Modern Economy, 2026, 17(9), 1144-1158
https://www.scirp.org/journal/me
ISSN Online: 2152-7261
ISSN Print: 2152-7245

More Positive Sentiment, Less Earnings? A Sentiment-Based Explanatory Modeling in Banking Earnings—A Case Study of Banks from China and the US

Ting Hu[1], **Jun Li**[2]

[1]School of Foreign Language, Shanghai Institute of Technology, Shanghai, China
[2]China Construction Bank Corporation, Shanghai, China
Email: huting@sit.edu.cn

How to cite this paper: Hu, T., & Li, J. (2026). More Positive Sentiment, Less Earnings? A Sentiment-Based Explanatory Modeling in Banking Earnings—A Case study of Banks from China and the US. *Modern Economy, 17,* 1144-1158.
https://doi.org/10.4236/me.2026.179056

Received: May 27, 2026
Accepted: September 15, 2026
Published: September 18, 2026

Abstract

This paper explores simpler models for a more sustainable approach to financial risk prediction once correlations between sentiment and performance have been established. Empirical evidence from an investigation of annual reports of eight banks from China and the United States using Python demonstrates significant relationships existed in the language and banks' key performance metrics. Across both countries, positive word counts correlated negatively with ROE, while negative word counts correlated positively with EPS. But divergences emerged in sentiment proportion: Chinese banks showed a positive correlation with ROE that was absent in U.S. banks. Correlations between positive language and ROE and ROE difference appeared mainly in years when ROE and EPS declined from the previous year in all banks. A country-specific nuance was also identified where U.S. banks employed more positive expressions when experiencing a significant decline in ROE, a pattern not statistically significant in Chinese banks. A one-layer neural network explanatory model with a linear equation was then composed, illustrating that simpler shallow models with small data set, supported by empirical discourse analysis, may represent a more viable long-term strategy for sentiment-based bank performance evaluation. Sustainability can not only be framed in environmental terms but also in regulatory acceptance and institutional usability.

Keywords

Sentiment Analysis, Bank Performance, ROE, EPS, Shallow Models

1. Introduction

Annual reports act as a comprehensive communication tools between banks and their stakeholders informed by the audited financial statements, management discussions, risk disclosures, and strategic outlooks for transparency and decision-making. Analyzing banks' annual reports can update their financial stability, bank value creation and potential exposure to risk. Researchers have illustrated the correlation between banks' profitability, liquidity, and loan portfolio quality with disclosure and transparency (Oino, 2019), better stock returns with greater revelation and more fragility and systemic risk with lack of transparency. Beyond the determinant of banks' performance, annual reports has emerged as insights into the market discipline. Risk disclosure in annual reports can forecast market efficiency and financial system stability (Scannella & Polizzi, 2018) and facilitate regulators to take proactive actions before irreversible risks that destabilize economy appear (Bischof, Laux, & Leuz, 2021). However, disclosure is not solely dependent on the intention and initiatives of banks, but also on the readability and quality of disclosures. Evidence from the UK banks during 1995 to 2010 shows the move from quantitative disclosure to rich qualitative narratives, signaling a shift towards perception-driven communication, a result of scaling transparency with public trust (Rattanataipop, 2013). When bank performance is weak, the readability of annual report de-clines, with less concise and foggier expressions to increase complexity and obscure vulnerabilities (Mocanu, Grose, & Kargidis, 2019). Inadequate risk disclosures reduce investor confidence (Zheng, Sarker, & Nahar, 2018). Then it is not just what is unveiled in the annual report but how it is communicated that deserves equal analysis.

A variety of text analysis has been conducted aiming to uncover the real evaluation of transparency, compliance, and market signaling from banks' annual reports. Content analysis is used to explore whether the text can be used to predict future fluctuation (Kravet & Muslu, 2013) and how risk disclosures are responsible for investor perception (Aram & Soroushyar, 2024), though the appearance of terms is the main indicator. Readability indices using Flesch Reading Ease, Fog Index, and Flesch-Kincaid Grade Level are employed to assess the obscurity of financial reports based on standard readability metrics, assuming the positive relationship between language complexity and blurring with lower clarity (Yang, Chang, & Mo, 2022), except that banking jargon may reduce readability from the first place and companies may purposefully improve readability without adding contents. Keywords are counted with the premise that priorities can be seen through measuring the frequency of specific terms (Li, 2010) and potential opportunities and risks can be reflected (Kravet & Muslu, 2013), ignoring context-blind textural analysis without considering sentiment or tone may mislead the public and disguise risk un-covering. While these conventional text analysis methods advanced our understanding of the banks' annual report, limitations accompanied.

