INTERNATIONAL JOURNAL OF COMPUTER AUDITING
Article

Text Mining and Machine Learning on 10-K Risk Factors and Net Income: Evidence from Apple

Sherry Huang[1] | **Shi-Ming Huang**[2]

[1]Jacksoft Commerce Automation, Ltd., Taiwan

[2]Department of Accounting and Information Technology, National Chung Cheng University, Taiwan

Correspondence
Shi-Ming Huang

Publication History:
Received 4 September 2025;
Revised 20 November 2025;
Accepted 2 December 2025

This study investigates whether risk factor disclosures in 10-K annual filings contain predictive signals about firms' future financial performance. Both structured (financial statements) and unstructured (risk factor narratives) data were analyzed using text mining and machine learning techniques. The JCAATs XBRL Connector and AI audit functions were employed to streamline data extraction, sentiment analysis, clustering, machine learning modeling, and SHAP-based interpretability. Sentiment scores and textual clusters were constructed as independent variables to explain subsequent-year net income. Empirical findings demonstrate that both sentiment and textual style are significant predictors of net income, supporting the view that risk disclosures provide forward-looking information. These findings extend language signal theory to the risk factor section and underscore the practical value of JCAATs in auditing and regulatory monitoring, highlighting how AI-driven text analytics can enhance disclosure assessment and financial supervision.

KEYWORDS
JCAATs, Text Mining, Machine Learning, Risk Factors, Financial Reporting, Explainable AI, Net Income Prediction, 10-K Analysis

Reference Format: Huang, S., & Huang, S. M. (2025). Text mining and machine learning on 10-K risk factors and net income: Evidence from Apple. International Journal of Computer Auditing, 7(1), 3–15.
https://doi.org/10.53106/256299802025120701002

1 | INTRODUCTION

The U.S. Securities and Exchange Commission (SEC) requires all publicly listed companies to file annual reports through the EDGAR (Electronic Data Gathering, Analysis, and Retrieval) system, which serves as the official online platform for corporate disclosures. Among these reports, the 10-K is the most comprehensive filing, containing audited financial statements, management discussion and analysis (MD&A), and detailed risk factor disclosures.

Since 2006, the SEC has expanded mandatory reporting requirements for risk factors. Beginning in 2011, filings were required to adopt XBRL (eXtensible Business Reporting Language), a standardized digital format that enables both structured data (e.g., income statements, balance sheets) and unstructured data (e.g., risk factors, MD&A) to be systematically retrieved and analyzed. In 2018, Inline XBRL (iXBRL) was introduced, and more recently, in 2022, the SEC mandated expanded ESG and climate-related disclosures. This evolution of SEC reporting (see Figure 1) has created an invaluable dataset for researchers and auditors, ensuring comparability, machine readability, and consistency across firms and over time.

Risk factor disclosures in 10-K filings are intended to inform investors and stakeholders about potential threats and uncertainties. While often regarded as compliance-oriented narratives, emerging research suggests that their sentiment and linguistic characteristics may provide predictive value for future firm performance. However, while prior studies have focused on MD&A and earnings calls, limited attention has been given to the predictive value of the risk factor section, despite its mandatory disclosure since 2006.

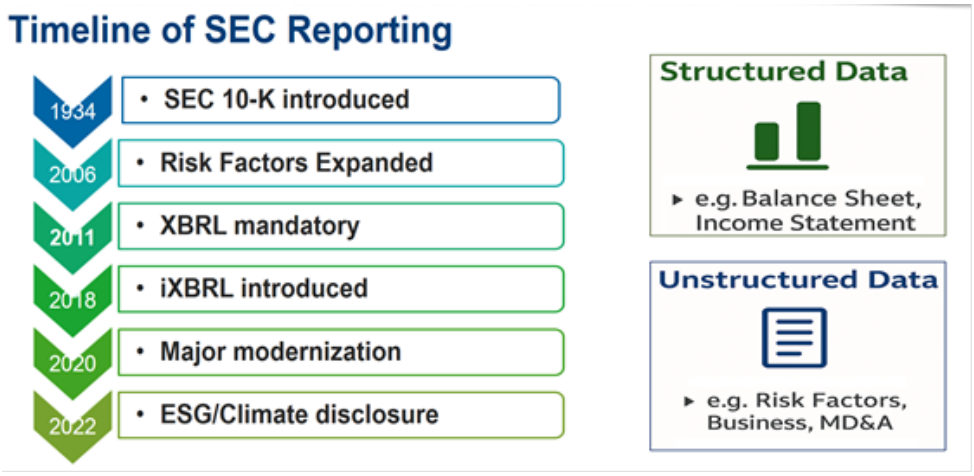


FIGURE 1 SEC Reporting Timeline (author's illustration based on SEC regulations))

This study applies text mining and machine learning to assess whether risk factor disclosures in Apple's 10-K filings (2010–2024) contain signals about subsequent-year net income. By integrating structured numerical data with unstructured textual information, the study establishes a replicable research framework that can be applied in both financial auditing and supervisory analytics. The analysis is conducted with the support of JCAATs (Jacksoft, 2025), an AI-based auditing platform that integrates XBRL data extraction, text mining, and machine learning capabilities.

