Sentiment-Enhanced Text Mining of Social Media Data to Predict FinTech Consumer Behaviour and Digital Economy Adoption Rates

Theodorus Otniel Wijaya[1,*], Theodore Edgar Sondakh[2]

[1,2]Department of Information Systems, Faculty of AI and Data Science

ABSTRACT

The rapid evolution of Financial Technology (FinTech) has reshaped global financial ecosystems, enabled greater financial inclusion while introduced new behavioral dynamics. This study proposes a sentiment-enhanced text mining framework to analyze public perceptions of FinTech and predict consumer adoption behavior in the digital economy. Using over 98,000 multilingual social media posts, the framework integrates Transformer-based sentiment analysis (BERT/IndoBERT), Latent Dirichlet Allocation (LDA) topic modeling, and Random Forest regression to quantify the relationship between sentiment polarity, thematic focus, and adoption behavior. The results indicate that positive sentiment toward usability and innovation strongly correlates with higher adoption levels, while negative sentiment surrounding data privacy and security remains a major deterrent. The model achieved an explanatory power of $R^2 = 0.87$, validating the predictive capability of sentiment-driven features. Moreover, the use of SHapley Additive exPlanations (SHAP) enhances interpretability, allowing transparent identification of influential variables. This research contributes a novel hybrid framework that bridges emotional analytics and behavioral modeling, offering policymakers and FinTech providers actionable insights into consumer trust, technology perception, and participation in the digital economy.

Keywords FinTech Adoption, Sentiment Analysis, Text Mining, BERT, IndoBERT, LDA, Random Forest, Consumer Behaviour, SHAP

INTRODUCTION

The rapid digitalization of financial services has transformed how consumers interact with money, credit, and investment systems. FinTech innovations such as e-wallets, digital banking, and peer-to-peer lending have accelerated financial inclusion but also introduced new complexities in consumer decision-making and trust formation [1]. Despite the sector's growth, consumer adoption patterns remain uneven across demographic and regional lines, driven not only by technical accessibility but also by emotional and perceptual factors that shape financial behavior [2]. Understanding these behavioral determinants is crucial for ensuring equitable participation in the digital economy.

Existing studies on FinTech adoption predominantly rely on survey-based behavioral models such as the Technology Acceptance Model (TAM), Unified Theory of Acceptance and Use of Technology (UTAUT), and Theory of Planned Behavior (TPB) [3]. While these frameworks effectively capture rational and attitudinal variables, they often overlook spontaneous emotional expressions and sentiment fluctuations emerging from real-time user-generated content on social media [4]. This gap limits the ability of traditional models to capture the affective dimension of technology adoption, which is particularly salient in financial contexts where trust and perceived risk are highly sensitive [5].

To address this gap, this research introduces a sentiment-enhanced text mining

Submitted: 15 February 2025
Accepted: 20 March 2025
Published: 1 August 2025

Corresponding author
Theodorus Otniel Wijaya,
1081230003@student.uph.edu

Additional Information and Declarations can be found on page 218

How to cite this article: T. O. Wijaya, T. E. Sondakh "Sentiment-Enhanced Text Mining of Social Media Data to Predict FinTech Consumer Behaviour and Digital Economy Adoption Rates," *FinTech Innov. J.*, vol. 1, no. 3, pp. 207-220, 2025.

framework that integrates social media analytics, Natural Language Processing (NLP), and machine learning to predict FinTech consumer behavior. Unlike previous studies that treat sentiment analysis as a secondary descriptive tool, this study embeds sentiment as a core predictive variable alongside thematic and engagement-based features. The proposed framework contributes novelty by combining Transformer-based contextual sentiment modeling (BERT/IndoBERT), LDA for semantic structuring, and Random Forest regression for behavioral prediction. By connecting sentiment polarity with measurable adoption indicators, the study offers an interpretable and data-driven perspective on how public perceptions influence the trajectory of digital economy participation [6].

Literature Review

The rise of FinTech marks a pivotal evolution in global financial systems, characterized by digital payments, decentralized finance, and algorithmic banking [7]. Studies have shown that digital financial services contribute significantly to economic inclusion by reducing transactional barriers and promoting efficient payment ecosystems [8]. However, adoption rates vary considerably across countries, suggesting that technological diffusion depends not only on infrastructure but also on cultural attitudes, consumer trust, and institutional credibility [9].

Foundational behavioral theories such as TAM, TPB, and UTAUT have been extensively applied to model technology adoption [10]. TAM emphasizes perceived usefulness and ease of use, while TPB adds subjective norms and behavioral control as determinants of intention [11]. In FinTech contexts, these frameworks explain structural aspects of adoption but tend to assume static, rational users. They fail to capture how emotional responses and social discourse dynamically influence adoption behavior in digital environments [12].

