

Multimodal Deep Learning for Enhancing Intangible Asset Valuation through Integration of Unstructured Data Sources and Market Sentiment Analysis: A Bibliometric Analysis

Hannan Afzal¹, Muhammad Qasim Ali^{2,*}, Kafait Ullah³, Muhammad Zubair Qasim⁴

¹Department of Business Administration, University of Sialkot, Sialkot, Punjab, Pakistan.

²Department of Education, University of Sialkot, Sialkot, Punjab, Pakistan.

³Faculty of Management and Administrative Sciences, University of Sialkot, Sialkot, Punjab, Pakistan.

⁴Department of Computer Science, Govt. High School 43/WB, Vehari, Punjab, Pakistan.

hannanafzal6@gmail.com¹, qasimvr@yahoo.com², kafiatullah@gmail.com³, zubairqasim825@gmail.com⁴

Abstract: The existing study examines the ongoing disconnect between conventional financial modeling and the governance of unstructured data, preventing the accurate valuation of intangible assets, a key component of the information economy. To improve valuation accuracy, the study uses multimodal deep learning methods and a scientometric evaluation to visualize the intellectual space and the discipline's development. A graphical representation of the domain architecture was generated using VOSviewer, which visualizes a large body of information gathered from major academic databases. The results indicate that the research is dominated by an Anglo-American research alliance that includes powerful players such as Google, Microsoft, Stanford, MIT, and the United States, which is leading in research volume, influence, and collaboration. At the same time, a new research center, spearheaded by China, is emerging in Asia. Analyses of citations point to the work of Erik Cambria in sentiment analysis and to that of Yoshua Bengio in deep learning. The keyword mapping helps identify key themes around big data, benchmark datasets, empirical performance metrics (accuracy and predictive power), and the significance of global scientific competition and collaboration facilitated by conferences. Generally, the research finds that although concentrated authorship is a faster method of innovation, it reduces diversity, and it is important to note that there is still a need to develop stronger AI systems capable of quantifying intangible assets.

Keywords: Intangible Asset Valuation; Unstructured Data; Sentiment Analysis; Deep Learning (DL); Financial Modeling; Scientometric Evaluation; Global Scientific Competition.

Received on: 19/09/2025, **Revised on:** 14/11/2025, **Accepted on:** 07/01/2026, **Published on:** 19/08/2026

Journal Homepage: <https://www.fmdbpublish.com/user/journals/details/FTSFDS>

DOI: <https://doi.org/10.69888/FTSFDS.2026.000740>

Cite as: H. Afzal, M. Q. Ali, K. Ullah, and M. Z. Qasim, "Multimodal Deep Learning for Enhancing Intangible Asset Valuation through Integration of Unstructured Data Sources and Market Sentiment Analysis: A Bibliometric Analysis," *FMDB Transactions on Sustainable Finance and Data Science*, vol. 1, no. 3, pp. 131–144, 2026.

Copyright © 2026 H. Afzal *et al.*, licensed to Fernando Martins De Bulhão (FMDB) Publishing Company. This is an open access article distributed under [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits copying, remixing, redistributing, transformation, and building upon the material in any medium so long as the original work is properly cited.

1. Introduction

Modern financial analysis and business strategy require precise valuation of intangible assets, including human capital, intellectual property, goodwill, and brand equity, which comprise a company's market value in a knowledge economy [24]. The mainstay of conventional approaches to value, such as the income, market, and cost methods, is organized, quantitative financial

*Corresponding author.

data [18]. Because they are subjective, changeable, and dependent on external factors, traditional models often fail to value intangibles adequately [27]. There is a substantial and persistent knowledge gap in the field of theoretically sophisticated financial models, as they do not account for the complex, unstructured data used to estimate an asset's value. It is difficult to integrate huge datasets, including news articles, social media sentiment, patent filings, CEO communication tones, and multimedia material, into credible valuation models, despite their potential to fill knowledge gaps [25]; [13]. High volume, diversity, velocity, and validity are the biggest challenges of the new data paradigm. The conventional analytical instruments work with numerical spreadsheets and databases. The data structure or design of unstructured data is not known. Early attempts to utilize this data often relied on manual content analysis or straightforward keyword counts, which are labor-intensive, not scalable, and lack an understanding of subtlety, irony, and contextual shifts [17].

This has led financial valuation experts to fail to reach a common methodology for estimating how factors such as a viral Twitter brand mention or an especially innovative technical feature in a patent document affect the ultimate value of a company's intangible assets. This weakness must be addressed by an efficient, automated analytical approach [20]. Due to these challenges, deep learning (DL) has revolutionized the extraction of patterns and hierarchical representations from complex, high-dimensional data [22]. This research summarizes and evaluates research in the rapidly emerging domains of intangible asset valuation, multimodal deep learning, and unstructured data analytics. The construction of MDL frameworks that connect non-traditional datasets with financial indicators to provide more accurate and complete valuation models. This data may include brand images, earnings call audio, social media feeds, and textual news. This review provides a basic introduction to the actual application of these state-of-the-art methodologies; readers may find it useful for improving their risk foresight, strategy formulation, and investment analysis skills. This line of research has the potential to improve market efficiency, reduce information asymmetry, and provide a more accurate representation of a company's intrinsic value relative to its market price by enabling the measurement of previously unquantifiable factors [26]. This review paper first provides a brief overview of the literature review methodology, then explains how multimodal learning and intangible asset appraisal work. It will then give a comprehensive overview of existing applications, problems, and future directions in the field.

2. Literature Review

2.1. Intangible Asset Valuation

In the knowledge-based economy, intangible asset assessment has moved from accounting to strategy [23]. Over 80% of the market value of S&P 500 enterprises comes from intangible assets, including intellectual property, brand equity, software, data, and human capital. Businesses in the modern day cannot be adequately studied using traditional financial reports that center on physical assets [8]. Intangible assets are challenging for standard-setters, investors, and appraisers to value because of their uncertain utility and the lack of clarity about when they will yield returns [24]. Modern researchers are working to improve methods for measuring these important assets, which significantly impact a company's value and competitive advantage. The next section covers basic valuation, current challenges, and new developments, including the use of unstructured data. The asset's worth is based on its initial investment and current output or replacement costs [28].

However, it works well for intangible assets that have just been acquired or for endeavors that do not yet have a market or a means to generate income, such as research and development (R&D) that is still in the development process. Comparable intangible asset market transactions are utilized in the market strategy, which contrasts with the market approach. The notion of substitution in economics implies that an informed investor would not pay more for an asset than the cost of a comparable item with the same level of utility [29]. An analysis of similar license agreements, sales, mergers, and acquisitions, along with adjustments for profitability, growth potential, and risk, is conducted to inform the market strategy. The utility of this method is limited because there is no active, transparent market for the majority of unique intangible assets. Additionally, it is challenging to obtain credible and comprehensive data on private transactions. On the other hand, its application is often limited to patents and trademarks with publicly documented licensing operations.

