and identifying overarching patterns and gaps in the field. Their inclusion provides a valuable foundation for understanding the methodological landscape and guiding future research directions.

Noting the above definitions, Fig. 8 highlights the consistency of more traditional techniques such as NLP, machine learning and neural networks, and deep learning in general. Interestingly, LLM have gained prominence from 2023, and dominated the research articles considered for 2024 and for the beginning of 2025. Most notably BERT made a large portion of these articles in 2023, and this was then succeeded by GPT in 2024. Fig. 9 compliments this analysis by presenting the headline categories discussed in Section 4.2 to the different methodologies represented in Fig. 8.

The network diagram in Fig. 9 shows the most popular technique for given purposes. For example, sentiment measurement broadly benefits from machine learning, neural networks, and BERT. BERT is more often used for sentiment measurement and measuring specific elements such as bias. Meanwhile, GPT is more popular in relation to creating explanatory factors using headline and headline generation. This suggests that the reduced frequency of BERT from 2023 to 2024, could be driven by a different outcome of research being prioritised for headlines in addition to GPT becoming more popular. Fake news, clickbait detection, and click rate appear to benefit from a broad range of techniques. Instead, headline generation is most benefited by the neural network class, with the exception of BERT.

Fig. 10 illustrates the popularity of methodologies in the reviewed articles, distinguishing between those within the banking sector and those from other fields. As shown in Fig. 10, there are more articles from outside the banking sector than within it, but the important observation lies in their relative proportions. Specifically, banking articles more frequently employ traditional methods such as NLP and machine learning. Where GPT has been applied in banking, it typically serves to generate explanatory variables from prompts based on financial headlines, while neural networks have been utilised to analyse financial domain-specific vocabulary and political biases. This suggests considerable potential for broader applications of advanced techniques such as GPT, and particularly BERT, within the banking context. In conclusion, the systematic review reveals a dynamic landscape of methodologies, with certain approaches gaining prominence in recent years, reflecting the evolving trends in the field.

4.3. Outcomes

The results of the articles reviewed reveal the various objectives pursued by the researchers, as depicted in Fig. 11. This links the outcome to how it relates to banking or headlines. There are three outcomes, better performance, analysis, and systematic review. Better performance relates to the researcher aiming for a higher goodness-of-fit or better out-of-sample performance. Analysis is where the researcher predominantly concludes on how the use of methods helped to better understand a particular problem. The systematic review refers to a synthesis of research with the intention of informing the public about a particular topic. Note that the allocation of the outcome was simplified, so while for example the outcome could technically be defined as applicability to a different language, if the goal is to improve the performance of the model, then the allocated outcome would be better performance.

Fig. 11 showcases the predominance of papers that focus on better performance, specifically to be able to measure elements of the text such as sentiment or bias, as well as the potential to create explanatory variables through headlines and headline generation. In contrast, the analysis focused on fake news, clickbait detection, and click rate and headline generation categories. Most systematic reviews focused on exploring broad techniques with potential applications. However, in the context of banking, only one such article focused specifically on anti-money laundering (AML). The research potential is in firstly more articles in the analysis and systematic review style in the area to take advantage of the better performance available to the underlying models such as GPT and BERT. Secondly, more research papers are required in the banking area; currently the focus is relatively limited, the most popular use case being financial headlines to improve the performance of prediction.

5. Discussion

In this section, we examine the essential aspects and findings derived from our comprehensive review of sentiment analysis. From the standpoint of identified research gaps, we observe that: (i) bankingspecific analyses remain limited in their engagement with recent advancements in textual or Natural Language Processing (NLP) analysis; (ii) there exists significant potential for methodologies incorporating recent advancements in Large Language Model (LLM) such as those in Generative Pre-trained Transformer (GPT) and Bidirectional Encoder Representations from Transformer (BERT); and (iii) existing research in text analysis has predominantly prioritised performance improvements (e.g. accuracy) over considerations of practical applicability and decision-making utility as shown in Section 4.3, which discussed research outcomes. These gaps are examined in this section and point to the need for more research in the banking domain that applies these tools in a practical, decision-focused context. This application is explored in more detail in the following section.