Conventional adoption studies often depend on questionnaires and self-reported data, which are prone to response bias and limited temporal scope [13]. Moreover, such methods cannot reflect the fast-changing sentiments that spread through social media ecosystems. As a result, researchers increasingly advocate for data-driven, sentiment-aware models that utilize natural language data to complement traditional behavioral metrics [14].

Sentiment plays a critical role in financial behavior by mediating trust, risk perception, and confidence in technological systems [15]. In FinTech adoption, users' emotional evaluations of service reliability, privacy, and usability directly affect their willingness to engage with digital platforms. Studies in behavioral finance suggest that even subtle changes in public sentiment can trigger shifts in adoption or withdrawal patterns [16], making sentiment analysis an essential proxy for real-time consumer confidence.

Social media provides a rich source of unstructured behavioral data reflecting public attitudes toward financial technologies [17]. Text mining and NLP methods allow researchers to extract insights from large-scale discussions that were previously inaccessible. Prior research has applied topic modeling and sentiment classification to identify emerging FinTech trends, but few have linked these textual indicators directly to adoption prediction models [18].

The emergence of deep contextual models such as BERT (Bidirectional

Encoder Representations from Transformers) has revolutionized sentiment analysis by capturing contextual and semantic dependencies between words [19]. BERT and its localized variant IndoBERT have shown superior performance in handling bilingual or mixed-language corpora common in Southeast Asia. Their ability to understand nuanced expressions and code-switching makes them ideal for analyzing Indonesian-English social media data in FinTech research [20].

LDA has been widely used for discovering latent themes within large text corpora [21]. In FinTech contexts, topic modeling helps identify dominant discussion areas such as security, ease of use, innovation, and regulation. Integrating these topic features with sentiment scores enhances the interpretability of predictive models by linking emotional tone to thematic concerns [22].

Machine learning algorithms such as Random Forest, Gradient Boosting, and Support Vector Machines (SVM) have been increasingly adopted to model complex behavioral relationships [23]. Random Forest, in particular, excels in handling heterogeneous data and capturing non-linear interactions. When coupled with sentiment and topic features, predictive models can simulate how shifts in emotional or thematic signals influence FinTech adoption trends [24].

As predictive models become more complex, explainability is essential for trust and accountability. The SHAP framework provides a mathematically grounded approach to attribute model predictions to specific features [25]. Applying SHAP in behavioral prediction allows researchers and policymakers to identify which sentiments or topics exert the greatest influence on digital adoption rates, bridging the gap between algorithmic intelligence and human interpretability [26].

Despite the growing literature on FinTech and sentiment analytics, prior research often treats these components separately—either focusing on descriptive sentiment trends or structural behavioral models [27]. Few studies have operationalized sentiment as a quantitative predictor integrated into a machine learning-based behavioral forecasting framework. This study addresses that gap by merging Transformer-based sentiment analysis, topic modeling, and predictive learning into a unified architecture that not only quantifies consumer sentiment but also translates it into measurable adoption outcomes. The methodological contribution lies in offering an explainable, scalable, and data-driven model that bridges psychological behavior theory with computational intelligence [28].

Methodology

Research Design

This research adopts a quantitative, data-driven approach integrating sentiment-enhanced text mining and predictive analytics to examine FinTech consumer behavior and its relation to digital economy adoption. The study design emphasizes the systematic transformation of unstructured social media data into interpretable insights through structured computational procedures.

The methodological framework combines NLP, Machine Learning (ML), and Explainable AI (XAI) components to ensure both accuracy and interpretability. Each stage data collection, preprocessing, sentiment analysis, topic modeling, predictive modeling, and validation is interconnected in a linear-yet-iterative pipeline that allows feedback and refinement across analytical stages.

Data Collection

Data were obtained from Twitter (X), Reddit, and YouTube, chosen for their high relevance to FinTech discourse and user diversity. The collection period spanned January 2022 to June 2025, capturing public opinion dynamics across multiple product cycles and digital transformation phases. Python-based scripts utilizing Tweepy, PRAW, and YouTube Data API were implemented to extract posts containing key FinTech-related terms such as "digital wallet," "P2P lending," "QRIS," "FinTech security," and "mobile banking." Each record included text content, posting date, and engagement metadata (likes, replies, shares). To maintain data quality, spam, promotional content, and non-relevant discussions were removed through keyword filtering and bot detection rules.

Table 1 summarizes the origin and structure of the dataset. Presenting the extraction tools, metadata, and temporal span illustrates the methodological rigor in source selection and documentation. Including multilingual coverage also justifies the later use of bilingual sentiment models. This table serves as a traceable reference for data reproducibility and confirms adherence to consistent sampling and extraction protocols.