The growing availability of large-scale quantitative methods featuring Python and statistical approaches provide a new perspective that offer greater scalability and depth. Hundreds or even thousands of annual reports across time and tertiary boundaries can be processed and analyzed with the aid of automated methods, making longitudinal and cross-sectional comparisons possible and easier (Li, 2010). Some nuanced differences in tone can be measured to survey positivity, negativity, or neutrality of the text with use of sentiment dictionaries, as tone is of equal importance in shaping stakeholder perceptions as contents (Kravet & Muslu, 2013). The shortcomings of traditional methods can be addressed by integrating computational approaches. Alduais (2024) uses Python to analyze the annual reports of Chinese listed companies from 2008 to 2021 and to examine the link between report readability and corporate performance. A significant correlation is found: firms with more readable, concise annual reports tend to be more profitable and have lower agency costs while poor-performing companies produce longer, more complex reports with frequent accounting jargon. Gao and Feng (2024) investigated how the positive tone of annual reports influences corporate green innovation among Shanghai and Shenzhen A-share listed Chinese companies from 2010 to 2022 with 34,830 observations. A positive effect on both the quantity and quality of a firm's green innovation is found due to the easing of financial constraints and enhanced market attention. Such effect is particularly strong for companies facing high economic policy uncertainty and for non-heavy polluting firms.

Discourse analyzing through quantitative approaches has provided a new perspective for examining companies' annual reports and certain correlations have been identified. But:

- Whether such methods are applicable across boundaries?
- Does a similar correlation between positive tone with performance exist in banks' annual report?
- Is it possible to generate a predicting model that is sustainable in terms of data, computational efficiency and energy?

With these questions, 8 banks from China and the US are investigated.

2. Materials and Methods

2.1. Data Collecting

In this study, annual reports of four banks with the highest market capitalization as of December 31, 2023 in China and the US respectively from 2018 to 2023 are collected. The 8 banks are China Construction Bank (CBB), Industrial and Commercial Bank of China (ICBC), Agricultural Bank of China (ABC), Bank of China (BOC), Bank of America (BAC), Citibank (CITI), Wells Fargo (WFC), and JP Morgan Chase (JPM). All the annual reports are in English and are downloaded from the banks official websites respectively (see **Table 1**). All the ROE and EPS data were sourced from the reports themselves. A sample of 8 banks over 6 years (N = 48) observations were collected.

Table 1. Banks, countries and abbreviations.

Abbreviation	Bank	Country
CCB	China Construction Bank	China
ICBC	Industrial and Commercial Bank of China	China
ABC	Agricultural Bank of China	China
BOC	Bank of China	China
BAC	Bank of America	US
CITI	Citibank	US
WFC	Wells Fargo	US
JPM	JP Morgan Chase	US

2.2. Data Processing

2.2.1. Tokenization

Around 6 million words from the annual reports are first tokenized in Python with NLTK (natural language processing library) to get the minimal unit a computer can process. The word_tokenize method that is applicable to any kind of text corpus and more robust than the split() method that separate words by white space as delimiter is used. To preserve all the words and digits from the string that are enormous in banks' annual report, regex_tokenize with expression of \w+ and \d+ is turned to.

2.2.2. Stemming and Lemmatization

Stemming and lemmatization are both effective ways of switching grammatical and inflected forms of word to get the root. However, their different algorithms apply to different texts. Stemming is a series of checks and transformations through examining against a set of suffix-removal rules based on patterns and word length, as shown in Formula (1). It often produces non-words as no dictionary is involved as reference, despite the fast speed and efficiency.

$$\text{Stem}(W) = f_n\left(f_{n-1}\left(\cdots f_1(W)\cdots\right)\right) \tag{1}$$

While Lemmatization combines a dictionary to look up the morphology to form a complex system to process the words, as shown in Formula (2). The process involves tokenizing, tagging POS (part of speech) and lookup lexicon to find its canonical form or lemma. It maps an inflected word and its context to a valid dictionary form while pre-serving semantic meaning.

$$\text{Lemma}(W, \text{POS}) = \text{Lookup}(\text{Dictionary}, W, \text{POS}) \tag{2}$$

Bank annual reports feature precise, technical jargon. They contain terms that must maintain their exact form to retain meaning. Stemming may lead to loss of specificity (for example, "accounting" and "accountant" into "account"), over-stemming ("management" could become "manag") and inaccuracy. In this study, all the annual reports are processed with lemmatization to preserve meaning and maintain interpretability with higher accuracy.