For valuing intangible assets that generate income, the income approach is the most frequently used and theoretically valid method. An object is evaluated according to the anticipated net economic advantages that it will provide throughout the course of its useful life [9]. In this area, the most important considerations are royalty relief and multi-period excess income. The concept of relief from royalty involves calculating the hypothetical royalty payments an owner can save by holding and then discounting them to present value [33]. The MEEM is a more sophisticated method because it distinguishes the net cash flows from the subject intangible asset by subtracting the costs of contributing assets from the costs of all other firm assets [28]. Although traditional income-based strategies are prevalent, measuring their effectiveness is difficult [24]. The validity of the method depends on the quality and objectivity of the inputs, which are often market data that are organized and collected internally [34].

2.1.1. Challenges in Valuing Intangible Assets

The lack of openness in market data makes it difficult to utilize for assessing benchmark pricing and value in this business. This encourages putting too much stock in theoretical frameworks and assumptions that have not been tested. This never-ending cycle causes poor choices. The absence of a liquid market makes it difficult to ascertain fair value and liquidity discounts. Because of problems with cash flow management. Connecting an intangible asset to its future cash flows is another issue. The process of value development in contemporary firms typically involves the collaboration of numerous tangible and intangible assets [4]. The process of separating the contribution of a software platform, a patent, or a trademark from a firm's overall performance is difficult and often leads to arbitrary decisions. Nevertheless, the assessment of adequate returns for other assets is subjective, which can lead to over- or under-valuation of the subject intangible. Contributory asset charges in techniques such as the MEEM try to solve this issue. When it comes to organizational capital and gathered labor, this attribution problem is far more severe.

Complexity may be found in both the regulatory and accounting frameworks. The recognition of the fair value of intangible assets acquired through a business combination is required by both IFRS and GAAP. However, intangible assets developed internally are normally expensed as incurred, according to International Accounting Standards Board guidance in 2020. When there is a significant disparity between a firm's book value and market value, often due to poor management, the financial statements become less meaningful to investors. Goodwill and indefinite-lived intangible impairment testing are both costly and challenging to perform annually. Many of its discounted cash flow models are sensitive to even the slightest assumptions [5]. Last but not least, rapid technological growth and economic obsolescence make value uncertain. In IT companies, intangible assets tend to have shorter lifespans. This is especially true for disruptive technologies and shifting customer preferences. It is risky to predict how the flow of financial resources will be in the future.

2.1.2. Development of Valuation Techniques

Computational technology and data analytics have helped firms in this sector develop solutions to conventional problems. The assessment models utilize additional unstructured data. Intangible value issues are shed some light on with newsletters, social media, patent applications, and commercial disclosures [26]. Qualitative input cannot be measured at scale using traditional quantitative methods. This data flood can be addressed with natural language processing and sentiment analysis. Sentiment analysis of reviews and social media discussions can provide market-driven metrics to assess a business's viability and consumer loyalty. These appraisals can affect royalty relief rates and income growth rates. The method offered a data-based analysis that was more detailed and broader in its scope than financial measures. Machine learning and deep learning models can transform value research by leveraging more accurate and sophisticated inputs derived from these enhanced data assessments. When applied to large datasets, they enabled researchers to identify complex, non-linear patterns that had not been detected by regression analysis. The intangible asset valuation could be estimated using a multimodal approach based on three sources: structured numerical data in the financial statements, unstructured language in the management discussion and analysis section of the annual report, and real-time Twitter sentiment about the organization's performance.

2.2. Multimodal Deep Learning

One of the challenges in the modern financial world is valuing intangible assets. Traditional accounting struggles to estimate their true value, and conventional valuation methods become less effective as the volume of unstructured data grows. Unstructured data includes public brand perception, IP strength, and management caliber. In the field of data science and artificial intelligence, multimodal deep learning (MDL) addresses this issue. Machine learning (ML) builds an adaptable analytical framework for new data [10]. This framework handles both structured and unstructured data, such as text, images, and audio. The method considers quantitative financial indicators. It also assesses intangible assets affected by qualitative elements, including executive statements, patent filings, news stories, and social media.

2.2.1. Fundamentals of Multimodal Deep Learning

In the first stage of data fusion, the raw or low-level data from various modalities is fed into a deep learning model. This model is vulnerable to noise and misalignment, but it may capture interactions at a fine grain. Although late fusion provides resilience, it may miss cross-modal linkages, as it processes each modality independently with separate models and then combines their high-level predictions or judgments at the output layer. Most often, intermediate fusion combines representations from intermediate layers of a deep learning architecture to simulate cross-modal interactions more flexibly and robustly [1]. In their 2019 study, Baltrušaitis et al. [4] used loss functions to align the embedding spaces of different modalities. This resulted in a unified representation that grouped concepts with similar semantic properties, making it easier to understand how they interact. By combining quantitative time-series data processed by LSTMs (Long Short-Term Memory networks) with semantic features

extracted from news headlines, more complex deep-learning integrations enable models to go beyond their initial use of bag-of-words representations for text [10].

Compared with unimodal benchmarks, these hybrid models consistently outperformed them. Risk management and the forecasting of corporate credit ratings are two areas where this technology has more sophisticated uses. A small number of accounting ratios are widely used in traditional credit scoring models; researchers state that multimodal approaches significantly expand the scope of this method by incorporating unstructured data, such as company earnings call transcripts, management discussion and analysis sections of annual reports, and even the body language and facial expressions of executives when they appear on stage. The ability to assign numerical values to what has been considered "soft" data is one of the primary ways MDL has enhanced financial analysis. Brand equity and market sentiment analysis may be the most pertinent use case for intangible asset appraisal. Numerous unstructured sources, such as social media, product reviews, and viral videos, shape the public's perception of a brand [31]. It is best to use multimodal models to investigate this setting. A dynamic, multi-dimensional brand health indicator is created from user-shared product images, reviews, and mention volume.

2.2.2. Advantages over Traditional Machine Learning Approaches

Multimodal deep learning offers several advantages over traditional machine learning approaches when valuing complex intangible assets. The majority of classical approaches rely on manually created structured features for their operations. This category includes combinations of SVMs, linear regression, and ensemble techniques such as random forests [16]. That reliance on feature engineering is severely constrained. When it comes to capturing more complex and advanced information, machine learning (MDL) models automate feature extraction through learning hierarchical representations from raw data. Another important perk is the potential for data complementarity. In practice, it is common for a single modality to provide an erroneous impression. When assessing the potential economic value of an invention, it might be more accurate to consider not only the textual description found in a patent document, but also pertinent scientific pictures or diagrams (image modality) and the funding history of the institution that created them (tabular data) [21]. Attempts to include such a diverse set of traits would overwhelm conventional models. To mimic the interaction among multiple modalities, MDL designs are tailored using its fusion procedures. The model can then use information from other modalities to fill in the gaps left by missing data from the first modalities. A more comprehensive and trustworthy assessment is the outcome for assets that lack an immediate market value.

Furthermore, MDL models are highly effective at handling conflicting or unclear data, which is common in financial and economic records. The quarterly results announcement's numerical figures may be positive, but the executive's spoken remark could be cautious or negative. Even though it uses no numbers or digits, a unimodal text model might nonetheless predict a drop in stock prices. Better predictions and the resolution of inconsistencies are achieved through the use of multimodal models that leverage both types of data [22]. MDL offers a distinctive capability to contextualize data by learning cross-modal connections. When read alongside influential sources, it can tell whether a comment is merely a warning or a serious one. These models may be able to more accurately portray semantic concepts by limiting noise in representation dynamics through a particular perspective. Collecting information from a larger dataset with more attributes might help achieve the goal. Their exceptional results across several metrics demonstrate their excellent quality. Consequently, intangible asset valuation models are less prone to overfitting past accounting records and can better manage new forms of unstructured data. The administration of many types of data, including social media data, is one such example. At the end of the day, it all adds up to a more precise framework for asset valuation that accounts for all the many value generators in the modern economy.

