risks

MDPI

Article

News Sentiment and Liquidity Risk Forecasting: Insights from Iranian Banks

Hamed Mirashk, Amir Albadvi *, Mehrdad Kargari and Mohammad Ali Rastegar

Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran 14117-13114, Iran; h.mirashk@modares.ac.ir (H.M.); m_kargari@modares.ac.ir (M.K.); ma_rastegar@modares.ac.ir (M.A.R.)
* Correspondence: albadvi@modares.ac.ir

Abstract: This study addresses the critical challenge of predicting liquidity risk in the banking sector, as emphasized by the Basel Committee on Banking Supervision. Liquidity risk serves as a key metric for evaluating a bank's short-term resilience to liquidity shocks. Despite limited prior research, particularly in anticipating upcoming positions of bank liquidity risk, especially in Iranian banks with high liquidity risk, this study aimed to develop an AI-based model to predict the liquidity coverage ratio (LCR) under Basel III reforms, focusing on its direction (up, down, stable) rather than on exact values, thus distinguishing itself from previous studies. The research objectively explores the influence of external signals, particularly news sentiment, on liquidity prediction, through novel data augmentation, supported by empirical research, as qualitative factors to build a model predicting LCR positions using AI techniques such as deep and convolutional neural networks. Focused on a semi-private Islamic bank in Iran incorporating 4,288,829 Persian economic news articles from 2004 to 2020, this study compared various AI algorithms. It revealed that real-time news content offers valuable insights into impending changes in LCR, particularly in Islamic banks with elevated liquidity risks, achieving a predictive accuracy of 88.6%. This discovery underscores the importance of complementing traditional qualitative metrics with contemporary news sentiments as a signal, particularly when traditional measures require time-consuming data preparation, offering a promising avenue for risk managers seeking more robust liquidity risk forecasts.

Keywords: banking liquidity risk; risk prediction; liquidity coverage ratio; sentiment analysis; natural language processing

check for updates

Citation: Mirashk, Hamed, Amir Albadvi, Mehrdad Kargari, and Mohammad Ali Rastegar. 2024. News Sentiment and Liquidity Risk Forecasting: Insights from Iranian Banks. *Risks* 12: 171. https://doi.org/10.3390/risks12110171

Academic Editor: Ramona Rupeika-Apoga

Received: 29 August 2024
Revised: 4 October 2024
Accepted: 12 October 2024
Published: 30 October 2024

1. Introduction

In the modern banking sector, the proactive management of risks is vital for financial institutions, with liquidity risk being particularly significant alongside credit, operational, and market risks (Zarei 2016; Tavana et al. 2018). Liquidity risk involves balancing long-term investment strategies with short-term obligations to shareholders and investors, which is crucial for a bank's stability. Mismanagement of liquidity can lead to bankruptcy or the erosion of a bank's viability. The 2007–2008 financial crisis highlighted the importance of liquidity in market efficiency. Liquidity risk encompasses various factors, including operational risk losses, credit quality decline, reliance on short-term loans, and susceptibility to external market influences. The effective management of liquidity risk is essential for a bank's stability and success (Matz 2006).

The discernible linkage between market risk factors and liquidity risk prompts an investigation into changes, sentiments, and financial behaviors within financial markets as potential indicators of liquidity risk. While previous studies have extensively examined bankruptcy prediction in banking or predominantly focused on fundamental factors (Doumpos et al. 2016; Elsas et al. 2010; Köhler 2015; Tasca et al. 2017; Manthoulis et al. 2020), financial market dynamics are often influenced by imbalanced information and behavioral idiosyncrasies, like market sentiment and speculative behavior (Calabrese and Giudici

2015). This interaction between market risk and behavioral dynamics underscores the significant impact of news and market events on liquidity risk. For instance, changes in external factors from market fluctuations affecting various asset types can lead to capital outflows from banks and liquidity risk (Campbell and Shiller 1986; MacKinlay 1997). This highlights the crucial role of news analysis in predicting stock prices and asset valuation. Additionally, economic conditions may result in loan defaults, creating credit risk that impacts liquidity risk in banks (Boguth et al. 2016).

The current approaches require extensive daily news analysis for manual market event prediction. Text mining, especially NLP[1], offers a promising solution. Text mining efficacy relies on chosen methods like TFIDF[2], bag-of-words, n-grams, and word embedding as text representation and machine learning algorithms (Hagenau et al. 2013; Schumaker and Chen 2009; Nam and Seong 2019; Nguyen et al. 2015; Kraus and Feuerriegel 2017). This study introduces a novel approach, integrating unique news augmentation with TFIDF and n-grams for machine learning-based prediction algorithms.

There are similar impacts across different languages given similar banking factors. For instance, news of a "terrorist attack on a military base" may influence liquidity risk regardless of language, as it affects customer sentiment towards deposit and investment safety. However, nuances may exist, requiring machine learning and AI algorithms trained with specific language data. While not exhaustive, this study aims to explore commonalities among languages and banking factors to facilitate future investigations with a broader scope.

Amidst the complexity of banking risk management, predicting liquidity risk position before it manifests in financial institutions becomes imperative. This prediction mandate extends beyond internal bank parameters, encompassing environmental factors like political decisions, macro-policy changes, and market events (Roeder et al. 2022). Consequently, the hypothesis posits that news sentiment influences liquidity risk, given its forward-looking nature, which necessitates empirical testing and validation.

The banking sector, particularly in emerging markets like Iran, faces significant challenges related to liquidity risk management, which can jeopardize financial stability and operational efficiency. Traditional approaches to liquidity risk assessment, such as the Liquidity Coverage Ratio (LCR), often rely on historical quantitative data and fail to capture real-time dynamics influenced by external factors, such as market sentiment and news events. This limitation is particularly pronounced in the context of Iranian banks, which operate under unique economic conditions exacerbated by sanctions and political volatility. Therefore, there is a pressing need for a more comprehensive framework that integrates qualitative signals—specifically news sentiment—into liquidity risk prediction models. This study aims to fill this gap by exploring how sentiment analysis can enhance the accuracy of liquidity risk forecasting, ultimately contributing to more robust risk management strategies in the banking sector. The significance of this research lies in its potential to provide banks with timely insights that can inform decision-making and regulatory compliance, thereby fostering greater resilience in the financial system.

This study aimed to predict LCR position (up, down, stable) rather than its exact value, employing a methodology distinct from traditional approaches, a facet largely overlooked in previous studies and hindering direct comparisons. External signals suggesting LCR fluctuations prompt an analysis of their liquidity impact. This study innovatively integrated external signals, exploring sentiment's influence on liquidity, supported by empirical research on market, political, and economic events assessment.

Through a literature review, we discuss previous efforts in applying sentiment analysis to finance and risk management and statistical methods and machine learning to liquidity risk assessment and prediction. The key literature inspiring this study includes Tavana's integration of machine learning methods for liquidity risk assessment (Tavana et al. 2018), An's prediction model for liquidity risk (An 2017), the ELECTRA model for evaluating banking feedback and news data through sentiment analysis (Mohanty and Cherukuri 2023), the impact of financial markets on risk considering sentiment (Paraboni et al. 2018), exposure to liquidity risk in Islamic banks (Khan et al. 2023), and the impact of NPL and credit risk on liquidity risk in Islamic banks (Hassan et al. 2018).

This study demonstrates an implicit relationship between the sentiment of news as a qualitative parameter and liquidity risk. Therefore, this qualitative factor is employed in predicting liquidity risk as a signal. Risk managers can use the proposed method to quickly assess liquidity positions, complementing conventional methods that require more time for computation and recommendation.

This paper's subsequent sections are structured as follows: the Section 2 provides a comprehensive review of the literature on sentiment analysis, especially in banking risk and related fields. The Section 3 explains the research data, methods, and framework utilized, outlining proposed models encompassing data, features, and implementation strategies. The following sections analyze results from the proposed methodology, evaluate applications and performance, and conclude with recommendations for future research.

2. Literature Review

2.1. Liquidity Risk Measurement

The contemporary financial landscape has placed a premium on banks' risk assessment, management, and compliance with evolving regulatory frameworks, including Basel I, II, and III, emphasizing the comprehensive evaluation of credit, market, and liquidity risks alongside operational and legal concerns. Liquidity risk, stemming from the protracted financing of bank assets through short-term debt, engenders an inherent mismatch between cash inflows and outflows, notably in financial products characterized by uncertain cash flow schedules (Jarrow and Deventer 1998; Kumar and Yadav 2013). The computation and assessment of liquidity risk predominantly entail scenario-based analyses, categorizing risk factors into two distinct groups: bank-specific and market-specific factors (Basel Committee on Banking Supervision 2010). Bank-specific indicators such as credit rating, operational losses, and market rumors interplay with market-specific elements like capital market disruptions and economic recessions to forecast liquidity risk scenarios, providing foresight into potential liquidity crises before they materialize (Tavana et al. 2018; Musakwa 2013; Diamond and Dybvig 1983).

2.2. Liquidity Coverage Ratio (LCR)

The Basel Committee on Banking Supervision has promulgated quantitative liquidity standards, notably the LCR, devised to gauge a bank's capacity to cover net cash flows over the ensuing 30 days from high-quality liquid assets, ensuring LCR $\geq$ 100% as defined in Equation (1) (Basel Committee on Banking Supervision 2008):

Quality

$$\begin{gathered} LCR = (Quality\ cash\ assets)/\left(Net\ outflows_{over\ the\ next\ 30\ days}\right) \geq 100\% \\ Net\ outflows_{over\ the\ next\ 30\ days} = Inflows_{over\ the\ next\ 30\ days} - Outflows_{over\ the\ next\ 30\ days} \end{gathered} \quad (1)$$

Three key factors in determining the liquidity coverage ratio (LCR) include the value of cash assets, the excess rate between liabilities and assets, and the division of requested deposits into long-term and short-term categories with respective coefficients (Tavana et al. 2018). Calculating and estimating these parameters requires significant effort monthly. However, with a historical record of the LCR, anticipating its fluctuations over time becomes possible, a focal point of this research.

This study aimed to predict monthly bank liquidity risk position using the preceding month's news data, necessitating a risk measure sensitive to short-term fluctuations. To fulfill this requirement, the LCR was selected owing to its forward-looking nature and data accessibility, particularly within Iranian banks. Based on this, other measures like NFSR, probabilistic models, and balance sheet ratios are not addressed in this study (Drehmann and Nikolaou 2013; Jobst 2014).

While not diminishing the efficacy of conventional liquidity risk measurement methods, this approach seeks to provide an overarching perspective on bank liquidity positions, particularly when traditional measures require time for data preparation and calculation.

Factors influencing liquidity risk fall into three main categories: bank-specific, microeconomic, and macroeconomic or external factors (Rostami 2015; Wójcik-Mazur and Szajt 2015; Alharbi 2017; Singh and Sharma 2016; Ahmad and Rasool 2017). Notably, liquidity risk assumes a pivotal role as a determining factor for other risks such as credit risk and a cornerstone of banking performance (Bissoondoyal-Bheenick and Treepongkaruna 2011; Arif and Anees 2012; Athanasoglou et al. 2008; Molyneux and Thornton 1992).

Factors influencing liquidity risk fall into three primary categories: bank-specific, microeconomic, and macroeconomic factors (Rostami 2015; Wójcik-Mazur and Szajt 2015; Alharbi 2017; Singh and Sharma 2016; Ahmad and Rasool 2017). Bank-specific factors typically include elements like capital adequacy, non-performing loans (NPLs), and return on assets (ROA). For instance, capital adequacy, measured by the capital adequacy ratio (CAR), provides a buffer against financial shocks, and banks with higher CAR generally exhibit greater liquidity stability (Tasnova 2022). NPLs negatively impact liquidity by reducing asset quality, while high ROA may incentivize banks to engage in more risky lending, potentially harming liquidity (Radovanov et al. 2023).

On the microeconomic front, bank size and operational efficiency play critical roles. Larger banks tend to have better access to funding sources, enabling them to maintain liquidity more effectively. However, during financial crises, even large banks may face liquidity constraints if they are highly leveraged (Radovanov et al. 2023; Pham and Pham 2022).

Macroeconomic factors such as GDP growth, inflation, and interest rate policies are external forces that can significantly impact liquidity risk. For instance, rising inflation often leads to tighter monetary policies, which restrict liquidity by increasing interest rates (Tasnova 2022). Conversely, lower economic growth can reduce loan demand, allowing banks to hold more liquid assets (Radovanov et al. 2023).

New studies also highlight the influence of regulatory frameworks such as Basel III, which mandates the Liquidity Coverage Ratio (LCR) and the Net Stable Funding Ratio (NSFR) to enhance banks' liquidity buffers (Radovanov et al. 2023).

This study focuses on estimating the effect of sentiment on the Liquidity Coverage Ratio (LCR), due to data accessibility limitations that prevent the inclusion of other liquidity risk indicators. Each study examines a specific case; future research may consider investigating the effects on other liquidity risk measures.

2.3. Prediction and Assessment of Liquidity Risk Using Machine Learning Methods

The assessment and prediction of liquidity risk have garnered attention through the application of machine learning (ML) techniques. Tavana et al. proposed a model integrating artificial neural networks and Bayesian networks to assess liquidity risk, while Guerra et al. leveraged ML for stress-testing and early warning system (EWS) scenarios (Tavana et al. 2018; Mohammad et al. 2020). An proposed a predictive model for financial liquidity risk using linear discriminant, probit, and logit models, emphasizing the significance of significant variables selection in firms' liquidity positions (An 2017).

Of particular note is the application of natural language processing (NLP) techniques, notably sentiment analysis, in the banking sector. Sentiment analysis elucidates sentiments from textual resources, offering applications ranging from market prediction to risk detection and forecasting in banking (Jiang and Lu 2020; Khurana et al. 2023; Solangi et al. 2018). Coleman et al. categorized sentiments into uncertainty, positive, and negative sentiments during the 2007–2008 crisis, elucidating how sentiments impact risk analysis (Nopp and Hanbury 2015). In the context of liquidity risk, this investigation categorizes sentiment into three distinct classifications: influential, non-influential, and counter-influential. These classifications delineate sentiment's capacity to respectively engender positive, neutral, or adverse effects on risk ratios. Keyword influence on liquidity risk was not manually calculated due to lacking lexicons.

Halls et al. conducted a review of various sentiment analysis methods, encompassing DNNs, recurrent neural networks (RNNs), convolutional neural networks (CNNs), and recursive neural networks (Do et al. 2019).

This study employed ML and AI to automate keyword sentiment analysis, understanding their impact using a black box method. Evaluating ML algorithms revealed results based on specific criteria with unseen data, identifying words attracting attention in liquidity risk contexts.

2.4. Studies on Sentiment Analysis in Risk Management

Sentiment analysis integration in banking risk management offers insights, highlighted by various studies presented in Table 1, revealing diverse methodologies and applications.

