DECODE: JURNAL PENDIDIKAN TEKNOLOGI INFORMASI
P-ISSN: 2775-2984 | E-ISSN: 2775-1813
Vol. 5, No 3, November 2025, Hal: 1169-1178
Doi: http://dx.doi.org/10.51454/decode.v5i3.1456
https://journal.umkendari.ac.id/index.php/decode

Reframing User Acceptance with User-Generated Content: Insights from Digital Banking in Indonesia

Herbert Siregar[1], Munir[2], Ade Sobandi[1], Lala Septem Riza[2], Samialloi Nusratullo[3]

[1]Management, Universitas Pendidikan Indonesia, Indonesia
[2]Computer Science Education, Universitas Pendidikan Indonesia, Indonesia
[3]Computer Science, Borough of Manhattan Community College, United States

Artikel Info

Keywords:

Digital Banking;
Machine Learning;
User Acceptance;
User Generated Content.

Article History:

Submitted: October 22, 2025
Accepted: November 25, 2025
Published: November 25, 2025

Abstract: This study examines user acceptance of BNI's newly launched digital banking application, WONDR, by leveraging user-generated content (UGC) obtained from the Google Play Store. A total of 60,434 user reviews were collected and analyzed using a supervised machine learning approach, specifically the Random Forest algorithm, to classify sentiments into positive and negative categories. The dataset underwent a rigorous preprocessing pipeline, including text normalization, tokenization, stopword removal, and stemming, followed by feature extraction using the Term Frequency–Inverse Document Frequency (TF–IDF) method. Model performance was evaluated through a 7-Fold Cross-Validation strategy and standard metrics such as accuracy, precision, recall, and F1-score, achieving strong and consistent results with an overall accuracy of 0.888. The findings indicate that UGC-based sentiment analysis provides a scalable and interpretable method for assessing user acceptance, offering actionable insights for strategic application development. Theoretically, this study extends traditional acceptance models (e.g., TAM and UTAUT) by operationalizing user perceptions through naturally occurring feedback. Practically, the proposed approach supports data-driven decision-making for enhancing digital service quality in the banking sector.

Corresponding Author:
Herbert Siregar
Email: herbert@upi.edu

INTRODUCTION

The concept of user acceptance has long been recognized as a critical determinant in the successful implementation of information technology. Through the Technology Acceptance Model (TAM), Davis et al. (1989) posited that technology adoption is largely shaped by two primary perceptions: perceived usefulness and perceived ease of use. Since then, research on user acceptance has expanded considerably with the development of various models. For example, the Unified Theory of Acceptance and Use of Technology (UTAUT), proposed by Venkatesh et al. (2003), incorporates additional dimensions such as social influence and facilitating conditions. In parallel, practical instruments such as the System Usability Scale (SUS) (Brooke, 1996) and the User Experience Questionnaire (UEQ) (Laugwitz et al., 2008) have been widely adopted to systematically assess usability and user experience. These models and tools have been applied across diverse domains, including education (Al-Adwan, 2020; Al-Adwan et al., 2023; Zogheib & Zogheib, 2024) and healthcare (Rouidi et al., 2022; Zin et al., 2023), consistently demonstrating their explanatory power in understanding technology adoption.

Nevertheless, traditional survey-based approaches have several limitations. According to Newby et al. (2003), the process of collecting data through questionnaires requires considerable time, cost, and effort to obtain a representative sample. Moreover, survey data are often distorted by response bias and social desirability bias, whereby respondents tend to provide answers perceived as socially acceptable rather than reflecting their actual experiences (Fricker & Schonlau, 2002). Another limitation is that surveys are unable to capture the dynamic nature of user opinions in real time, even though digital applications are typically updated in rapid cycles that may cause user perceptions to shift over time.