2.3. Unstructured Data in Business Valuation

There has been a dramatic increase in unstructured data, or data that is not organized according to any particular standard. With the rise of the digital era, this trend has intensified. It accounts for more than 80 percent of all business data and contains significant information that has not yet been used to assess a firm's actual value [13]. Four. Text, photos, audio, and video are examples of unstructured data formats. Conventional methods of valuing a company need to be abandoned since intangible assets, such as a company's reputation, intellectual property, customer sentiment, and management talent, are becoming increasingly important. First, researchers will discuss the most significant challenges associated with processing this data. Next, it will investigate how unstructured data might enhance estimates of intangible assets. The extensive and varied network of unstructured data sources provides essential information for asset evaluation. The researchers mentioned that text is the predominant kind of data [26]. Disclosures made by the corporation, news articles, analyst reports, and regulatory filings are all included in this data. News and the opinions of financial experts can affect a company's valuation [30]. Earnings call transcripts and other company contacts are excellent predictors of future success because they reveal management's plan, confidence, and market response. Two of the most important categories are social media and internet data. Interactions with customers and the value of the brand are crucial in today's highly competitive business environment.

People's opinions, levels of contentment, and impressions of brands can be gleaned from real-time data collected from X (formerly Twitter), Facebook, professional forums, and product review websites (such as Yelp and Amazon). Compared with polls, which are prone to error and time-consuming, researchers argue that this data provides a more precise indicator of brand performance for companies that interact with consumers or technology [15]. The importance of unstructured data in the media industry is growing. The operations of a corporation can be identified through images and videos. It is possible to capture sales and foot traffic from satellite images of parking lots at retail centers. Integrating unstructured data improves the evaluation of intangible assets. Accuracy, timeliness, and predictive power all increase despite the difficulties that are encountered. The primary advantage is the increase in time value. A retrospective analysis is the only option because financial data is disclosed on a monthly or annual basis. Real-time unstructured data streams, such as social media sentiment, news feeds, and site visitor data, help bridge the gap between the two metrics. Reducing the time between the initial view of a firm and its evaluation may lead to the creation of more dynamic models. This model can forecast rapid market reactions to events such as product recalls, management appointments, and geopolitical developments. Measuring intangible value drivers is another advantage of leveraging unstructured data, such as product evaluations, social media posts, and emails.

According to Zambonino-Torres et al. [7], valuers may learn more about customer sentiment and brand loyalty by analyzing consumers' evaluations and social media interactions. The conventional accounting rules make this task challenging. Accurately valuing a business requires knowledge not only of the firm's financial realities but also of its competitive advantages and strategic positioning. The use of unstructured data improves both the risk and value models. The concerns managers have about hazards not highlighted in the financial statements are often evident in earnings call transcripts. The research found that media analysis can reveal new challenges in social governance, law, and the environment that affect financial flows. By using these qualitative risk indicators, valuers can assess an organization's risk and sustainability. Additionally, they allow modifications to discount rates and scenario probabilities. In general, unstructured data helps to diminish the presence of information that is unequally distributed and offers substantial insights. In most businesses, the most important data is scattered across various unstructured sources rather than consolidated in financial reports. Complex data analysis can provide analysts and investors with a competitive advantage by enabling them to identify undervalued assets [6]. The favorable reception a specialized technology receives in online expert forums may indicate its economic feasibility before it has a direct impact on the firm's bottom line. These dispersed signals will be collected and analyzed by multimodal deep learning algorithms, thereby enhancing the appraisal of intangible assets and revealing a firm's true value.

2.4. Market Sentiment Analysis

Financial experts have always been wary about the use of quantitative methods for assessing intangible assets. Several models fail to account for the fact that value is subjective and complex. Universities have access to bulk data, including publicly available research, social media posts, and financial news, which may be used to gather market statistics and participants' perspectives [26]. Therefore, the value of intangible trademarks and creative concepts is often subjective. Entrepreneurship, growth, and intellectual capital are only a few of the numerous aspects. This finding disproves assumptions and deepens our understanding of the broader investigation. Sentiment research broadens valuation frameworks beyond economic measurement to consider market psychology, which includes the collective views and behaviors of investors. Keeping tabs on market sentiment impacts the valuation and pricing of assets. Furthermore, researchers will look into its potential integration with advanced financial models. Market sentiment analysis is an essential component of contemporary accounting and finance, as you will come to realize by the conclusion of the study. The fundamental premise of market sentiment research is to use subjective data analyzed by computer algorithms to determine the global investor attitude towards a certain asset, market, or financial environment. Fundamental value and the current sentiments of buyers and sellers, including greed and fear, are among the many elements that impact market pricing.

Irrational bursts of optimism or severe pessimism might result from these emotions. Prospect theory and other early studies in behavioral finance lay the theoretical framework for the idea that investor psychology consistently deviates from rationality. Given its ability to convert qualitative, unstructured input into a quantitative, organized metric, sentiment analysis is a powerful tool for analytical models. Some examples of such sources are Bloomberg, Reuters, and X (previously Twitter), a social media platform. Removing stop words, tokenizing, and lemmatizing the text are aspects of data cleaning after collection. When choosing between lexicon, machine learning, and deep learning as methods, it is essential to consider factors such as computational cost, interpretability, and accuracy [19]. Although smaller, faster models would be better suited for high-frequency trading, a more intricate, fine-tuned model may be necessary for a thorough valuation report, given its improved accuracy. Future financial modeling and valuation assignments will rely on sentiment indicators, which in turn depend on the reliability of the sentiment research process. The mentioned variations are driven by sentiment, an essential component of this model. Moreover, market recovery is strongly correlated with widespread pessimism, and subsequent market returns are weaker when investor optimism is high [3]. There is a strong correlation between the lack of actual cash flows and the reliance on subjective narratives and perceptions among investors in growth companies, start-ups, and firms with intangible assets.

The value of these assets might be dictated more by sentiment than by traditional fundamentals in the short term. In other words, a change in optimistic sentiment, stemming from a successful product launch or favorable regulatory news, can significantly increase the perceived value of a patent portfolio. While a social media event has the mass potential to erode brand equity swiftly, it is an important intangible asset [14]. Specifically, our findings demonstrate that sentiment is not merely ambient noise but an important component that prices have not fully incorporated into long-term intangible investment cash flows. Executives responsible for funding companies' long-term strategies should understand the weight of emotion. To better understand how investors perceive their strategies and communications, managers of organizations with numerous intangible assets may benefit from sentiment research. Integrating market sentiment data into financial modeling frameworks to increase value has been the final test. Achieving integration requires little effort when sentiment indicators are fed into prediction algorithms. To forecast market returns using a time-series regression model, the researcher suggests using a sentiment score derived from news headlines as an independent variable [30]. In addition to P/E ratios and market capitalization scales. Market volatility that cannot be explained or predicted by historical volatility is a phenomenon that GARCH and other volatility forecasting models attempt to address by utilizing sentiment indexes. Particularly when dealing with intangible, mostly unpredictable asset values, this approach considers how investor sentiment affects the asset's systematic risk.