Table 1. Summary of sentiment analysis research in financial and risk management.

Year	Area	Article Title	Method
2023	Risk Management	Liquidity Risk Prediction Using News Sentiment Analysis	Using DSR and Machine Learning Approach for Liquidity Risk Prediction (Mirashk et al. 2023)
2015	Risk management	Detecting Risks in the Banking System by Sentiment Analysis	Using the sentiment analysis approach to identify risk in the bank (Solangi et al. 2018)
2019	Risk management	Discovering bank risk factors from financial statements	Identifying risk factors from financial statements using text mining (Wei et al. 2019)
2014	Risk management	Risk reporting by German banks	Review of financial reports by financial banks: using standard methods (Schlueter et al. 2014)
2018	Risk management	Text Mining and Reporting Quality in German Banks	Checking the quality of financial reports using sentiment analysis in German banks (Fritz and Tőws 2018)
2018	Social banking	Application of Social Media Analytics in the Banking Sector	Social network and trend analysis are used to investigate the consumer's view in the payment industries (Manzira and Bankole 2018)
2018	Investment	The impact of Media Sentiment on Firm Risk	Analyzes the positive relationship between media sentiment and a company's future stock returns (Huang et al. 2018)
2021	Market risk	The Liquidity Dimensions to Sentiment Analysis through Microblogging Data	Finding pessimistic content (blog) increases trading costs, illiquidity, and price dispersion (Guijarro et al. 2021)
2018	Market risk	The relationship between Sentiment and Risk in Financial Markets	Investigating the effects of financial markets on risk with the sentiment analysis approach (Paraboni et al. 2018)
2019	Bank health	Using Annual Report Sentiment as a Proxy for Financial Distress in U.S. Banks	Using annual analytical reports as a proxy to detect financial crises in American banks (Gandhi et al. 2019)

Table 1. *Cont.*

Year	Area	Article Title	Method
2018	Bank health	Text Mining and Reporting Quality in German Banks	Investigating the quality of financial reports using sentiment analysis (Fritz and Tőws 2018)
2019	Bank health	Discovering bank risk factors from financial statements based on a new text mining algorithm	Comprehensively reveals a bank's risk factors from the textual risk disclosures reported in the financial statements (Wei et al. 2019)
2019	Financial sentiment analysis	Financial Sentiment Analysis with Pre-trained Language Models	The FinBERT language model, designed based on BERT, addresses financial sentiment analysis challenges (Araci 2019)
2023	Financial sentiment analysis	Sentiment Analysis on Banking Feedback and News Data using Synonyms and Antonyms	Evaluates sentiment scores of words, rephrases using synonyms/antonyms, and uses the ELECTRA model for SA (Mohanty and Cherukuri 2023)
2022	Financial sentiment analysis	Aspect-Level Sentiment Analysis Algorithm based on BERT for Multi-Domain Texts	BERT-based algorithm for aspect-level sentiment analysis across domains, leveraging source domain data to enhance target domain analysis (Liu and Zhao 2022)
2022	Financial sentiment analysis	Deep Learning-based Sentiment Analysis of Financial Statements	A sentiment analysis method for financial statements using deep learning and domain adaptation (Shao and Chen 2022)
2023	Financial sentiment analysis	Explainable hybrid word representations for sentiment analysis of financial news	Improve SA with explainable hybrid word representation, addressing class imbalance and integrating embeddings (Adhikari et al. 2023)
2021	Financial sentiment analysis	Analyzing DistilBERT for Sentiment Classification of Banking Financial News	Compares fine-tuned DistilBERT and TF-IDF with supervised machine learning classifiers for sentiment classification (Dogra et al. 2021)

Texts offer diverse content, including explanations and opinions, posing challenges for sentiment analysis. Regulators seek objective risk reporting, yet Webb et al. found subjectivity in 44% of news sentences (Wiebe et al. 2001).

A study by Feroz Khan et al. indicates that the non-performing loans (NPLs) of Islamic banks are 4 to 5 times higher than the standards set by international finance experts (Khan et al. 2023). This finding suggests that Islamic banks are more susceptible to risks, particularly credit and liquidity risks, which need to be managed with greater care. Another study by Kabir Hassan et al. shows a positive relationship between liquidity risk and credit risk in conventional banks, but a negative relationship in Islamic banks. Initially, lower liquidity risk may improve stability. However, as bank management takes on more risk to increase profitability, this offsets the initial positive impact, leading to increased instability and affecting other risks in a chain reaction (Hassan et al. 2018). Based on the stochastic behavior of liquidity risk in Islamic banks and higher levels of risk, this study aims to predict liquidity risk to prevent further consequences.

The amalgamation of sentiment analysis within banking risk management has significantly progressed risk detection and prediction, while also unveiling new avenues for analyzing liquidity risk factors. Leveraging sentiments from diverse textual sources facilitates a deeper understanding of the interplay between sentiments and liquidity risk factors, enhancing risk assessments and predictive models in the banking sector. Factors influencing liquidity risks include forecasting financial crises using sentiment analysis from news and social networks (Ormerod et al. 2015), integrating sentiments into credit risk assessment for historical evaluation (Erlwein-Sayer and Yu 2020), predicting bank interest rates by analyzing online news (Kumar and Yadav 2013), assessing political impacts on stock and bond fluctuations (Erlwein-Sayer 2018), and indirectly addressing market risk through sentiment analysis of news and market events (Paraboni et al. 2018). The literature review underscores the multifaceted nature of liquidity risk, affirming the integration of machine learning and natural language processing techniques for refined risk assessment and prediction in banking.

3. Research Method

The outline of this research is shown in Figure 1. Sources of input data included news and data of a bank LCR. Features were then extracted from the text data. In this step, using feature extraction methods in text mining, qualitative features were extracted from textual news data and these features were examined for use in the next step.

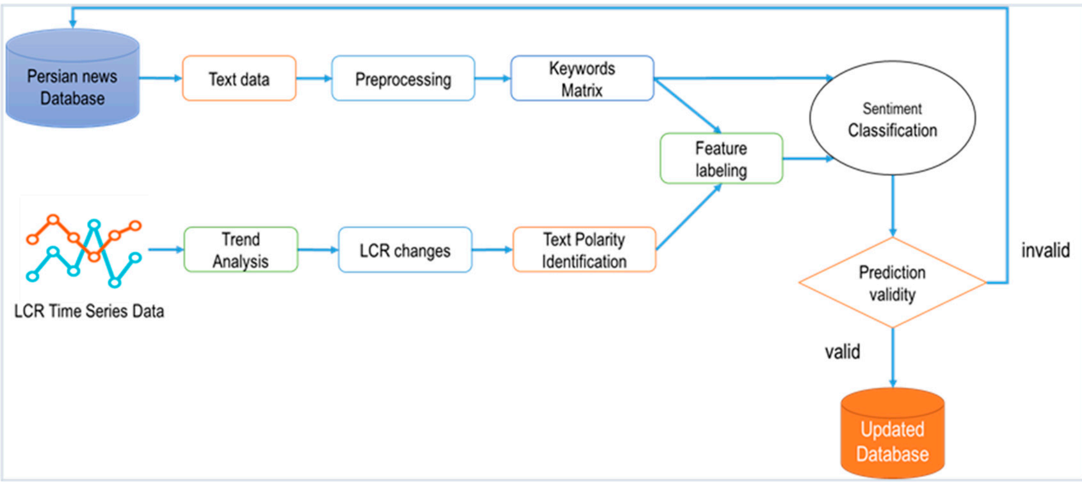


Figure 1. Overview of the research.

The general outline of the research methodology comprises several key steps aimed at understanding and predicting liquidity risk through sentiment analysis of textual data.

1. Data Collection: Quantitative and qualitative data were sourced from a bank for liquidity data and a Fars digital news agency for textual news data, primarily focusing on economic and political spheres.
2. Text Data Preparation: For qualitative data, the HAZM NLP library was employed for the Persian language. Textual data underwent initial processing, which included normalizing and refining formats and word forms, clearing texts, classifying texts, extracting and validating, removing duplicate words, polarity detection, labeling, sentiment analysis, result validation, and presentation and comparison of findings. The quantitative data obtained over time for the LCR from a sample bank required quantification to integrate it into the sentiment prediction model. For instance, if the LCR of the current month was 90% and the LCR of the previous month was 80%, there was a 10% increase in LCR, thus determining the label for that month as influential (1).

3. Qualitative Feature Extraction: Features were extracted from processed news texts to establish correlations with liquidity risk.
4. Sentiment Analysis Model Construction: Conventional machine learning and deep learning methods were utilized to construct a sentiment analysis model, with various algorithms tested and the optimal one selected. Validation criteria including accuracy, precision, and recall were employed to assess and compare outcomes.
5. Prediction and Model Evaluation: Liquidity risk was predicted and evaluated using the chosen method, with validation criteria such as accuracy and visualizations employed for comparison with actual values.

These steps adhered to the research roadmap aiming to identify qualitative variables impacting liquidity risk prediction, providing insights into the role of sentiment analysis in this domain.

3.1. Research Variables

Banking standards like Basel and CAMELS, along with previous research, categorize influential variables into economic indicators, internal bank metrics, and qualitative factors like sentiment, indirectly impacting liquidity risk. As shown in Figure 2, the internal factors of a bank, which were referred to in previous sections as microeconomic or controlled factors (static variables), and external factors, which were either implicit or time-bound (dynamic variables), have a direct effect on liquidity risk. Therefore, they can be used to predict liquidity risk.

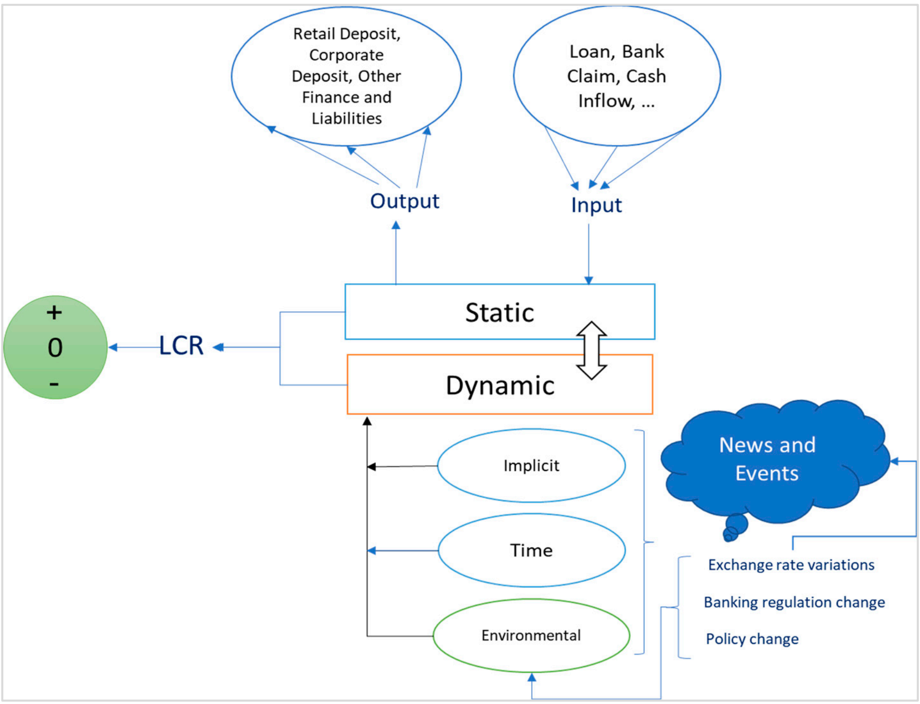


Figure 2. The effect of qualitative variables as a dynamic influencing factor in predicting liquidity risk.

This study aimed to investigate the impact of qualitatively dependent variables, previously overlooked, on liquidity risk, noting their dynamic nature and interrelation with internal and external factors, ultimately employing machine learning models to predict liquidity coverage ratio (LCR) position using dynamic variables extracted from news and events.

3.2. Data Collection

Research data were classified into two categories of dependent and independent variables. As depicted in Table 2, the dependent variables were the news qualitative features and the independent variables were the liquidity risk ratios; in this regard, how to collect qualitative data and then the dependent or predictor variable were explained.

Table 2. Research data.

Variable Name	Variable Type	Type of Data	Data Time	Source of Data
Liquidity coverage ratio	Quantitative variable	Bank liquidity risk data	April 2004–November 2020	A semi-private sector bank in Iran
News quality index	Qualitative variable	News	April 2004–November 2020	Fars News Agency

3.2.1. Collecting Quantitative Data

Regarding quantitative data, as mentioned in the table above, the liquidity coverage ratio index for a bank was considered, which was calculated and reported on a monthly basis. As stated in the previous study, the bank risk index is an indicator with a history, which means that it shows the current or past status of the bank (Nopp and Hanbury 2015).

The data on liquidity risk from April 2004 to November 2020 were obtained from the Central Bank of Iran for a semi-private bank (excel file containing LCR for each month). Semi-private banks offer advantages resembling private banks, catering to various customers and providing diverse products. They handle high transaction volumes and invest in non-banking markets. However, they face state bank-like regulations, including those from central banks, affecting loans, interest rates, and anti-money laundering measures. Due to data scarcity, the analysis focused on one semi-private bank, representative of others, to understand liquidity dynamics. To ensure data validation and reliability, the Central Bank of Iran operates a supervision office. This office employs both onsite and offsite procedures to monitor, assess, and confirm the reports provided by banks. These procedures adhere to the supervisory review and evaluation process mandated by BASEL guidelines, upon which we relied for our analysis.

Figure 3 shows the trends of the liquidity risk index of the bank from 2004 to 2020. In this chart, the trend of changes in the liquidity risk of the bank in question is clear and can be extracted.

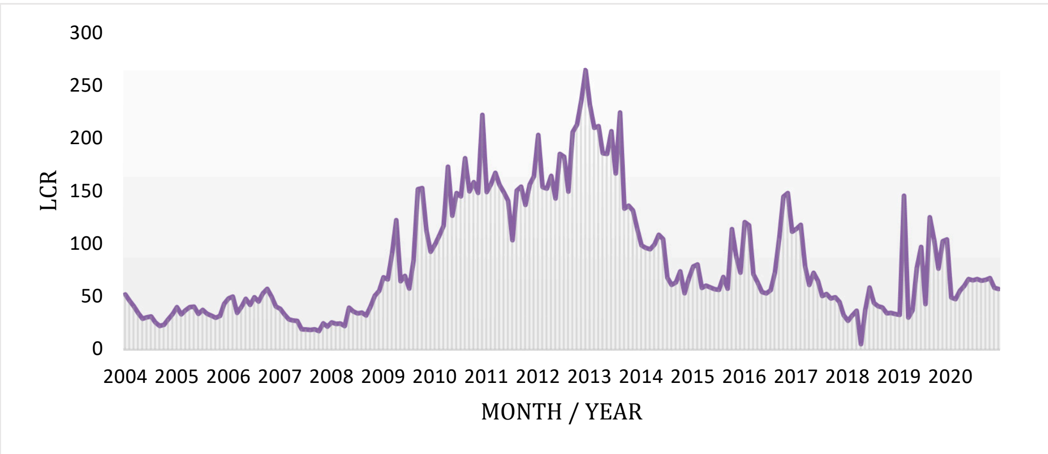


Figure 3. Trend chart of the liquidity risk index of the bank evaluated during the study period (LCR, or liquidity coverage ratio, is a liquidity risk indicator defined by the Basel Committee on Banking Supervision).