As an alternative, user-generated content (UGC) in the form of application reviews on the Google Play Store has emerged as a valuable data source for assessing user acceptance. According to Zaghloul et al. (2024), such reviews are voluntarily written based on real experiences, thereby providing opinions that are both authentic and rich in emotional nuance. Several studies have underscored the strategic value of analyzing app reviews. For instance, Alanzi (2021) emphasized that feedback from the Google Play Store can be leveraged to identify an application's strengths and weaknesses, while Alsaleh et al. (2025) demonstrated that sentiment analysis of reviews supports the evaluation of usability aspects. Furthermore, Fei et al. (2022) developed a predictive approach that uses reviews to estimate app ratings, highlighting that UGC can function as a continuous "natural survey."

To process large volumes of review data, machine learning approaches are particularly relevant. Bashiri and Naderi (2024) define sentiment analysis as the process of extracting opinions from text and classifying them into distinct categories, such as positive, neutral, or negative. Various algorithms have been employed in prior studies, including Naïve Bayes (Danyal et al., 2024) and Support Vector Machines (SVM) (Kurani et al., 2023). However, Random Forest provides several advantages over other algorithms. Li et al. (2022) emphasize that Random Forest performs well with high-dimensional data, yields stable results, and is relatively robust against overfitting. A comparative study by Neogi et al. (2021) further demonstrated that Random Forest achieved higher accuracy than SVM in opinion classification within the context of Twitter data. These findings highlight the potential of Random Forest for analyzing digital banking application review data.

Although the literature on user acceptance is extensive, important gaps remain to be addressed. Most studies on technology adoption continue to rely on classical survey-based models such as TAM, UTAUT, SUS, and UEQ, with limited exploration of user-generated content as an indicator of acceptance. Research on sentiment analysis in the context of digital applications in Indonesia—particularly within state-owned banks such as BNI—also remains scarce. Moreover, few studies have explicitly integrated machine learning–based analytics with conceptual frameworks of user acceptance to evaluate the adoption of digital banking applications.

Based on this background, the present study investigates the newly launched WONDR BNI application (https://www.bni.co.id/id-id) as a case study. From the 639,000 user reviews available on the Google Play Store (https://play.google.com/store/search?q=bni+wondr&c=apps&hl=id), a total of

60,434 reviews were collected and classified into positive and negative categories of user acceptance using the Random Forest algorithm. The objectives of this study are threefold: (1) to develop a machine learning–based classification model for user acceptance, (2) to evaluate the model's performance using accuracy, precision, recall, and F1-score metrics, and (3) to describe the distribution of user acceptance as an indicator of the application's success. This research is expected to contribute theoretically by extending the literature on user acceptance through a natural, data-driven approach and practically by providing BNI with insights into actual user perceptions to support strategic planning for future digital application development.

METHOD

Figure 1 presents the research methodology workflow used to achieve the study's objectives and address the research questions, starting from dataset preparation through to deployment and predictive analysis.

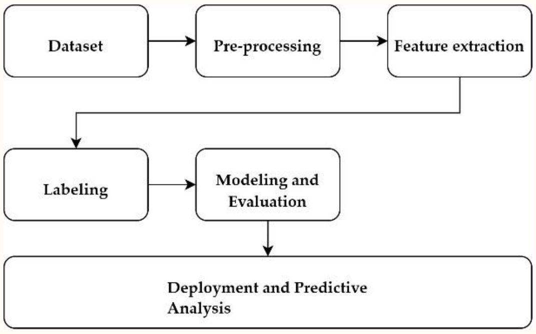

Figure 1. Research Methodology workflow

Dataset

This study adopts a quantitative research design employing supervised machine learning, as outlined by Ren et al. (2023), to evaluate user acceptance of BNI's WONDR application. Secondary data were obtained from the Google Play Store by collecting user-generated reviews published up to August 2025. These reviews were chosen because they represent authentic, voluntarily expressed user experiences following direct interaction with the application. The data were compiled into a structured dataset using the Google Play Scraper library in Python, with the following parameters: id = 'id.bni.wondr', lang = 'id', count = 65,000, and country = 'id'. The collected dataset was exported in .CSV format and contained three primary attributes: review text, star rating, and publication timestamp.