A consistently negative mood signals a higher risk, so discount cash flows from intangible assets at a higher rate. The researcher connects high terminal growth rates to increased asset confidence and optimism [9]. Utilize emotional research and market comparables in the market approach to explain value multiples. Review the tone of competitor press coverage to determine if a premium price-to-sales ratio is reasonable before applying it to your own product. By combining structured numerical data with unstructured textual and audiovisual content, our state-of-the-art integrations leverage multimodal deep learning frameworks. According to research conducted in the IFRS Foundation [20], a unified system can simultaneously process stock prices, management comments, and analysis from annual reports, and the tone of executives' voices on earnings calls. When it comes to a more nuanced and practical valuation, the model suggests that the market may respond differently to a cautiously communicated good earnings surprise by the CEO compared to an enthusiastically proclaimed identical surprise. Despite its promising future, incorporating sentiment research into financial models is far from straightforward. The market may already have reacted when sentiment is examined, making it a coincidental or even lagging indicator. Furthermore, data snooping and overfitting are issues that might arise when a model performs well with historical data but struggles with sample-typical data. The "black-box" nature of many deep learning algorithms makes them unsuitable for evaluation.

3. Materials and Methods

3.1. Research Methods

Research in developing areas, including e-commerce, science and technology, and business administration, might benefit from bibliometric analysis, which can reveal important themes, partnerships, research trends, and core ideas Ellegaard and Wallin [12]. To aid strategic decision-making, this data-driven approach combines mapping with performance analysis to show how a region has changed and the conceptual relationships that exist within it. Research in data science, artificial intelligence, and finance increasingly relies on bibliometric methods. Data interchange patterns and transdisciplinary links are the focus of the investigations [34]. Such approaches to assessing intangible assets allow for better decision-making by yielding deeper insights. The use of multimodal analytics enables researchers to do this by uncovering patterns in the research, creating maps of subject clusters, and providing clarity on intellectual structures.

3.2. Sources, Keywords, and Search Strings

The current study was carried out by the researchers using bibliometric, quantitative, and descriptive research methodologies. Dimensions is a database that contains research articles, policy research articles, and peer-reviewed literature. Data for this study were extracted from the above-mentioned database. Citation analysis, co-authorship mapping, and keyword co-occurrence are among the techniques used in quantitative bibliometric research to identify trends [11]. By correlating citation strength with frequency, these methodologies assess the research's impact and the emphasis on its themes. The provided parameters guarantee the study's validity and openness. This study examines the integration of sentiment analytics with unstructured datasets into intangible asset appraisal, focusing on the intellectual and collaborative framework that shapes this process. Author information, departmental and institutional affiliations, citation counts, and keyword counts are all part of the structured metadata that the study aims to address. Due to these characteristics, the study may be easily reproduced. After that, the researcher identifies three primary indicators: quantity, quality, and structural dimensions; and explains how bibliometric analysis is understood in this setting [2]. The number of publications, citation impact, and connections among publications, authors, and subject areas can be measured using various metrics. The Dimensions database was the primary tool for this research's keyword and data source searches. Using specific keyword searches, the researchers uncovered pertinent information, including 'market sentiment', 'multimodal deep learning', 'unstructured data analysis', 'intangible asset

appraisal’, and ‘intangible asset valuation’. These questions sought to identify the channels through which AI development has entered financial analytics.

3.3. Scope and Framework

Three essential dimensions, market sentiment analysis, multimodal deep learning, and intangible asset valuation, were used to carry out the current study. To identify interdisciplinary connections among AI, behavioral finance, and company/organizational value in 2023, the researchers employed triangulation [27]. Within its bibliometric framework, the study tracks the development of research topics and intellectual groups, as well as the distribution of key methodological approaches and scientific discoveries within this field. Design components included scientific mapping, which traces the evolution of research and the relationship between concepts over time. By integrating networks of keywords, co-occurrence, citation trajectories, and co-authorship, scientific mapping provides a methodical perspective on the development and progress of themes. Assists in the identification of both central and peripheral, but rapidly expanding, research fields that might impact the use of AI on intangible assets in the near future.

3.4. Analytical Strategy

Microsoft Excel was used to fine-tune and pre-process the 690-article dataset for bibliometric mapping and performance analysis. After filtering, the dataset contained only peer-reviewed journal articles and papers written in English. Author productivity, publication trends, and citation frequency were among the performance measures computed after data cleaning and normalization. The network visualization was inspired by VOSviewer, a tool that lets users link keywords, citations, and authorship/co-authorship. As a result, researchers identified themes and followed intellectual structures [32]. Researchers are given a systematic way to investigate asset value using this methodology. This enhances their ability to compare and contrast quantitative market data with unstructured sources such as language, images, and emotions. Finding important research institutions and trends becomes easier with AI, and researchers show how AI improves corporate finance analytics, helping fill knowledge gaps.

3.5. Time Frame

To maintain relevance, the researchers considered only publications from 2014 to 2024. To ensure our study is up to date, researchers focused on this time frame because that’s when intangible asset valuation and market sentiment analysis saw major advances.

3.6. Data Analysis

The bibliometric analysis of the collected data is as follows: The bibliometric data highlights key contributors to multidisciplinary work on multimodal deep learning, unstructured data, and market sentiment (Figure 1).

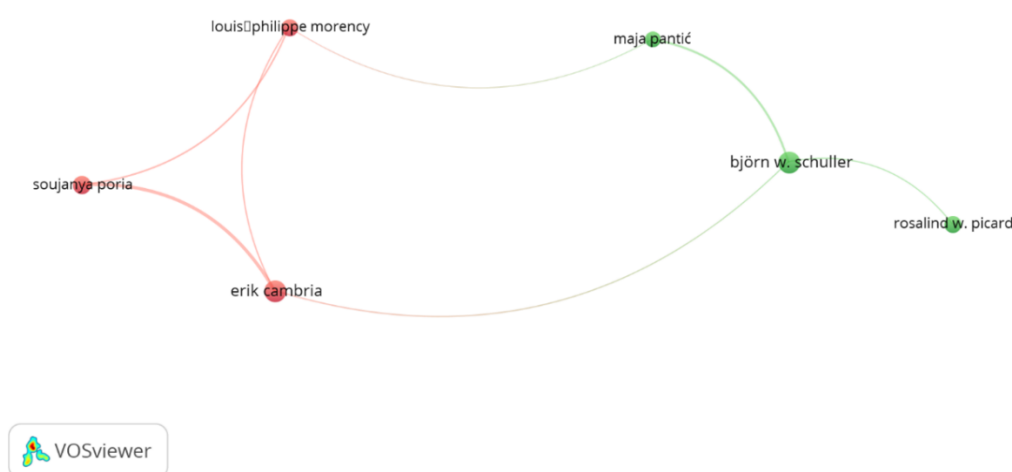


Figure 1: A VOSviewer network map illustrating co-authorship among authors

Notably, only 16 of 34,276 authors have published 20 or more articles, indicating a high concentration and specialization among the top authors. Using these numbers, researchers can see that the two main author profiles are very different. Yoshua Bengio