3.2.2. Collection of Qualitative Data

This research focused on digital Persian news from the Fars News online agency, highlighting economic and political subjects potentially influencing liquidity risk. By analyzing these articles, we can extract sentiment—positive, negative, or neutral—reflecting public perception towards the bank's liquidity. News data were crawled from the Fars News Agency from April 2004 to November 2020 and maintained in an SQL server database containing news summaries, types, bodies, titles, and dates. Data were gathered using web scraping methods, employing the Get request HTTP command through the Beautiful Soap (BS4) library in Python to retrieve news and extract relevant content such as news titles, summaries, and descriptions. The code was developed in the Python language, and cloud server resources were utilized for website crawling.

Keywords were extracted from news items, particularly in economic, political, and financial realms, by combining the three fields of summary, title, and body of the news, while also exploring the impact of censorship on Persian news. While economic and financial news generally evades censorship, political censorship seeks to mitigate economic consequences or boost public confidence. Machine learning was employed to assess the influence of manipulated news on bank liquidity, aiding in predicting future outcomes. Fars News Agency was chosen due to its diverse news sources, including political parties, enabling the examination of manipulated news effects on liquidity risk amidst censorship, alongside its considerable volume and diversity of news.

As depicted in Figure 4, with a total of 4,288,829 news items collected before preprocessing, their forward-looking nature (Fritz and Tőws 2018) underscores their relevance to liquidity risk, reflecting trends across the years.

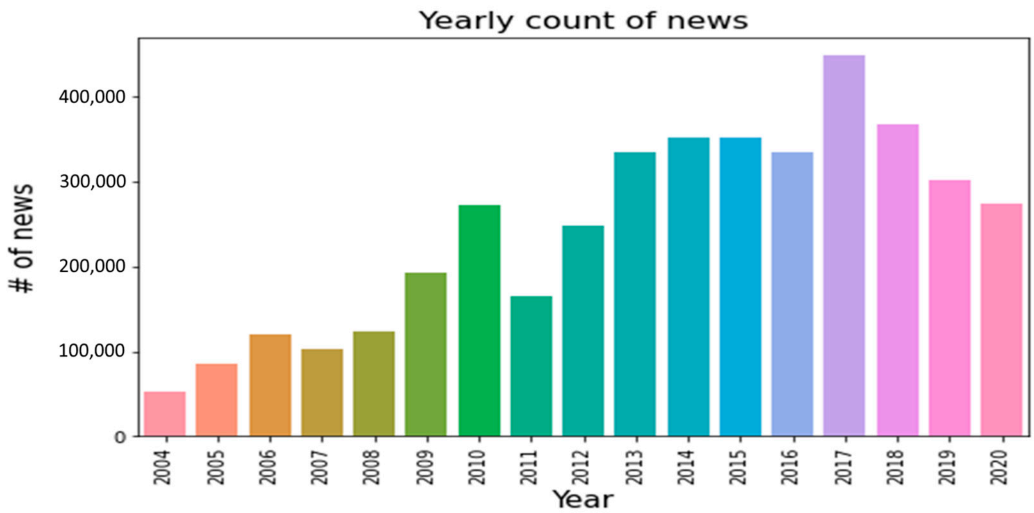


Figure 4. Chart of news collected by different years.

Liquidity risk is influenced by both current and past conditions of a bank, while news sentiment can predict future liquidity risk; hence, changes in liquidity risk from the previous month serve as labels to determine the polarity of news samples, for instance, an increase in liquidity risk results in a positive label for the previous month's news sample.

3.3. Pre-Processing of Collected Data

The data underwent preprocessing for both qualitative (textual) and quantitative segments. Qualitative data, comprising news titles, summaries, and bodies, were merged into cohesive news samples. Quantitative data, initially divided into three or five main classes, were categorized monthly, labeling liquidity risk based on each month's changes: it could be either upward (1), downward (−1), or constant (0). Post-preprocessing, the dataset comprised approximately 494,650 samples, constituting a near-balanced dataset. Algorithm 1 shows the preprocessing phase.

Algorithm 1. Preprocessing	
1:	Notation: Input: Text File (Persian News Dataset)
2:	Output: Sentiment (Positive, Neutral, Negative)
3:	Begin
4:	Read text data from Persian news dataset (combine summary, body, and lead for each news item)
5:	Remove URL, numbers, punctuation
6:	Standardization and tokenization of sentences into words
7:	Words -> remove stop words
8:	Add words from each sentence into preprocessed library
9:	Stemming and lemmatizing words into their root forms
10:	Add stemmed words from each sentence into processed library to be labeled

In the realm of Persian language analysis, textual data preprocessing is a meticulous process aimed at refining and standardizing text for optimal analysis. This involves employing a combination of regular expressions and replacement dictionaries for format standardization, followed by text normalization through techniques like stemming and lemmatization facilitated by the HAZM library. Punctuation removal and whitespace management are seamlessly executed using Python libraries and the Normalizer function, respectively. Tokenization, conducted by HAZM, divides the text into tokens, facilitating subsequent analysis. Furthermore, specialized algorithms like Krautz (Estahbanati and Javidan 2011) and Taghva (Taghva et al. 2005), based on established algorithms like Porter (Porter 1997) and Lancaster (Khyani and Siddhartha 2021), are employed for accurate stemming, while lemmatization utilizes lexical dictionaries to map words to their base form, preserving crucial linguistic nuances. This comprehensive preprocessing lays a robust foundation for nuanced insights into Persian language-specific textual data.

3.4. Feature (Keyword) Engineering

In the quest for efficient information retrieval within Persian language text, this study employed robust methodologies to extract keywords essential for comprehension. The procedure consisted of the following steps, as explained in Algorithm 2:

Algorithm 2. Feature engineering	
1:	Features <- extract features (keywords) using TF-IDF and N-gram algorithms.
2:	Matrix of weighted words <- Create a matrix of features, with each row representing a weighted keyword of a sentence.
3:	For each month in LCR time series data
4:	$LCR_{changes} = LCR_{current\ month} - LCR_{previous\ month}$
5:	For each row (keywords of a news) in Matrix
6:	Label each row with LCR changes of corresponding month
7:	$positive(1)\ if\ LCR_{changes} > 0$
8:	$neutral(0)\ if\ LCR_{changes} = 0$
9:	$negative(-1)\ if\ LCR_{changes} < 0$
10:	The news matrix is divided into n groups, with each group consisting of m news items.
11:	For each m items (news) in Matrix in same month
12:	$Augmented\ sample = merge\ (m\ items)$
13:	$augmented\ sample\ lable = merge\ (labels)\ which\ is\ label\ of\ that\ month$
14:	Hyper tune m as an important hyperparameter for best classification

Feature extraction: Leveraging statistical techniques such as TF-IDF and N-Gram analysis, the text underwent meticulous scrutiny to unveil meaningful linguistic patterns. Collocations or NGram analyses discern cohesive word combinations, revealing nuanced semantic structures beyond individual words, while TF-IDF evaluation highlights the significance of words within their contextual framework. The procedural framework began with feature extraction using the TfidfVectorizer function, meticulously configured with optimized hyperparameters to strike a balance between relevance and representation. By segmenting news text into distinct training, validation, and testing sets, this study ensured comprehensive coverage of the dataset.

The extraction process prioritized salient attributes present in an optimal frequency range, guided by parameter settings like min_df (0.008) and max_df (0.05), alongside the ngram_range (set at (1, 2)). Through the consolidation of news titles, leads, and bodies, the text was prepared for in-depth analysis and model development. This meticulous approach set the stage for insightful exploration within the Persian language domain, laying a robust foundation for subsequent analytical endeavors.

Labeling data and sentiment score: In this study, news data labeling was pivotal, employing a sentiment analysis approach to discern sentiment polarity. Building on prior works by (Groth and Muntermann 2011; Kogan et al. 2009), and (Hájek and Olej 2013), that utilized machine learning and quantitative metrics for labeling, this study focused on banking sentiment analysis. Three sentiment classes—positive, negative, and neutral (as explained in Equations (2) and (3))—and five sentiment classes (as explained in Equations (2) and (4)) were employed to gauge sentiment scores, as established in previous studies (Nopp and Hanbury 2015) utilizing quantitative indicators as changes in liquidity risk index. This index, with classifications into three and five classes, was analyzed alongside textual news data from 2004 to 2020.

$$LCR_{changes} = LCR_{current\ month} - LCR_{previous\ month} \tag{2}$$

$$three\ classes \begin{cases} positive(1)\ if\ LCR_{changes} > 0 \\ neutral(0)\ if\ LCR_{changes} = 0 \\ negative(-1)\ if\ LCR_{changes} < 0 \end{cases} \tag{3}$$

$$five\ classes \begin{cases} Significant\ increase\ (2)\ if\ LCR_{changes} > 25\% \\ positive(1)\ if\ 0 > LCR_{changes} > 25\% \\ neutral(0)\ if\ LCR_{changes} = 0 \\ negative(-1)\ if\ 0 > LCR_{changes} > -25\% \\ Significant\ negative(-2)\ if\ LCR_{changes} < -25\% \end{cases} \tag{4}$$

Notably, news records were labeled with sentiment for each month, and a novel approach merged news with the same label, enhancing generalizability for robust analysis. The approach is illustrated in Figure 5.

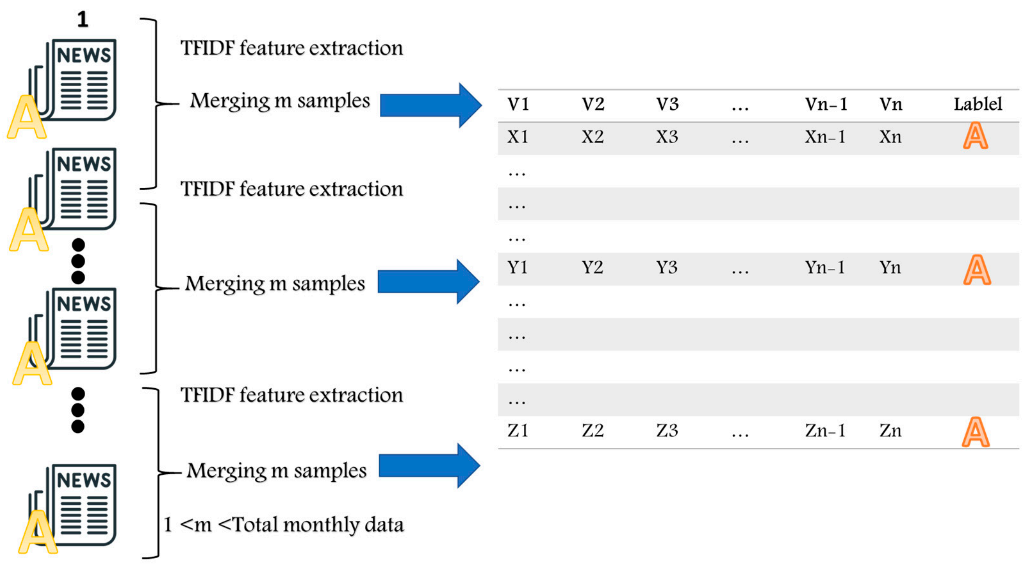


Figure 5. Extraction and integration of features in the features matrix.

Feature augmentation: This study introduces a novel data augmentation approach to merge keyword features from multiple news pieces within a monthly timeframe. By consolidating m news articles per month, a more comprehensive representation of data is achieved, acknowledging the dynamic nature of sentiment across time. Subsequently, by choosing a value for m as a consolidation number, ranging between 1 and the total number of news items (n) for that month, m news articles are merged (Equation (5)).

$$0 < m < n \tag{5}$$

Leveraging TF-IDF and N-Gram algorithms on aggregated news facilitates the selection of nuanced features, contributing to a richer dataset for machine learning models. The resulting input feature matrix condenses to n/m per sample, where m serves as a hyperparameter influencing model performance. Optimal parameters, such as 300 and 500, were determined through rigorous evaluation to shape subsequent analyses and ensure robust outcomes.

Feature selection: The abundance of features relative to observations can lead to redundancy, inefficiency, or extraneousness, exacerbating the issue. Various feature selection methods, categorized as "filter", "wrapper", or "embedded" offer distinct approaches to mitigate these challenges (Feki et al. 2012). In this study, the wrapper method was adopted for feature evaluation, intertwining feature selection with model prediction. By configuring the parameter 'm' as described earlier and analyzing the results, the most optimal feature subset was determined, contributing to more refined and efficient model training and evaluation.

3.5. Model Implementation of Sentiment Analysis

Sentiment analysis is a technique used to determine the emotional tone behind a series of words. In our study, we applied natural language processing (NLP) techniques to classify the sentiment of news articles. For example, positive news about economic growth would likely correlate with improved liquidity, while negative news about market instability could indicate potential liquidity challenges.

After data preprocessing, a keyword-based sentiment analysis approach employing machine learning methods was utilized to predict polarity based on inclusive labels of positive, negative, and neutral sentiments, with a focus on analyzing the impact of news sentiment labels on liquidity risk. As depicted in Algorithm 3, various risk prediction methods, including logistic regression (Lanine and Vennet 2006), support vector machines (Vapnik 1995), deep neural networks, and convolutional neural networks (Gouvêa and Bacconi 2021), were explored to analyze the relationship between liquidity position and news sentiment.