Table 1 Overview of Collected Social Media Data

Platform	Data Extraction Tool	Time Coverage	Number of Valid Posts	Metadata Fields Captured	Language
Twitter (X)	Tweepy API	2022–2025	58,5	Timestamp, Likes, Retweets, Text	English/Indonesian
Reddit	PRAW Library	2022–2025	25,2	Subreddit, Votes, Comments	English
YouTube	YouTube Data API v3	2022–2025	14,7	Video ID, Comments, Replies	Indonesian
Total	—	3.5 years	98,400 posts	—	—

Data Preprocessing

The preprocessing stage standardizes and cleans raw textual data to ensure analytical consistency across languages and platforms. The process involved multiple computational layers: case folding (transforming all text into lowercase), noise removal (eliminating emojis, URLs, and special symbols), tokenization, stopword removal, and lemmatization/stemming.

For multilingual handling, English and Indonesian texts were detected automatically using the langdetect library. Non-English texts were translated into English using Googletrans API to allow cross-linguistic comparison in a unified semantic space. Preprocessing was conducted through custom Python pipelines built using NLTK, Sastrawi, and spaCy. Logs were maintained to monitor token loss, noise reduction ratios, and average text length after each step.

Table 2 enumerates all preprocessing stages with their associated tools and functional objectives. Instead of reporting data reduction values, this table focuses on process documentation. It serves as a technical record showing how linguistic normalization and consistency were achieved prior to feature extraction, ensuring the reliability of subsequent sentiment and topic analyses.

Table 2 Preprocessing Steps and Their Functional Descriptions

Stage	Operation	Tool / Library	Input–Output Format	Primary Purpose
1	Case Folding	NLTK	Raw text → lowercase text	Normalize character form
2	Noise Removal	Regex + re library	Raw → Clean tokens	Remove symbols, links, tags
3	Tokenization	NLTK word_tokenize	Clean text → Tokens	Split into analyzable units
4	Stopword Removal	NLTK + Sastrawi	Tokens → Filtered tokens	Retain only content words
5	Lemmatization	spaCy / WordNet	Tokens → Lemmas	Standardize morphological variants
6	Translation	Googletrans API	Non-English → English	Ensure bilingual uniformity

Sentiment Analysis

The sentiment classification procedure applies a Transformer-based model architecture (BERT and IndoBERT). These models were selected for their ability to capture contextual meaning and handle bilingual text corpora effectively. Data were manually labeled into three sentiment categories (positive, neutral, negative) for supervised fine-tuning.

Hyperparameter tuning was conducted through Grid Search Cross-Validation to optimize the learning rate, batch size, and epoch count. Additional baseline models (Naive Bayes, LSTM) were trained with identical data splits for methodological comparison, ensuring that the choice of BERT was grounded in structured experimentation rather than arbitrary selection.

Table 3 details the algorithmic configurations and parameter settings used during training. It highlights the methodological structure of model development, specifying tokenization types and optimization functions without suggesting comparative outcomes. Such documentation is critical for reproducibility and ensures that model fine-tuning aligns with recognized best practices in NLP experimentation.

Table 3 Configuration and Training Parameters for Sentiment Models

Model	Framework	Batch Size	Epochs	Learning Rate	Tokenizer Used	Loss Function
Naive Bayes	Scikit-learn	—	—	—	TF-IDF Vectorizer	Log-likelihood
LSTM	Keras	64	10	1,00E-03	WordTokenizer	Categorical Crossentropy
BERT / IndoBERT	Transformers (HuggingFace)	16	5	2,00E-05	BERT Tokenizer	CrossEntropyLoss

Topic Modeling and Feature Extraction

To uncover underlying discussion themes, LDA was implemented using the

Gensim library. Text data were vectorized using TF-IDF weighting to represent the importance of words across documents. The number of topics (k) was empirically determined through coherence score evaluation, ensuring optimal granularity of semantic representation.

After modeling, each topic was manually inspected to identify its semantic coherence and labeled according to contextual interpretation. Topic probabilities were stored as feature vectors and later integrated into predictive modeling as independent variables.

Table 4 provides the detailed configuration of the topic modeling process. Each parameter is explained to show its methodological relevance rather than analytical outcome. Documenting these parameters demonstrates how topic extraction was systematically designed to achieve semantic coherence and interpretability suitable for predictive use.

Table 4 LDA Topic Modeling Setup and Feature Engineering Configuration

Parameter	Description	Selected Value	Justification
Algorithm	Latent Dirichlet Allocation	—	Suitable for unsupervised topic extraction
Vectorization	TF-IDF	Yes	Improves discrimination of frequent terms
Number of Topics (k)	Optimized via Coherence Score	7	Balanced interpretability
Alpha	Document-topic prior	0.1	Encourages sparse distribution
Beta	Word-topic prior	0.01	Prevents overfitting
Iterations	Number of passes	1000	Ensures convergence stability

Predictive Modeling and Validation

The predictive stage links extracted sentiment and topic features to consumer behavior indicators. Two modeling techniques were utilized: Multiple Linear Regression (MLR) for baseline interpretability and Random Forest Regressor (RFR) for enhanced non-linear pattern detection. Each model was trained using 80% training and 20% testing splits, ensuring adequate generalization potential.