(198,707 citations), Erik Cambria (18,352), and Soujanya Poria (15,118) are definitely foundational characters in core AI and sentiment analysis. They are among the writers in the first group with extremely high citation counts. Tirunillai and Tellis [31] are likely central nodes in the research network due to their high total link strength and high number of citations. There is evidence of this in the second file's co-occurrence cluster, which places them in the same intellectual group as Arevalo et al. [1], all of whom are well-known figures in the fields of multimodal affective computing and sentiment analysis. Those in the second category include writers with atypical patterns, such as Parr [29], who have large numbers of documents (82 and 38, respectively) and many citations. They may be part of strong collaboration networks in highly applied or niche technical disciplines, given their high total link strengths of 37.00 apiece. On the flip side, this dataset shows that certain productive authors, such as Gandomi and Haider [13], and the highly cited Ellegaard and Wallin [12], have no link strength. This would mean that the main research talks in this bibliometric network on the intersection of multimodal learning and asset valuation don't cite their work as much as they should, despite its general influence and foundational status. In the end, the analysis suggests a field grounded in the theoretical and methodological foundations of a tightly knit core group in multimodal sentiment analysis. This core group is additionally supported by other researchers who work closely together, and the field as a whole is characterized by a highly selective author base that produces substantial scholarly work.

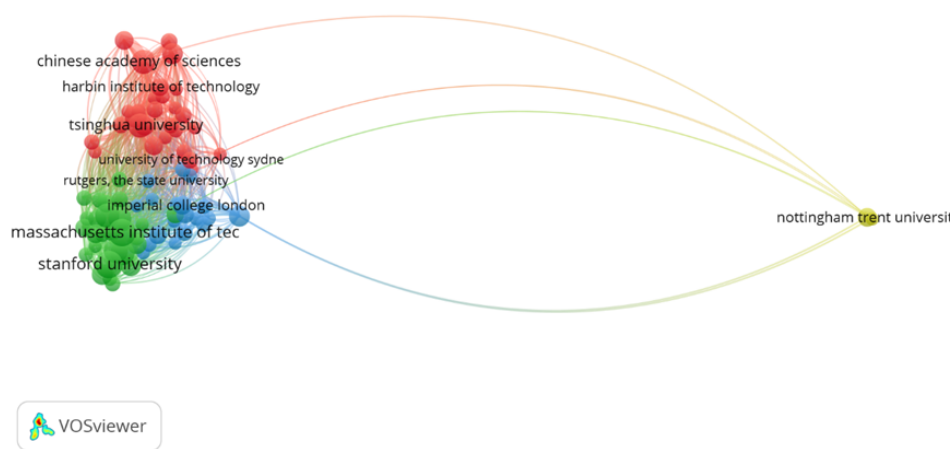


Figure 2: A VOSviewer network map illustrating scientific collaboration links among organizations

The data used for the bibliometric network analysis appear in Figure 2. This network highlights the most important firms in the field. Of the 5,606 organizations mentioned, only 77 have published 40 or more papers, indicating a narrow focus on high-output research within this multidisciplinary area. Elite Anglo-American institutions dominate the hierarchy, as seen in the visualization and data table. Stanford University and MIT rank first and second in both volume and influence, with 181 and 172 publications, and citation counts of 293,880 and 193,892. These two nodes are crucial to the research discourse because of their high “total link strength” ratings (82 and 87), indicating that they collaborate widely. Tech behemoths like Microsoft and Google lead the way in the private sector. Particularly noteworthy are Google’s research outputs, which have garnered 374,153 citations in more than 99 publications. There is a strong presence of Chinese universities alongside these leaders. Among the most prolific institutions in China, Tsinghua University has 139 publications, and the Chinese Academy of Sciences (CAS) has 138. Their citation counts are high, but they lag behind those of Western colleges and corporations. This disparity could be due to the Chinese contributions’ relative youth and continued success in accruing citations, or to an emphasis on publishing quantity rather than quality. Additionally, prominent academic institutions such as Carnegie Mellon University, the University of California, Berkeley, and Oxford University also have extensive networks of connections and references.

Academic institutions such as Nottingham Trent University, Imperial College London, the University of Technology Sydney, and Rutgers within the threshold group reflect the worldwide reach of research in this field, which integrates data science, artificial intelligence, and finance to varying degrees. The countries and regions worldwide dominated global research output, with Asia making a significant contribution. This conclusion is based on bibliometric data from 54 countries with at least 20 publications. The US leads with 3,708 papers and 3.94 million citations, showing its broad influence. The UK has been a major contributor to this field with 1,172 publications and over one million citations. China has fewer citations than the UK, indicating a lower impact on publications. The world's most productive paper producer is China (1,874). Besides the top three — France, Italy, Spain, and Germany — other countries produce hundreds of documents and form strong collaboration and citation links. Asian countries and areas are remarkable. The strong collaborative networks of mainland China, Hong Kong SAR, and Singapore are evident in their high total link strength relative to their document counts. The Asia-Pacific research community

benefits from South Korea, Japan, India, Taiwan, and Australia. Although less popular, South Asian and Middle Eastern nations such as Iran, Turkey, Pakistan, and Saudi Arabia participate (Figure 3).

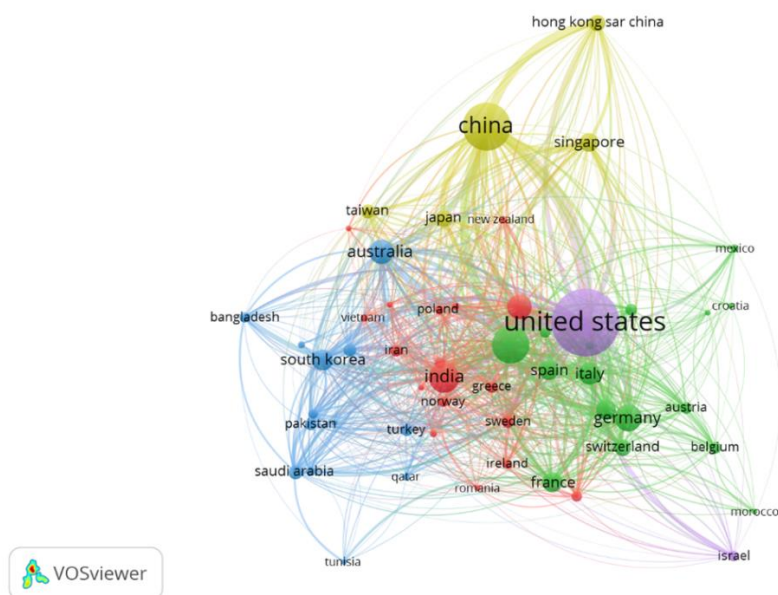


Figure 3: A VOSviewer network map illustrating authorship among countries

The research demonstrates that the integration of multimodal deep learning into financial valuation is a fast-growing, internationally collaborative area, led by the US and UK. Asia is the field's hub. Examining the top 16 most frequently referenced writers (34,276) reveals a distinct group of researchers. Twenty documents with fifty citations are needed to identify notable and prolific writers. When seen through VOSviewer's network visualization capabilities, the field is dominated by a small number of highly cited individuals. The greatest number of citations (198,707) goes to Bengio, whose work has had a significant impact on deep learning. Among the many authors who have made significant contributions to the field of emotional computing and multimodal data studies are Erik Cambria (18,352 citations), Soujanya Poria (15,118 citations), Louis-Philippe Morency (15,335 citations), and Björn W. Schuller (12,531 citations) (Figure 4).