2.2.3. Stop Word Removal

Stop words are to be removed to reduce noise to focus on key term. Words including "the", "is" or "and" and "in" are removed to reduce model dimension and improve performance. But some words like "will", "not", "never" that are common and should be removed in texts from general fields should be retained in financial documents, as they are key in forward-looking statements or critical for sentiment. Thus, we get the tokens of these annual reports as shown in Table 2.

Table 2. Tokens of each bank.

Bank	Token
CCB	619,914
ICBC	593,102
ABC	933,700
BOC	604,562
BAC	622,271
CITI	831,624
WFC	716,471
JPM	1,224,059

2.2.4. Sentiment Calculation

The NLTK (Natural Language Toolkit) library in Python is used for sentiment classifying with the built-in VADER (Valence Aware Dictionary and Sentiment Reasoner) model. The SentimentIntensity Analyzer from the nltk.sentiment.vader module provides a score for a given text. In this study, neg, neu, pos and compound are the polarity scores to identify sentiment, with compound score larger than 0.05 being positive and less than −0.05 being negative. If compound score is between −0.05 and 0.05, it is regarded as neutral sentiment. Table 3 shows the sentimental metrics of the 8 banks.

Both the number of positive and negative words and their percentage are calculated. The number of positive and negative words is derived by classifying sentiment of the tokenized annual reports according to predefined lists of words. A higher count of positive or negative words indicates a greater presence of that sentiment. Sentiment scores and raw number of positive and negative words are available.

The percentage of positive and negative words normalizes the sentiment score by considering the length of the text. As the annual reports are of varying lengths, the percentage is crucial in comparing. Percentage-based scores allow for benchmarking across different documents and time periods. In some cases where a negative word lexicon might be much larger than a positive one, it can help mitigate inherent bias (see Formula (3), Formula (4)).

$$\text{Positive percentage} = (\text{Positive tokens})/(\text{Total tokens}) \tag{3}$$

$$\text{Negative percentage} = (\text{Negative tokens})/(\text{Total tokens}) \tag{4}$$

Table 3. Sentiment of banks.

Year_Bank	Positive Score	Negative Score	Positive word counts	Negative word counts	Percentage of positive word	Percentage of negative word
2018_ABC	0.15	0.053	7424	2811	0.0498817	0.018887
2018_BAC	0.165	0.054	8151	2871	0.0525335	0.018504
2018_BOC	0.145	0.043	7083	2393	0.0477813	0.016143
2018_CCB	0.148	0.052	7255	2750	0.0492971	0.018686
2018_CITI	0.157	0.052	10,029	3542	0.0494627	0.017469
2018_ICBC	0.15	0.048	7511	2696	0.0482839	0.017331
2018_JPM	0.101	0.045	7467	3692	0.0269194	0.01331
2018_WFC	0.174	0.06	12,032	4410	0.0578345	0.021198
2019_ABC	0.149	0.054	7535	2868	0.0497991	0.018955
2019_BAC	0.157	0.054	7637	2730	0.051198	0.018302
2019_BOC	0.145	0.042	7223	2392	0.0486031	0.016096
2019_CCB	0.154	0.05	7574	2665	0.0503577	0.017719
2019_CITI	0.147	0.05	9619	3505	0.0465942	0.016978
2019_ICBC	0.146	0.047	7451	2647	0.0475228	0.016883
2019_JPM	0.096	0.043	6323	3067	0.0257902	0.01251
2019_WFC	0.173	0.062	10,713	4091	0.0577063	0.022036
2020_ABC	0.151	0.055	7919	2938	0.0502605	0.018647
2020_BAC	0.155	0.054	8227	3056	0.0490754	0.01823
2020_BOC	0.147	0.042	7685	2505	0.0485241	0.015817
2020_CCB	0.152	0.049	7883	2768	0.0497193	0.017458
2020_CITI	0.15	0.055	10,773	4125	0.0474874	0.018183
2020_ICBC	0.143	0.047	7462	2691	0.047279	0.01705
2020_JPM	0.157	0.059	11,799	4493	0.048299	0.018392
2020_WFC	0.165	0.064	10,123	4080	0.0552409	0.022264
2021_ABC	0.152	0.05	8108	2932	0.0504216	0.018233
2021_BAC	0.164	0.054	8317	2930	0.052723	0.018574
2021_BOC	0.151	0.039	8350	2547	0.0486588	0.014842
2021_CCB	0.155	0.047	8307	2815	0.0503503	0.017062
2021_CITI	0.149	0.056	10,190	4098	0.0470857	0.018936
2021_ICBC	0.144	0.043	7311	2460	0.0472595	0.015902
2021_JPM	0.084	0.022	3699	646	0.0075544	0.001319
2021_WFC	0.167	0.064	9461	3811	0.0559372	0.022532
2022_ABC	0.143	0.05	7946	2867	0.0502969	0.018148
2022_BAC	0.164	0.054	8331	2937	0.0523577	0.018458
2022_BOC	0.152	0.043	8441	2831	0.0489277	0.01641

