Aspect-Based Sentiment Analysis in Identifying Factors Causing Technostress in Fintech Users Using Naïve Bayes Algorithm

Yanuar Taruna Lutfi[1], Muhardi Saputra[2], and Riska Yanu Fa'rifah[3]

[1,2,3] Faculty of Industrial and System Engineering, Telkom University, Faculty of Engineering, Bandung, Indonesia
[1] hanstarunal@student.telkomuniversity.ac.id

Abstract. Technology has revolutionized finance through fintech, simplifying transactions and access to financial services. In Indonesia, fintech, particularly e-wallets, has experienced rapid growth. However, these technological advancements also present challenges such as technostress, which affects user behavior. Although OVO dominated the e-wallet market in 2021, user ratings declined in 2022, possibly related to news of complaints related to OVO services on the Google Play Store. This research aims to identify key aspects through data mining, using Aspect-Based Sentiment Analysis (ABSA) using Naïve Bayes with these factors, and determine the causes of technostress among OVO users. In addition, this study also evaluated text transformation methods: TF-IDF and Bag-of-Words (BoW). LDA-based topic modeling revealed 5 clusters with 4 core topics: features, access, services, and security. In general sentiment analysis, a data sharing ratio of 70:30 yielded the highest accuracy of 94.3% for TF-IDF and 95.25% for BoW compared to 75:25 and 80:20. BoW excels in terms of accuracy and prediction quality without overfitting. However, the TF-IDF model had difficulty with the prediction of positive reviews. The causes of technostress among OVO users were related to the transfer process (feature aspect), login problems (access aspect), concerns about customer service quality (service aspect), and satisfaction with security (security aspect).

Keywords: e-wallet, technostress, TF-IDF, BoW, ABSA, LDA, Naïve Bayes.

1 Introduction

In this digital era, various aspects of human life have been revolutionized by technology, even financial services [1]. Fintech or financial technology is a tangible form of revolution in the field of financial services that provides convenience in the transaction process that was originally done manually. The positive impact of fintech in the economy such as easy access to financial services through digital platforms [2]. The adoption rate of fintech continues to increase globally, led by countries such as China and India [3].

M. D. Sulistiyo and R. A. Nugraha (eds.), *Proceedings of the International Conference on Enterprise and Industrial Systems (ICOEINS 2023)*, Advances in Economics, Business and Management Research 270,
https://doi.org/10.2991/978-94-6463-340-5_10

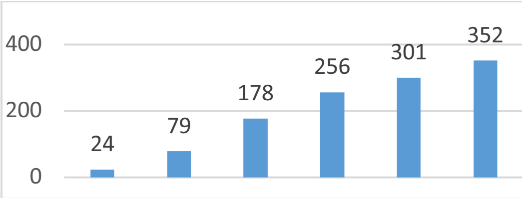

Fig. 1. Data Members of AFTECH 2016- Q4 2021

This is in line with what is happening in Indonesia, where fintech is also growing rapidly. This is characterized by the growing number of fintech players, accompanied by a significant contribution to the economy, the data of which is shown in Figure 2. The high investment in this field gave birth to unicorns that have an important role. This is in line with the adoption of fintech by the public, which can be seen from the value of transactions using one of the fintechs, namely electronic money or e-wallets such as QRIS, which is able to exceed the target. The popularity of using e-wallets is increasing, which can be seen from the number of application users such as Gopay, OVO, ShopeePay, LinkAja, and Dana [4]. The market share percentage is shown in Figure 2.