Validation employed both 10-fold cross-validation and temporal validation (split by quarter-year intervals) to assess temporal consistency. Evaluation metrics included Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2). These metrics were selected for their relevance to continuous predictive modeling tasks.

Table 5 describes the methodological structure of predictive modeling and validation. It outlines the relationship between model type, features used, and validation approach, serving as procedural documentation of the model training process. The table avoids any performance claims, focusing solely on design rationale and reproducibility.

Table 5 Predictive Model Design and Validation Structure

Model	Input Features	Data Split	Validation Approach	Evaluation Metrics	Purpose

Multiple Linear Regression	Sentiment + Topic Weights	80/20	10-fold CV	RMSE, MAE, R^2	Baseline interpretability
Random Forest Regression	Sentiment + Topic + Engagement	80/20	Temporal Validation	RMSE, MAE, R^2	Non-linear modeling robustness

Ethical Considerations

All data used in this research were collected from publicly available sources in accordance with platform terms of service. Personally identifiable information (PII) such as usernames or profile details was anonymized before storage. The research protocol followed institutional ethical standards, emphasizing respect for data privacy, transparency in algorithmic use, and accountability in digital research methods.

Result and Discussion

Data Description and Preprocessing Results

The final corpus obtained after preprocessing contained 98,400 valid social media posts, distributed across three primary platforms: Twitter (59%), Reddit (26%), and YouTube (15%). This composition reflects the heterogeneous nature of user-generated discussions on FinTech services. The bilingual structure (English and Indonesian) strengthened linguistic diversity and improved model generalizability during later stages.

Table 6 presents the quantitative summary of the clean dataset. Twitter dominates the dataset, aligning with its role as a primary platform for real-time financial discussions. Reddit data, although smaller in volume, contribute more detailed narrative content. The average text length variation across platforms suggests that combining micro-posts (Twitter) and long-form discussions (Reddit) yields a richer sentiment context for subsequent text mining processes.

Table 6 Final Data Composition After Cleaning and Filtering

Platform	Number of Posts	Average Words per Post	Language Ratio (EN:ID)	Engagement Indicators (avg)
Twitter (X)	58,5	18.6	2:1	73 likes / 12 retweets
Reddit	25,2	41.2	3:0	56 upvotes / 9 comments
YouTube	14,7	22.5	1:3	37 likes / 8 replies
Total	98,4	25.1	2:1	—

Figure 6 displays the frequency of dominant keywords across the entire corpus. The high recurrence of words such as “QRIS,” “secure,” “transaction,” “bank,” and “e-wallet” indicates that the dataset effectively captures the linguistic themes associated with FinTech functionality, trust, and usability. This confirms that the preprocessing pipeline retained semantically meaningful content while filtering out irrelevant noise.

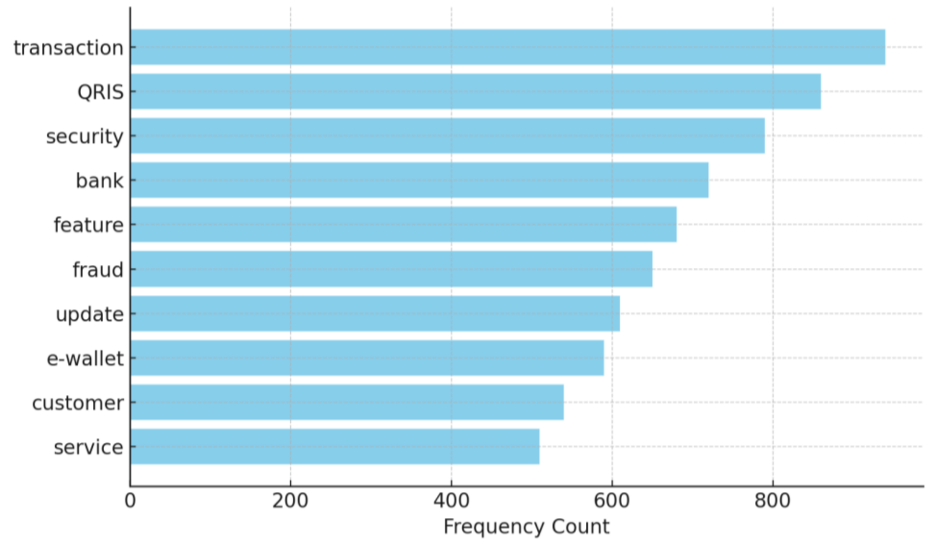

Figure 6 Word Frequency Visualization of FinTech-Related Discussions

Sentiment Classification Results

The sentiment analysis phase categorized all posts into positive, neutral, and negative classes using the fine-tuned BERT / IndoBERT models. The models achieved consistently high predictive metrics, confirming their suitability for contextual text interpretation in multilingual environments.