Continued

2022_CCB	0.16	0.046	10,075	3312	0.0514908	0.016927
2022_CITI	0.15	0.057	10,428	4164	0.0474177	0.018934
2022_ICBC	0.145	0.045	7716	2633	0.0484153	0.016521
2022_JPM	0.159	0.059	10,516	4003	0.0486395	0.018515
2022_WFC	0.167	0.068	9118	3805	0.0561765	0.023443
2023_ABC	0.143	0.049	7857	2830	0.0499672	0.017998
2023_BAC	0.165	0.055	8555	3023	0.0526232	0.018595
2023_BOC	0.149	0.044	8199	2841	0.0481348	0.016679
2023_CCB	0.152	0.045	9858	3223	0.050364	0.016466
2023_CITI	0.152	0.058	10,716	4362	0.0481644	0.019606
2023_ICBC	0.149	0.046	8026	2795	0.0485871	0.01692
2023_JPM	0.16	0.056	11,864	4291	0.0486365	0.017591
2023_WFC	0.169	0.067	9105	3722	0.0574999	0.023505

2.2.5. Metrics of Bank Performance

ROE (Return on Equity), ROI (Return on Investment) and EPS (Earnings Per Share) are profitability metrics to measure is a measure a company's profitability in relation to the money shareholders, the return on any type of investment, and a company's profit al-located to each outstanding share of common stock. ROE is a key indicator of how efficiently management uses equity to generate profits while ROI evaluates the profitability of an investment by comparing its gain or loss to its cost. ESP evaluates a company's performance from investors' perspective (Banker, Chang, Janakiraman & Konstans, 2004). See Formulas (5)-(7).

$$\text{ROE} = (\text{Net Income})/(\text{Total Shareholders' Equity}) \tag{5}$$

$$\text{ROI} = (\text{Net Profit from investment} - \text{Cost of Investment})/(\text{Cost of Investment}) \tag{6}$$

$$\text{EPS} = (\text{Net Income} - \text{Preferred Dividends})/(\text{Average outstanding Common Shares}) \tag{7}$$

However, ROI is less suitable as a key indicator of bank's performance measuring since it doesn't account for the unique capital structure of banks that rely heavily on deposits as liabilities, not just on investor capital. Unlike non-financial firms, banks prefer to retain earnings to grow their equity, and a bank's ability to issue equity signal financial weakness and dilute existing shareholder value which prioritize its internal capital generation. Therefore, banks usually target ROE through issuing new shares when their earnings growth is low to avoid negative market signals and potential costs associated with external equity issuance (Pennacchi & Santos, 2021).

This mirrors the study that found weak relationship between the same-year values for ROE and Spreads to shareholder value. ROE doesn't contribute significantly to shareholders' returns while Spreds, EPS and DPS, were major benefactor to the correlation found in stepwise regression (De Wet & Du Toit, 2007).

Nevertheless, ROE can be a key component of a business model and a significant predictor for future risk of certain banks, especially during periods of financial instability. The correlation between high pre-crisis ROE and high risk is unique to the banking industry and is not observed in non-financial firms. Banks may not correctly adjust capital charges for the risks of their assets, leading to an artificially inflated ROE and over-investment in risky assets. The compensation structure for bank executives, closely tied to ROE, may incentivize hidden, latent risks to boost their pay (Moussu & Petit-Romec, 2017).

Therefore, none of these metrics alone is sufficient for analyzing a bank's business, but a comprehensive investigation of ROE and ESP may provide more insights.

2.2.6. Correlation Coefficient

As exploring whether there are patterns unseen in banks' annual reports in terms of sentiment and banks' performance is our aim in the study, the correlation between the two are investigated in SPSS. Pearson and Spearman are the two most common correlation coefficients used to measure the relationship between two variables. They produce a value between −1 and +1 to indicate the strength and direction of a relationship based on different assumptions and are appropriate for different types of data.

Pearson Correlation Coefficient (r)

Pearson's correlation assumes the data is normally distributed and measures the linear relationship between two continuous variables based on their covariance, as shown in Formula (8).

$$r=\frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2}} \tag{8}$$

where n is the sample size, x_i and y_i are the individual data points, and $\bar{x}$ and $\bar{y}$ are the sample means.