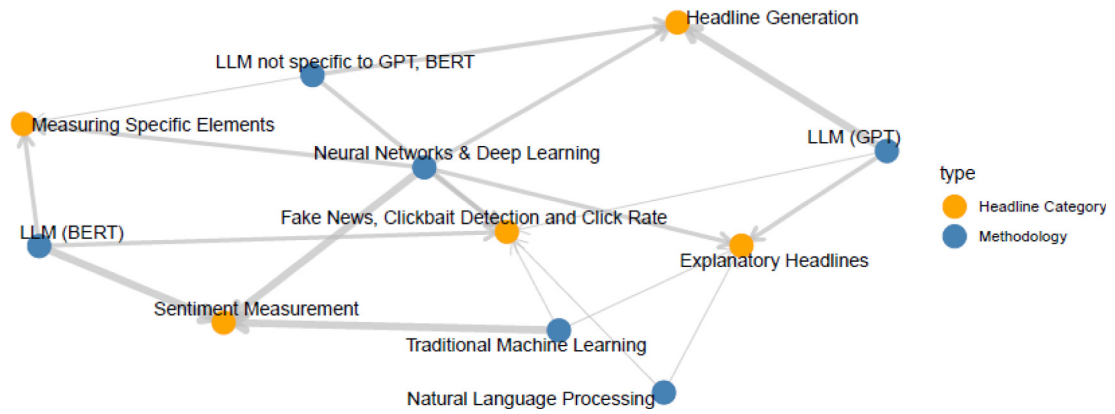

Fig. 9. Network diagram between headline category and methodology.

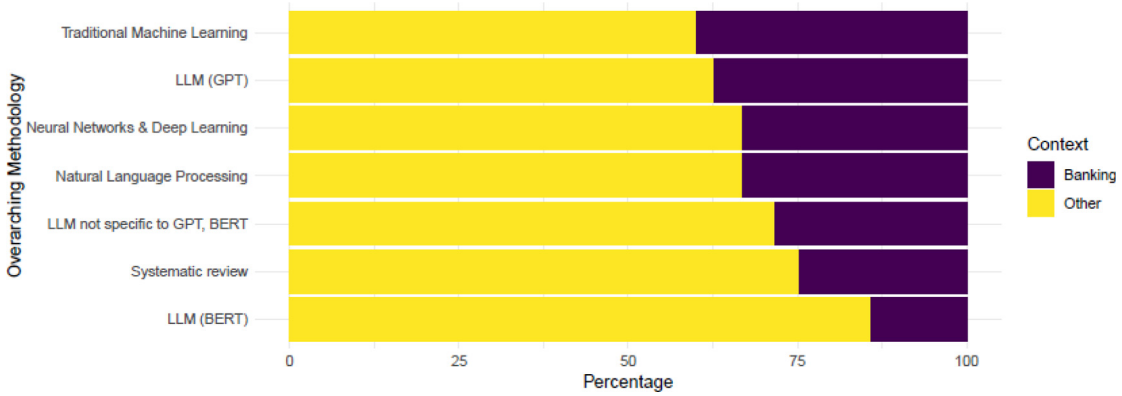

Fig. 10. Methodology popularity over time.

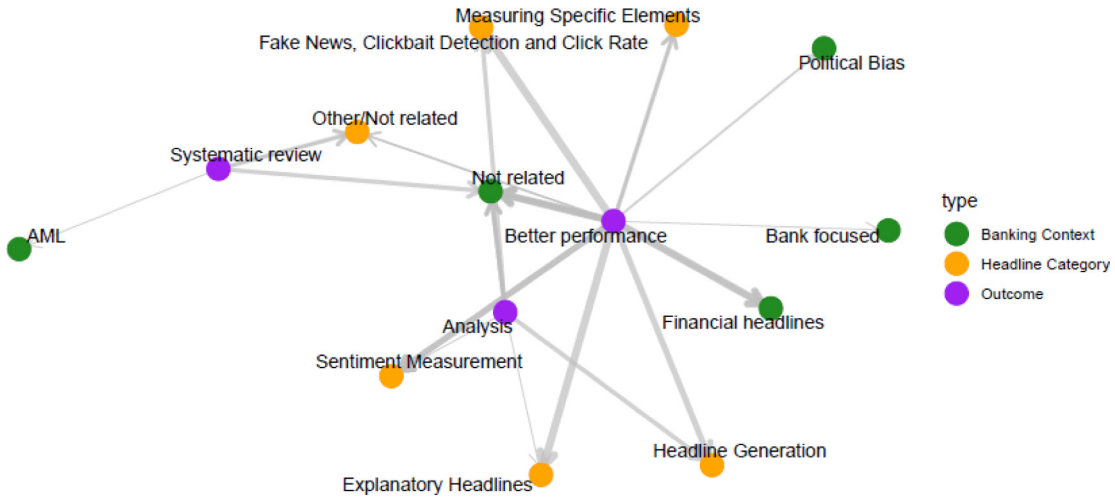

Fig. 11. Network diagram between banking relationship, headline relationship and research outcome.

5.1. Limited banking analysis

Our review revealed a notable lack of existing literature on the use of sentiment analysis within the banking industry, specifically related to headlines. Although sentiment analysis has been thoroughly investigated in multiple fields, including social networks and healthcare, its implementation in the banking sector remains limited. This shortcoming highlights the opportunity for additional research in this domain, as understanding public sentiment towards the banking industry is vital for informed decision making and ensuring financial stability.