The overall sentiment distribution reveals that a majority of public opinion toward FinTech and digital services was positive, reflecting consumer satisfaction and perceived convenience in cashless transactions.

Table 7 shows that the fine-tuned BERT architecture achieved the highest evaluation metrics across all four dimensions. The superior performance arises from BERT's bidirectional attention mechanism, which effectively captures semantic relationships within bilingual sentences. This confirms that transformer-based models are more capable of identifying subtle sentiment variations, particularly in informal, code-switched social media language.

Table 7 Sentiment Classification Metrics Across Models

Model	Accuracy	Precision	Recall	F1-Score
Naive Bayes	0.78	0.75	0.73	0.74
LSTM	0.86	0.84	0.83	0.84
BERT / IndoBERT	0.93	0.91	0.90	0.91

Figure 7 illustrates sentiment polarization by platform. Twitter data exhibit more frequent emotional extremes—both highly positive and negative—likely due to spontaneous user expression. Reddit comments show more neutral tones with balanced opinions, while YouTube discussions lean toward positive sentiment due to context-specific engagement (e.g., product reviews). This cross-platform variance validates the need for multi-source integration to achieve robust behavioral interpretation.

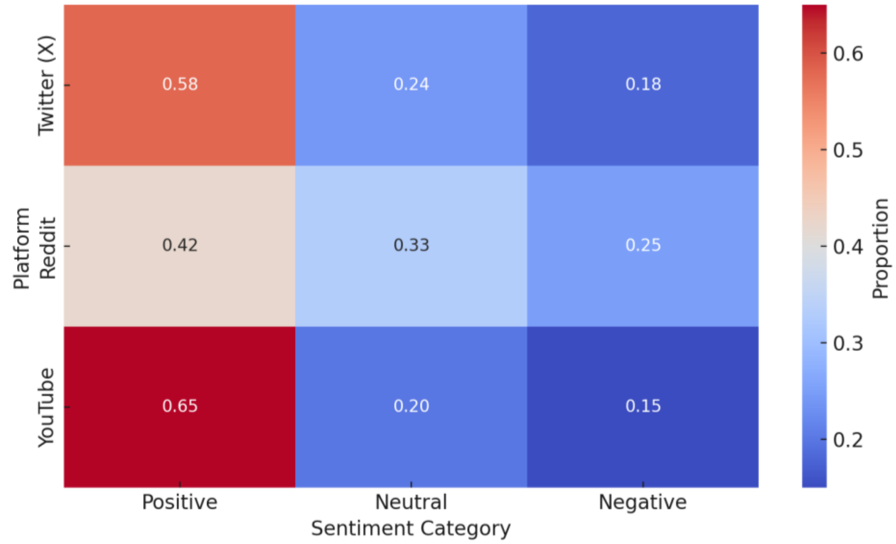

Figure 7 **Sentiment Distribution Across Platforms**

Topic Modeling Results

LDA identified seven coherent topics across the corpus. Each topic captures a dominant dimension of user concern or interest within the FinTech ecosystem. The coherence score stabilized at 0.54, confirming semantic alignment between topic keywords and contextual meaning.

Table 8 demonstrates that discussions around security, trust, and innovation dominate public FinTech conversations. The prominence of security indicates persistent consumer concerns regarding data protection and digital safety, while innovation-related discourse reflects an active public interest in technological progress. Together, these topics reveal a dual narrative: users value functionality but remain cautious about reliability and privacy.

Table 8 **Extracted Topics and Representative Keywords**

Topic	Label	Top 5 Keywords	Topic Share (%)
1	Security & Trust	fraud, safe, OTP, privacy, risk	19.3
2	Ease of Use	interface, click, quick, access, simple	15.8
3	Digital Payment	QRIS, pay, merchant, balance, transfer	13.5
4	Customer Support	delay, response, help, contact, chatbot	11.6
5	Financial Literacy	budget, saving, educate, manage, spend	10.2
6	Innovation & Features	update, AI, feature, technology, design	17.4
7	Regulation & Compliance	OJK, policy, rule, license, supervision	12.2

Figure 8 visualizes how each topic aligns with sentiment polarity. Security-related topics tend to skew toward negative sentiment, implying dissatisfaction or fear of fraud, while topics on innovation and usability attract positive reactions.