Pre-processing

Before analysis, the raw data underwent systematic pre-processing to ensure accuracy and reliability (Manyol et al., 2022). The text was cleaned by converting all characters to lowercase and removing punctuation, numbers, and non-essential symbols using efficient and flexible regular expressions. Tokenization was then applied to split the text into individual words, enabling structured linguistic representation and forming the basis for meaningful feature extraction in later modelling.

Common words (stopwords) were removed to improve the signal-to-noise ratio, and lexical normalization was performed through stemming with the Sastrawi library — a widely used Indonesian language processor (Suhaeni et al., 2025). This step reduced morphological variations by converting words to their root forms, ensuring semantic consistency across the corpus.

Feature extraction

Feature extraction was performed using the Term Frequency–Inverse Document Frequency (TF-IDF) method (Chawla et al., 2023), a statistical technique that evaluates the importance of a word within a document relative to the entire corpus. The calculation integrates the frequency of a word in a document (term frequency) with its distinctiveness across the collection (inverse document frequency).

As a result, words that occur too frequently are assigned lower weights, whereas more distinctive and informative words receive higher weights. The TF–IDF weighting scheme is mathematically defined as $\text{TF–IDF}(t, d) = \text{TF}(t, d) \times \log\left(\frac{N}{\text{DF}(t)}\right)$, where $\text{TF}(t, d)$ denotes the frequency of term t in document d; N represents the total number of documents in the corpus; and $\text{DF}(t)$ refers to the number of documents containing the term t.

Labeling

Following data cleansing and feature extraction, we retained only a subset of the original dataset to ensure a balanced distribution. The resulting score distribution, ranging from 1 to 5, comprised 11,014 entries with a score of 1, 2,678 entries with a score of 2, 3,021 entries with a score of 3, 3,920 entries with a score of 4, and 15,000 entries with a score of 5. For sentiment classification purposes, scores from 1 to 3 were grouped into the negative class (assigned label -1), while scores of 4 and 5 were categorized as the positive class (assigned label 1) (Bahtiar et al., 2023). Based on this grouping, the final distribution consisted of 16,713 entries (46.90%) in the negative class and 18,920 entries (53.10%) in the positive class. This proportion indicates that the class distribution has achieved a reasonably balanced state, which is considered suitable and uncommon for further analytical purposes (Bae et al., 2021).

Modeling and Evaluation

The classification framework was developed using the Random Forest algorithm, selected for its proven robustness in handling high-dimensional textual data, its consistent generalization performance, and its inherent resistance to overfitting (Xin & Ren, 2022). A total of 80% of the dataset—encompassing both positive and negative sentiment classes—was allocated for model training. The algorithm was configured with 400 decision trees and no predefined depth limitation, allowing it to effectively capture complex and nonlinear relationships within the data. Computational efficiency was enhanced through parallel execution across all available CPU cores, while a fixed random seed ensured full reproducibility of results. To mitigate class imbalance, the model employed automated class-weight adjustments within each bootstrap iteration, thereby promoting equitable learning between majority and minority categories.

Model validation was performed using a 7-Fold Cross-Validation strategy (Chamorro-Atalaya et al., 2023), in which the dataset was systematically partitioned into seven equally sized subsets, each representing approximately 14% of the total data volume. During each iteration, one subset served as the validation fold while the remaining six were used for training, ensuring comprehensive model evaluation across all data segments. Performance was assessed using widely recognized evaluation metrics (Tharwat, 2021), including accuracy $\left(\frac{TP+TN}{TP+TN+FP+FN}\right)$, precision $\left(\frac{TP}{TP+FP}\right)$, recall $\left(\frac{TP}{TP+FN}\right)$, and F1-score $\left(2 \cdot \frac{\text{Precision}\cdot\text{Recall}}{\text{Precision}+\text{Recall}}\right)$. Among these, the F1-score was prioritized as the primary indicator of model performance, as it provides a balanced measure of precision and recall—particularly under conditions of class imbalance.