This is evident in thirteen articles related either directly or indirectly to banking for the search terms considered. Of these articles, five focused on the use of textual data to create an explanatory variable to improve model performance, and four of these articles considered sentiment measurement. A wordcloud is presented in Fig. 12 to highlight the areas of interest based on the summary of purpose of the articles performed during the systematic review, it excludes headlines, analysis and reviews, among others, to ensure that the diagram is informative.

The stock prediction, politics, money laundering and cryptocurrency as key words showcases the relation to banking and financial services.

Fig. 12. Word cloud for themes relating banking related papers.

There is not as much focus on concepts such as strategy, marketing, or pricing which papers outside banking have considered. Furthermore, analysis has focused on the sentiment most prominently. With the advent of textual data analysis, it would be expected that more decision support software would be developed with the intention of leveraging this information. Despite the current limitations in academic research in this regard, there are notable areas of opportunity that can benefit from headline analysis.

5.2. GPT, new and emerging

Our review, particularly as reflected in the methodologies described in Fig. 8, demonstrates the increasing relevance of GPT in 2024 compared to 2023. This trend showcases that either GPT will continue to be more popular in 2025, or it could also suggest that it could be surpassed by an alternative LLM method such as Large Language Model Meta AI (Llama). This suggests that this is an evolving area and that it is important to note the available techniques.

Furthermore, some of these papers on GPT instead focused on some of the shortcomings. For example using GPT to derive explanatory variables from headlines has a risk of the GPT being based on the data after the event itself occurred. Therefore, providing it with a future view is not valuable for back-testing. Additionally, GPT might use information it knows in general to fill in the gaps, and not consider the information at face value, encouraging the researcher to ignore known organisations for example. Others talk about the potential copyright issues of GPT in how the model itself is trained.

Despite this, the shortcomings of the method will have increasingly better solutions, as is evident from the results of the researchers focused on better performance in Fig. 11. It is important that a researcher for this space uses these methodologies while recognising weaknesses. The refined GPT language modelling capabilities could be instrumental in deciphering often ambiguous and jargon-laden financial news, which is crucial for banks in assessing market sentiments and consumer confidence.

5.3. BERT is a complementary versatile approach

BERT and GPT are both transformer-based language models, but they differ fundamentally in architecture and usage. BERT is a bidirectional encoder model, trained to understand the context of both directions in a sentence simultaneously, making it well suited for classification, extraction, and sentencelevel understanding tasks. In contrast, GPT is a unidirectional decoder model, trained in a left-to-right manner to predict the next word in a sequence, enabling it to excel in text generation, completion, and creative language tasks.

Fig. 9 showcased this, as BERT has not been used as much to create explanatory headlines or to generate headlines for the sample considered as the main technique. A future research area of interest might be an ensemble that leverages the advantages of BERT and GPT, for example a system that uses BERT for the classification of review or headline data and GPT to provide decision support through proposed actions to improve sentiment.

The prevalence and versatility of BERT, especially in the analysis of small text formats such as headlines, is evident from the findings in Section 4.2. In the banking context, this means leveraging BERT's robust contextual understanding of BERT to analyse headlines that often carry significant weight in influencing market movements and shaping public opinion. However, as shown in Fig. 10, the application in banking is limited compared to other techniques.

The ability of BERT to discern nuanced sentiments in financial news can be a crucial tool for banks to predict market trends and understand public perception towards financial products and policies. This is particularly relevant in today's fast-paced information age, where headlines can dramatically sway investor behaviour and customer trust.

5.4. Research less focused on application

The results of the reviewed articles indicate a predominant focus on showcasing the superiority of proposed methods over previous approaches through metrics such as accuracy and F1 value performance metrics. This is highlighted in Fig. 11. The purpose of the researchers has been to demonstrate that their chosen methods perform better in varied textual tasks, thereby contributing to the development of more accurate models. However, there is room for further research that goes beyond performance improvement and explores the broader application of sentiment analysis, such as in the banking industry.

6. Future perspective

This section builds on the identified research gaps by examining how recent advancements in Large Language Model (LLM) and Natural Language Processing (NLP) can be applied to banking-specific sentiment analysis. Our findings indicate that headlines, as concise encapsulations of critical information, hold untapped potential for banking analysis. Without limiting the view to banking, the general area of text analysis is growing as captured in Fig. 4 and discussed in Section 3.4. The succinct nature of headlines makes them ideal for rapid sentiment analysis, providing immediate insight into market trends and public opinion. Using sentiment analysis for headlines can aid banks in real-time decision making, risk assessment, and trend forecasting. Future research should explore how banks can integrate headline analysis into their data analytics frameworks to enhance responsiveness and strategic planning.