2 | LITERATURE REVIEW AND HYPOTHESES

Prior research has established that linguistic features embedded in corporate disclosures influence investor judgment, risk perception, and firm valuation (Li, 2010; Loughran & McDonald, 2011; Feldman et al., 2010). In line with language signal theory, narrative elements in financial reports convey implicit forward-looking information that may reveal managerial expectations, risk assessments, and strategic priorities. Given that 10-K filings contain extensive qualitative content, text mining and computational linguistic techniques have become essential tools for extracting and analyzing disclosure-based signals.

Recent studies (2022–2025) have increasingly applied topic modeling, semantic clustering, and predictive machine learning to examine how narrative disclosures reflect firm risks and future outcomes. For example, Bogachek et al. (2025) demonstrate that thematic content extracted from full 10-K narratives substantially improves the prediction of tax outcomes, highlighting that qualitative information distributed throughout a filing holds incremental predictive value.

Similarly, Fengler and Phan (2025) use word-embedding–based clustering to identify interpretable themes in MD&A sections and show that variation in topic prevalence and sentiment is associated with differences in stock market reactions. At the broader market level, Uberti-Bona Marin et al. (2025) document a large increase in AI-related risk disclosures in SEC filings, illustrating that firms increasingly rely on narrative reporting to communicate emerging technological uncertainties.

In parallel, recent research has increasingly applied machine-learning techniques to analyze 10-K disclosures. These studies demonstrate that linguistic uncertainty, tone, and semantic patterns extracted through ML-based text-processing pipelines can serve as predictive signals for important firm outcomes (Kacar, Yumusak,& Kodaz, 2023). Survey evidence further highlights that ML approaches have become essential for identifying financial risks embedded in narrative disclosures (Zavitsanos et al., 2024). Collectively, these findings indicate that machine-learning analysis of 10-K narratives provides informative signals about future firm performance and potential downside risks.

A related methodological stream highlights the importance of interpretability when applying machine learning to financial reporting. As machine learning models become more embedded in accounting and auditing applications, researchers increasingly emphasize the need for transparent and traceable model outputs. Zhang, Cho, and Vasarhelyi (2022) introduce Explainable Artificial Intelligence (XAI) techniques—such as LIME and SHAP—to auditing contexts and illustrate how these tools enhance the interpretability of prediction models while supporting audit documentation requirements. Similarly, recent systematic reviews highlight that explainability techniques such as feature importance, SHAP, and rule-based methods have become central to interpreting black-box financial models. These methods improve transparency in tasks such as credit risk assessment, stock price prediction, and fraud detection, underscoring the expanding role of XAI in financial decision-making (Černevičienė & Kabašinskas, 2024). These developments emphasize that XAI is essential for ensuring accountability, auditability, and regulatory compliance when applying machine learning to narrative disclosures.

Overall, the literature suggests that qualitative content in 10-K filings contains economically meaningful information, and that combining text mining with interpretable machine learning methods yields reliable and actionable insights. Building on these findings, this study examines whether the sentiment and textual style of 10-K risk factor disclosures provide predictive signals regarding subsequent firm performance.

H1: The sentiment of risk factor disclosures is significantly associated with subsequent-year net income.

H2: The textual style of risk factor disclosures, represented by semantic clusters, provides predictive value for subsequent-year net income.

3 | RESEARCH METHODOLOGY

3.1 | Research Framework

This study employed JCAATs (Jacksoft Computer-Assisted Auditing Tools), an AI-powered auditing platform, to support each stage of the analysis. JCAATs integrates an XBRL connector for structured data extraction, natural language processing (NLP) modules for sentiment analysis, clustering algorithms for text style classification, and built-in machine learning functions such as Random Forest, Support Vector Machines (SVM), and Logistic Regression. In addition, JCAATs incorporates explainable AI techniques, such as SHAP, to ensure transparency and interpretability of model predictions. The automation and integration provided by JCAATs substantially reduced research time and enhanced replicability, making it a practical tool for both academic research and professional audit practices.