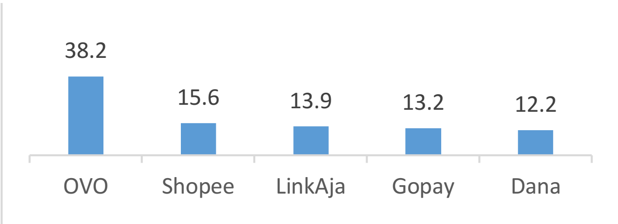

Fig. 2. Market Share Percentage

However, technological development does not only have a positive impact, it also has a negative impact that can make users experience pressure or stress [5]. This phenomenon is referred to as “technostress”, which arises from users' difficulties in adapting to technological change or new technology [6]. According to Weil and Rosen, technostress can cause anxiety and have a negative impact on behavior, thoughts, and aspects of human psychology stemming from the use of technology either directly or indirectly [7]. Fintech technology is not an exception, fintech applications can cause technostresses where users can experience difficulties in adapting to changes in application features, complex interfaces, fear of the risk of privacy violations, or even limited customer support [5].

In understanding technostress many studies have been conducted using surveys, but the limitations of this method raise some concerns [8]. Machine Learning (ML) can be used as a solution to address these concerns [9]. User behavior, as well as aspects that contribute to technostress, can be identified using ML. Aspect Based Sentiment Analysis (ABSA) is an appropriate approach to identifying sentiment related to certain aspects of text [10].

In addition, the abundance of data, data management, and data analytics are increasingly important and increasingly necessary, in analyzing text data it is important to determine the events, concepts, or topics of a document, but for computer programs only given raw text without understanding the context of the document which makes topic modeling important [11]. Latent Dirichlet Allocation (LDA) is a popular probability model for extracting topics from document collections [12] In addition, Naïve Bayes can be used as a sentiment analysis algorithm, which has proven effective in previous studies [13].

This study aims to identify and understand the aspects that influence and can lead to user technostress in using fintech, especially in the OVO digital wallet application. With a deeper understanding, it is expected to identify and address technostress problems experienced by users, to improve user experience and maintain a wider user base.

2 Literature Review

2.1 E-wallet

Carrying cash in a digital format is the definition of an electronic wallet (e-wallet). E-wallet applications contain a variety of required information such as name, account number, or phone number. The money in digital format will be transferred to another e-wallet instead of using a plastic card, the transaction is done online. The virtual wallet or electronic wallet (e-wallet) mechanism will be filled by the user (deposit) with the desired value. When an e-wallet transaction is made, it will be debited according to the nominal in the transaction [14].

2.2 Technostress

The development of technology has a negative impact, one of which is technostress, which is the pressure felt by users to use technology [5]. Technostress can be triggered because users experience difficulties in adapting to new technology, it can cause anxiety and have a negative impact on the mind, behavior, and psychological aspects of users [6]. According to Weil and Rosen, raging technostress can cause anxiety and have a negative impact on behavior, thoughts, and aspects of human psychology that stem from the use of technology either directly or indirectly [7].

3 Methodology

3.1 Latent Dirichlet Allocation (LDA)

In topic modeling, one of the most used algorithms is Latent Dirichlet Allocation (LDA). Being categorized as an unsupervised learning algorithm, LDA can generate topic clusters that contain word distributions into keywords or words that are distinctive and can represent the cluster. This can be done because LDA uses a probabilistic generative model. When classifying documents in large data sets, this algorithm can outperform other topic modeling algorithms [15]. In addition, LDA is a Bayesian model with a three-level hierarchy, where each item in the collection is considered a finite mixture of several topics [16].

3.2 SentiStrength

Sentistrength is a classification algorithm that analyzes sentiment strength in short text using a lexicon-based approach, which is a list of words that have a certain sentiment value or strength. It additionally makes use of rules and additional linguistic records. Not only that, according to psychological research humans can feel positive and negative emotions simultaneously to a certain extent, therefore Sentistrength uses a dual-scale (positive-negative) system [17]. Sentistrength generates positive and negative scores with a range of values from 1 to 5. A score of 1 means that the sentence contains no positive or negative sentiment, while a score of 5 means that the sentence has a highly positive sentiment or a highly negative sentiment [18].