Through a synthesis of current literature, this chapter highlights the evolving methodological landscape of sentiment analysis and its relevance to the specific needs of the banking sector. It emphasises the importance of adapting recent innovations, particularly large language models, to suit domain-specific applications. The analysis provides a foundation for future academic inquiry, especially in the addressing of persistent challenges such as sentiment detection. It also identifies a notable gap in the use of headline data within financial services and highlights the limited focus on explainable artificial intelligence (Explainable Artificial Intelligence (XAI)) in existing research, suggesting a clear area for further exploration. From a managerial perspective, the findings offer practical information by illustrating how headline data have been used in related fields and how similar approaches could support more informed data-driven decision making in dynamic financial environments. To achieve this, this section is divided into (i) the underutilisation of LLM within banking; (ii) the opportunities presented by headline-based sentiment analysis; and (iii) the growing relevance of these tools for real-time decision-making and strategic forecasting.

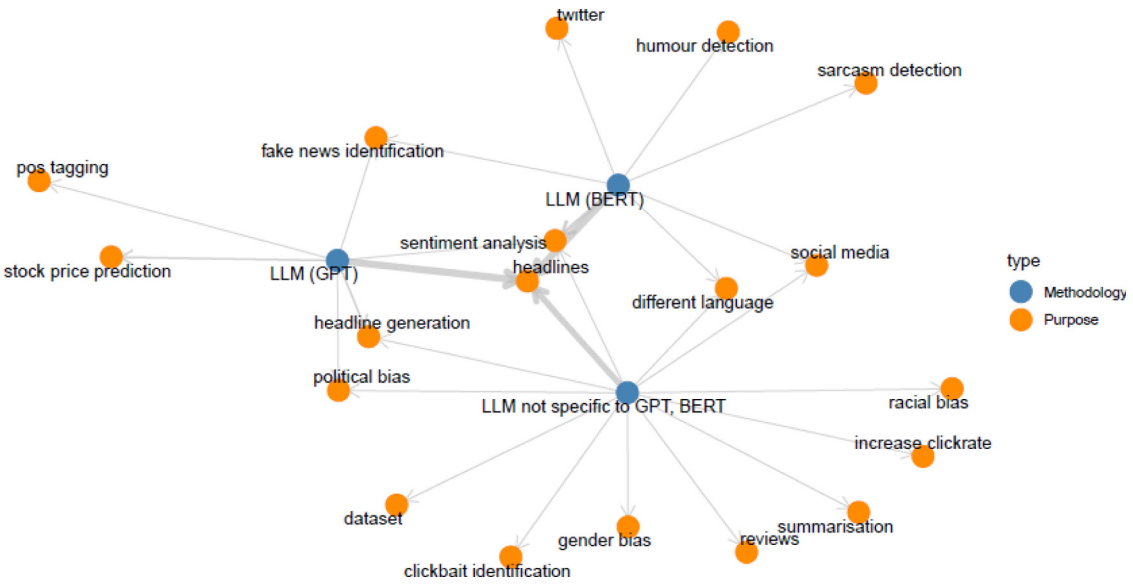

Fig. 13. Network diagram between LLM methodology and paper purpose.

6.1. LLM has large potential

The systematic review, particularly in Sections 3 and 4.2, underscores the potential of LLM in sentiment analysis, despite their current underutilisation in banking. This is further established in Section 5.3 which highlights the popularity of Bidirectional Encoder Representations from Transformer (BERT) in recent years. The advanced capabilities of LLM, especially in capturing and generating nuanced language, open new frontiers for the analysis of the banking sector. Future research could focus on harnessing models like Generative Pre-trained Transformer (GPT) for more in-depth understanding of market sentiments, customer opinions, and economic trends. In particular, GPT is currently underused in the literature as discussed in Section 5.2, suggesting further potential even outside of banking.

The research establishes different LLM used for different purposes in the context of headlines. To help visualise which LLM has been used for different areas, Fig. 13 is provided.

Within Fig. 13, it is possible to note that sentiment analysis, humour detection, and sarcasm detection leverage BERT, while fake news identification could utilise either BERT or GPT. In the case of bias, broader methods are used, as is evident from the links to LLM not specific to GPT or BERT. This breakdown can help guide to know what has been considered so it might be used for business purposes, or alternatively, what has not been considered so that it might define a potential research opportunity.

The adaptation of LLM to banking-specific needs could revolutionise the way financial institutions interact with and understand their market environment. This potential, coupled with the increasing availability of sophisticated NLP tools, sets the stage for significant advances in this field over the next decade.