Spearman's Correlation Coefficient (r)

Spearman's correlation is a non-parametric alternative to Pearson's, which measures the relationship where the variables tend to move in the same direction, but not necessarily at a constant rate, as shown in Formula (9).

$$\rho=1-\frac{6\sum d_i^2}{n\left(n^2-1\right)} \tag{9}$$

where d_i is the difference between the ranks of each pair of observations, and n is the number of pairs.

Either sentimental data or ROE and ESP are non-linear continuous data, Spearman's correlation is appropriate to be used to run a bivariate correlation.

3. Results

3.1. Preliminary Sentiment and Performance Associations

Initial exploratory analysis investigated the concurrent relationship between sentiment word counts and key performance indicators, Return on Equity (ROE) and Earnings Per Share (EPS). To mitigate the risk of Type I errors arising from multiple comparisons, Benjamini-Hochberg False Discovery Rate (FDR) corrections were applied to all initial bivariate tests. Significant negative correlations were observed between ROE and both positive and negative sentiment word counts (FDR-adjusted $p < 0.05$). However, a longer report will naturally contain more sentiment-charged words simply due to its size. When normalizing for report length using sentiment word percentages, pooled bivariate tests yielded limited significance across the aggregate dataset. Sub-group analysis revealed that for U.S. banks, a higher percentage of positive language was significantly associated with a negative ROE difference (FDR-adjusted $p = 0.048$). This indicates that U.S. banks employ more positive expressions when experiencing a significant decline in ROE. This strategic use of language highlights a justificatory discourse commonly employed to maintain stakeholder confidence during periods of financial underperformance.

3.2. Panel Data Regression Analysis

To rigorously evaluate the explanatory power of sentiment language on financial performance—and to account for the repeated annual observations inherent in the dataset ($N = 48$)—a Two-Way Fixed Effects Panel OLS regression was conducted. This approach supersedes pooled bivariate correlations by controlling for unobserved bank-specific entity effects and year-specific macroeconomic shocks, ensuring that within-bank clustering and serial dependence are properly addressed.

Table 4. Two-way fixed effects panel OLS estimation for EPS.

Parameter	Coefficient	Std. Err.	T-stat	*P*-value	Lower CI (95%)	Upper CI (95%)
Intercept	0.8420	0.3110	2.7074	0.0104	0.2078	1.4762
Positive Pct	8.4512	3.6210	2.3340	0.0248*	1.0664	15.836
Negative Pct	−4.1205	2.4501	−1.6818	0.1012	−9.1174	0.8764

a. $N = 48$ bank-year observations (8 entities, 6 time periods). Standard errors are clustered at the bank level. *$p < 0.05$.

As shown in Table 4, the model utilizes 48 bank-year observations with standard errors clustered at the bank level to ensure robust inference. The analysis reveals a statistically significant positive association between the percentage of positive sentiment words and EPS ($\beta = 8.4512$, $p = 0.0248$, 95% CI: [1.0664, 15.8360]). Conversely, the negative sentiment percentage did not yield a statistically significant impact on EPS at the $\alpha = 0.05$ level ($p = 0.1012$). These fixed-effects results confirm that the sentiment-performance correlations are not merely artifacts of unobserved bank characteristics or specific annual economic conditions, providing a robust foundation for subsequent neural network modeling.

3.3. Robustness Check: Finance-Specific Lexicon Benchmarking

Recognizing that standard sentiment analyzers (e.g., VADER) may misclassify technical banking jargon—such as "liability," "debt," or "tax"—as inherently negative, a robustness check was performed utilizing the Loughran-McDonald (LM) finance-specific dictionary. The annual reports were re-tokenized and scored against the LM lexicon to properly handle technical banking terms. During this phase, total tokens were explicitly calculated for each document-year as the denominator for percentage normalization, ensuring mathematical consistency across all observations. The explanatory relationships identified in the primary panel regression remained consistent when utilizing the LM dictionary. This confirms that the observed sentiment-performance associations are driven by genuine management tone rather than artifactual misclassifications of standard financial terminology.

4. Discussion and Modeling

Compared with non-financial firms, whose reports typically center on operational risks, supply chain issues, and market competition, bank annual reports are highly "financialized", using jargon related to complex financial products, asset management, and financial engineering, oriented toward financial market actors (Cheng & Ho, 2017). However, it does not mean there is no sentiment in bank annual reports. While all annual reports have a promotional dimension, bank reports often employ a more explicit justificatory or self-appraisal discourse (Durst, 2013). Our panel data regression confirms this dynamic, demonstrating a significant positive association between the percentage of positive sentiment and EPS after controlling for unobserved bank and year effects.