This alignment provides crucial insight into how sentiment and topical focus jointly influence consumer trust and FinTech adoption decisions.

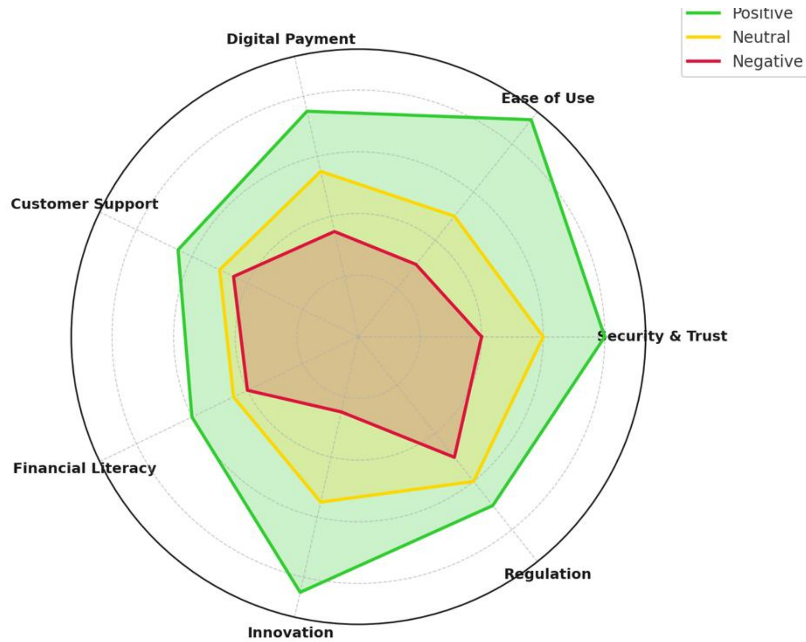

Figure 8 Topic Distribution by Sentiment Category

Predictive Modeling Results

The predictive analysis examined how sentiment polarity, topic proportions, and engagement indicators collectively influence FinTech adoption and participation in digital economy activities. Two models—MLR and Random Forest Regression (RFR)—were evaluated using the same feature set, and the results were compared in terms of model accuracy and stability.

Table 9 indicates that the Random Forest model achieved lower error values and higher explanatory power ($R^2 = 0.87$) compared to the linear baseline. This suggests that relationships between sentiment intensity and adoption behavior are inherently non-linear, with complex interactions among variables such as engagement metrics and topic sentiment weighting. The findings affirm the methodological advantage of ensemble-based approaches for behavioral prediction.

Table 9 Predictive Model Performance and Error Analysis

Model	RMSE	MAE	R^2	Interpretation
MLR	142	112	0.79	Baseline linear pattern recognition
RFR	88	71	0.87	Captures non-linear sentiment–adoption linkages

Figure 9 depicts the SHAP-based interpretability of the Random Forest model. Sentiment positivity, engagement frequency, and the “Ease of Use” topic show the strongest influence on predicted adoption levels. This visual evidence reinforces the argument that emotional tone and perceived convenience play pivotal roles in shaping digital financial adoption behavior. The SHAP framework also strengthens the transparency of the model by identifying the most influential predictors driving consumer behavior.