6.2. Headlines complement banking analysis

The headlines offer a unique advantage for sentiment analysis in the banking industry. As shown in Section 5.2, there are several papers that have used headline data for financial headlines, where these papers tended to focus on cryptocurrency or stock markets predominantly. However, where headlines were used for explanatory purpose per Section 4.2, six articles directly leveraged this, with five considered directly related to banking.

This is a research opportunity, as headlines provide a concise summary of news and events, making them a valuable source of information for financial institutions. Banks can leverage headline sentiment analysis to gain insight into public perception, predict market trends, and inform strategic decisionmaking. This advantage highlights the potential for further research and practical applications in utilising headlines for banking analysis. To provide a parallel scenario, consider Khalil et al. [51] who look at optimising the hybrid process of the AI-assisted Letter of Credit exam to improve trade finance. They are able to leverage artificial intelligence to, among other purposes, reduce operational risks. If headline data could be used similarly, this could benefit other areas and could be practically integrated into business as usual in banking decision making.

Despite this, one of the challenges in sentiment analysis is one, especially when applied to headlines, is the limited amount of text available for analysis as headlines are inherently concise. Therefore, extracting meaningful sentiment from such a brevity can be challenging. Researchers need to develop innovative techniques to overcome this limitation and ensure the precision of sentiment analysis in short text formats. Noting this, with the advancement of techniques such as BERT discussed in Section 5.3 and the potential driven by GPT as discussed in Section 5.2, the tools for doing so are becoming more advanced, and therefore this limitation is becoming less prevalent.

6.3. Dynamic research area

The dynamic and ever-changing nature of NLP and sentiment analysis in banking, highlighted throughout this review, underscores the need for continuous adaptation and innovation in this field. Rapid advancements after 2018 have revolutionised sentiment analysis, making it more accessible and relevant for banking applications. As the field continues to evolve, it will be crucial for both academia and industry to stay abreast of the latest developments, ensuring that methodologies and applications remain cutting-edge and relevant to the changing landscape of banking and finance. Fig. 10 showcases this change as some methods fade out of use or see limited use while new methods such as BERT and GPT gain popularity.

Given the context of financial services, it is expected that in the future as LLM are adopted for decision making, there will also be a desire for greater XAI which will allow a greater understanding of the rationale behind the output behind LLM. A paper that highlights the trends in XAI is Saranya and Subhashini [52], which highlights examples of how this could be used in social media that have similarities to headline data. Although the systematic review did not explicitly

identify articles focused on this topic, the concept of XAI is, from our perspective, closely aligned with the rising studies on LLM and the ability to trust its results.

7. Conclusion

In summary, this systematic review has thoroughly examined the application of sentiment analysis in the context of news headlines, particularly its relevance in the banking and finance sector. Our analysis reveals the evolution of sentiment analysis methodologies over time, ranging from traditional machine learning approaches to advanced deep learning models such as Bidirectional Encoder Representations from Transformer (BERT) and Generative Pre-trained Transformer (GPT).

Furthermore, we have highlighted the diverse applications of sentiment analysis within banking, including headline generation, sentiment measurement, fake news detection, political bias analysis, and classification tasks. These applications underscore the adaptability of sentiment analysis techniques for informing decision making in the financial industry.

However, despite progress in sentiment analysis methodologies and their deployment in related fields, their direct integration into banking and finance remains limited. This research gap presents an exciting opportunity for future investigations to explore the practical implementation of sentiment analysis within financial institutions.

In conclusion, this systematic review contributes to our understanding of sentiment analysis within the banking and finance domain. It emphasises the importance of further research to bridge the gap between sentiment analysis and banking practices. As the financial landscape continues to evolve, sentiment analysis remains a valuable tool to gain insight into market trends and risk assessment.

Declaration of competing interest

No conflict of interest is declared.

Data availability

Data will be made available on request.

References

[1] Luis N. Zuñiga-Morales, Jorge Angel González-Ordiano, J. Emilio Quiroz-Ibarra, César Vil-Ĺanueva Rivas, Machine learning framework for country image analysis, J. Comput. Soc. Sci. 7 (1) (2024) 523–547, http://dx.doi.org/10.1007/s42001-023-00246-3.

[2] Nabanita Das, Bikash Sadhukhan, Rajdeep Chatterjee, Satyajit Chakrabarti, Integrating sentiment analysis with graph neural networks for enhanced stock prediction: A comprehensive survey, Decis. Anal. J. 10 (2024) http://dx.doi.org/10.1016/j.dajour.2024.100417.

[3] Daniil Karzanov, Headline-driven classification and local interpretation for market outperformance and low-risk stock prediction, Comput. Econ. 64 (2) (2024) 769–788, http://dx.doi.org/10.1007/s10614-023-10449-5.