Meanwhile, banks usually aim to construct an image of stability and trustworthiness, a crucial aspect of a confidence-based business model. This involves attributing successes to management foresight and ethical values while framing negative events as unavoidable consequences of external economic conditions. This justificatory strategy explains our finding that U.S. banks employ significantly more positive language when experiencing a decline in ROE. This strategic use of language is highly pronounced in the banking sector due to its inherent vulnerability to market sentiment and public perception, where management attempts to linguistically offset poor quantitative performance.

Since sentiment from banks' annual reports provides valuable "soft information" that traditional financial metrics do not unveil, creating a statistical model that maps sentiment to concurrent bank performance is highly valuable. Academic evidence shows that the tone and sentiment expressed in annual reports can explain financial performance (Iqbal, Sohail, & Malik, 2023) and stock returns (Hájek & Boháčová, 2016), improve risk assessment (Gandhi, Loughran, & Mcdonald, 2018), and provide early warning signals for distress (Huang et al., 2022) and supervisory insights (Cowhey, Lee, Spiller, & Vojtech, 2022; Nopp & Hanbury, 2015). Rather than strictly predicting future outcomes, mapping management's implicit signals, risks, and confidence levels to same-year performance helps investors, regulators, and policymakers build robust associative early warning systems and reduce information asymmetry.

Neural network models are often used to obtain more accurate modeling. Yet, creating deep Convolutional Neural Network (CNN) models from thousands of annual reports often results in "black boxes" where the decision-making process is opaque, carries a higher risk of overfitting, and requires massive data preprocessing and computational power (Huang et al., 2022). If we shift from just "complexity" to "sustainability" in explanatory modeling of banks' performance using sentiment, simpler neural networks with shallow architectures may outperform complex ones. In financial forecasting contexts, data samples are often relatively small, noisy, and volatile. Deeper models face a high risk of overfitting, while simpler NNs (Neural Networks) can generalize more reliably (Zhang et al., 2022). A one-layer NN enjoys computational sustainability in which training is lightweight, requiring minimal processing time and no specialized hardware (Singh et al., 2023). This brings lower energy consumption, faster deployment, and greater scalability for practical use in financial institutions. A shallow NN also provides clearer transparency: the weights connecting predictors to the hidden layer can be directly examined, offering insights into the relationships between financial sentiment indicators and performance outcomes (Mhlanga, 2021). A one-layer NN is more resilient to changing input patterns and easier to maintain over time, which is a crucial consideration in financial research where consistent and interpretable models are favored (Shui et al., 2022), especially in terms of institutional usability.

Therefore, building upon the explanatory relationships established in our panel regression, a feed-forward artificial neural network (ANN) was constructed using SPSS. The network comprised one input layer, a single hidden layer, and one output node. The input layer included one categorical factor variable (EPS_P_N, coded as positive or negative) and one continuous covariate (positive_percent). The hidden layer consisted of one processing unit with a hyperbolic tangent activation function, while the output layer applied a linear (identity) activation function (see Formula (10)). The model employed the sum of squares as the error function, and covariates were standardized prior to training (see Formula (11)). The parameter estimates indicate that the model captures both categorical and continuous influences (see Formula (12)).

Model construct:

- Hidden layer with 1 unit:

$$H1 = \tanh\left(b_h + w_1 + X_1 + w_2 + X_2 + w_3 \cdot X_3\right) \tag{10}$$

where:

$$b_h = -0.505\left(\text{bias for hidden node}\right)$$

$$X_1 = \left[\text{EPS\_P\_N} = \text{Negative}\right], \text{weight} = -0.466$$

$$X_2 = [\text{EPS\_P\_N} = \text{Positive}], \text{weight} = 0.817$$

$$X_3 = \text{positive\_percent}, \text{weight} = -0.301$$

- Hidden to output layer:

$$\hat{Y} = b_o + w_o \cdot H1 \tag{11}$$

where:

$$b_o = 0.309$$

$$w_o = 1.344$$

- The final model is:

$$\begin{aligned}\widehat{\text{EPS\_difference}} = 0.309 + 1.344 \cdot \tanh\big(&-0.505 - 0.466 \cdot D_{Neg} \\ &+ 0.817 \cdot D_{Pos} - 0.301 \cdot \text{Positive\_percent}\big)\end{aligned} \tag{12}$$

where D_{Neg} and D_{Pos} are dummy variables for EPS_P_N category.

The model performance is quite robust, with a training relative error of 0.406 and a markedly lower testing relative error of 0.150, which demonstrates that the shallow architecture generalizes well to unseen data and avoids overfitting.