In the final evaluation phase, model bias and overfitting potential were examined by comparing the aggregated performance of the 7-Fold Cross-Validation with the results obtained from training on the complete 80% dataset. The absence of significant deviation between these evaluations indicated satisfactory model stability and generalization capability. Consequently, the optimized Random Forest model, along with the corresponding TF-IDF feature representations, was stored for future deployment and downstream predictive analysis.

In this phase, the remaining 20% of the dataset, which was not used during the training process, was employed for prediction. The previously trained model and extracted feature representations were applied to this unseen data to generate class predictions. Model performance was evaluated using the same set of metrics applied during the validation stage, supplemented by the presentation of a confusion matrix to assess the model's overall accuracy, classification precision, and the degree of consistency between predicted and actual labels (Li et al., 2023).

RESULT AND DISCUSSION

Result

The 7-Fold Cross-Validation results summarized in Table 1 demonstrate the model's consistent performance across all folds. Accuracy values ranged from 0.870 to 0.888, while the corresponding F1-macro scores followed a nearly identical pattern, indicating balanced precision and recall across classes. The mean cross-validation accuracy was 0.880±0.0060.880, and the mean F1-macro score was 0.880±0.0060.880, reflecting both stability and reliability in model generalization. The low standard deviation across folds further suggests that the Random Forest classifier achieved robust predictive performance with minimal variance, confirming that the model effectively avoided overfitting and maintained consistent accuracy across different data partitions.

Table 1. Internal Validation Result

Fold	Accuracy	F1 macro
1	0.877	0.877
2	0.886	0.886
3	0.883	0.883
4	0.874	0.874
5	0.870	0.870
6	0.888	0.888
7	0.882	0.882
CV Accuracy (mean±sd): 0.880 ± 0.006		
CV F1-macro (mean±sd): 0.880 ± 0.006		

As shown in Table 2, the evaluation results indicate that the Random Forest classifier demonstrates strong and balanced predictive performance across both sentiment categories. The model effectively distinguishes between negative and positive classes, showing high precision and recall values that reflect a well-calibrated trade-off between accuracy and sensitivity. Overall, the model achieves consistent results across all metrics, confirming its robustness and reliability in handling moderately imbalanced datasets. These findings suggest that the model generalizes well to unseen data and is suitable for practical sentiment classification tasks requiring both interpretability and stable performance.

Table 2. Final Model Metrics

Description	Precision	Recall	F1-score	Support
Class Negatif (-1)	0.847	0.920	0.882	1337
Class Positive (1)	0.923	0.853	0.887	1514
accuracy			0.884	2851
macro avg	0.885	0.886	0.884	2851
weighted avg	0.887	0.884	0.884	2851

Based on the developed model, the results presented in Table 3 and Figure 2 demonstrate that the Random Forest classifier delivers highly reliable and balanced predictive performance. The confusion matrix indicates that most instances were correctly classified, with only a small proportion of misclassifications between the negative and positive classes. The model exhibits strong precision and recall across both categories, reflecting its ability to minimize false positives and false negatives simultaneously.

Overall, the accuracy rate approaching 0.89 confirms the model's stability and robustness in handling large-scale text data with moderate class imbalance. The similarity between the macro- and weighted-average scores suggests that the model performs consistently across classes without bias toward any particular category. These results collectively indicate that the Random Forest model generalises well to unseen data and provides a dependable framework for sentiment classification tasks requiring both precision and interpretability.