[4] Basant Agarwal, Financial sentiment analysis model utilizing knowledge-base and domain-specific representation, Multimedia Tools Appl. 82 (6) (2023) 8899–8920, http://dx.doi.org/10.1007/s11042-022-12181-y.

[5] Kerstin Denecke, Daniel Reichenpfader, Sentiment analysis of clinical narratives: A scoping review, J. Biomed. Informatics 140 (2023) http://dx.doi.org/10.1016/j.jbi.2023.104336.

[6] Sezai Tunca, Violetta Wilk, Bulent Sezen, Defining virtual consumerism through content and sentiment analyses, Cyberpsychol., Behav. Soc. Netw. 26 (3) (2023) 198–213, http://dx.doi.org/10.1089/cyber.2022.0079.

[7] Na Li, Rita Yi Man Li, Qi Yao, Lingxi Song, Jirawan Deeprasert, Housing safety and health academic and public opinion mining from 1945 to 2021: Prisma, cluster analysis, and natural language processing approaches, Front. Public Heal. 10 (2022) http://dx.doi.org/10.3389/fpubh.2022.902576.

[8] Ratnesh Kumar Sharma, Gnana Bharathy, Faezeh Karimi, Anil V. Mishra, Mukesh Prasad, Thematic analysis of big data in financial institutions using nlp techniques with a cloud computing perspective: A systematic literature review, Inf. (Switzerland) 14 (10) (2023) http://dx.doi.org/10.3390/info14100577.

[9] Thierry Warin, Aleksandar Stojkov, Machine learning in finance: A metadata-based systematic review of the literature, J. Risk Financ. Manag. 14 (7) (2021) http://dx.doi.org/10.3390/jrfm14070302.

[10] Fernando G.D.C. Ferreira, Amir H. Gandomi, Rodrigo T.N. Cardoso, Artificial intelligence applied to stock market trading: A review, IEEE Access 9 (2021) 30898–30917, http://dx.doi.org/10.1109/ACCESS.2021.3058133.

[11] Robert Dale, Gpt-3: What's it good for? Nat. Lang. Eng. 27 (1) (2021) 113–118, http://dx.doi.org/10.1017/S1351324920000601.

[12] Mateusz Szczepański, Marek Pawlicki, Rafa l Kozik, Michal Choraś, New explainability method for bert-based model in fake news detection, Sci. Rep. 11 (1) (2021) http://dx.doi.org/10.1038/s41598-021-03100-6.

[13] Muhammad Awais, Rao Muhammad Adeel Nawab, Abstractive text summarization for the Urdu language: Data and methods, IEEE Access 12 (2024) 61198–61210, http://dx.doi.org/10.1109/ACCESS.2024.3378300.

[14] Jin Shao, Jingke Hong, Meiping Wang, Xianzhu Wang, New energy vehicles sales forecasting using machine learning: The role of media sentiment, Comput. Ind. Eng. 201 (2025) http://dx.doi.org/10.1016/j.cie.2025.110928.

[15] Anamika Anamika, Sowmya Subramaniam, Do news headlines matter in the cryptocurrency market? Appl. Econ. 54 (54) (2022) 6322–6338, http://dx.doi.org/10.1080/00036846.2022.2061904.

[16] Han Zhang, Zongxi Li, Haoran Xie, Raymond Y.K. Lau, Gary Cheng, Qing Li, Dian Zhang, Leveraging statistical information in fine-grained financial sentiment analysis, World Wide Web 25 (2) (2022) 513–531, http://dx.doi.org/10.1007/s11280-021-00993-1.

[17] Feng Ma, Zhichong Lyu, Haibo Li, Can chatgpt predict Chinese equity premiums? Financ. Res. Lett. 65 (2024) http://dx.doi.org/10.1016/j.frl.2024.105631.

[18] Elsa Scola, Isabel Segura-Bedmar, Sarcasm detection with bert; [detección de sarcasmo con bert], Proces. Del Leng. Nat. 67 (2021) 13–25, http://dx.doi.org/10.26342/2021-67-1.

[19] Wolfgang Wahlster, Understanding computational dialogue understanding, Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. 381 (2251) (2023) http://dx.doi.org/10.1098/rsta.2022.0049.

[20] Peter J. Cobb, Large language models and generative ai, oh my! Adv. Archaeol. Pr. 11 (3) (2023) 363–369, http://dx.doi.org/10.1017/aap.2023.20.

[21] Steve Rathje, Dan-Mircea Mirea, Ilia Sucholutsky, Raja Marjieh, Claire E. Robertson, Jay J. Van Bavel, Gpt is an effective tool for multilingual psychological text analysis, Proc. Natl. Acad. Sci. USA 121 (34) (2024) http://dx.doi.org/10.1073/pnas.2308950121.