A key component in this study is the 10-K XBRL EDGAR Connector embedded within JCAATs, which enables the direct retrieval of structured financial statements and unstructured textual disclosures from the SEC EDGAR database. This feature ensured that Apple's 10-K reports from 2010 to 2024 were systematically extracted, standardized, and integrated into the subsequent text mining and machine learning pipeline.

Building on these capabilities, the research framework followed a four-stage process (Figure 2

1. Data Collection: Extraction of both structured and unstructured information from SEC 10-K filings.
2. Data Preparation: Sentiment analysis, text clustering, and construction of the dependent variable.
3. Machine Learning Modeling: Application and comparison of predictive algorithms.
4. Explainable AI Analysis: Evaluation of feature contributions using SHAP.

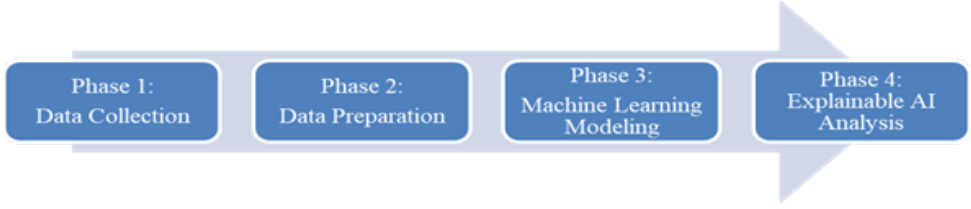


FIGURE 2 Research Process Flow

3.2 | Data Collection

The empirical analysis focuses on Apple Inc.'s 10-K filings from 2010 to 2024, covering a total of 15 reports. Using the JCAATs 10-K XBRL EDGAR Connector, two types of data were retrieved:

- Structured data: net income figures extracted from the income statement.
- Unstructured data: narrative text from the risk factor sections of the 10-K reports.

The connector enabled consistent and systematic retrieval of both datasets, which were then processed for subsequent analysis. Although Apple has filed annual 10-K reports since its IPO in 1980, this study focuses specifically on the 2010–2024 period, which coincides with the SEC's mandatory risk factor disclosure requirements and the adoption of XBRL reporting standards.

SEC 10-K filings have been provided in Inline XBRL (iXBRL) since 2018, embedding HTML and XBRL tags directly within the narrative text. This hybrid structure complicates text mining because disclosures are often fragmented by nested tags, markup elements, and presentation rules. Prior to 2018, however, filings were submitted in HTML or traditional (non-inline) XBRL formats, creating an additional challenge for researchers working with multi-year datasets. As a result, many studies rely on third-party vendors that reprocess SEC filings into simplified text files. Although convenient, such reliance introduces risks, including loss of original context, inconsistencies created by undocumented preprocessing steps, and potential deviations from official EDGAR data. These issues can undermine the replicability, accuracy, and validity of text-based analyses.

To address these challenges and to ensure a consistent processing framework across the entire 2010–2024 sample, this study employs the JCAATs XBRL Connector, which directly retrieves filings from the SEC's EDGAR system. For post-2018 submissions, the Connector automatically parses iXBRL, removes embedded HTML/XBRL tags, and reconstructs the narrative flow while preserving paragraph boundaries, semantic meaning, and table structures. For pre-2018 filings, the Connector detects the legacy HTML or traditional XBRL format and applies the appropriate extraction procedure to isolate narrative sections using the same segmentation logic. This unified approach produces both clean and readable unstructured textual data and well-organized structured quantitative data, ensuring comparability across all filing years.

Figure 3 illustrates the JCAATs system interface, showing how the platform connects directly to EDGAR to obtain official filings. This direct connection eliminates the data-integrity risks inherent in third-party sources and provides a reliable mechanism for processing both structured and unstructured information extracted from diverse SEC filing formats.

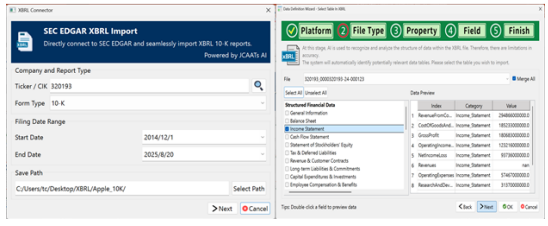


FIGURE 3 XBRL Connector Data Extraction from SEC EDGAR

3.3 | Data Preparation

The dependent variable (Y) was defined as next-year net income, coded as an increase or decrease relative to the current year. The independent variables (X) consisted of two components. First, sentiment scores were extracted from the risk-factor narratives using the VADER (Valence Aware Dictionary and Sentiment Reasoner) algorithm (Hutto & Gilbert, 2014), with additional lexicon adjustments implemented within JCAATs to better capture the characteristics of financial text. Second, textual-style features were generated using K-means clustering, with the optimal number of clusters selected based on the silhouette score

To ensure methodological consistency, both sentiment scores and cluster assignments were produced using JCAATs' built-in text-mining modules, which standardize tokenization, sentiment computation, text vectorization, and clustering. Figure 4 presents the workflow used to generate the independent variables: the SENTIMENT command produces the sentiment-score dataset, the TEXT CLUSTER command generates the textual-style cluster dataset, and the two outputs are merged using the JOIN command (key = source_file) to create the final research dataset.