Table 3. Performance Metrics of Model Predictions

Description	**Precision**	**Recall**	**F1-score**	**Support**
Class Negatif (-1)	0.856	0.914	0.884	3343
Class Positive (1)	0.919	0.864	0.891	3784
accuracy			0.888	7127
macro avg	0.888	0.889	0.888	7127
weighted avg	0.890	0.888	0.888	7127

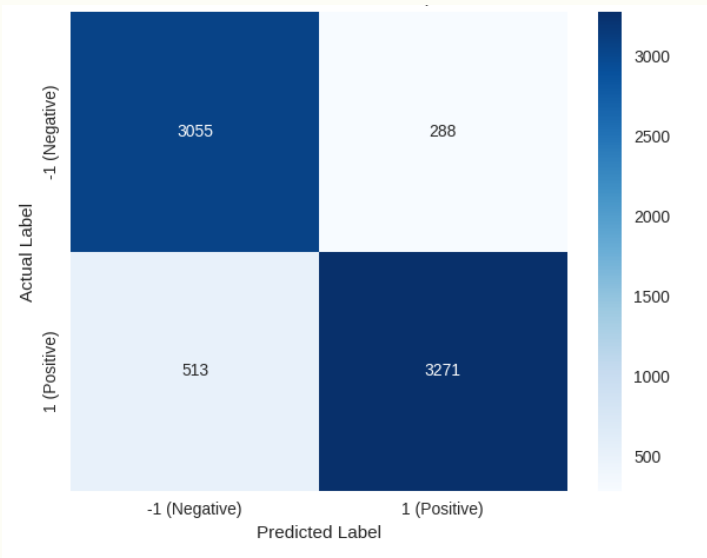

Figure 2. Confusion Matrix of Model Predictions

]To facilitate the visualization of frequently occurring terms within each class, a wordcloud representation was employed. The wordcloud in Figure 3 reveals a concentration of user complaints regarding mobile application performance, particularly in financial services. Dominant terms such as “masuk” (login), “verifikasi” (verification), “gagal” (failed), “transaksi” (transaction), and “gangguan” (interference) indicate recurring technical disruptions that hinder access and system reliability. These issues are further compounded by negative qualifiers such as “tidak” (not), “susah” (difficult), “sering” (often), “terus” (continuously), and “lama” (long), which reflect persistent dissatisfaction and emotional strain among users.

Figure 3. Word cloud of negative sentiment showing terms indicating user dissatisfaction

Subsequently, the wordcloud in Figure 4 reflects a predominantly positive user experience with the mobile application, emphasizing ease of use, responsiveness, and practical value. Prominent terms such as "mudah" (easy), "bagus" (good), "membantu" (helpful), and "sangat" (very) suggest high levels of user satisfaction and perceived utility. These expressions are reinforced by additional descriptors like "cepat" (fast), "praktis" (practical), "simpel" (simple), and "menarik" (interesting), which collectively indicate that the application is intuitive, efficient, and engaging.

The presence of emotionally affirming terms such as "mantap" (great) and "terimakasih" (thank you) further highlights spontaneous appreciation, often observed in unsolicited user reviews. Moreover, the inclusion of "transaksi" (transaction) and "aplikasinya" (the application) suggests that core functionalities—likely associated with financial services—are operating smoothly and effectively meeting user expectations.

Figure 4. Word cloud of positive sentiment showing terms reflecting user satisfaction

Discussion

The findings indicate that the Random Forest algorithm holds significant potential for sentiment classification based on user-generated content (UGC), particularly for measuring the acceptance of newly developed applications and evaluating users' initial experiences. This approach is especially valuable when an efficient, interpretable, and easily deployable model is required in business and organizational contexts.

However, model quality is not determined solely by the choice of algorithm but also by the rigour applied during the data preprocessing stage. Processes such as text cleaning, normalization, tokenization, stopword removal, and stemming—when executed systematically and consistently—have a substantial influence on the quality of feature representation. Well-structured preprocessing ensures clean and standardized data, enabling the model to learn sentiment patterns more effectively.

Moreover, as highlighted by Gasparetto et al. (2022), careful attention must be paid to the labelling system, particularly when employing semi-automated approaches such as score conversion. Inconsistent or biased labelling can compromise model validity and lead to misleading interpretations of user sentiment. Therefore, a hybrid approach that combines semi-automated and manual labelling is recommended to achieve higher precision through human oversight.