[22] Dattatray Vishnu Kute, Biswajeet Pradhan, Nagesh Shukla, Abdullah Alamri, Deep learning and explainable artificial intelligence techniques applied for detecting money laundering-a critical review, IEEE Access 9 (2021) 82300–82317, http://dx.doi.org/10.1109/ACCESS.2021.3086230.

[23] Miehleketo Mathebula, Abiodun Modupe, Vukosi Marivate, Fine-tuning retrieval-augmented generation with an auto-regressive language model for sentiment analysis in financial reviews, Appl. Sci. (Switzerland) 14 (23) (2024) http://dx.doi.org/10.3390/app142310782.

[24] Drew Gorenz, Norbert Schwarz, How funny is chatgpt? A comparison of human- and a.i.produced jokes, PLoS One 19 (7 July) (2024) http://dx.doi.org/10.1371/journal.pone.0305364.

[25] Noureen Fatima, Sher Muhammad Daudpota, Zenun Kastrati, Ali Shariq Imran, Saif Hassan, Nouh Sabri Elmitwally, Improving news headline text generation quality through frequent pos-tag patterns analysis, Eng. Appl. Artif. Intell. 125 (2023) http://dx.doi.org/10.1016/j.engappai.2023.106718.

[26] Rajeev Kumar Singh, Sonia Khetarpaul, Rohan Gorantla, Sai Giridhar Allada, Sheg: summarization and headline generation of news articles using deep learning, Neural Comput. Appl. 33 (8) (2021) 3251–3265, http://dx.doi.org/10.1007/s00521-020-05188-9.

[27] Irene Benedetto, Luca Cagliero, Michele Ferro, Francesco Tarasconi, Claudia Bernini, Giuseppe Giacalone, Leveraging large language models for abstractive summarization of Italian legal news, Artif. Intell. Law (2025) http://dx.doi.org/10.1007/s10506-025-09431-3.

[28] Xiao Fang, Shangkun Che, Minjia Mao, Hongzhe Zhang, Ming Zhao, Xiaohang Zhao, Bias of ai-generated content: an examination of news produced by large language models, Sci. Rep. 14 (1) (2024) http://dx.doi.org/10.1038/s41598-024-55686-2.

[29] Petre Breazu, Napoleon Katsos, Chatgpt-4 as a journalist: Whose perspectives is it reproducing? Discourse Soc. 35 (6) (2024) 687–707, http://dx.doi.org/10.1177/09579265241251479.

[30] Deeksha Chandola, Akshit Mehta, Shikha Singh, Vinay Anand Tikkiwal, Himanshu. Agrawal, Forecasting directional movement of stock prices using deep learning, Ann. Data Sci. 10 (5) (2023) 1361–1378, http://dx.doi.org/10.1007/s40745-022-00432-6.

[31] Jiangwei Liu, Xiaohong Huang, Forecasting crude oil price using event extraction, IEEE Access 9 (2021) 149067–149076, http://dx.doi.org/10.1109/ACCESS.2021.3124802.

[32] Paul Glasserman, Caden Lin, Assessing look-ahead bias in stock return predictions generated by gpt sentiment analysis, J. Financ. Data Sci. 6 (1) (2024) 25–42, http://dx.doi.org/10.3905/jfds.2023.1.143.

[33] Lu Zhou, Peng Wang, Huijun Zhang, Shengbo Wu, Tao Zhang, A multiscale interactive attention short text classification model based on bert, IEEE Access 12 (2024) 160992–161001, http://dx.doi.org/10.1109/ACCESS.2024.3478781.

[34] Ayoub Jannani, Soukaina Bouhsissin, Nawal Sael, Faouzia Benabbou, Topic modeling and sentiment analysis of arabic news headlines for a societal well-being scoring and monitoring system: Moroccan use case, IEEE Access 13 (2025) 26345–26364, http://dx.doi.org/10.1109/ACCESS.2025.3538888.

[35] Dharmendra Dangi, Sonal Telang Chandel, Dheeraj Kumar Dixit, Suvarna Sharma, Amit Bhagat, An efficient model for sentiment analysis using artificial rabbits optimized vector functional link network, Expert Syst. Appl. 225 (2023) http://dx.doi.org/10.1016/j.eswa.2023.119849.

[36] Ronghao Pan, José Antonio García-Díaz, Francisco Garcia-Sanchez, Rafael Valencia-García, Evaluation of transformer models for financial targeted sentiment analysis in Spanish, PeerJ Comput. Sci. 9 (2023) http://dx.doi.org/10.7717/peerj-cs.1377.

[37] Rodrigo Gutiérrez Benítez, Alejandra Segura Navarrete, Christian Vidal-Castro, Claudia. Martínez-Araneda, Guide for the application of the data augmentation approach on sets of texts in Spanish for sentiment and emotion analysis, PLoS One 19 (9) (2024) http://dx.doi.org/10.1371/journal.pone.0310707.