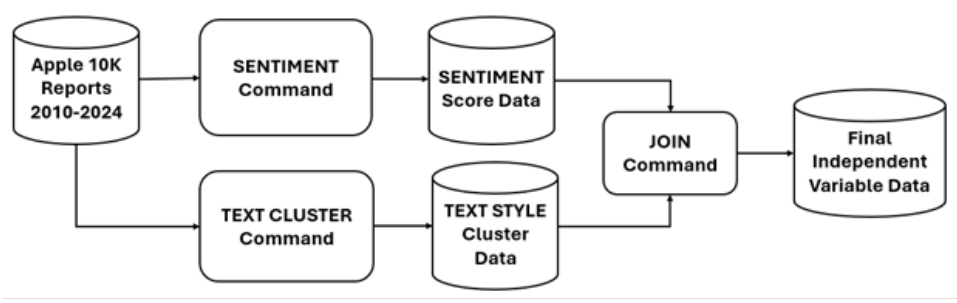


FIGURE 4 Workflow for Generating Independent Variables Using JCAATs

Figure 5 displays the actual JCAATs system output, including the computed sentiment fields (positive, negative, neutral, compound), the derived cluster labels, and the distance values that characterize each text segment's position within its cluster. This screenshot is included to document the exact system settings and outputs used in the study, providing transparency and enabling other researchers to replicate the sentiment and clustering procedures without ambiguity.

3.4 | Machine Learning

To ensure that machine-learning procedures are transparent and reproducible, this study uses the built-in ML pipeline provided by JCAATs, which is implemented using Python's scikit-learn module (Pedregosa et.al. 2011). The pipeline guides users through each preprocessing and modeling step in a standardized manner. Three classification algorithms were implemented: Support Vector Machine (SVM), Random Forest, and Logistic Regression. These models represent three major methodological families, namely margin-based, ensemble-based, and linear approaches. This design enables a systematic comparison of predictive performance across different modeling paradigms. Model performance was evaluated using the F1-score as the primary metric, with precision and recall serving as supplementary indicators. To address class imbalance, the Synthetic Minority Oversampling Technique (SMOTE) was applied (Chawla, Bowyer, Hall, & Kegelmeyer, 2002), and all models were trained and validated using five-fold cross-validation with one-hot

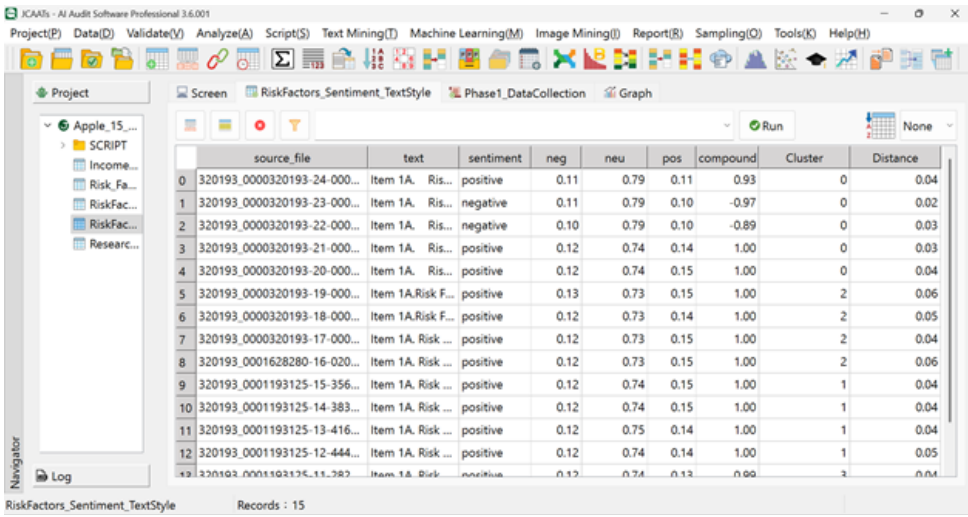


FIGURE 5 Text Mining Outputs (Sentiment Scores and Risk Factor Clusters)

encoding for categorical features

Figure 6 illustrates the machine-learning workflow embedded in JCAATs and displays the exact configuration settings used in this study, including model selection, target-field assignment, missing-value handling, encoding options, SMOTE thresholds, and cross-validation folds. Unlike code-based implementations, where preprocessing choices may vary across researchers, the JCAATs interface enforces a consistent and fully documented pipeline. Users explicitly select the target variable and features, choose the algorithm, and configure all preprocessing parameters through the graphical interface. These interface controls ensure that the entire machine-learning process, from data preparation to model evaluation, is transparent, standardized, and easily replicable, reducing the ambiguity and undocumented variation often encountered in custom-coded workflows.