A. Logistic regression classifier: A widely used algorithm in sentiment analysis for banking risks like credit and default risks employs a binary dependent variable and classified or continuous independent variables, constituting a special case of the generalized linear model (Equations (6)–(8)). This study utilized logistic classification for sentiment analysis, where the hyperparameter C, representing the inverse of regularization power, was crucial in mitigating overfitting by imposing costs to optimize parameter values (Equation (9)). Through parameter tuning, particularly setting parameter C to minimize prediction errors, the logistic regression model effectively learned to predict sentiment based on input data.

$$Z = \beta' X = \ln\left\{\frac{p\,(X)}{1 - p\,(X)}\right\} \tag{6}$$

$$\beta' = (\beta_0, \beta_1, \ldots \beta_n\} \tag{7}$$

$$p(X) = E\,(Y = 1|X) = \frac{1}{a + e^{-(a+bX)}} \tag{8}$$

$$C = b = (X'X)^{-1} X'y \tag{9}$$

Algorithm 3. Sentiment model and prediction algorithm.

samples <- Using Matrix of features with labels
$train\ data = 70\%\ of\ samples$
$validation\ data = 15\%\ of\ samples$
$test\ data = 15\%\ of\ samples$
A. logistic regression classifier
    Train with $train\ data$
    Hyperparameter tuning with $validation\ data$
    Evaluate the result with $test\ data$
B. SVM classifier
    Train with $train\ data$
    Hyperparameter tuning with $validation\ data$
    Evaluate the result with $test\ data$
C. DNN classifier
    Train with $train\ data$
    Hyperparameter tuning (neural network architecture) with $validation\ data$
    Evaluate the result with $test\ data$
B. CNN classifier
    Train with $train\ data$
    Hyperparameter tuning (neural network architecture) with $validation\ data$
    Evaluate the result with $test\ data$
Compare the results of different algorithms and select the best method.
Assess the best result in a real-case scenario to evaluate the applicability of the method.

B. Multi-class support vector machine classifier:

The SVM classifier operates on the principle of linear determination within the search space, aiming to best separate different classes by specifying the separator margin. Text data fit ideally within the support vector machine classifier due to the nature of text, where multiple features are often insignificant (Equation (10)). It separates data by a hyperplane in the input space using optimization techniques, which maximizes the margin between classes and minimizes the classification error (Equation (11)). The optimization problem is expressed as follows:

$$\max_{w}(w^{*}, b^{*}) = d = \frac{2}{||w||} + c.\sum_{i=1}^{n} \varepsilon_i \tag{10}$$

$$s.t:$$

$$\begin{cases} (w.x + b) \geq 1,\ \forall x\ of\ class\ A \\ (w.x + b) \leq -1,\ \forall x\ of\ class\ B \end{cases} \tag{11}$$

This classifier has seen diverse applications, including in the classification of user comments based on their quality (Vapnik 1995). This study leveraged the multi-class support vector machine classification method for sentiment analysis, yielding satisfactory results. Notably, the C parameter, a critical parameter in this algorithm, guided SVM optimization by determining how much misclassification of each training example needed to be avoided. Tuning this parameter led to improved analysis outcomes (Chen and Tseng 2011).

C. Artificial deep neural network:

Deep learning, an evolving field within machine learning, encompasses a comprehensive approach to learning representation levels, addressing intricate nonlinear challenges in various domains like natural language processing and image analysis (LeCun et al. 2015). Recent advancements have fueled the prominence of deep neural networks, which differ from traditional machine learning in their inherent capacity to extract features across multiple levels (Yuan et al. 2018). Illustrating a multilayer deep learning classifier model, this study employed a feed-forward neural network structure with dense layers interconnected to facilitate data dimensionality reduction. Nonlinear activation functions, such as ReLU, enabled the network to capture complex patterns within the data.

Parameters like "Unit" for neuron count and "Loss" for quantifying disparity between predicted and actual outputs, alongside "Optimizer" for weight adjustment, play pivotal roles. Considering (labeling) three-class and five-class polarity, this research developed deep learning models tailored to both m = 300 and m = 500 for data augmentation, catering to diverse sentiment analysis needs spanning from positive and negative sentiments to more nuanced categories. Figure 6 delineates the deep neural network layer architecture for both three-class and five-class polarity modes, encompassing 2563 attribute matrix keywords in the input layer. Each network had five layers, with the last layer acting as the final classifier for the designated classes.

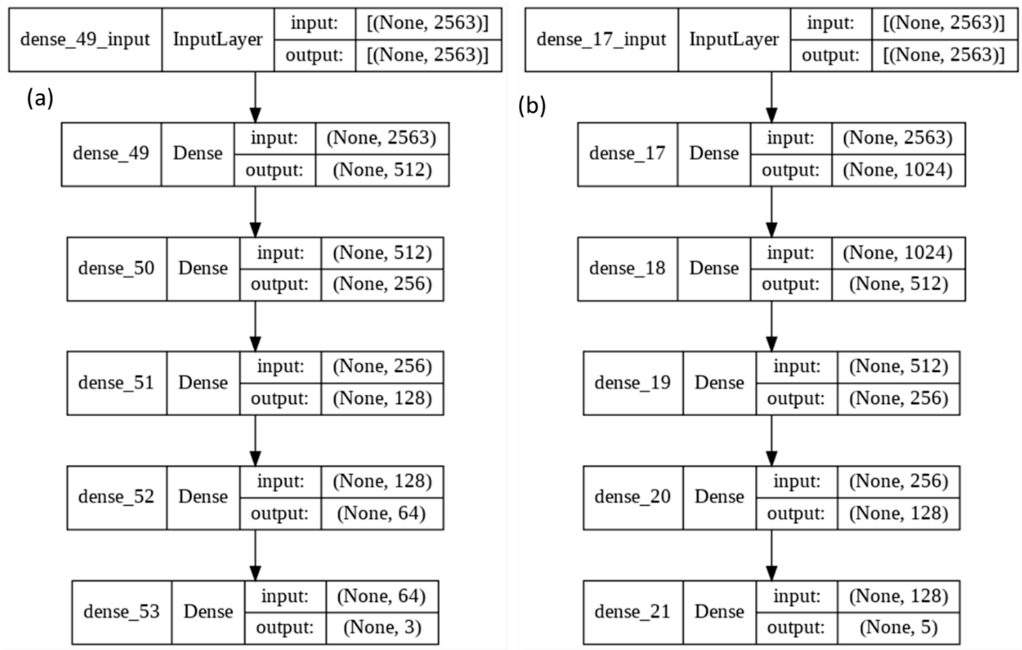


Figure 6. Architecture of neural network layers in two states: (**a**) 3-class polarity; (**b**) 5-class polarity.

D. Convolutional neural network:

In the realm of deep learning for text classification, the convolutional neural network (CNN) emerges as a standout approach, leveraging convolutional filters to autonomously extract relevant features crucial for various tasks, including sentiment classification. Notably, studies have demonstrated CNN's ability to achieve impressive performance without intricate hyperparameter tuning, making it adaptable across diverse text analysis endeavors (Read 2005; Zhang et al. 2015; Kim and Jeong 2019). The general architecture of CNNs, depicted in Figure 7, comprises layers such as convolutional, pooling, and dense layers, each playing distinct roles in feature extraction, downsampling, and classification (Thakkar and Chaudhari 2021; LeCun et al. 2015).

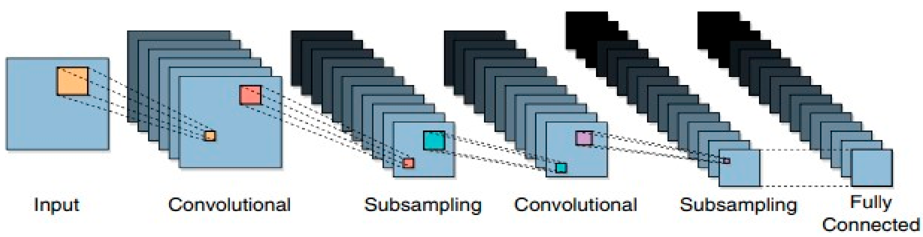


Figure 7. General model of convolutional learning network structure (Thakkar and Chaudhari 2021).

This study utilized one-dimensional convolutional and max pooling layers alongside dense layers, underscoring the CNN's effectiveness in automating feature detection and classification tasks. The hyperparameters "Filters" and "Kernel_size" controlled the number of detected features and the convolutional window's length in the Conv1D layer, respectively, while "Pool_size" defined the dimensions of the max pooling window in the MaxPooling1D layer for downsampling the feature matrix without trainable parameters.

These models (C, D) automatically learn and extract features from the input data, which helps improve the prediction accuracy compared to traditional methods.

3.6. Evaluation Criteria for Prediction

This study utilized randomization techniques to select training, testing, and validation data and evaluated algorithmic predictions using various classifier validation criteria, including the confusion matrix (Table 3), precision, recall, accuracy, F-score, Kappa, area under ROC curve (AUC), and ROC curve. Equations (12)–(17) with news data merged into a feature matrix associated with liquidity risk trend labels. The data were divided into training (70%), validation (15%), and testing (15%) sets, and classifier performance was assessed through confusion matrices and associated criteria, alongside calculations for sensitivity, specificity, accuracy, F1-score, and AUC values, with the kappa criterion comparing algorithm performance against random classification. The AUC was used to evaluate and compare different classifiers, indicating model accuracy. The kappa criterion compared the classification algorithm against a random classifier to assess multi-class algorithm performance (Sokolova et al. 2006; Juba and Le 2019; Pham and Ho 2021; Lasko et al. 2005; Zweig and Campbell 1993; Hanley and Mcneil 1982).

$$Recall = SN = \frac{TP}{TP + FN} \tag{12}$$

$$Precision = SP = \frac{TN}{TN + FP} \tag{13}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{14}$$

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{15}$$

$$\{(SN\{c\}, 1 - SP(c)) : -\infty < C < +\infty\} \tag{16}$$

$$AUC = \int_0^1 ROC(t)_{dt}\, dt \tag{17}$$

Table 3. Confusion matrix (Zhang and Liu 2021).

Actual/True Class		0–Positive	1–Negative
Predicted class	0–True	True Positive (TP)	False Positive (FP)
	1–False	False Negative (FN)	True Negative (TN)

4. Experimental Results

In this section, this study's outcomes based on liquidity risk data from a bank and news collected from Fars News Agency are discussed. News data were consolidated into 300 and 500 pieces, forming the final feature matrix and labels representing the bank's liquidity risk positions in both three-class and five-class forms. Four data input cases were established for classification algorithms (Table 4), elucidating the outcomes in this segment. The table presents various data options fed into the learning model, varying sample numbers, consolidation types, and number of classes.

Table 4. Input data modes present the input data options to the learning model.

Number of Classes	Type of Augmentation	Number of Samples	Train Samples	Test Samples	Validation
$l = 3$	$m = 300$	14,169	9918	2125	2126
$l = 5$	$m = 300$	14,169	9918	2125	2126
$l = 3$	$m = 500$	8540	5978	1281	1281
$l = 5$	$m = 500$	8540	5978	1281	1281

4.1. Data Visualization

The t-SNE algorithm, employed as an unsupervised method for visualizing high-dimensional data, facilitated the understanding of data organization. By applying TF-IDF followed by t-SNE, with perplexity as a crucial parameter, the resulting visualization in Figure 8 illustrates the varying degrees of separation for both three-class and five-class modes, offering insights into their classification capacity and separability. Considering that using m as a batch size meant merging and combining every m news item to generate a sample, we had labels a and b with a batch size of 300 and labels c and d with a batch size of 500. Therefore, Figure 8 shows that the classification and separability of m = 300 (labels a and b) as the news batch size were more pronounced.

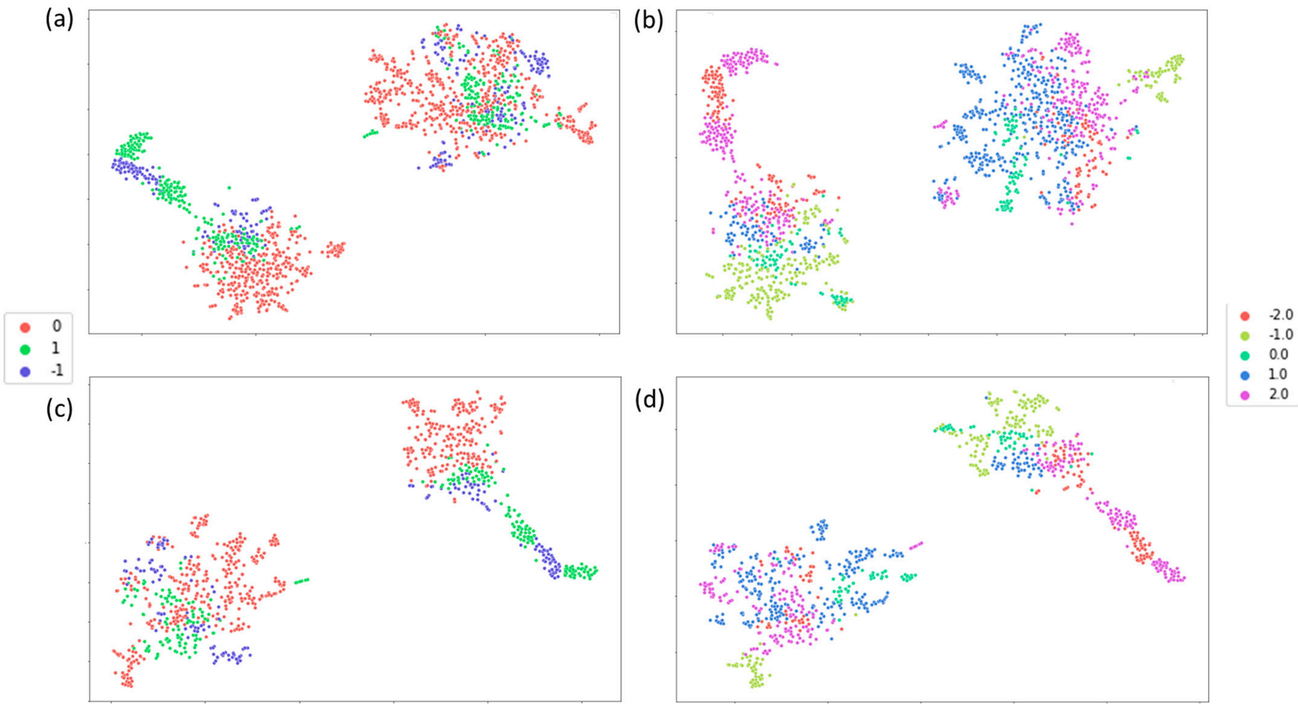


Figure 8. T-SNE diagram and the degree of separation of 3 and 5-class modes (the dimensions of the diagram were obtained after reducing the T distribution dimension so that these dimensions could be mapped to the input properties, but, in the diagram, only two dimensions are mapped. (**a**) T-SNE diagrams related to 3-class and combination model of 300; (**b**) T-SNE diagrams related to 5-class and combination model 300; (**c**) T-SNE diagrams related to 3-class and combination model 500; (**d**) T-SNE diagrams related to 5-class and combination (m) model 500. (t-distributed stochastic neighbor embedding is a non-linear algorithm that reduces high-dimensional data for easier human observation and analysis).

4.2. Adjustment of Hyperparameters

This section focuses on optimizing hyperparameters for classification algorithms using different data augmentation parameters. The comprehensive tables in Appendix A detail the fine-tuning of hyperparameters for each algorithm, with logistic regression achieving the best validation accuracy of 71.58% for both three-class and five-class modes. SVM attained an accuracy of 86.42% for the three-class mode, while the deep neural network performed well, with approximately 89.7% accuracy in both modes. However, the CNN's validation accuracy slightly decreased to 84.07% in the three-class mode. Overall, these findings show that the DNN, followed by the SVM, outperformed other algorithms based on their validation accuracy in predicting liquidity risk trends. Appendix A thoroughly describes the hyperparameters used for training, validation, and fine-tuning the model. It also includes the structure and architecture of the designed model, specifically for the CNN and DNN.