The capability of machine learning to process large-scale data and transform it into predictive models allows UGC, such as customer reviews, to be analyzed rapidly and accurately. The model developed in this study achieved strong performance, with an accuracy of 0.888, precision of 0.888, recall of 0.889, and an F1-score of 0.888—results that surpass those reported by Khan et al. (2022), who obtained an accuracy of 0.847, precision of 0.840, recall of 0.850, and an F1-score of 0.844. This comparison confirms the competitive performance of the proposed model in classifying user sentiment, particularly within the context of UGC analysis in digital service environments.

In the context of the rapid growth of Indonesia's banking industry (Santoso et al., 2021), which increasingly relies on digital applications, the implementation of UGC solutions represents a timely and effective approach for capturing user feedback. The classification of such responses plays a critical role in enabling organizations to identify both the strengths and weaknesses of their applications. Furthermore, user feedback serves as a dynamic indicator of progress over time, particularly in anticipating risks that may hinder the achievement of strategic organizational goals. These responses also facilitate meaningful interaction between organizations and their customers, especially when integrated with analytical Customer Relationship Management (CRM) systems. As demonstrated by Suhaeni et al., (2025), such systems can effectively extract relevant information to support decision-making processes.

CONCLUSION

This study developed a machine learning–based classification model to predict user acceptance of a digital application, assessed through accuracy, precision, recall, and F1-score. The classification results and acceptance distribution offer meaningful indicators of perceived application performance. Theoretically, the study extends TAM and UTAUT by showing how perceived usefulness, ease of use, performance expectancy, and effort expectancy can be inferred from naturally occurring user feedback. Practically, the model provides BNI with data-driven insights that support strategic decisions for improving digital services.

Despite its contributions, this study has limitations. Model performance depends on the quality and representativeness of user-generated content, which may reflect bias or incomplete information. The framework also relies solely on textual features and excludes contextual variables—such as demographics or longitudinal usage—that could strengthen predictive accuracy. Furthermore, although the study aligns with TAM and UTAUT, it does not empirically validate these constructs through structured instruments. Finally, the findings are specific to BNI's application and may not generalize to other platforms without further testing.

REFERENCES

Al-Adwan, A. S. (2020). Investigating the drivers and barriers to MOOCs adoption: The perspective of TAM. *Education and Information Technologies*, *25*(6), 5771–5795. https://doi.org/10.1007/s10639-020-10250-z

Al-Adwan, A. S., Li, N., Al-Adwan, A., Abbasi, G. A., Albelbisi, N. A., & Habibi, A. (2023). Extending The Technology Acceptance Model (TAM) To Predict University Students' Intentions To Use Metaverse-Based Learning Platforms. *Education and Information Technologies*, *28*(11), 15381–15413. https://doi.org/10.1007/s10639-023-11816-3

Alanzi, T. (2021). A Review Of Mobile Applications Available In The App And Google Play Stores Used During The COVID-19 Outbreak. *Journal of Multidisciplinary Healthcare*, *14*, 45–57. https://doi.org/10.2147/JMDH.S285014

Alsaleh, N., Alnanih, R., & Alowidi, N. (2025). Hybrid Deep Learning Approach For Automating App Review Classification: Advancing Usability Metrics Classification With An Aspect-Based Sentiment Analysis Framework. *Computers, Materials and Continua*, *82*(1), 949–976. https://doi.org/10.32604/cmc.2024.059351

Bae, S. Y., Lee, J., Jeong, J., Lim, C., & Choi, J. (2021). Effective Data-Balancing Methods For Class-Imbalanced Genotoxicity Datasets Using Machine Learning Algorithms And Molecular Fingerprints. *Computational Toxicology*, *20*. https://doi.org/10.1016/j.comtox.2021.100178