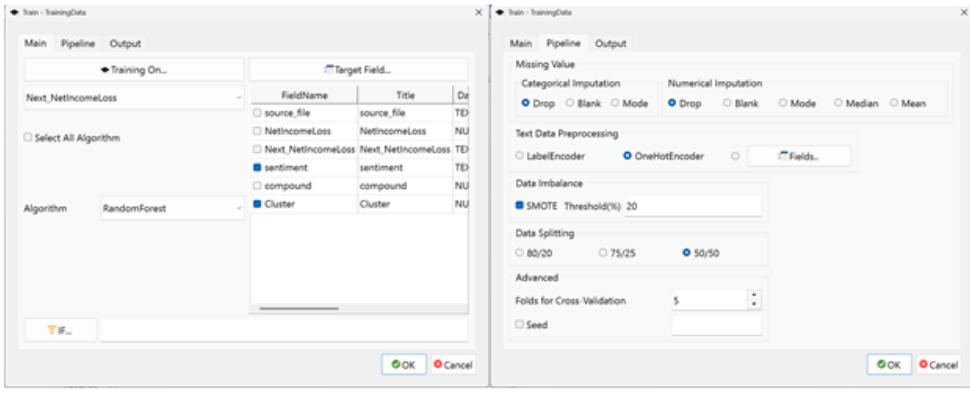


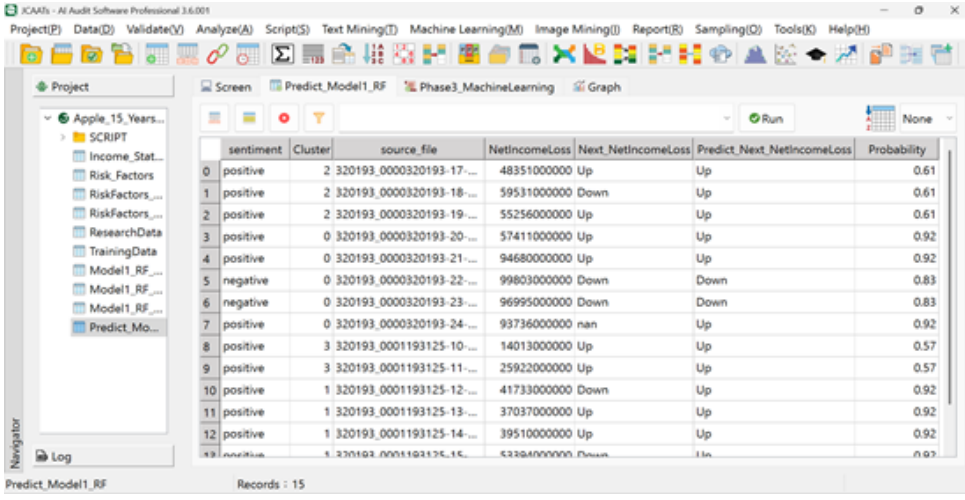
FIGURE 6 Machine Learning Workflow in JCAATs

3.5 | Explainable AI

To ensure interpretability of the predictive models, the study adopted a two-step process. First, the Predict function in JCAATs was used to generate model outputs, including both the predicted class (increase or decrease in next-year net income) and the associated probability scores (see Figure 7). This step provided a direct view of the model's decision-making confidence.

As an additional validation of predictive effectiveness, the model's out-of-sample prediction for Apple Inc. in August 2025 indicated an expected increase in net income for fiscal year 2025. When Apple subsequently released its official 2025 Form 10-K, the reported net income increased substantially from USD 93,736 million in 2024 to USD 112,010 million in 2025. This real-world outcome is consistent with our model's prediction and provides further evidence of the model's correctness and practical applicability.