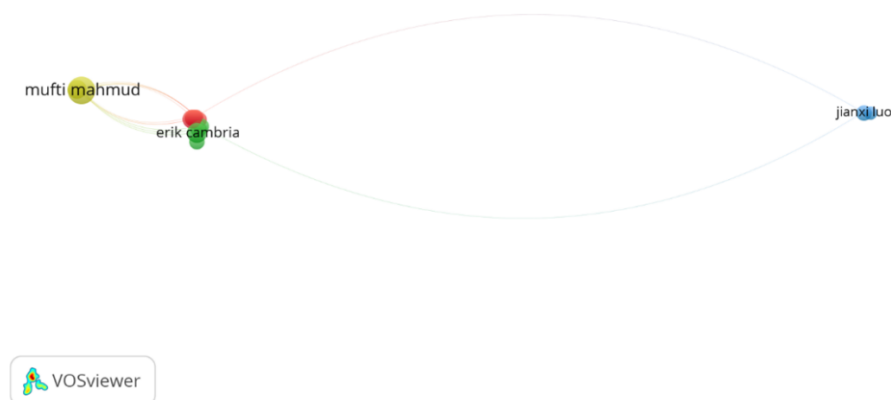


Figure 4: A VOSviewer network map illustrating citations of authors

Two groups—one focused on foundational deep learning and the other on emotional computing—show the diversity and value of knowledge in this research area. Data reveals another group of authors who shape the field through community and collaboration. Authorship linkages form a prolific and well-connected group. Brain science, artificial intelligence, and data processing leaders Ding [10], with 403 and 394 documents, and the greatest co- Although undercited, Secundo et al. [25] and Wood are notable for their numerous connections and may focus on engineering design and innovation, making them part of an emerging cluster. These top-cited AI pioneers, core affective computing academics, and important application specialists collaborate to develop multimodal deep learning for complicated valuation problems. Although undercited, Inkinen [8] and

Wood are notable for their numerous connections and may focus on engineering design and innovation, making them part of an emerging cluster. These top-cited AI pioneers, core affective computing academics, and important application specialists collaborate to develop multimodal deep learning for complicated valuation problems. Figure 5 presents bibliometric data on the top organizations in the field of "Multimodal Deep Learning for Enhancing Intangible Asset Valuation." US and Chinese institutions are highly competitive, and elite universities dominate (Figure 6).

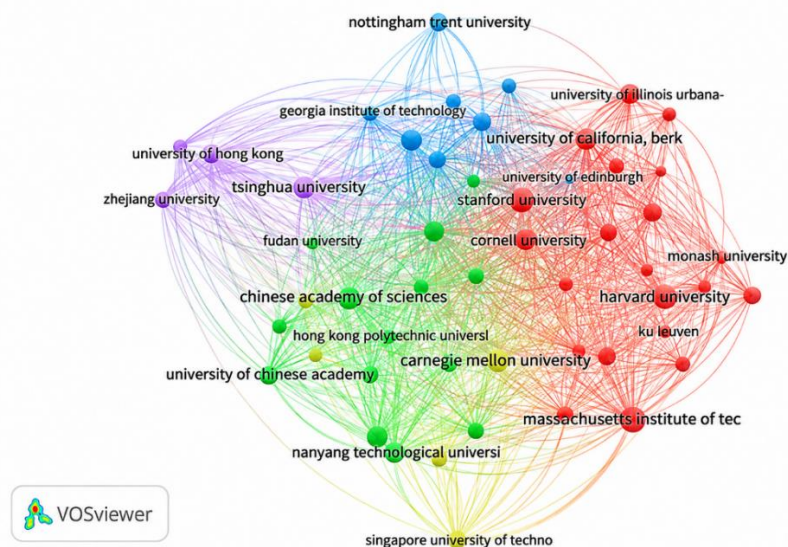


Figure 5: A VOSviewer network map illustrating citations of organizations

The 52 organizations that met the strict threshold of 50 documents and 50 citations out of 5,606 are highlighted according to their professional influence. Stanford University (293,880 citations), Cornell University (234,713 citations), and the University of California, Berkeley (362,180 citations) stand out among the top American colleges according to citation metrics, indicating their significant impact. With 374,153 citations, Google is just one example of the power of corporate research; Microsoft is another. Tsinghua University (88,134 citations), the Chinese Academy of Sciences (44,772 citations), and Fudan University have made significant contributions to this topic. These universities have highlighted Asian and American leadership. Hong Kong's University of Science and Technology and Nanyang Technological University are just two of many Asian institutions that contribute to Asia's growing clout.

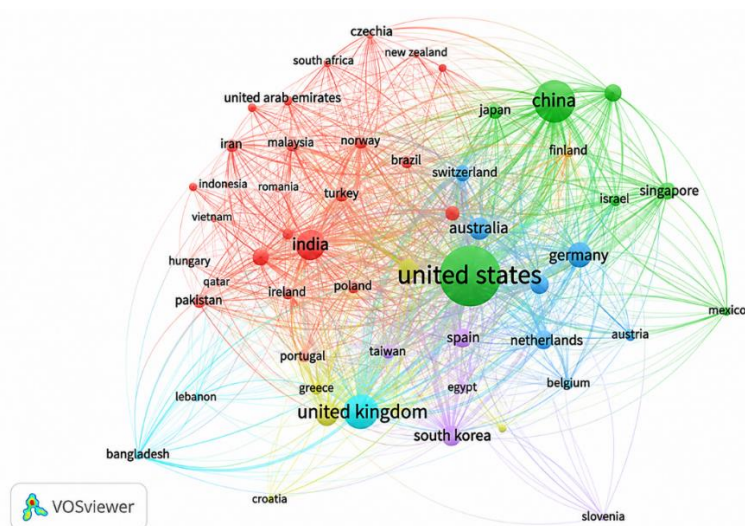


Figure 6: A VOSviewer network map illustrating citations of countries

The topic's global importance is further shown by the substantial roles played by European universities such as the University of Edinburgh, KU Leuven, and the University of Oxford. Most of the key improvements have been made by 52 renowned

universities. This indicates that, although the discipline is still in its early stages, well-established centers in computer science and data analytics are driving progress. This worldwide, collaborative, and competitive network is pioneering methods for evaluating intangible assets using complex, multimodal data. Bibliometric data from 54 nations show a distinct hierarchy of influence and contribution, based on at least 20 papers and 50 citations. With 3,5,916 links, 3,941,669 citations, and 3,708 publications, the United States is by far the leader. This establishes the United States as the leading publisher and a key location for international research collaborations (Figure 7).