4.3. Evaluation of the Results

We compared the results of different algorithms to identify which provided the most reliable predictions. This comparison is crucial because it helps in understanding how well each model can adapt to changing market conditions.

The analysis of confusion matrix results from the logistic regression, support vector machine (SVM), deep neural network (DNN), and convolutional neural network (CNN) algorithms, detailed in Appendix B, revealed varying accuracies across different combinations. Notably, combination type m = 300 consistently outperformed m = 500 across logistic regression and SVM methods. Similarly, for deep feed-forward and CNN methods, architectural differences led to varying accuracies, with m = 300 consistently showing superior performance. These findings underscore the significance of neural configurations and classifiers in accuracy, emphasizing the need for architectural tuning for optimal performance. Overall, SVM followed by DNN emerged as the top-performing algorithms based on the confusion matrix results.

The Tables 5–8 detail the performance metrics of various machine learning models, including Logistic Regression (Table 5), Support Vector Machine (SVM) (Table 6), Deep Feed-Forward Neural Network (Table 7), and Convolutional Neural Network (CNN) (Table 8), for predicting liquidity risk. Here's a detailed explanation of each:

Table 5. Results of different validation criteria extracted from logistic regression method.

Logistic Regression										
Triple										
Bin Size	**C**	**Solver**	**Max_Iter**	**Val_acc**	**Test_acc**	**Precision**	**Recall**	**F1 Score**	**Cohens Kappa**	**ROC AUC**
300	10	lbfgs	1000	71.58	68.81	68.61	68.61	68.65	52.55	86.02
500	10	lbfgs	1000	71.58	70.73	70.68	70.73	70.56	55.51	87.89
Quintuple										
300	17	lbfgs	1000	62.45	64.02	64.03	64.02	63.98	54.79	89.38
500	20	lbfgs	1000	68.31	67.06	67.06	67.06	67.02	58.69	91.06

Table 6. Results of various validation criteria extracted from the support vector machine.

SVM										
Triple										
Bin Size	**C**	**Solver**	**Max_Iter**	**Val_acc**	**Test_acc**	**Precision**	**Recall**	**F1 Score**	**Cohens Kappa**	**ROC AUC**
300	5	rbf	2000	86.42	86.74	86.75	86.74	86.73	79.91	95.17
500	5	rbf	2000	86.42	87.04	87.01	87.04	87.01	80.38	95.37
Quintuple										
300	15	rbf	2000	82.78	85.23	85.4	85.23	85.28	81.48	96.02
500	5	rbf	2000	86.1	86.1	86.1	86.1	86.09	82.57	96.41

Table 7. Results of validation criteria obtained from the deep feed-forward learning network method.

Feed-Forward Neural Network													
Triple													
Bin Size	**Dense Units**	**Activation**	**Optimizer**	**Batch Size**	**Epochs**	**Val_acc**	**Test_acc**	**Precision**	**Recall**	**F1 Score**	**Balance Accuracy**	**Cohens Kappa**	**ROC AUC**
300	512-256-128-64	relu	adam	8	29	83.11	84.1	83.56	84.07	84.1	84.07	75.88	95.44
500	512-256-256	relu	rms	64	43	88.91	88.6	88.5	88.63	88.6	88.56	82.75	97.56
Quintuple													
300	1024-512-256-128	relu	rms	64	19	84.85	88.29	88.23	88.5	88.29	88.31	85.31	98.53
500	512-256-128	relu	rms	64	42	87.35	82.51	82.19	83.42	82.51	82.3	78.04	97.22

Table 8. Results of validation criteria obtained from the convolutional deep learning network method.

CNN															
Triple															
Bin Size	**Filter s**	**Kernel Size**	**Pool Size**	**Dense Units**	**Activation**	**Optimizer**	**Batch Size**	**Epochs**	**Balance Accuracy**	**Test_acc**	**Precision**	**Recall**	**F1 Score**	**Cohens Kappa**	**ROC AUC**
300	64-128-256	3	2	256	relu	adam	8	25	82.0	81.7	81.7	82.0	81.7	81.7	72.3
500	64-128	3	2	128	relu	adam	32	21	84.0	83.2	82.7	83.3	83.2	83.2	74.6
Quintuple															
300	128-256-512	3	2	512	relu	rms	16	16	80.2	81.1	80.9	81.4	81.1	81.2	76.3
500	64-128-256	3	2	256	relu	adam	64	30	80.9	82.2	82.1	82.3	82.2	82.2	77.7

Table 5 showcases the performance of the Logistic Regression model across different parameter configurations, focusing on two bin sizes (300 and 500) and two classification modes: “Triple” and “Quintuple.” The metrics used for evaluation include validation accuracy, test accuracy, precision, recall, F1 score, Cohen’s Kappa, and ROC AUC. The highest validation accuracy of 71.58% is achieved with the “Triple” mode and a bin size of 500. In the Quintuple mode, although validation accuracy slightly decreases, the model achieves a higher ROC AUC of 91.06, indicating good predictive capacity. Overall, the Logistic Regression model provides moderate performance, particularly in handling imbalanced datasets, as reflected by Cohen’s Kappa scores.

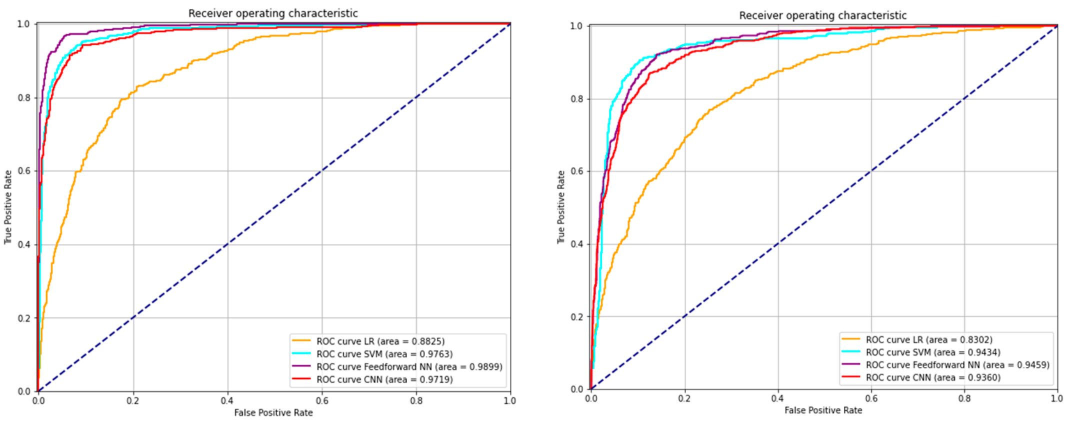
Table 6 presents the results of the Support Vector Machine (SVM) model, which outperforms Logistic Regression in both classification accuracy and stability. The SVM achieves a validation accuracy of up to 86.42% in the Triple mode with a bin size of 500, and a test accuracy of 87.04%, suggesting that SVM is highly effective for predicting liquidity risk. The ROC AUC reaches an impressive 96.41%, showcasing the model’s ability to distinguish between classes accurately. Other key metrics, such as precision and recall, are consistently high across different configurations, demonstrating the robustness of SVM in handling complex financial data.

The results in Table 7 reveal that the Deep Feed-Forward Neural Network (DNN) is the best-performing model for liquidity risk prediction. The DNN achieves a validation accuracy of 88.91% in the Triple mode with a bin size of 500, and its ROC AUC score reaches up to 98.53% in the Quintuple mode, highlighting its superior classification capability. The F1 scores and Cohen's Kappa metrics further affirm the model's accuracy and reliability, making it ideal for dealing with imbalanced data. The DNN's multi-layered architecture and optimization lead to highly accurate predictions, surpassing both Logistic Regression and SVM models.

Table 8 summarizes the performance of the Convolutional Neural Network (CNN), which, while slightly trailing behind DNN and SVM, still delivers solid results. The best validation accuracy of 84% is achieved in the Triple mode with a bin size of 500, and the ROC AUC peaks at 77.7%. Although the CNN model does not outperform the DNN, it still shows strong potential for sentiment-based liquidity risk prediction, particularly in complex, large-scale data environments. The precision and recall metrics remain consistently high, supporting its use in tasks requiring automated feature extraction from textual data.

The examination of various performance criteria revealed significant insights: the feed-forward neural network demonstrated exceptional accuracy, precision, recall, and F1-score, particularly achieving 88.6% accuracy in the three-class model and outperforming other methods like SVM. Additionally, it excelled in kappa and balance accuracy metrics, indicating robustness in handling unbalanced data. The ROC-AUC metric further underscored the superiority of the feed-forward neural network, with an outstanding 98.53% score in the five-class model. Collectively, these findings position the feed-forward neural network as the optimal method for predicting bank liquidity risk using qualitative news data, followed by SVM, convolutional neural network, and logistic regression.

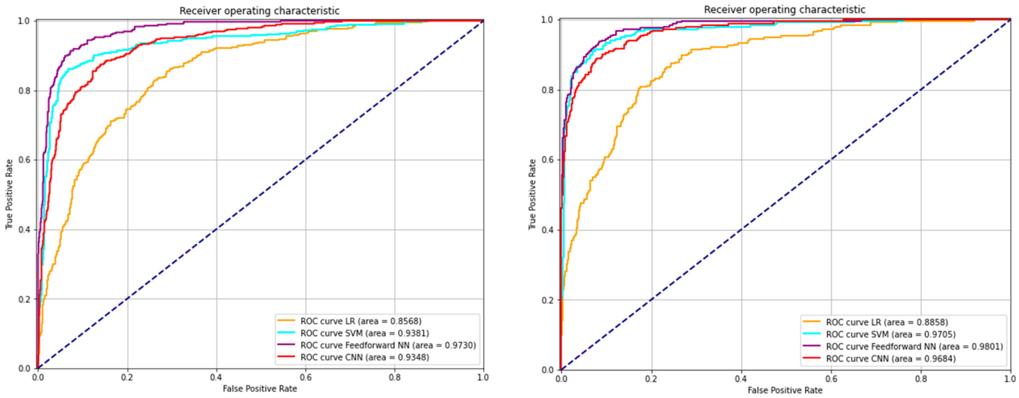
An in-depth analysis of ROC curves was essential for evaluating the methodologies employed in thid study. Figures 9 and 10 showcase ROC curves for various combinations, with Figure 9a highlighting the superior performance of the deep feed-forward neural network in the five-class mode with a combination of m = 300. Conversely, Figure 9b demonstrates comparable outcomes between the support vector machine and deep feed-forward network in the three-class mode with the m = 300 combination. In Figure 10, both the support vector machine and deep feed-forward network exhibit similar, yet notably better, results compared to other methods in both three-class and five-class modes.



(**a**) (**b**)

Figure 9. (**a**) ROC curve associated with m = 300 and l = 5; (**b**) ROC curve associated with m = 300 and l = 3.

According to BASEL, there are two areas for LCR: one is above the 100% threshold, considered a safe zone, and the other is below the threshold, indicating critical risk. This study's practical application was validated by the risk management office of the bank, confirming the model's efficacy in predicting fluctuations between safe and critical risk zones as per BASEL guidelines, thus affirming its successful practical utility.



(**a**) (**b**)

Figure 10. (**a**) ROC curve associated with m = 500 and l = 3; (**b**) ROC curve associated with m = 500 and l = 5. (Receiver operating characteristic curve is a graph showing the performance of a classification model at all classification thresholds).

5. Discussion

This study aimed to assess the impact of qualitative parameters on predicting bank liquidity risk, employing news data alongside a bank's liquidity index to build and evaluate models using test data. The top-performing model, a deep neural network, achieved an impressive 88.6% accuracy, supported by metrics like precision, recall, F1-score, and ROC-AUC. Predicting liquidity risk remains challenging, with limited prior research compared to credit or market risk, making the method proposed in this study a practical and cost-effective approach to assess a bank's risk position. Despite the uniqueness of this study's qualitative approach limiting comparisons with other studies, employing ANOVA and logit/probit models (An 2017) revealed prediction accuracies of 88.8% for DNN and 87% for SVM, surpassing a neural network model by Tavana et al. (2018) that achieved around 70%. A comparison with classic algorithms showcased the superior accuracy of the DNN and SVM models, underlining their potential for liquidity risk prediction.

In a similar vein, other studies have explored stock price prediction using sentiment analysis, exhibiting varied accuracies. A comparative analysis in Table 9 illustrates the accuracies of various methods in predicting financial risk on stock prices, while Table 10 details the accuracies achieved in this study, with DNN demonstrating the highest accuracy at 88.6%. These comparisons provide valuable insights into the effectiveness of different approaches and affirm the robust performance of the DNN and SVM models in predicting liquidity risk.

Table 9. Comparison of the results obtained from the analysis of news data to predict financial risk index and stock price (Adhikari et al. 2023).

No.	Author	Dataset	Effect	Feature Type	Methods	Accuracy
1	Schumaker and Chen 2009	US financial news	Stock prices (intraday)	Noun phrases	SVM	58.2%
2	Schumaker and Chen 2009	US financial news	Stock prices (intraday)	Noun phrases	SVR	59.0%
3	Groth and Muntermann 2011	German adhoc announcements	Stock prices (daily)	Bag-of-words	SVM	56.5%

Table 9. *Cont.*

No.	Author	Dataset	Effect	Feature Type	Methods	Accuracy
4	Mittermayer 2004	US financial news	Stock prices (daily)	Bag-of-words	SVM	-
5	Wüthrich et al. 1998	Worldwide general news	Stock prices (daily)	Bag-of-words	K-NN, ANNs, naïve Bayes	Not comparable
6	Li 2008	US corporate filings	Stock prices (daily)	Bag-of-words	Naïve Bayes	Not available
7	Antweiler and Frank 2004	US message postings	Stock prices (intraday) and volatility	Bag-of-words	Bayes, SVM	Not available
8	Das and Chen 2007	US message postings	Stock and index prices (daily)	Bag-of-words	Combinations	Not comparable
9	Tetlock 2007	US financial news	Stock prices (daily)	Bag-of-words	Ratio of negative words	Not available
10	Groth and Muntermann 2011	German adhoc announcements	Intraday market risk	Bag-of-words	SVM	Not comparable
11	Butler and Kešelj 2009	US annual reports	1-year market drift	N-Gram	Proprietary distance measure	Not comparable

Table 10. Comparison of the results obtained from the analysis of news data to predict financial risk in this study.