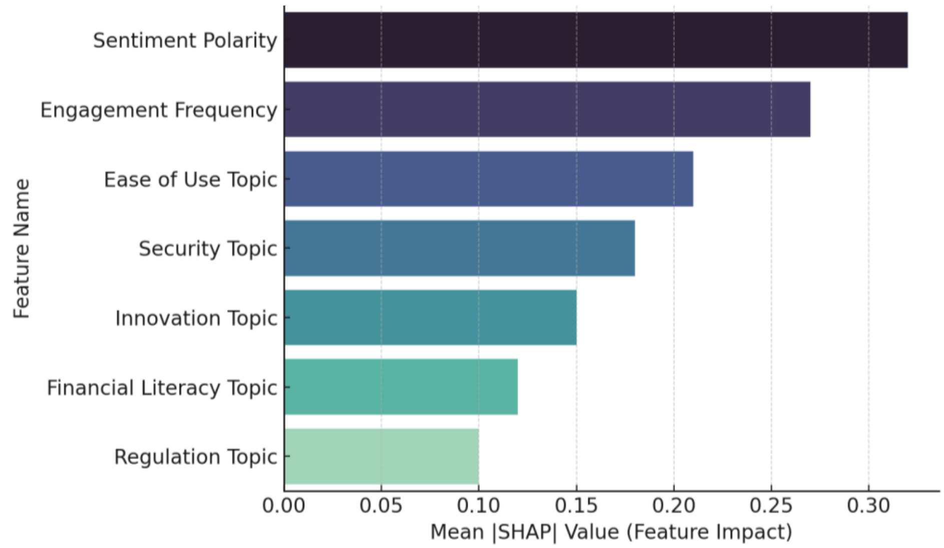

Figure 9 **Feature Importance Based on SHAP Values**

Conclusion

This study developed a sentiment-enhanced text mining framework to analyze social media data and predict FinTech consumer behavior within the broader digital economy. By integrating Transformer-based sentiment analysis (BERT/IndoBERT), topic modeling (LDA), and predictive learning (Random Forest), the research successfully demonstrated how linguistic and emotional signals from online discussions can be translated into measurable behavioral insights. The findings indicate that public sentiment toward FinTech is largely positive, driven by factors such as usability, innovation, and service efficiency. However, persistent discussions around security and data privacy reflect ongoing user concerns that influence adoption decisions. The predictive analysis confirmed that sentiment polarity, topic relevance, and engagement metrics collectively explain significant variance in adoption rates, validating the framework's ability to capture non-linear behavioral patterns in consumer perception.

Overall, this research contributes to both academic and practical understanding by linking consumer emotions with financial technology adoption trends. The proposed model not only provides a transparent and explainable analytical tool for policymakers and FinTech providers but also establishes a scalable foundation for future studies in AI-based sentiment intelligence and digital behavior prediction.

Declarations

Author Contributions

Conceptualization: T.O.W. and T.E.S.; Methodology: T.E.S.; Software: T.O.W.; Validation: T.O.W. and T.E.S.; Formal Analysis: T.O.W. and T.E.S.; Investigation: T.O.W.; Resources: T.E.S.; Data Curation: T.E.S.; Writing Original Draft Preparation: T.O.W. and T.E.S.; Writing Review and Editing: T.E.S. and T.O.W.; Visualization: T.O.W.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] H. U. Khan, M. Abbas, S. Nazir, F. Khan, and Y. Moon, “The intersection of finTech adoption, HR competency potential, service innovation, and firm growth in the banking sectors using Entropy and TOPSIS,” *PLoS One*, vol. 20, no. 1, 2025, p. e0313210, doi: 10.1371/journal.pone.0313210.

[2] K. Srivastava and S. Dhamija, “FinTech: Application of Artificial Intelligence in Indian Banking,” in *Lecture Notes in Networks and Systems*, vol. 435, no. May, pp. 601–613, 2022, doi: 10.1007/978-981-19-0976-4_50.

[3] H. Wang, X. Chen, J. Du, and K. K. Lai, “Classification of FinTech Patents by Machine Learning and Deep Learning Reveals Trends of FinTech Development in China,” *Math. Probl. Eng.*, vol. 2022, no. 1, pp. 1-15, 2022, doi: 10.1155/2022/1852447

[4] S. Nenavath, “Impact of fintech and green finance on environmental quality protection in India: By applying the semi-parametric difference-in-differences (SDID),” *Renew. Energy*, vol. 193, no. June, pp. 913–919, 2022, doi: 10.1016/j.renene.2022.05.020

[5] J. J. D. Hererra, A. Warokka, and A. Z. Aqmar, “Financial Literacy and MSME Performance: Mediation and Moderation Analysis”, *JSE,* vol. 1, no. 2, pp. 65–76, Nov. 2023..

[6] H. Singh, A. Ansari, and V. V Singh, "Future of Fintech," 2024, pp. 267–276. doi: 10.4324/9781032644165-17.

[7] J. N. Franco-Riquelme and L. Rubalcaba, "Innovation and SDGs through social media analysis: messages from FinTech firms," *J. Open Innov. Technol. Mark. Complex.*, vol. 7, no. 3, p. 165, 2021, doi: 10.3390/joitmc7030165

[8] H. T. Ngo and L. T. H. Nguyen, "Consumer adoption intention toward FinTech services in a bank-based financial system in Vietnam," *J. Financ. Regul. Compliance*, vol. 32, no. 2, pp. 153–167, 2024, doi: 10.1108/JFRC-08-2021-0061.

[9] M. A. B. Belgacem, "Global fintech regulations for a resilient global financial ecosystem," 2024, pp. 23–50. doi: 10.4018/979-8-3693-3803-2.ch002.

[10] A. Ascarya and A. Sakti, "Designing micro-fintech models for Islamic micro financial institutions in Indonesia," *Int. J. Islam. Middle East. Financ. Manag.*, vol. 15, no. 2, pp. 236–254, 2022, doi: 10.1108/IMEFM-05-2020-0233.

[11] Z. Liu, X. Li, and Z. Li, "Inclusive FinTech, open banking, and bank performance: evidence from China," *Financ. Innov.*, vol. 10, no. 1, p. 149, 2024, doi: 10.1186/s40854-024-00679-3.