Figure 7: Map based on text data from the bibliographic database (Title and Abstract Field)

Ultimately, 211 key concepts were defined, shedding light on the field's core problems. According to the paper's title, the most frequently used terms are "sentiment" (1.82354), "large language model" (1.2298), "LLM" (1.8578), and "dataset" (1.7257). This shows that researchers' growing interest in advanced AI, particularly LLMs, for analyzing unstructured data and market sentiment is evident. "Model" (1.0494), "Framework" (0.8709), "Prediction" (0.865), and "Feature" (1.7342) are comparable methodological terms that further bolster this argument. A major influence on research direction is the academic conference. (The abundant usage of terms such as "international conference" (246 times), "proceedings" (231 times), and "volume" (226 times) lends credence to this thrust. This corpus appears to have been extracted from conference proceedings, as evidenced by words such as "ECML PKDD" and "European conference." Commonly used terms include "full paper" and "short paper" as well. Model validation and empirical evaluation are given great importance in the research. This is demonstrated by the high relevance scores of the terms "experiment" (2.0035), "performance" (1.9906), "benchmark" (2.2835), and "accuracy" (2.0995). This field of introspection pursues the voids. The words "limitation" (1.949) and "challenge" (123) prove this. The phrase "future direction" (1.5833) implies possible next steps. Using advanced learning algorithms to handle data value concerns, this work paints a picture of a dynamic, real-world field. Worldwide computer science conferences are the primary venues for disseminating and discussing new information.

Extensive scientometric methods (quantitative techniques for analyzing scientific literature, such as mapping patterns in published academic work) formed the analytical basis of this bibliometric study, which aimed to map the intellectual and

collaborative contours of the research subject. In this investigation, VOSviewer—a specialized program for creating and viewing bibliometric networks (i.e., graphical representations of relationships between scientific publications)—served as the principal instrument. The papers included in this research centered on asset valuation methods that combined multimodal deep learning (the use of artificial neural networks to analyze multiple types of data, such as text and images, simultaneously) with unstructured data integration (incorporating data not organized in a predefined manner, such as text in social media posts or images) and market sentiment analysis (evaluating attitudes and trends in financial markets using computational methods that process large textual datasets). The analysis was conducted using a comprehensive dataset that was sourced from esteemed academic databases. During this time, citation analysis was used to discern between seminal and more recent works by counting the number of times other publications cited a paper. This method measures the impact of research. The field's dominating vocabulary, methodological trends, and basic research challenges were also identified by a co-occurrence analysis (tracking how often specific terms or keywords appear together in publications) of keywords and concepts extracted from abstracts and titles. When these approaches were combined, they produced a data-driven picture of a dynamic, rapidly growing multidisciplinary research environment.

Building upon these bibliometric methods, the results of the citation and co-authorship analyses show that this area is home to a very small number of highly specialized thinkers. Based on an analysis of 34,276 authors, just 16 have written 20 or more documents, indicating that prolific contributions are quite rare. Different domains of impact might be categorized under this limited category. The first group is renowned for its groundbreaking work. An early and influential figure in deep learning, Sun et al. [6] has amassed 198,707 citations for his work. Beatty et al. [5] laid the theoretical groundwork for this topic, and much of the discipline is based on his work. Erik Cambria, Baltrušaitis et al. [4] are well-established figures in the field of emotional computing and sentiment analysis, as evidenced by their high citation counts and strong overall link strength. They are perfectly aligned and form a second cluster. Their cooperative network serves as the theoretical foundation for analyzing unstructured data and drawing conclusions about market sentiment. Authors like Donthu et al. [11] are part of a third profile. These writers are very productive and have large networks of collaborators; nevertheless, they may work in more applied technical fields, which shows that the field is branching out into real-world problems. Extending the analysis from individuals to institutions, the organizational environment of this study is dominated by a collaboration of prominent Anglo-American academic institutions and top technical corporations, which clearly establishes a hierarchy of influence and output. Experts in the field are carrying out this study. The fact that only 77 of 5,606 organizations met the stringent requirement of 40 publications suggests that only a handful produce the bulk of high-quality research. As vital nodes with extensive networks of connections, a few universities—including Stanford, MIT, and Berkeley—are at the vanguard of academic impact and volume. Google and Microsoft lead industry research. For example, Google earned 374,153 citations from just 99 papers, demonstrating the power of practical applications.

This field relies on public-private synergy, with academic and corporate R&D working together to drive innovation. Despite its many problems, the United States remains the leading center for national-level research and the world's top research institution. Intellectual circles place a high value on American opinion. With 3,708 articles, 35,916 collaborative links, and 3.94 million citations, this work is quite impressive. The United Kingdom is second only to the United States in terms of size and importance as a secondary center. China has 1,874 publications, more than any other country, while having a lower total citation count than the UK. This chasm suggests that, whether due to a focus on regional publishing or a more recent launch, the potential worldwide impact of quantitative strategy output has not been fully explored. The rapid emergence of the Asian research hub is being propelled by strong collaborative networks among Singapore, Hong Kong, and China. The field is grounded in fundamental AI research and defined by an exclusive, global network of top universities. Authors, groups, and countries are all part of these citation studies and analyses. The author's citation map makes it easy to see which schools of thought are influencing this field, from Bengio's deep learning theory to the work of Cambria and Poria in applied emotional computing. Google, Stanford, and UC Berkeley have the critical mass to produce research with substantial influence, as shown by their citation dominance. Someone at the very top of the company is clearly involved in this. National citation landscape data shows that, aside from the technically savvy United States and United Kingdom, countries like Canada, Germany, and Switzerland also contribute significantly. While emerging research nations are making their voices heard, the data demonstrate that the established knowledge economies continue to dominate the scientific community. Through this prism of social structure, research methods can be examined. You can learn a lot about the field's methodological aims and how information is presented by looking at the titles and abstracts.

Artificial intelligence engineers are constantly looking for better ways to handle and leverage massive, unstructured datasets. "Sentiment," "dataset," "benchmark," and "large language model" (LLM) are all verbs that support this method. The terms "experiment," "accuracy," and "prediction" all refer to actions and facts that have been proven through experiments. The research paradigm advances when all models and frameworks are carefully evaluated using the most up-to-date standards. Rather than publishing in journals, computer science conferences disseminate information through means such as "international conference," "proceedings," and "ECML PKDD" (European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases). It is unexpected to find the field's primary platform for discussion. To stay ahead of the

competition, the culture places a premium on quick feedback and information sharing, and technology is continually evolving. This bibliometric study sheds light on a research field that is both strategically vital and constantly evolving. The preceding insights show that practical innovation is being propelled by industry. In contrast, a small but influential group of authors and institutions is driving the integration of multimodal deep learning for intangible asset assessment. Within the United States of America and a small handful of other countries lies the bulk of this core. The present research goal is to study the performance of massive language models on complex valuation tasks, with deep learning and sentiment analysis providing the theoretical underpinnings. In all of these fields, the theoretical underpinnings are well-established. The field's reliance on conference findings is a sign of its roots in the community of computer scientists and of its rapid development. On the flip side, the concentration of power suggests potential bottlenecks and the need for more intellectual diversity.

4.1. Recommendations

Take these steps to strengthen the field's development:

- Connect core AI clusters directly with experts in financial economics, accounting, and regulatory policy. Improve the real-world use of valuation models.
- Develop standardized, open-source benchmarks and datasets for intangible asset valuation. Ensure model evaluation is reproducible and transparent.
- Prioritize journal publications and long-term impact studies to support conference research. Deepen validation and track model performance and economic effects over time.