No.	Author	Dataset	Effect	Feature Type	Methods	Accuracy
1	Mirashk et al.	Iran Fars News (daily)	LCR monthly	N-Gram -TFIDF	Feed-forward neural network	88.6%
2	Mirashk et al.	Iran Fars News (daily)	LCR monthly	N-Gram -TFIDF	SVM	87.04%
3	Mirashk et al.	Iran Fars News (daily)	LCR monthly	N-Gram -TFIDF	Convolutional neural network	83.29%
4	Mirashk et al.	Iran Fars News (daily)	LCR monthly	N-Gram -TFIDF	Logistic regression	70.73%

Table 9 compares the results of various studies that applied different machine learning techniques to predict financial risk and stock prices. The accuracies vary significantly, with methods like SVM and SVR performing moderately well, achieving accuracy levels ranging from 56.5% to 59%. The studies focus on different financial data, including US financial news and corporate filings, with techniques like Bag-of-Words and Noun Phrase extraction. The comparison demonstrates the challenges and limitations in predicting financial risks using traditional methods.

Table 10 presents the results of the current study, showcasing the performance of different models—DNN, SVM, CNN, and Logistic Regression—in predicting liquidity coverage ratio (LCR) from Iranian news data. The DNN model stands out, achieving the highest accuracy of 88.6%, followed by the SVM model at 87.04%. The CNN and Logistic Regression models perform relatively lower but still offer valuable predictive capabilities. These results highlight the effectiveness of advanced deep learning techniques, particularly in handling complex and dynamic financial data for liquidity risk prediction.

The superior performance of deep learning models, particularly convolutional neural networks (CNNs), can be attributed to their capacity to capture complex patterns in high-dimensional data. Unlike traditional models, CNNs automatically extract relevant features from textual data, which allows them to better understand nuanced sentiment signals. This capability is crucial in the context of liquidity risk, where sentiment can rapidly shift in response to market events. Traditional models, such as logistic regression, demonstrated lower predictive accuracy. This can be attributed to their reliance on manual their inability to process unstructured text data effectively. The limited capacity of these models to capture the dynamic nature of market sentiment highlights the necessity for more sophisticated methodologies in risk prediction.

6. Conclusions and Future Works

The contribution of this study to the field of banking and finance is primarily its novel integration of sentiment analysis into liquidity risk prediction models. This approach introduces a qualitative parameter—news sentiment—into the traditionally quantitative realm of liquidity risk management. By focusing on the Liquidity Coverage Ratio (LCR) and employing advanced machine learning techniques, the study enhances the accuracy of predicting liquidity positions in banks. This methodology, which accounts for external signals like political and economic events, offers a more dynamic and real-time tool for risk managers. Moreover, the research has significant policy implications for the banking sector, especially in high-risk environments like Iran. It provides a complementary tool that can aid regulatory bodies in monitoring liquidity risk, enabling more timely interventions. This approach is particularly valuable for Islamic banks, where stricter regulations and unique risk profiles exist. By complementing traditional risk measures, the study suggests a shift in how liquidity risk could be managed, offering both academic and practical advancements in financial risk prediction.

Our study underscores the promising potential of integrating sentiment analysis for forecasting bank liquidity risk position, showcasing commendable results. The core strength of our approach lies in the thoughtful selection of feature extraction and selection methods, elevating classification accuracy and bolstering the efficacy of sentiment analysis. Our innovative concept of incorporating news data for constructing features has proven instrumental, significantly complementing conventional sentiment analysis methodologies. This novel feature selection method has substantially enhanced classification accuracy, mitigated noise, and minimized adverse effects inherent in machine learning-driven textual news classification, leading to more precise predictions.

Incorporating the $m = 500$ and $m = 300$ news augmentation models for feature selection enabled our models to achieve over 64% accuracy, reaching a pinnacle of 88.6% accuracy in predicting three-class tags. For optimal implementation, datasets with substantial news content and a clear index for prediction are pivotal. The validity of our findings was reaffirmed through a separate validation dataset, affirming the robustness and replicability of our approach.

The findings of this study have significant policy implications for risk management in Iran's banking sector. Given the high liquidity risk faced by Iranian banks, exacerbated by economic sanctions and volatile political environments, the incorporation of real-time news sentiment into liquidity risk forecasting can offer regulators and banks an early warning system. This would allow banks to anticipate liquidity shortages and take preemptive actions, such as adjusting reserves or managing loan portfolios, to mitigate risks. Furthermore, the study demonstrates the need for the Central Bank of Iran and other regulatory bodies to consider qualitative signals, like news sentiment, in their supervisory frameworks, which can enhance the robustness of liquidity stress tests and improve the overall stability of the banking system. Implementing such models could also facilitate compliance with Basel III regulations, particularly regarding liquidity coverage ratio requirements, by offering a more dynamic and responsive risk monitoring tool.

In conclusion, this study operated under the assumption that its experimental framework relied on Iranian banking data due to limitations in accessing international liquidity risk and related news data in Iran. While recognizing the necessity for validation with international data, such as from US or European banks, there exists a foundational rationale justifying the proposed approach, which can be substantiated through further international research efforts. Additionally, drawing from the insights of Mohammad et al. (2020), Islamic banks are identified as more susceptible to liquidity risk due to stricter capital regulations and credit risk, alongside significant impacts of long-term debts. Focusing on Iranian banks, which face heightened liquidity challenges, this study introduced and validated a novel approach utilizing news sentiment as a critical qualitative factor for assessing liquidity risk in high-risk banking environments. Although there is potential for similar risk selectivity among other high-risk banks internationally, additional studies are imperative to confirm this generalization.

Considering that some news in Iran is censored, there may be a bias in news orientations based on government tendencies. Although this study acknowledges the presence of such orientations, the impact of censorship on liquidity risks is evident. Further studies could investigate the effects of unbiased news sentiment on liquidity risks. Another limitation of this study is the analysis of news sentiment in isolated periods. There might be a cumulative effect of news sentiment across sequential periods. Future research could explore whether a series of news items can produce butterfly effects or unprecedented conditions using time series analysis methods.

Author Contributions: The authors confirm that contributions to this paper as follows: Conceptualization, A.A. and H.M.; methodology, H.M.; software, H.M.; validation, H.M., M.A.R. and M.K.; formal analysis, H.M.; investigation, H.M. and A.A.; resources, M.K.; data curation, H.M.; writing—original draft preparation, H.M.; writing—review and editing, A.A. and M.A.R.; visualization, H.M.; supervision, A.A. and M.K.; project administration, A.A.; funding acquisition, none. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data that support the findings of this study are available from https://drive.google.com/drive/folders/1m047ZnH9gAD6mrOXtJXwHuTRzKqg8y4X?usp=drive_link (accessed on 1 January 2024) and news is available from https://farsnews.ir/archive (accessed on 1 January 2024), but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are, however, available from the authors upon reasonable request and with the permission of the Central Bank of Iran. A source code for is also provided https://drive.google.com/drive/folders/1m047ZnH9gAD6mrOXtJXwHuTRzKqg8y4X?usp=drive_link (accessed on 1 January 2024). The source code contains all the custom computer code used to generate the results that are reported in this paper and central to its main claim.

Acknowledgments: We would like to express our sincere gratitude to Mohammad Talebi for his invaluable support, guidance, and for providing access to data from the Central Bank of Iran.

Conflicts of Interest: The authors declare no conflicts of interest. There are no funders that had a role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A. The Hyperparameter Tuning of Algorithms

Table A1. Comparison table and adjustment of hyperparameters related to logistic regression algorithm.

Classifier		Logistic Regression							
Hyperparameters	Bin Length	L = Triple				L = Quintuple			
		C	Solver	Max_Iter	Val_acc	C	Solver	Max_Iter	Val_acc
Values	M = 300	1	lbfgs	1000	66.82	1	lbfgs	1000	57.98
		5	lbfgs	1000	71.51	10	lbfgs	1000	61.79
		10	lbfgs	1000	71.58	15	lbfgs	1000	62.35
		15	lbfgs	1000	71.04	17	lbfgs	1000	62.4
		20	lbfgs	1000	70.41	19	lbfgs	1000	62.45
	M = 500	1	lbfgs	1000	66.82	1	lbfgs	1000	57.77
		5	lbfgs	1000	71.51	5	lbfgs	1000	65.11
		10	lbfgs	1000	71.58	10	lbfgs	1000	66.82
		15	lbfgs	1000	71.04	15	lbfgs	1000	67.29
		20	lbfgs	1000	70.41	20	lbfgs	1000	68.31
		40	lbfgs	1000	70.88	40	lbfgs	1000	68.07
		70	lbfgs	1000	70.02	70	lbfgs	1000	68.31

Table A2. Comparison and adjustment of parameters related to the support vector machine algorithm.

Classifier		SVM							
Hyperparameters	Bin Length	L = Triple				L = Quintuple			
		C	Solver	Max_Iter	Val_acc	C	Kernel	Max_Iter	Val_acc
Values	M = 300	10	rbf	2000	86.18	10	rbf	2000	82.64
		15	rbf	2000	83.29	15	rbf	2000	82.78
		12	rbf	2000	82.92	20	rbf	2000	82.73
	M = 500	1	rbf	2000	77.52	1	rbf	2000	79.39
		5	rbf	2000	86.42	5	rbf	2000	86.1
		10	rbf	2000	86.18	10	rbf	2000	85.64
		15	rbf	2000	85.79	15	rbf	2000	85.25
		20	rbf	2000	86.03	20	rbf	2000	85.25
		30	rbf	2000	86.1	30	rbf	2000	85.25

Table A3. Comparison and adjustment of parameters related to deep neural network algorithm.

Feed-Forward Neural Network													
Hyperparameters	Bin Length	L = Triple						L = Quintuple					
		Dense Units	Activation	Optimizer	Batch Size	Epochs	Val_acc	Dense Units	Activation	Optimizer	Batch Size	Epochs	Val_acc
Values	300	1024-512-256-128	relu	rms	8	32	81.69	1024-512-256-128	relu	rms	16	17	82.73
		512-256-128-64	relu	adam	8	29	83.11	1024-512-512-256	relu	adam	16	42	80.14
		512-256-128-64	relu	adam	16	45	82.92	512-256-64	relu	rms	16	19	81.08
								1024-512-256-128	relu	rms	64	19	84.85
								1024-512-256-128	relu	rms	128	21	84.8
								1024-512-256-128	relu	rms	32	15	84.33
	500	512-256-128-64	relu	adam	8	14	82.12	512-256-128-64	relu	rms	64	42	87.2
		512-256-64	relu	rms	8	24	84	256-128-64	relu	rms	32	18	81.03
		1024-512-128	relu	rms	8	17	83.22	1024-512-256-128	relu	rms	64	25	86.42
		512-256-128	relu	rms	16	24	85.25	512-256-128-64	relu	adam	32	61	82.9
		512-256-256	relu	rms	32	20	86.18	512-256-128	relu	rms	64	42	87.35
		512-256-256	relu	rms	64	23	86.96						
		512-256-256	relu	rms	128	57	89. 31						
		512-256-256	relu	rms	256	69	89.7						
		512-256-256	relu	rms	512	68	88.29						

Table A4. Comparison and adjustment of parameters related to CNN algorithm.

	CNN																	
	Triple									**Quintuple**								
Bin Length	**Filters**	**Kernel Size**	**Pool Size**	**Dense Units**	**Activation**	**Optimizer**	**Batch Size**	**Epochs**	**Val_acc**	**Filters**	**Kernel Size**	**Pool Size**	**Dense Units**	**Activation**	**Optimizer**	**Batch Size**	**Epochs**	**Val_acc**
300	32-64-128	3	2	128	Relu	adam	8	29	79	128-256-512	3	2	512	relu	rms	16	16	80
	64-128-256	3	2	256	Relu	adam	8	25	82	64-128-256	3	2	256	relu	rms	8	26	77
	16-32-64-128	3	2	128	Relu	adam	8	27	77	64-64-128	3	2	128	relu	rms	16	20	77
										32-64-128	3	2	128	relu	rms	64	15	77
500	32-64-128	3	2	128	Relu	adam	64	29	81	64-128	3	2	128	relu	adam	32	20	77
	32-32-64	3	2	64	Relu	adam	128	33	79	32-64-128	3	2	128	relu	adam	64	19	75
	32-64	3	2	64	Relu	adam	32	43	82	16-32-64	3	2	64	relu	adam	32	28	73
	64-128	3	2	128	Relu	adam	32	21	84	64-128-256	3	2	256	relu	adam	64	30	80
	128-256	3	2	256	Relu	adam	32	35	83	64-128-128	3	2	128	relu	adam	64	23	77

Appendix B. Confusion Matrices of the Tuned Algorithms

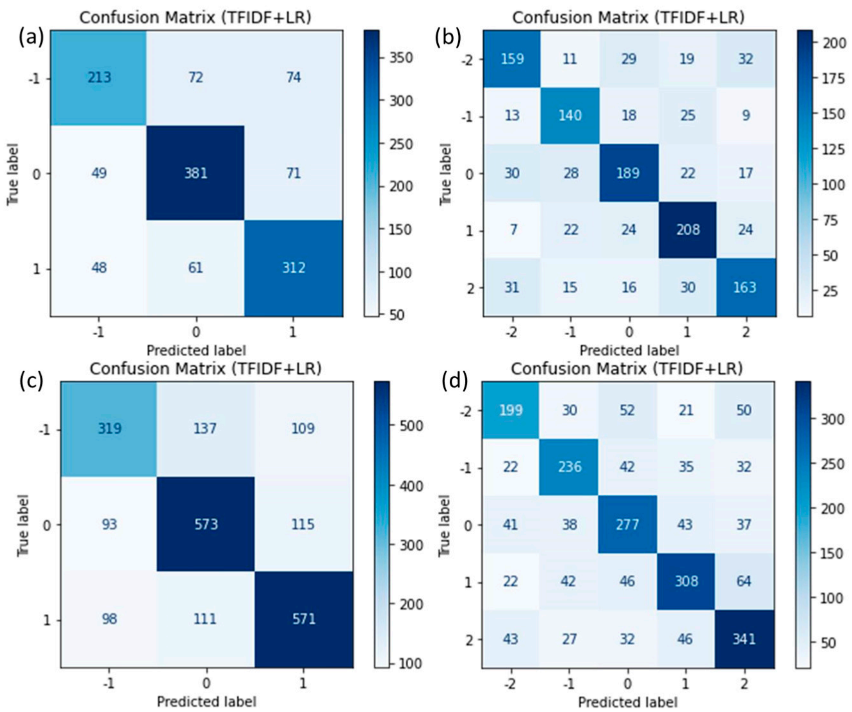


Figure A1. Confusion matrix result of logistic regression method. (**a**) Confusion matrix obtained from 3-class l = 3, combination type m = 500, and c = 10; (**b**) confusion matrix obtained from l = 5, m = 500, and c = 20; (**c**) confusion matrix resulting from l = 3, m = 300, and c = 10; (**d**) confusion matrix resulting from l = 5, m = 300, and c = 17.