Bahtiar, S. A. H., Dewa, C. K., & Luthfi, A. (2023). Comparison of Naïve Bayes and Logistic Regression in Sentiment Analysis on Marketplace Reviews Using Rating-Based Labeling. *Journal of Information Systems and Informatics*, *5*(3), 915–927. https://doi.org/10.51519/journalisi.v5i3.539

Bashiri, H., & Naderi, H. (2024). Comprehensive Review And Comparative Analysis Of Transformer Models In Sentiment Analysis. *Knowledge and Information Systems*, *66*(12), 7305–7361. https://doi.org/10.1007/s10115-024-02214-3

Brooke, J. (1996). SUS—a Quick And Dirty Usability Scale. *Usability Evaluation in Industry*, *189*(194), 4–7. https://hell.meiert.org/core/pdf/sus.pdf

Chamorro-Atalaya, O., Arévalo-Tuesta, J., Balarezo-Mares, D., Gonzáles-Pacheco, A., Mendoza-León, O., Quipuscoa-Silvestre, M., Tomás-Quispe, G., & Suarez-Bazalar, R. (2023). K-Fold Cross-Validation through Identification of the Opinion Classification Algorithm for the Satisfaction of University Students. *International Journal of Online and Biomedical Engineering*, *19*(11), 140–158. https://doi.org/10.3991/ijoe.v19i11.39887

Chawla, S., Kaur, R., & Aggarwal, P. (2023). Text Classification Framework For Short Text Based On TFIDF-FastText. *Multimedia Tools and Applications*, *82*(26), 40167–40180. https://doi.org/10.1007/s11042-023-15211-5

Danyal, M. M., Khan, S. S., Khan, M., Ullah, S., Ghaffar, M. B., & Khan, W. (2024). Sentiment Analysis Of Movie Reviews Based On NB Approaches Using TF–IDF And Count Vectorizer. *Social Network Analysis and Mining*, *14*(1), 87. https://doi.org/10.1007/s13278-024-01250-9

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). Technology Acceptance Model. *J Manag Sci*, *35*(8), 982–1003.

Fei, H., Chua, T.-S., Li, C., Ji, D., Zhang, M., & Ren, Y. (2022). On The Robustness Of Aspect-Based Sentiment Analysis: Rethinking Model, Data, And Training. *ACM Transactions on Information Systems*, *41*(2), 1–32.

Fricker, R. D., & Schonlau, M. (2002). Advantages And Disadvantages Of Internet Research Surveys: Evidence From The Literature. *Field Methods*, *14*(4), 347–367. https://doi.org/10.1177/152582202237725

Gasparetto, A., Marcuzzo, M., Zangari, A., & Albarelli, A. (2022). A Survey On Text Classification Algorithms: From Text To Predictions. *Information*, *13*(2), 83.

Khan, L., Amjad, A., Afaq, K. M., & Chang, H.-T. (2022). Deep Sentiment Analysis Using CNN-LSTM Architecture Of English And Roman Urdu Text Shared In Social Media. *Applied Sciences*, *12*(5), 2694.

Kurani, A., Doshi, P., Vakharia, A., & Shah, M. (2023). A Comprehensive Comparative Study Of Artificial Neural Network (ANN) And Support Vector Machines (SVM) On Stock Forecasting. *Annals of Data Science*, *10*(1), 183–208. https://doi.org/10.1007/s40745-021-00344-x

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction And Evaluation Of A User Experience Questionnaire. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *5298 LNCS*, 63–76. https://doi.org/10.1007/978-3-540-89350-9_6

Li, C., Li, L., Zheng, J., Wang, J., Yuan, Y., Lv, Z., Wei, Y., Han, Q., Gao, J., & Liu, W. (2022). China's Public Firms' Attitudes Towards Environmental Protection Based On Sentiment Analysis And Random Forest Models. *Sustainability (Switzerland)*, *14*(9), 5046. https://doi.org/10.3390/su14095046

Li, J., Sun, H., & Li, J. (2023). Beyond Confusion Matrix: Learning From Multiple Annotators With Awareness Of Instance Features. *Machine Learning*, *112*(3), 1053–1075. https://doi.org/10.1007/s10994-022-06211-x