Second, the study employed SHAP (Shapley Additive Explanations), a widely used method in explainable AI (XAI), to quantify the marginal contribution of each feature to the prediction outcome. SHAP values enabled an evaluation of the relative importance of sentiment scores and textual style clusters, thereby enhancing transparency and allowing stakeholders to understand the rationale behind the model's outputs.



	sentiment	Cluster	source_file	NetIncomeLoss	Next_NetIncomeLoss	Predict_Next_NetIncomeLoss	Probability
0	positive	2	320193_0000320193-17-...	48351000000	Up	Up	0.61
1	positive	2	320193_0000320193-18-...	59531000000	Down	Up	0.61
2	positive	2	320193_0000320193-19-...	55256000000	Up	Up	0.61
3	positive	0	320193_0000320193-20-...	57411000000	Up	Up	0.92
4	positive	0	320193_0000320193-21-...	94680000000	Up	Up	0.92
5	negative	0	320193_0000320193-22-...	99803000000	Down	Down	0.83
6	negative	0	320193_0000320193-23-...	96995000000	Down	Down	0.83
7	positive	0	320193_0000320193-24-...	93736000000	nan	Up	0.92
8	positive	3	320193_0001193125-10-...	14013000000	Up	Up	0.57
9	positive	3	320193_0001193125-11-...	25922000000	Up	Up	0.57
10	positive	1	320193_0001193125-12-...	41733000000	Down	Up	0.92
11	positive	1	320193_0001193125-13-...	37037000000	Up	Up	0.92
12	positive	1	320193_0001193125-14-...	39510000000	Up	Up	0.92

FIGURE 7 Text Mining Outputs (Sentiment Scores and Risk Factor Clusters)

4 | RESULTS

4.1 | Descriptive Findings

Textual clustering analysis was applied to the risk factor narratives in Apple's 10-K filings from 2010 to 2024. Using the silhouette score as a criterion for model selection, the optimal number of clusters was determined to be K = 4. This partitioning provided not only the best fit in terms of clustering validity but also produced categories that aligned meaningfully with regulatory and macroeconomic developments over the sample period.

The four clusters correspond to distinct phases in the evolution of Apple's risk factor disclosures (See Table 1).

Table 1: Risk Factors Text Style Cluster

Cluster	Year	Text Style
3	2010–2011	Short, traditional, SEC did not require detailed risk disclosure.
1	2012–2015	Inclusion of compliance and regulatory risks following the Dodd-Frank Act.
2	2016–2019	Expanded to global supply chain, tax risks, and U.S.-China trade issues.
0	2020–2024	Increased focus on COVID-19, geopolitical tensions, digital security, and ESG.

These findings suggest that the clustering approach not only captured statistically coherent groupings but also reflected substantive shifts in disclosure practices shaped by regulatory, economic, and geopolitical contexts.

4.2 | Model Performance

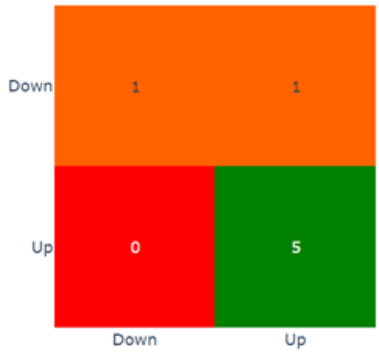
Among the machine learning algorithms tested, Random Forest achieved the best overall performance relative to Support Vector Machine (SVM) and Logistic Regression. As shown in Figure 7, the Random Forest classifier obtained an F1-score of 0.84, with balanced precision (0.88) and recall (0.86), resulting in an overall accuracy of 0.86. These results indicate that the ensemble-based approach was more effective in capturing the nonlinear relationships between textual features and financial outcomes. In particular, ensemble methods are well-suited for small samples and nonlinear patterns, making Random Forest a robust choice for this study's dataset.

In addition to these single-split results, the model's mean accuracy across five-fold cross-validation was 0.73. This lower value, compared with the hold-out accuracy of 0.86, indicates variability in performance across different training and test partitions. The discrepancy is largely attributable to the relatively small sample size of 15 annual reports, suggesting that while Random Forest captured strong predictive patterns in certain folds, its generalization remains sensitive to data partitioning.

The confusion matrix (Figure 8, left panel) further illustrates the model's predictive capability. The Random Forest correctly identified all five cases of next-year income increases and one case of decreases, with only a single misclassification. This suggests that the model was particularly effective in predicting positive earnings outcomes while still retaining reasonable sensitivity to negative outcomes.

Incorporating sentiment scores and textual cluster variables significantly enhanced predictive performance compared

with using numerical financial data alone. Feature importance analysis (Figure 8, right panel) highlights that negative sentiment (0.31) and positive sentiment (0.19) were among the most influential predictors, followed by cluster variables that captured thematic disclosure patterns (0.18–0.10). These findings support the hypothesis that risk factor narratives contain forward-looking signals relevant to subsequent firm profitability.