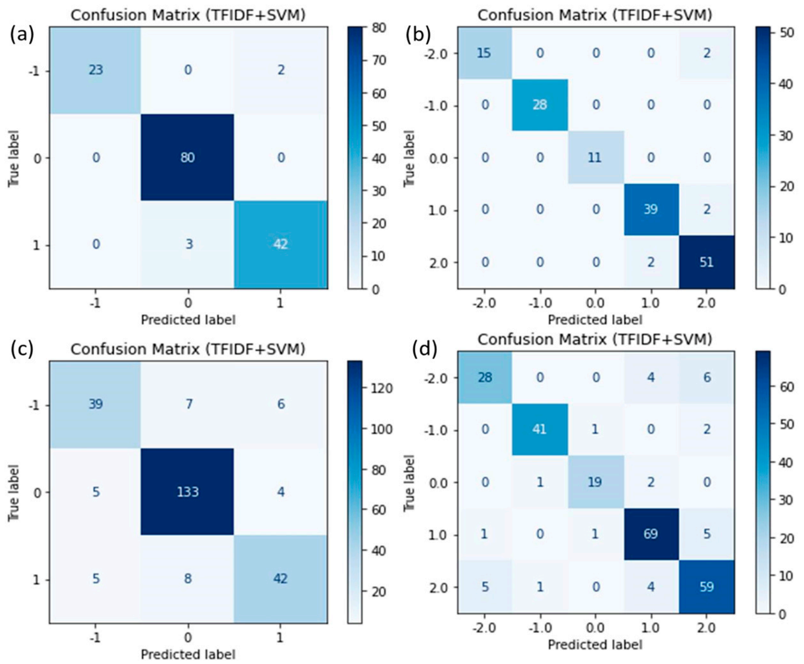


Figure A2. Confusion matrix results of the support vector machine method. (**a**) Confusion matrix resulting from 3-class l = 3, combination type m = 500, and c = 5; (**b**) confusion matrix resulting from l = 5, m = 500, and c = 5; (**c**) confusion matrix resulting from l = 3, m = 300, and c = 5; (**d**) confusion matrix resulting from l = 5, m = 300, and c = 15.

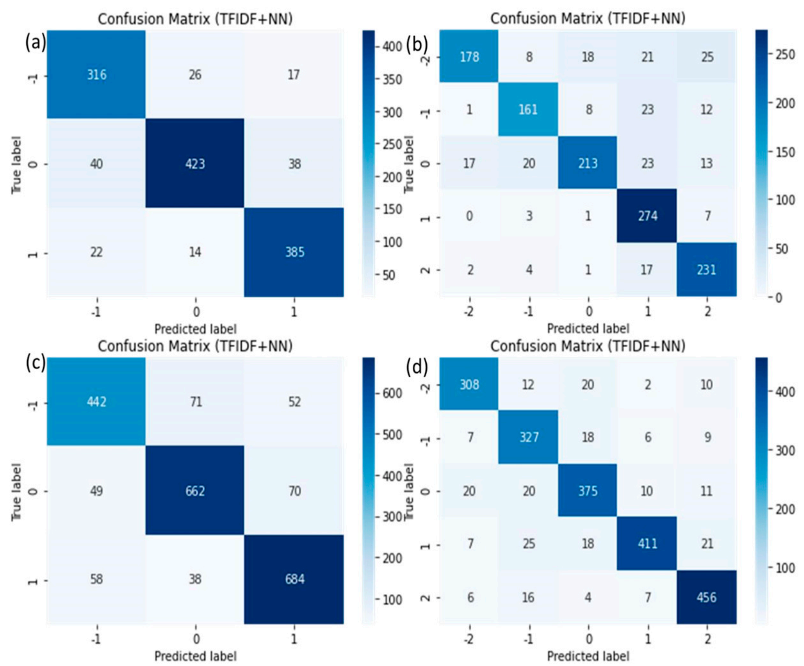


Figure A3. Confusion matrix results of deep feed-forward method. (**a**) Confusion matrix resulting from 3-class l = 3, type of combination m = 500; (**b**) confusion matrix resulting from l = 5, m = 500; (**c**) confusion matrix resulting from l = 3, m = 300; (**d**) confusion matrix resulting from l = 5, m = 300.

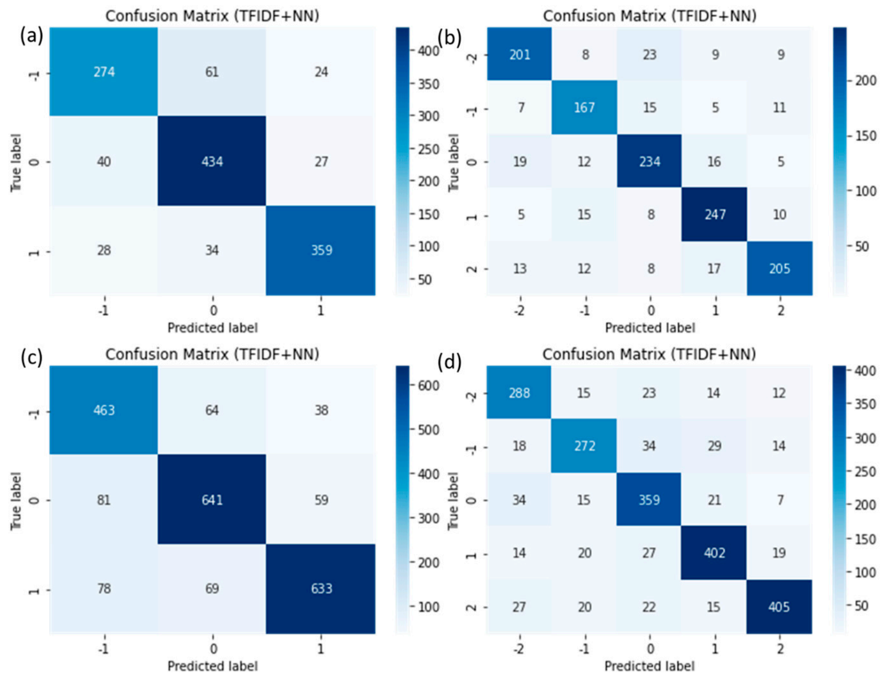


Figure A4. Confusion matrix results of convolutional neural network method. (**a**) Confusion matrix resulting from 3-class l = 3, type of combination m = 500; (**b**) confusion matrix resulting from l = 5, m = 500; (**c**) confusion matrix resulting from l = 3, m = 300; (**d**) confusion matrix resulting from l = 5, m = 300. Results of the test data using the CNN algorithm.

Notes

1 Natural language processing.

2 Term frequency inverse document frequency.

References

Adhikari, Surabhi, Surendrabikram Thapa, Usman Naseem, Hai Ya Lu, Gnana Bharathy, and Mukesh Prasad. 2023. Explainable Hybrid Word Representations for Sentiment Analysis of Financial News. *Neural Networks* 164: 115–23. [CrossRef] [PubMed]

Ahmad, Farooq, and Nasir Rasool. 2017. Determinants of Bank Liquidity: Empirical Evidence from Listed Commercial Banks with SBP. *Journal of Economics and Sustainable Development* 8: 47–55.

Alharbi, Ahmad. 2017. Determinants of Banks Liquidity: Evidence from OIC Countries. *Journal of Economic and Administrative Sciences* 33: 164–77. [CrossRef]

An, Chang-Ho. 2017. A Study on Estimation of Financial Liquidity Risk Prediction Model Using Financial Analysis. Available online: https://api.semanticscholar.org/CorpusID:173987578 (accessed on 1 January 2018).

Antweiler, Werner, and Murray Frank. 2004. Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. *Journal of Finance* 59: 1259–94. [CrossRef]

Araci, Dogu. 2019. FinBERT: Financial Sentiment Analysis with Pre-Trained Language Models. *arXiv* arXiv:1908.10063. Available online: https://api.semanticscholar.org/CorpusID:201646244 (accessed on 1 January 2020).

Arif, Ahmed, and Ahmed Nauman Anees. 2012. Liquidity Risk and Performance of Banking System. *Journal of Financial Regulation and Compliance* 20: 182–95. [CrossRef]

Athanasoglou, Panayiotis, Sophocles Brissimis, and Matthaios Delis. 2008. Bank-Specific, Industry-Specific and Macroeconomic Determinants of Bank Profitability. *Journal of International Financial Markets, Institutions and Money* 18: 121–36. [CrossRef]

Basel Committee on Banking Supervision. 2008. *Principles for Sound Liquidity Risk Management and Supervision*. Basel: BIS.org, September.

Basel Committee on Banking Supervision. 2010. *Basel III: International Framework for Liquidity Risk Measurement Standards and Monitoring*. Basel: BIS.org.

Bissoondoyal-Bheenick, Emawtee, and Sirimon Treepongkaruna. 2011. An Analysis of the Determinants of Bank Ratings: Comparison across Ratings Agencies. *Australian Journal of Management* 36: 405–24. [CrossRef]

Boguth, Oliver, Murray Carlson, Adlai Fisher, and Mikhail Simutin. 2016. Horizon Effects in Average Returns: The Role of Slow Information Diffusion. *Review of Financial Studies* 29: 2241–81. [CrossRef]

Butler, Matthew, and Vlado Kešelj. 2009. Financial Forecasting Using Character N-Gram Analysis and Readability Scores of Annual Reports. In *Advances in Artificial Intelligence*. Edited by Yong Gao and Nathalie Japkowicz. Berlin/Heidelberg: Springer, pp. 39–51. [CrossRef]

Calabrese, Raffaella, and Paolo Giudici. 2015. Estimating bank default with generalised extreme value regression models. *Journal of the Operational Research Society* 66: 1783–92. [CrossRef]

Campbell, John, and Robert Shiller. 1986. *Cointegration and Tests of Present Value Models*. Working Paper Series, Working Paper No. 1885. Cambridge: National Bureau of Economic Research, April. [CrossRef]

Chen, Chien Chin, and You-De Tseng. 2011. Quality Evaluation of Product Reviews Using an Information Quality Framework. *Decision Support Systems* 50: 755–68. [CrossRef]

Das, Sanjiv, and Mike Chen. 2007. Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. *Management Science* 53: 1375–88. [CrossRef]

Diamond, Douglas, and Philip Dybvig. 1983. Bank Runs, Deposit Insurance, and Liquidity. *Journal of Political Economy* 91: 401–19. Available online: http://www.jstor.org/stable/1837095 (accessed on 12 October 2024). [CrossRef]

Do, Hai Ha, PWC Prasad, Angelika Maag, and Abeer Alsadoon. 2019. Deep Learning for Aspect-Based Sentiment Analysis: A Comparative Review. *Expert Systems with Applications* 118: 272–99. [CrossRef]

Dogra, Varun, Aman Singh, Kavita Verma, Noor Jhanjhi, and Maria Talib. 2021. Analyzing DistilBERT for Sentiment Classification of Banking Financial News. Paper presented at 2021 International Conference on Advances in Computing, Communications and Informatics, Kochi, India, October 21–23; Singapore: Springer, pp. 501–10. [CrossRef]

Doumpos, Michael, Chrysovalantis Gaganis, and Fotios Pasiouras. 2016. Bank Diversification and Overall Financial Strength: International Evidence. *Financial Markets, Institutions & Instruments* 25: 169–213. [CrossRef]

Drehmann, Mathias, and Kleopatra Nikolaou. 2013. Funding Liquidity Risk: Definition and Measurement. *Journal of Banking & Finance* 37: 2173–82. Available online: https://www.sciencedirect.com/science/article/pii/S0378426612000039 (accessed on 1 January 2021). [CrossRef]

Elsas, Ralf, Andreas Hackethal, and Markus Holzhäuser. 2010. The anatomy of bank diversification. *Journal of Banking & Finance* 34: 1274–87. [CrossRef]

Erlwein-Sayer, Christina. 2018. Macroeconomic News Sentiment: Enhanced Risk Assessment for Sovereign Bonds. *Risks* 6: 141. [CrossRef]

Erlwein-Sayer, Christina, and Xiang Yu. 2020. Sentiment Analysis for Credit Risk and Portfolio Construction. Available online: https://bigdatafinance.eu/wp/wp-content/uploads/2017/07/Sentiment-Analysis-for-Credit-Risk-and-Portfolio-Construction.pdf (accessed on 1 January 2021).

Estahbanati, Somayye, and Reza Javidan. 2011. A New Stemmer for Farsi Language. Paper presented at 2011 CSI International Symposium on Computer Science and Software Engineering (CSSE), Tehran, Iran, June 15–16; pp. 25–29. [CrossRef]

Feki, Asma, Anis Ben Ishak, and Saber Feki. 2012. Feature Selection Using Bayesian and Multiclass Support Vector Machines Approaches: Application to Bank Risk Prediction. *Expert Systems with Applications* 39: 3087–99. [CrossRef]

Fritz, David, and Eugen Tőws. 2018. Text Mining and Reporting Quality in German Banks—A Cooccurrence and Sentiment Analysis. *Universal Journal of Accounting and Finance* 6: 54–81. [CrossRef]

Gandhi, Priyank, Tim Loughran, and Bill McDonald. 2019. Using Annual Report Sentiment as a Proxy for Financial Distress in U.S. Banks. *Journal of Behavioral Finance* 20: 424–36. [CrossRef]

Gouvêa, Maria, and Eric Bacconi. 2021. Credit Risk Analysis Applying Logistic Regression, Neural Networks and Genetic Algorithms Models. *International Journal of Advanced Engineering Research and Science* 8: 198–209. [CrossRef]

Groth, Sven, and Jan Muntermann. 2011. An Intraday Market Risk Management Approach Based on Textual Analysis. *Decision Support Systems* 50: 680–91. [CrossRef]

Guijarro, Francisco, Ismael Moya-Clemente, and Jawad Saleemi. 2021. Market Liquidity and Its Dimensions: Linking the Liquidity Dimensions to Sentiment Analysis through Microblogging Data. *Journal of Risk and Financial Management* 14: 394. [CrossRef]

Hagenau, Michael, Michael Liebmann, and Dirk Neumann. 2013. Automated News Reading: Stock Price Prediction Based on Financial News Using Context-Capturing Features. *Decision Support Systems* 55: 685–97. [CrossRef]

Hanley, James, and Barbara Mcneil. 1982. The Meaning and Use of the Area Under a Receiver Operating Characteristic (ROC) Curve. *Radiology* 143: 29–36. [CrossRef] [PubMed]

Hassan, Kabir, Ashraf Khan, and Andrea Paltrinieri. 2018. Liquidity Risk, Credit Risk and Stability in Islamic and Conventional Banks. *Research in International Business and Finance* 48: 17–31. [CrossRef]

Hájek, Petr, and Vladimír Olej. 2013. Evaluating Sentiment in Annual Reports for Financial Distress Prediction Using Neural Networks and Support Vector Machines. *Communications in Computer and Information Science* 384: 1–10. [CrossRef]

Huang, Jiexiang, Helen Roberts, and Eric Tan. 2018. The Impact of Media Sentiment on Firm Risk, Corporate Investment and Financial Policies. DecisionSciRN: Leadership & Decision-Making (Topic). Available online: https://api.semanticscholar.org/CorpusID:169821512 (accessed on 1 January 2018).