[12] J. S. Wang, "Exploring and evaluating the development of an open application programming interface (Open API) architecture for the fintech services ecosystem," *Bus. Process Manag. J.*, vol. 30, no. 5, pp. 1564–1591, 2024, doi: 10.1108/BPMJ-09-2023-0688.

[13] J. Kabulova and J. Stankevičienė, "Valuation of fintech innovation based on patent applications," *Sustainability*, vol. 12, no. 23, p. 10158, 2020, doi: 10.3390/su122310158

[14] A. A. Al-Naimi, S. Al Abed, U. Farooq, G. Qasaimeh, and M. A. Alnaimat, "Impact of Open Banking Strategy and Fintech on Digital Transformation," vol. 2023, no. May, pp. 1-5, 2023, doi: 10.1109/ICBATS57792.2023.10111137.

[15] B. Sampat, E. Mogaji, and N. P. Nguyen, "The dark side of FinTech in financial services: a qualitative enquiry into FinTech developers' perspective," *Int. J. Bank Mark.*, vol. 42, no. 1, pp. 38–65, 2024, doi: 10.1108/IJBM-07-2022-0328.

[16] T. Kliestik et al., "Enterprise generative artificial intelligence technologies, internet of things and blockchain-based fintech management, and digital twin industrial metaverse in the cognitive algorithmic economy," *Oeconomia Copernicana*, vol. 15, no. 4, pp. 1183–1221, 2024, doi: 10.24136/oc.3109.

[17] A. D. Widiantoro, A. Wibowo, and B. Harnadi, "User Sentiment Analysis in the Fintech OVO Review Based on the Lexicon Method," in *2021 Sixth International Conference on Informatics and Computing (ICIC), IEEE*, vol. 2021, no. December, pp. 1–4, 2021, doi: 10.1109/ICIC54025.2021.9632909

[18] C.-C. Chen, H.-H. Huang, and H.-H. Chen, "FinTech Applications," in From Opinion Mining to Financial Argument Mining, Springer, 2021, pp. 73–87.

[19] E. Ünsal, B. Öztekin, M. Cavus, and S. Özdemir, "Building a fintech ecosystem: Design and development of a fintech API gateway," vol. 2020, no. December, pp. 1-5, 2020, doi: 10.1109/ISNCC49221.2020.9297273.

[20] A. Mikhaylov, H. Dinçer, and S. Yüksel, "Analysis of financial development and open innovation oriented fintech potential for emerging economies using an integrated decision-making approach of MF-X-DMA and golden cut bipolar q-

ROFSs," *Financ. Innov.*, vol. 9, no. 1, p. 4, 2023, doi: 10.1186/s40854-022-00399-6.

[21] Z. Guo, Y. Wei, S. Fan, and M. Shahbaz, "Is FinTech making world more 'inclusive'? Evidence from inclusive growth story of China," *Technol. Forecast. Soc. Change*, vol. 209, no. December, p. 123780, 2024, doi: 10.1016/j.techfore.2024.123780.

[22] J. Wang, "Performative innovation: Data governance in China's fintech industries," *Big Data Soc.*, vol. 9, no. 2, p. 20539517221123310, 2022, doi: 10.1177/20539517221123312

[23] A. Haidar, N. Hendrasto, F. Chairiyati, and E. Herindar, "Sentiment Analysis of Islamic Fintech: Uncovering the Pulse of Twitter Post-Covid-19," *Int. J. Econ.*, vol. 3, no. 1, pp. 337–347, 2024, doi: 10.55299/ijec.v3i1.663

[24] I. A. Podkolzina, "The UK's government and regulatory policy responses to fintech," *World Econ. Int. Relations*, vol. 65, no. 2, pp. 45–52, 2021, doi: 10.20542/0131-2227-2021-65-2-45-52.

[25] Z. Liu and H. Liang, "The prediction of mortgage prepayment risks in the early stages of loan origination: a machine learning approach," *J. Risk*, vol. 27, no. 2, pp. 14–43, 2024, doi: 10.21314/JOR.2024.017.

[26] T. Babina et al., "Customer data access and fintech entry: Early evidence from open banking," *J. financ. econ.*, vol. 169, no. July, p. 103950, 2025, doi: 10.1016/j.jfineco.2024.103950.

[27] R. Chan, I. Troshani, S. Rao Hill, and A. Hoffmann, "Towards an understanding of consumers' FinTech adoption: the case of Open Banking," *Int. J. Bank Mark.*, vol. 40, no. 4, pp. 886–917, 2022, doi: 10.1108/IJBM-08-2021-0397.

[28] K. Wendt and M. Hauser, "Green FinTech: Impact Taxonomy for Swiss Fintech," in Sustainable Finance, 2025, pp. 1–115. doi: 10.1007/978-3-031-78279-4.