Name	Value	Indicator	Model
Accuracy	0.8571428571	Metric	RandomForest
Precision_weight	0.8809523810	Metric	RandomForest
Recall_weight	0.8571428571	Metric	RandomForest
F1_weight	0.8398268398	Metric	RandomForest
sentiment_negative	0.3063552856	Importance	RandomForest
sentiment_positive	0.1877401957	Importance	RandomForest
Cluster_0	0.1827703047	Importance	RandomForest
Cluster_1	0.1482147448	Importance	RandomForest
Cluster_2	0.0986421167	Importance	RandomForest
Cluster_3	0.0762773524	Importance	RandomForest
Mean_Accuracy	0.7333333333	Metric	RandomForest

FIGURE 8 Confusion Matrix & Model Comparison Results

4.3 | Explainable AI Results

To further interpret the Random Forest model, SHAP (Shapley Additive Explanations) analysis was performed. The SHAP summary plot (Figure 9) illustrates the relative importance of sentiment variables and text clusters in predicting subsequent-year net income.

The results indicate that negative sentiment was the most influential predictor, with SHAP values ranging up to 0.3, suggesting a strong association between negative risk disclosures and decreases in future profitability. Positive sentiment, while also contributing to predictions, had a comparatively smaller but still notable effect.

Beyond sentiment, the SHAP analysis highlights that the textual style of Apple's risk factor disclosures has become increasingly informative over time. Earlier reports (2010–2011) contained short and generic narratives with limited predictive value, whereas more recent disclosures (2020–2024) emphasized COVID-19, geopolitical risks, digital security, and ESG considerations. The relatively higher SHAP contributions of Cluster 0 indicate that these contemporary narratives are not only more detailed but also provide clearer forward-looking signals regarding financial outcomes. Quantitatively, Cluster 0 reached SHAP values of approximately 0.18, higher than Cluster 1–3 (ranging from 0.15 to 0.08), confirming its comparatively stronger predictive influence.

Overall, the SHAP analysis confirms that the predictive capacity of the model is driven not only by sentiment polarity but also by the underlying semantic structures of risk factor narratives. This reinforces both the value of textual analysis in financial forecasting and the notion that corporate risk factor disclosures have evolved into a more substantive and decision-relevant component of 10-K reports.

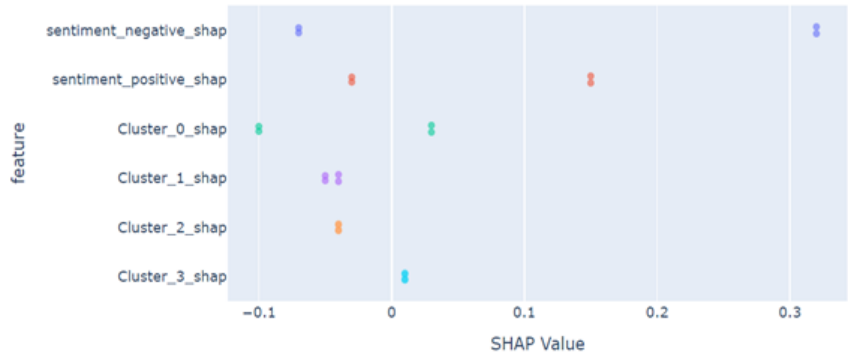


FIGURE 9 SHAP Summary Plot

5 | DISCUSSION

The results of this study demonstrate both theoretical and practical implications. In practice, JCAATs shows strong potential for industry applications. For auditors, the platform automates the process of analyzing complex financial narratives, thereby reducing manual effort and improving efficiency in identifying potential risks. Regulators may also benefit from using JCAATs to monitor disclosure quality across industries, with explainable AI methods such as SHAP ensuring transparency and accountability in the supervisory process. For academic researchers, the integration of text mining, machine learning, and explainability within a single platform lowers technical barriers, enabling faster prototyping and reproducible studies. Collectively, these contributions highlight the role of JCAATs as a bridge between advanced computational techniques and the practical needs of auditing, regulation, and financial analysis

The findings further confirm that 10-K risk factor narratives, beyond their compliance-oriented function, contain significant forward-looking signals. This reinforces the language signal theory (Li, 2010; Loughran & McDonald, 2011) and supports recent work emphasizing the value of narrative disclosures in predicting financial outcomes (Uberti-Bona Marin et al., 2025; TinyXRA, 2025). Specifically, the results suggest that financial text analytics should be integrated into auditing and supervisory workflows, as the sentiment and style of corporate disclosures provide predictive information relevant to firm performance.

Implications extend across multiple stakeholders. Auditors can proactively identify financial risks through automated text analysis. Regulators may incorporate linguistic indicators into their assessment of disclosure quality. Investors stand to benefit from combining textual signals with traditional quantitative ratios, thereby gaining richer insights into firm risk and performance prospects.