[38] Jyoti Prakash Singh Supriya, Gunjan Kumar, Identification of clickbait news articles using sbert and correlation matrix, Soc. Netw. Anal. Min. 13 (1) (2023) http://dx.doi.org/10.1007/s13278-023-01162-0.

[39] Mahmood Farokhian, Vahid Rafe, Hadi Veisi, Fake news detection using dual bert deep neural networks, Multimedia Tools Appl. 83 (15) (2024) 43831–43848, http://dx.doi.org/10.1007/s11042-023-17115-w.

[40] Ganesh Gopal Devarajan, Senthil Murugan Nagarajan, Sardar Irfanullah Amanullah, S.A. Sahaaya Arul Mary, Ali Kashif Bashir, Ai-assisted deep nlp-based approach for prediction of fake news from social media users, IEEE Trans. Comput. Soc. Syst. 11 (4) (2024) 4975–4985, http://dx.doi.org/10.1109/TCSS.2023.3259480.

[41] Iker García-Ferrero, Begoñ Altuna, Noticia: A clickbait article summarization dataset in spanish; [noticia: Un dataset para el resumen de artículos clickbait en espanõl], Proces. Del Leng. Nat. (73) (2024) 191–207, http://dx.doi.org/10.26342/2024-73-14.

[42] Matthew R. DeVerna, Harry Yaojun Yan, Kai-Cheng Yang, Filippo Menczer, Fact-checking information from large language models can decrease headline discernment, Proc. Natl. Acad. Sci. USA 121 (50) (2024) http://dx.doi.org/10.1073/pnas.2322823121.

[43] Ashish Bajaj, Dinesh Kumar Vishwakarma, Non-alpha-num: a novel architecture for generating adversarial examples for bypassing nlp-based clickbait detection mechanisms, Int. J. Inf. Secur. 23 (4) (2024) 2711–2737, http://dx.doi.org/10.1007/s10207-024-00861-9.

[44] Nguyen Nguyen, Joseph Johnson, Michael Tsiros, Unlimited testing: Let's test your emails with ai, Mark. Sci. 43 (2) (2024) 419–439, http://dx.doi.org/10.1287/mksc.2021.0126.

[45] Jingyi Qiu, Russell Golman, Curiosity in news consumption, Appl. Cogn. Psychol. 38 (2) (2024) http://dx.doi.org/10.1002/acp.4195.

[46] Farah Shatnawi, Malak Abdullah, Mahmoud Hammad, Mahmoud Al-Ayyoub, Comprehensive study of pre-trained language models: detecting humor in news headlines, Soft Comput. 27 (5) (2023) 2575–2599, http://dx.doi.org/10.1007/s00500-022-07573-z.

[47] Priya Goel, Rachna Jain, Anand Nayyar, Shruti Singhal, Muskan Srivastava, Sarcasm detection using deep learning and ensemble learning, Multimedia Tools Appl. 81 (30) (2022) 43229–43252, http://dx.doi.org/10.1007/s11042-022-12930-z.

[48] Dilip Kumar Sharma, Bhuvanesh Singh, Saurabh Agarwal, Nikhil Pachauri, Amel Ali Alhussan, Hanaa A. Abdallah, Sarcasm detection over social media platforms using hybrid ensemble model with fuzzy logic, Electron. (Switzerland) 12 (4) (2023) http://dx.doi.org/10.3390/electronics12040937.

[49] Swati Swati, Dunja Mladenic, Marko Grobelnik, An inferential commonsense-driven framework for predicting political bias in news headlines, IEEE Access 11 (2023) 80886–80902, http://dx.doi.org/10.1109/ACCESS.2023.3298877.

[50] Swati Swati, Adrian Mladenić Grobelnik, Dunja Mladenić, Marko Grobelnik, A commonsenseinfused language-agnostic learning framework for enhancing prediction of political bias in multilingual news headlines, Knowl.-Based Syst. 277 (2023) http://dx.doi.org/10.1016/j.knosys.2023.110838.

[51] Mounaf Asaad Khalil, Majed Hadid, Regina Padmanabhan, Adel Elomri, Laoucine Kerbache, An integrated artificial intelligence and optimization model for operational efficiency and risk reduction in letter of credit examination process, Decis. Anal. J. 14 (2025) http://dx.doi.org/10.1016/j.dajour.2025.100552.

[52] A. Saranya, R. Subhashini, A systematic review of explainable artificial intelligence models and applications: Recent developments and future trends, Decis. Anal. J. 7 (2023) http://dx.doi.org/10.1016/j.dajour.2023.100230.