5. Conclusions

This study investigated the explanatory relationship between sentiment language in annual reports and key financial performance metrics—Return on Equity (ROE) and Earnings Per Share (EPS)—across eight major banks in China and the United States. By transitioning from traditional exploratory analysis to a rigorous Two-Way Fixed Effects Panel OLS regression, this research established a statistically significant positive association between the percentage of positive sentiment and EPS, effectively controlling for unobserved bank-specific characteristics and annual macroeconomic shocks. Crucially, a robustness check utilizing the finance-specific Loughran-McDonald lexicon confirmed that these linguistic patterns reflect genuine management tone rather than the misclassification of standard banking terminology.

A notable strategic nuance emerged when analyzing periods of financial underperformance. The analysis revealed that U.S. banks exhibit a significant tendency to employ more positive expressions when experiencing a decline in ROE. This highlights a justificatory discourse specific to the banking sector, where management leverages optimistic language to offset poor quantitative performance and maintain stakeholder confidence.

The established associative relationships present a compelling opportunity to develop statistical models based on linguistic features extracted from corporate disclosures. Leveraging these findings, this study constructed a deliberately shallow, one-layer feed-forward neural network. By incorporating EPS polarity and the percentage of positive sentiment as transparent inputs, the model achieved robust performance while avoiding the opaque "black box" limitations typical of deep learning architectures. Its lower risk of overfitting, reduced computational burden, and enhanced interpretability make it particularly well-suited for small-scale, structured financial data environments. Ultimately, this demonstrates that sustainability in financial modeling encompasses not only energy efficiency but also regulatory transparency and institutional usability.

Author Contributions

Conceptualization, Hu.T.; methodology, Hu.T. and Li.J.; software, Hu.T. and Li.J.; validation, Hu.T. and Li.J.; formal analysis, Hu.T.; investigation, Hu.T.; resources, Li.J.; data curation, Hu.T. and Li.J.; writing original draft preparation, Hu.T. and Li.J.; writing, review and editing, Hu.T.; project administration, Hu.T.; funding acquisition, Hu.T. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Alduais, F. (2024). Textual Analysis of the Annual Report and Corporate Performance: Evidence from China. *Journal of Financial Reporting and Accounting, 22,* 1221-1252. https://doi.org/10.1108/jfra-04-2022-0129

Aram, M., & Soroushyar, A. (2024). Comparing the Informed and Noise Investors' Perception of the Tone of Financial Statements and Its Impact on Stock Returns: A Text-Mining Approach. *International Journal of Finance & Managerial Accounting, 9,* 157-172. http://www.ijfma.ir/article_22716.html

Banker, R. D., Chang, H., Janakiraman, S. N., & Konstans, C. (2004). A Balanced Scorecard Analysis of Performance Metrics. *European Journal of Operational Research, 154,* 423-436. https://doi.org/10.1016/s0377-2217(03)00179-6

Bischof, J., Laux, C., & Leuz, C. (2021). Accounting for Financial Stability: Bank Disclosure and Loss Recognition in the Financial Crisis. *Journal of Financial Economics, 141,* 1188-1217. https://doi.org/10.1016/j.jfineco.2021.05.016

Cheng, W., & Ho, J. (2017). A Corpus Study of Bank Financial Analyst Reports: Semantic Fields and Metaphors. *International Journal of Business Communication, 54,* 258-282. https://doi.org/10.1177/2329488415572790

Cowhey, M., Lee, S. J., Spiller, T. P., & Vojtech, C. M. (2022). Sentiment in Bank Examination Reports and Bank Outcomes. *Finance and Economics Discussion Series* 2022-077 (pp. 1-46), Federal Reserve Board. https://doi.org/10.17016/feds.2022.077

De Wet, J. H. V. H., & Du Toit, E. (2007). Return on Equity: A Popular, but Flawed Measure of Corporate Financial Performance. *South African Journal of Business Management, 38,* 59-69. https://doi.org/10.4102/sajbm.v38i1.578

Durst, S. (2013). An Exploratory Study of Intangibles Risk Disclosure in Annual Reports of Banking Companies from the UK, US, Germany and Italy-Some Descriptive Insights. *Financial Reporting, No. 1,* 81-120. https://doi.org/10.3280/fr2013-001005

Gandhi, P., Loughran, T., & McDonald, B. (2018). Using Annual Report Sentiment as a Proxy for Financial Distress in U.S. Banks. *Journal of Behavioral Finance, 20,* 424-436. https://doi.org/10.1080/15427560.2019.1553176

Gao, Y., & Feng, J. (2024). Annual Report Text's Positive Tone and Corporate Green Innovation: Evidence from China. *PLOS ONE, 19,* e0306184. https://doi.org/10.1371/journal.pone.0306184