6 | CONCLUSION, LIMITATIONS, AND FUTURE RESEARCH

This research highlights the broader value of JCAATs as a financial analytics platform. By providing an end-to-end environment—from data collection to model interpretability—JCAATs enables practitioners, regulators, and academics to conduct rigorous analysis without requiring advanced programming skills. This lowers the entry barrier for applying

AI in auditing and ensures that explainable, data-driven insights can be embedded into decision-making processes. The versatility of JCAATs positions it not only as a research tool but also as a scalable solution for industry adoption.

Empirical findings demonstrate that Apple's 10-K risk factor disclosures from 2010–2024 significantly predict subsequent-year net income through sentiment and linguistic style analysis. By applying machine learning models and explainable AI methods, this study establishes a replicable framework for financial text analytics that integrates structured and unstructured information sources

Despite these contributions, the study has several limitations. First, the analysis was restricted to a single firm, Apple Inc., which may limit the generalizability of the findings to other industries or companies with different disclosure practices. Second, the dataset covered a relatively short period of 15 years (2010–2024), which, while sufficient to capture major regulatory and economic events, may not fully reflect long-term disclosure dynamics. Third, the sentiment analysis relied on adapted lexicon-based methods, which may not capture the full nuance of financial narratives despite domain-specific adjustments.

Future research should therefore expand the sample to include multiple firms across industries, extend the temporal coverage, and explore alternative sentiment modeling approaches such as deep learning or large language models. In addition, examining other disclosure sections—such as Management Discussion and Analysis (MD&A)—would provide complementary narrative signals. Finally, expanding the framework to international reporting regimes would shed light on how cultural and regulatory differences influence disclosure practices and their predictive capacity.

References

1. Bogachek, O., De Vito, A., Demeré, P., & colleagues. (2025). Using narrative disclosures to predict tax outcomes. Review of Accounting Studies. Advance online publication. https://doi.org/10.1007/s11142-025-09914-3.
2. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321–357. https://doi.org/10.1613/jair.953.
3. Černevičienė, J., & Kabašinskas, A. (2024). Explainable artificial intelligence (XAI) in finance: A systematic literature review. Artificial Intelligence Review, 57, 216. https://doi.org/10.1007/s10462-024-10854-8
4. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321–357. https://doi.org/10.1613/jair.953
5. Feldman, R., Govindaraj, S., Livnat, J., & Segal, B. (2010). Management's tone change, post earnings announcement drift and accruals. Review of Accounting Studies, 15(4), 915–953. https://doi.org/10.1007/s11142-009-9106-2
6. Fengler, M. R., & Phan, T. M. (2025). Unveiling themes in 10-K disclosures: A new topic modeling perspective. International Review of Financial Analysis, 103, 104121. https://doi.org/10.1016/j.irfa.2025.104121
7. Hutto, C., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the 8th International AAAI Conference on Weblogs and Social Media (pp. 216–225). AAAI.
8. Jacksoft. (2025). JCAATs (Jacksoft Computer-Assisted Auditing Tools) [Computer software]. Jacksoft. https://www.jacksoft.com.tw/
9. Kacar, M. S., Yumusak, S., & Kodaz, H. (2023). A methodology on converting 10-K filings into a machine learning dataset and its applications. IEICE Transactions on Information and Systems, E106.D(4), 477–487. https://doi.org/10.1587/transinf.2022IIP0001
10. Li, F. (2010). The information content of forward-looking statements in corporate filings—A naïve Bayesian ma-

chine learning approach. Journal of Accounting Research, 48(5), 1049–1102. https://doi.org/10.1111/j.1475-679X.2010.00382.x

11. Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. Journal of Finance, 66(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

12. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

13. Uberti-Bona Marin, L. G., Rijsbosch, B., Spanakis, G., & Kollnig, K. (2025). Are companies taking AI risks seriously? A systematic analysis of companies' AI risk disclosures in SEC 10-K forms. arXiv preprint arXiv:2508.19313. https://doi.org/10.48550/arXiv.2508.19313

14. Zhang, C. (A.), Cho, S., & Vasarhelyi, M. (2022). Explainable artificial intelligence (XAI) in auditing. International Journal of Accounting Information Systems, 46, 100572. https://doi.org/10.1016/j.accinf.2022.100572

15. Zavitsanos, E., Spyropoulou, E., Giannakopoulos, G., & Paliouras, G. (2024). Machine learning for identifying risk in financial statements: A survey. ACM Computing Surveys, 57(9), Article 220, 1–37. https://doi.org/10.1145/3723157