Acknowledgment: The authors sincerely acknowledge the academic support, research facilities, and collaborative environment provided by the University of Sialkot and Govt. H/S 43/WB, which greatly contributed to the successful completion of this research work.

Data Availability Statement: The datasets and supporting materials generated during this study are available from the corresponding author upon reasonable request, subject to institutional policies and applicable ethical requirements.

Funding Statement: This research was conducted without financial support from any government, commercial, private, or non-profit funding agency.

Conflicts of Interest Statement: The authors declare that they have no financial, professional, or personal conflicts of interest related to this research, and all referenced sources have been appropriately acknowledged.

Ethics and Consent Statement: The study was carried out in accordance with accepted ethical standards, with informed consent obtained from all participants and appropriate measures taken to protect confidentiality and privacy.

References

1. J. Arevalo, T. Solorio, M. Montes-y-Gómez, and F. A. González, "Gated multimodal units for information fusion," *arXiv preprint arXiv:1702.01992*, 2017. [Accessed by 07/07/2025].
2. M. Aria and C. Cuccurullo, "bibliometrix: An R-tool for comprehensive science mapping analysis," *Journal of Informetrics*, vol. 11, no. 4, pp. 959–975, 2017.
3. M. Baker and J. Wurgler, "Investor sentiment in the stock market," *Journal of Economic Perspectives*, vol. 21, no. 2, pp. 129–152, 2007.
4. T. Baltrušaitis, C. Ahuja, and L. P. Morency, "Multimodal machine learning: A survey and taxonomy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 2, pp. 423–443, 2019.
5. A. Beatty, S. Liao, and J. P. Weber, "How Good Is Goodwill Accounting? Comparative Evidence on Post-FAS 141(R) Acquired Intangibles Accounting," *The Accounting Review*, vol. 101, no. 3, pp. 1–37, 2026.
6. Y. Sun, L. Liu, Y. Xu, X. Zeng, Y. Shi, H. Hu, J. Jiang, and A. Abraham, "Alternative data in finance and business: Emerging applications and theory analysis (review)," *Financial Innovation*, vol. 10, no. 12, pp. 1–32, 2024.
7. M. J. Zambonino-Torres, J. M. Coello-Viejó, and S. C. Zambonino-Torres, "Strategic value driven by artificial intelligence in global businesses: A bibliometric and qualitative analysis of the most influential literature," *Frontiers in Artificial Intelligence*, vol. 9, no. 4, pp. 1–19, 2026.
8. H. Inkpen, "Review of empirical research on intellectual capital and firm performance," *Journal of Intellectual Capital*, vol. 16, no. 3, pp. 518–565, 2015.
9. A. Damodaran, "Investment valuation: Tools and techniques for determining the value of any asset," 3rd ed., *John Wiley & Sons*, Hoboken, New Jersey, United States of America, 2012.

10. X. Ding, Y. Zhang, T. Liu, and J. Duan, "Deep learning for event-driven stock prediction," in *Proc. 24th International Joint Conference Artificial Intelligence (IJCAI 2015)*, Buenos Aires, Argentina, 2015.
11. N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, "How to conduct a bibliometric analysis: An overview and guidelines," *Journal of Business Research*, vol. 133, no. 9, pp. 285–296, 2021.
12. O. Ellegaard and J. A. Wallin, "The bibliometric analysis of scholarly production: How great is the impact?" *Scientometrics*, vol. 105, no. 7, pp. 1809–1831, 2015.
13. A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *International Journal of Information Management*, vol. 35, no. 2, pp. 137–144, 2015.
14. D. Garcia, "Sentiment during recessions," *The Journal of Finance*, vol. 68, no. 3, pp. 1267–1300, 2013.
15. M. Ghiassi, J. Skinner, and D. Zimbra, "Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network," *Expert Systems with Applications*, vol. 40, no. 16, pp. 6266–6282, 2013.
16. T. Hastie, R. Tibshirani, and J. Friedman, "The Elements of Statistical Learning: Data Mining, Inference, and Prediction," 2nd ed., *Springer*, New York, United States of America, 2009.
17. J. Hirschberg and C. D. Manning, "Advances in natural language processing," *science*, vol. 349, no. 6245, pp. 261–266, 2015.
18. T. Hulugh and C. O. Ikeh, "Artificial intelligence in product management: Automating roadmap prioritization through sentiment analysis and customer feature demand modeling," *International Journal of Science and Research Archive*, vol. 15, no. 1, pp. 1455–1474, 2025.
19. C. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, no. 1, pp. 216–225, 2014.
20. IFRS Foundation, "IAS 38 Intangible Assets," *IFRS Foundation*, London, United Kingdom, 2020.
21. L. Kogan and D. Papanikolaou, "Technological innovation, intangible capital, and asset prices," *Annual Review of Financial Economics*, vol. 11, no. 12, pp. 221–242, 2019.
22. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 5, pp. 436–444, 2015.
23. B. Lev, "The deteriorating usefulness of financial report information and how to reverse it," *Accounting and Business Research*, vol. 48, no. 5, pp. 465–493, 2018.
24. B. Lev and F. Gu, "The End of Accounting and the Path Forward for Investors and Managers." *John Wiley & Sons*, Hoboken, New Jersey, United States of America, 2016.
25. G. Secundo, P. Del Vecchio, J. Dumay, and G. Passiante, "Intellectual capital in the age of Big Data: establishing a research agenda," *Journal of Intellectual Capital*, vol. 18, no. 2, pp. 242–261, 2017.
26. T. Loughran and B. McDonald, "Textual analysis in accounting and finance: A survey," *Journal of Accounting Research*, vol. 54, no. 4, pp. 1187–1230, 2016.
27. S. Chen, S. Ren, and Q. Zhang, "Hybrid architectures that combine LLMs and predictive analytics for next-generation financial modeling," *Mathematical Modeling and Algorithm Application*, vol. 6, no. 2, pp. 31–43, 2025.
28. R. F. Reilly and R. Schweihs, "The Handbook of Business Valuation and Intellectual Property Analysis," *McGraw-Hill Professional*, New York, United States of America, 2004.
29. R. L. Parr, "Intellectual Property: Valuation, Exploitation, and Infringement Damages," *John Wiley & Sons*, Hoboken, New Jersey, United States of America, 2018.
30. P. C. Tetlock, "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," *The Journal of Finance*, vol. 62, no. 3, pp. 1139–1168, 2007.
31. S. Tirunillai and G. J. Tellis, "Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent Dirichlet allocation," *Journal of Marketing Research*, vol. 51, no. 4, pp. 463–479, 2014.
32. D. Wong, "VOSviewer," *Technical Services Quarterly*, vol. 35, no. 2, pp. 219–220, 2018.
33. D. Pastor, J. Glova, F. Lipták, and V. Kováč, "Intangibles and methods for their valuation in financial terms: Literature review," *Intangible Capital*, vol. 13, no. 2, pp. 387–410, 2017.
34. W. Wang, D. Tran, and M. Feiszli, "What makes training multi-modal classification networks hard?" in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, Washington, United States of America, 2020.

Publisher's Note: The publisher remains impartial concerning jurisdictional claims in published maps and institutional affiliations. Responsibility for the content rests entirely with the authors and does not necessarily reflect the publisher's perspectives.