Hájek, P., & Boháčová, J. (2016). Predicting Abnormal Bank Stock Returns Using Textual Analysis of Annual Reports—A Neural Network Approach. In C. Jayne, & L. Iliadis (Eds.), *Communications in Computer and Information Science* (pp. 67-78). Springer International Publishing. https://doi.org/10.1007/978-3-319-44188-7_5

Huang, B., Yao, X. C., Luo, Y., & Li, J. (2022). Improving Financial Distress Prediction Using Textual Sentiment of Annual Reports. *Annals of Operations Research, 330,* 457-484. https://doi.org/10.1007/s10479-022-04633-3

Iqbal, J., Sohail, M. K., & Malik, M. K. (2023). Predicting the Future Financial Performance of Islamic Banks: A Sentiment Analysis Approach. *International Journal of Islamic and Middle Eastern Finance and Management, 16,* 1287-1305. https://doi.org/10.1108/imefm-07-2022-0267

Kravet, T., & Muslu, V. (2013). Textual Risk Disclosures and Investors' Risk Perceptions. *Review of Accounting Studies, 18,* 1088-1122. https://doi.org/10.1007/s11142-013-9228-9

Li, F. (2010). The Information Content of Forward-Looking Statements in Corporate Filings—A Naïve Bayesian Machine Learning Approach. *Journal of Accounting Research, 48,* 1049-1102. https://doi.org/10.1111/j.1475-679x.2010.00382.x

Mhlanga, D. (2021). Financial Inclusion in Emerging Economies: The Application of Machine Learning and Artificial Intelligence in Credit Risk Assessment. *International Journal of Financial Studies, 9,* Article 39. https://doi.org/10.3390/ijfs9030039

Mocanu, M., Grose, C., & Kargidis, T. (2019). Readability of Operational Risk Disclosures of Banks. *Studies in Business and Economics, 14,* 108-116. https://doi.org/10.2478/sbe-2019-0047

Moussu, C., & Petit-Romec, A. (2017). ROE in Banks: Performance or Risk Measure? Evidence from Financial Crises. *Finance, 38,* 95-133. https://doi.org/10.3917/e.fina.382.0095

Nopp, C., & Hanbury, A. (2015). Detecting Risks in the Banking System by Sentiment Analysis. In L. Màrquez, C. Callison-Burch, & J. Su (Eds.), *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing* (pp. 591-600). Association for Computational Linguistics. https://doi.org/10.18653/v1/d15-1071

Oino, I. (2019). Do Disclosure and Transparency Affect Bank's Financial Performance? *Corporate Governance: The International Journal of Business in Society, 19,* 1344-1361. https://doi.org/10.1108/cg-12-2018-0378

Pennacchi, G. G., & Santos, J. A. C. (2021). Why Do Banks Target ROE? *Journal of Financial Stability, 54,* Article 100856. https://doi.org/10.1016/j.jfs.2021.100856

Rattanataipop, P. (2013). *Risk Disclosures in the Annual Reports of UK Banks, 1995-2010.* Doctoral Dissertation, Newcastle University.

Scannella, E., & Polizzi, S. (2018). Market Risk Disclosure in Banking: An Empirical Analysis on Four Global Systemically Important European Banks. *Journal of Banking Regulation, 19,* 87-100. https://doi.org/10.1057/s41261-017-0039-y

Shui, X., Zheng, H., Xia, X., Yang, F., Wang, W., & Yu, Y. (2022). Diffraction Model-Informed Neural Network for Unsupervised Layer-Based Computer-Generated Holography. *Optics Express, 30,* Article 44814. https://doi.org/10.1364/oe.474137

Singh, G., Rani, L., Sarangi, P. K., Sharma, K., Sahoo, A. K., & Malhotra, R. (2023). CNN-RNN Based Hybrid Deep Learning Model for Predicting Fluctuations in the Stock Market. *2023 5th International Conference on Advances in Computing, Communication Control and Networking* (ICAC3N) (pp. 417-422). IEEE.

https://doi.org/10.1109/icac3n60023.2023.10541397

Yang, F., Chang, C. C., & Mo, L. (2022). An Introduction to Multilevel Analyses of Text Readability in Accounting and Finance. *Journal of Emerging Technologies in Accounting, 19,* 187-197. https://doi.org/10.2308/jeta-19-09-02-37

Zhang, H., Wu, C., Zhang, Z., Zhu, Y., Lin, H., Zhang, Z. et al. (2022). ResNeSt: Split-Attention Networks. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (CVPRW) (pp. 2735-2745). IEEE. https://doi.org/10.1109/cvprw56347.2022.00309

Zheng, C., Sarker, N., & Nahar, S. (2018). Factors Affecting Bank Credit Risk: An Empirical Insight. *Journal of Applied Finance & Banking, 8,* 45-67.