Jarrow, Robert, and Donald van Deventer. 1998. The Arbitrage-Free Valuation and Hedging of Demand Deposits and Credit Card Loans. *Journal of Banking & Finance* 22: 249–72. [CrossRef]

Jiang, Kai, and Xi Lu. 2020. Natural Language Processing and Its Applications in Machine Translation: A Diachronic Review. Paper presented at 2020 IEEE 3rd International Conference of Safe Production and Informatization (IICSPI), Chongqing, China, November 28–30; pp. 210–14. [CrossRef]

Jobst, Andreas. 2014. Measuring Systemic Risk-Adjusted Liquidity (SRL)—A Model Approach. *Journal of Banking & Finance* 45: 270–87. [CrossRef]

Juba, Brendan, and Hai Le. 2019. Precision-Recall versus Accuracy and the Role of Large Data Sets. *Proceedings of the AAAI Conference on Artificial Intelligence* 33: 4039–48. [CrossRef]

Khan, Md, Md Sumon Ali, Md Naiem Hossain, and Mithun Bairagi. 2023. Determinants of Non-performing Loans in Conventional and Islamic Banks: Emerging Market Evidence. *Modern Finance* 1: 56–69. [CrossRef]

Khurana, Diksha, Aditya Koli, Kiran Khatter, and Sukhdev Singh. 2023. Natural Language Processing: State of the Art, Current Trends, and Challenges. *Multimedia Tools and Applications* 82: 3713–44. [CrossRef]

Khyani, Divya, and Siddhartha. 2021. An Interpretation of Lemmatization and Stemming in Natural Language Processing. *Shanghai Ligong Daxue Xuebao/Journal of University of Shanghai for Science and Technology* 22: 350–57.

Kim, Hannah, and Young-Seob Jeong. 2019. Sentiment Classification Using Convolutional Neural Networks. *Applied Sciences* 9: 2347. [CrossRef]

Kogan, Shimon, Dimitry Levin, Bryan Routledge, Jacob Sagi, and Noah Smith. 2009. Predicting Risk from Financial Reports with Regression. Paper presented at Conference on Empirical Methods in Natural Language Processing (EMNLP), Singapore, August 6–7; pp. 272–80. [CrossRef]

Köhler, Matthias. 2015. Which banks are more risky? The impact of business models on bank stability. *Journal of Financial Stability* 16: 195–212. [CrossRef]

Kraus, Mathias, and Stefan Feuerriegel. 2017. Decision Support from Financial Disclosures with Deep Neural Networks and Transfer Learning. *Decision Support Systems* 104: 38–48. [CrossRef]

Kumar, Manish, and Ghanshyam Chand Yadav. 2013. Liquidity Risk Management in Bank: A Conceptual Framework. *Global Journal of Engineering Science and Research Management* 7: 3.

Lanine, Gleb, and Rudi Vander Vennet. 2006. Failure Prediction in the Russian Bank Sector with Logit and Trait Recognition Models. *Expert Systems with Applications* 30: 463–78. [CrossRef]

Lasko, Thomas, Jui Bhagwat, Kelly Zou, and Lucila Ohno-Machado. 2005. The Use of Receiver Operating Characteristic Curves in Biomedical Informatics. *Journal of Biomedical Informatics* 38: 404–15. [CrossRef]

LeCun, Yann, Bengio, and Geoffrey Hinton. 2015. Deep Learning. *Nature* 521: 436–44. [CrossRef]

Li, Feng. 2008. The Determinants and Information Content of the Forward-Looking Statements in Corporate Filings—A Naïve Bayesian Machine Learning Approach. *Journal of Accounting Research* 48: 1049–102. [CrossRef]

Liu, Ning, and Jianhua Zhao. 2022. A BERT-Based Aspect-Level Sentiment Analysis Algorithm for Cross-Domain Text. *Computational Intelligence and Neuroscience* 2022: 8726621. [CrossRef]

MacKinlay, Craig. 1997. Event Studies in Economics and Finance. *Journal of Economic Literature* 35: 13–39.

Manthoulis, Georgios, Michalis Doumpos, Constantin Zopounidis, and Emilios Galariotis. 2020. An ordinal classification framework for bank failure prediction: Methodology and empirical evidence for US banks. *European Journal of Operational Research* 282: 786–801. [CrossRef]

Manzira, Francis, and Felix Bankole. 2018. Application of Social Media Analytics in the Banking Sector to Drive Growth and Sustainability: A Proposed Integrated Framework. Paper presented at 2018 Open Innovations Conference (OI), Johannesburg, South Africa, October 3–5; pp. 223–33. [CrossRef]

Matz, Leonard. 2006. Scenario Analysis and Stress Testing. In *Liquidity Risk Measurement and Management*. Hoboken: John Wiley & Sons, Ltd., pp. 37–63. [CrossRef]

Mirashk, Hamed, Amir Albadvi, and Mohammad Ali Rastegar. 2023. Liquidity Risk Prediction Using News Sentiment Analysis. *Journal of Information and Communication Technology* 55: 258–73.

Mittermayer, Marc-Anton. 2004. Forecasting intraday stock price trends with text mining techniques. Paper presented at 37th Annual Hawaii International Conference on System Sciences, Big Island, HI, USA, January 5–8. [CrossRef]

Mohammad, Sabri, Mehmet Asutay, Rob Dixon, and Elena Platonova. 2020. Liquidity Risk Exposure and Its Determinants in the Banking Sector: A Comparative Analysis between Islamic, Conventional and Hybrid Banks. *Journal of International Financial Markets, Institutions and Money* 66: 101196. [CrossRef]

Mohanty, Aniruddha, and Ravindranath Cherukuri. 2023. Sentiment Analysis on Banking Feedback and News Data Using Synonyms and Antonyms. *International Journal of Advanced Computer Science and Applications* 14: 934–44. [CrossRef]

Molyneux, Philip, and John Thornton. 1992. Determinants of European Bank Profitability: A Note. *Journal of Banking & Finance* 16: 1173–78. [CrossRef]

Musakwa, Fidelis. 2013. Measuring Bank Funding Liquidity Risk. University of the Witwatersrand, April 30. Available online: https://www.actuaries.org/lyon2013/papers/AFIR_Musakwa.pdf (accessed on 30 April 2023).

Nam, KiHwan, and NohYoon Seong. 2019. Financial News-Based Stock Movement Prediction Using Causality Analysis of Influence in the Korean Stock Market. *Decision Support Systems* 117: 100–12. [CrossRef]

Nguyen, Thien Hai, Kiyoaki Shirai, and Julien Velcin. 2015. Sentiment Analysis on Social Media for Stock Movement Prediction. *Expert Systems with Applications* 42: 9603–11. [CrossRef]

Nopp, Clemens, and Allan Hanbury. 2015. Detecting Risks in the Banking System by Sentiment Analysis. Paper presented at 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP), Lisbon, Portugal, September 17–21; pp. 591–600. [CrossRef]

Ormerod, Paul, Rickard Nyman, and David Tuckett. 2015. Measuring Financial Sentiment to Predict Financial Instability: A New Approach Based on Text Analysis. *arXiv* arXiv:1508.05357.

Paraboni, Ana, Marcelo Righi, Kelmara Vieira, and Vinícius Silveira. 2018. The Relationship between Sentiment and Risk in Financial Markets. *BAR—Brazilian Administration Review* 15: e170055. [CrossRef]

Pham, Hai Nam, and Thi Hong Nhung Pham. 2022. Determinants of bank liquidity: Evidence from Vietnam. In *Prediction and Causality in Econometrics and Related Topics*. Edited by Nguyen Ngoc Thach, Doan Thanh Ha, Nguyen Duc Trung and Vladik Kreinovich. Cham: Springer International Publishing, pp. 185–96. [CrossRef]

Pham, Xuan, and Tin Ho. 2021. Using Boosting Algorithms to Predict Bank Failure: An Untold Story. *International Review of Economics & Finance* 76: 40–54. [CrossRef]

Porter, Martin. 1997. An Algorithm for Suffix Stripping. *Program* 40: 211–18. [CrossRef]

Radovanov, Boris, Nada Milenković, Branimir Kalaš, and Aleksandra Marcikić Horvat. 2023. Do the same determinants affect banks' profitability and liquidity? Evidence from West Balkan countries using a panel data regression analysis. *Mathematics* 11: 4072. [CrossRef]

Read, Jonathon. 2005. Using Emoticons to Reduce Dependency in Machine Learning Techniques for Sentiment Classification. In Annual Meeting of the Association for Computational Linguistics. Available online: https://api.semanticscholar.org/CorpusID:12596803 (accessed on 1 July 2005).

Roeder, Jan, Matthias Palmer, and Jan Muntermann. 2022. Data-Driven Decision-Making in Credit Risk Management: The Information Value of Analyst Reports. *Decision Support Systems* 158: 113770. [CrossRef]

Rostami, Malihe. 2015. CAMELS' Analysis in Banking Industry. *Global Journal of Engineering Science and Research Management* 2: 10–26.

Schlueter, Tobias, Thomas Hartmann-Wendels, Tim Weber, and MSc Zander. 2014. Die Risikoberichterstattung Deutscher Banken: Erhebung des Branchenstandards. *Schmalenbachs Zeitschrift für Betriebswirtschaftliche Forschung* 66: 386–427. [CrossRef]

Schumaker, Robert, and Hsinchun Chen. 2009. Textual Analysis of Stock Market Prediction Using Breaking Financial News: The AZFin Text System. *ACM Transactions on Information Systems* 27: 12. [CrossRef]

Shao, Chen, and Xiaochen Chen. 2022. Deep-Learning-Based Financial Message Sentiment Classification in Business Management. *Computational Intelligence and Neuroscience* 2022: 3888675. [CrossRef]

Singh, Anamika, and Anil Kumar Sharma. 2016. An Empirical Analysis of Macroeconomic and Bank-Specific Factors Affecting Liquidity of Indian Banks. *Future Business Journal* 2: 40–53. [CrossRef]

Sokolova, Marina, Nathalie Japkowicz, and Stan Szpakowicz. 2006. Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. In *AI 2006: Advances in Artificial Intelligence, Lecture Notes in Computer Science*. Berlin: Springer, vol. 4304, pp. 1015–21. [CrossRef]

Solangi, Yasir Ali, Zulfiqar Ali Solangi, Samreen Aarain, Amna Abro, Ghulam Ali Mallah, and Asadullah Shah. 2018. Review on Natural Language Processing (NLP) and Its Toolkits for Opinion Mining and Sentiment Analysis. Paper presented at 2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences (ICETAS), Bangkok, Thailand, November 22–23; pp. 1–4. [CrossRef]

Taghva, Kazem, Russell Beckley, and Mohammad Sadeh. 2005. A Stemming Algorithm for the Farsi Language. Paper presented at International Conference on Information Technology: Coding and Computing (ITCC'05)—Volume II, Las Vegas, NV, USA, April 4–6, vol. 1, pp. 158–62. [CrossRef]

Tasca, Paolo, Stefano Battiston, and Andrea Deghi. 2017. Portfolio diversification and systemic risk in interbank networks. *Journal of Economic Dynamics and Control* 82: 96–124. [CrossRef]

Tasnova, Nishat. 2022. Impact of bank-specific and macroeconomic determinants on banks' liquidity: An empirical study on listed commercial banks in Bangladesh. *Finance & Economics Review* 4: 11–24. [CrossRef]

Tavana, Madjid, Amir-Reza Abtahi, Debora Di Caprio, and Maryam Poortarigh. 2018. An Artificial Neural Network and Bayesian Network model for liquidity risk assessment in banking. *Neurocomputing* 275: 2525–54. [CrossRef]

Tetlock, Paul. 2007. Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *Journal of Finance* 62: 1139–68. [CrossRef]

Thakkar, Ankit, and Kinjal Chaudhari. 2021. A Comprehensive Survey on Deep Neural Networks for Stock Market: The Need, Challenges, and Future Directions. *Expert Systems with Applications* 177: 114800. [CrossRef]

Vapnik, Vladimir. 1995. *The Nature of Statistical Learning Theory*. New York: Springer.

Wei, Lu, Guowen Li, Xiaoqian Zhu, and Jianping Li. 2019. Discovering Bank Risk Factors from Financial Statements Based on a New Semi-Supervised Text Mining Algorithm. *Accounting & Finance* 59: 1519–52. [CrossRef]

Wiebe, Janyce, Theresa Wilson, and Matthew Bell. 2001. Identifying Collocations for Recognizing Opinions. Paper presented at ACL-01 Workshop on Collocation: Computational Extraction, Analysis, and Exploitation, Toulouse, France, July 7; pp. 24–31.

Wójcik-Mazur, Agnieszka, and Marek Szajt. 2015. Determinants of Liquidity Risk in Commercial Banks in the European Union. *Argumenta Oeconomica* 35: 25–47. [CrossRef]

Wüthrich, Beat, Vincent Cho, Steven Leung, D. Permunetilleke, Kavitha Sankaran, and Jian Zhang. 1998. Daily Stock Market Forecast from Textual Web Data. Paper presented at SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics, San Diego, CA, USA, August 6, vol. vol. 3, pp. 2720–25.

Yuan, Zhigang, Sixing Wu, Fangzhao Wu, Junxin Liu, and Yongfeng Huang. 2018. Domain Attention Model for Multi-Domain Sentiment Classification. *Knowledge-Based Systems* 155: 1–10. [CrossRef]

Zarei, Ajmad. 2016. Measurement of credit risk portfolio of bank by method of coefficient variation bound. *Sains Malaysiana* 11: 3908–13.

Zhang, Xiang, Junbo Zhao, and Yann LeCun. 2015. Character-Level Convolutional Networks for Text Classification. In *Advances in Neural Information Processing Systems*. vol. 28, Conference on Neural Information Processing Systems (NeurIPS), Montreal, QC, Canada. Available online: https://proceedings.neurips.cc/paper_files/paper/2015/file/250cf8b51c773f3f8dc8b4be867a9a02-Paper.pdf (accessed on 1 March 2015).

Zhang, Yihao, and Xiaoyang Liu. 2021. Learning Attention Embeddings Based on Memory Networks for Neural Collaborative Recommendation. *Expert Systems with Applications* 183: 115439. [CrossRef]

Zweig, Mark H., and Gregory Campbell. 1993. Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clinical Chemistry* 39: 561–77. [CrossRef]