Manyol, M., Eke, S., Massoma, A. J. M., Biboum, A., & Mouangue, R. (2022). Preprocessing Approach for Power Transformer Maintenance Data Mining Based on k-Nearest Neighbor Completion and Principal Component Analysis. *International Transactions on Electrical Energy Systems*, *2022*, 1–10.

https://doi.org/10.1155/2022/8546588

Neogi, A. S., Garg, K. A., Mishra, R. K., & Dwivedi, Y. K. (2021). Sentiment Analysis And Classification Of Indian Farmers' Protest Using Twitter Data. *International Journal of Information Management Data Insights*, *1*(2), 100019. https://doi.org/10.1016/j.jjimei.2021.100019

Newby, R., Watson, J., & Woodliff, D. (2003). SME Survey Methodology: Response Rates, Data Quality, And Cost Effectiveness. *Entrepreneurship Theory and Practice*, *28*(2), 163–172. https://doi.org/10.1046/j.1540-6520.2003.00037.x

Ren, Z., Wang, S., & Zhang, Y. (2023). Weakly Supervised Machine Learning. *CAAI Transactions on Intelligence Technology*, *8*(3), 549–580. https://doi.org/10.1049/cit2.12216

Rouidi, M., Elouadi, A. E., Hamdoune, A., Choujtani, K., & Chati, A. (2022). TAM-UTAUT And The Acceptance Of Remote Healthcare Technologies By Healthcare Professionals: A Systematic Review. *Informatics in Medicine Unlocked*, *32*, 101008. https://doi.org/10.1016/j.imu.2022.101008

Santoso, W., Sitorus, P. M., Batunanggar, S., Krisanti, F. T., Anggadwita, G., & Alamsyah, A. (2021). Talent Mapping: A Strategic Approach Toward Digitalization Initiatives In The Banking And Financial Technology (Fintech) Industry In Indonesia. *Journal of Science and Technology Policy Management*, *12*(3), 399–420. https://doi.org/10.1108/JSTPM-04-2020-0075

Suhaeni, C., Kamila, S. A., Fahira, F., Yusran, M., & Alfa Dito, G. (2025). Exploring a Large Language Model on the ChatGPT Platform for Indonesian Text Preprocessing Tasks. *Indonesian Journal of Statistics and Its Applications*, *9*(1), 100–116. https://doi.org/10.29244/ijsa.v9i1p100-116

Tharwat, A. (2021). Classification Assessment Methods. *Applied Computing and Informatics*, *17*(1), 168–192. https://doi.org/10.1016/j.aci.2018.08.003

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance Of Information Technology: Toward A Unified View. *MIS Quarterly: Management Information Systems*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Xin, Y., & Ren, X. (2022). Predicting Depression Among Rural And Urban Disabled Elderly In China Using A Random Forest Classifier. *BMC Psychiatry*, *22*(1), 118. https://doi.org/10.1186/s12888-022-03742-4

Zaghloul, M., Barakat, S., & Rezk, A. (2024). Predicting E-Commerce Customer Satisfaction: Traditional Machine Learning Vs. Deep Learning Approaches. *Journal of Retailing and Consumer Services*, *79*, 103865. https://doi.org/10.1016/j.jretconser.2024.103865

Zin, K. S. L. T., Kim, S., Kim, H. S., & Feyissa, I. F. (2023). A Study On Technology Acceptance Of Digital Healthcare Among Older Korean Adults Using Extended TAM (Extended Technology Acceptance Model). *Administrative Sciences*, *13*(42), 1-18 . https://doi.org/10.3390/admsci13020042

Zogheib, S., & Zogheib, B. (2024). Understanding University Dtudents' Sdoption Of Chatgpt: Insights From TAM, SDT, And Beyond. *Journal of Information Technology Education: Research*, *23*, 1-14. https://doi.org/10.28945/5377