3.3 Naïve Bayes

Naïve Bayes is a simple probabilistic technique that performs item classification based on the application of Bayes' Theorem with a strong assumption of independence among features. This makes the calculation process more efficient, but the assumption also limits its applicability. A small amount of data and the estimated variance for each class is sufficient to train the naïve Bayes classification algorithm instead of the entire covariance matrix. The naïve Bayes classification has been proven to work well in many real-world classification applications under certain conditions despite its seemingly simple assumptions. The advantages of Naïve Bayes classification are the need for small training data and sufficient classification power to overlook the shortcomings of the underlying naïve probability model [19]

3.4 TF-IDF

Term Frequency-Inverse Document Frequency (TF-IDF) is an important technical term in text processing, TF-IDF is used to identify essential words in a document [20]. TF-IDF performs the function of breaking down input sentences into proper instructions and retrieving solution facts from the dataset by using vectorizing all sentence entries. This entails calculating the frequency of incidence of terms or words inside the sentence [21].

3.5 Bag of Word (BoW)

The Bag-of-Words (BoW) method was originally introduced in the context of text search to be used for text document analysis [22]. BoW is an approach used to perform text-to-numeric conversion in the form of a matrix, where each row in the matrix represents one text or document [20]. In this method, the text is broken down into smaller units, such as words or tokens, and then the frequency of occurrence of each word in each text is calculated. In this BoW method, it does not pay attention to the order of words in the text which makes information about the grammar and context of word order not considered. The BoW method provides a solid foundation for further text processing using statistical and machine learning methods and opens up opportunities for further exploration in the development of more advanced text processing techniques [22].

3.6 Aspect Based Sentiment Analysis

In Aspect Based Sentiment Analysis (ABSA) there are two main tasks, namely aspect extraction and sentiment identification. In aspect extraction there are three main categories of techniques proposed based on previous research, these techniques include unsupervised learning, semi-supervised learning, and supervised learning. This research uses topic modelling with LDA algorithm. The next task in ABSA is sentiment identification, which involves determining the polarity value of a sentiment for each aspect that has been extracted. The value indicates the positive or negative intensity of the opinion about that aspect. there are many techniques that have been proposed for sentiment identification. In this research, Naïve Bayes is used for sentiment identification.

3.7 Research workflow

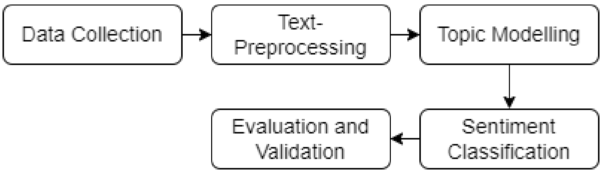

Fig. 3. Resarch workflow

Various steps were carried out in this research, including the steps depicted in Figure 3, namely:

1. Data Collection
   Researchers use the scraping method to collect data from the google play store using python and the google play scraper library. The results obtained 100,000 reviews with a vulnerable time from 2018 to 2023.
2. Text Preprocessing
   To ensure the research results are relevant, data filtering is carried out. Researchers limited the vulnerable reviews analyzed from January 1, 2022, to March 14, 2023, to 33,120 reviews. After that, researchers split the sentences. To ensure that each sentence has only one type of sentiment, researchers separate sentences with punctuation marks (.) and conjunctions with contrasting sentiment meanings. In Indonesian, these sentiments include the words “tapi”,

"namun", and "tetapi". Furthermore, case folding to equalize the entire text into lowercase, then efforts are made to process the reader to focus on sentiment words and their aspects such as removal of punctuation, removal of numbers, spelling correction, stemming, and stopwords.

3. Topic Modelling
   After preprocessing the text. One of the tasks in ABSA is aspect extraction, this research will be done with topic modeling. In this topic modeling is done using the LDA algorithm, to get the topics or aspects contained in the reviews.
4. Sentiment Classification
   After getting the next aspect to transform the text into a numeric matrix, this process is done with two methods, namely TF-IDF and BoW. This is to be able to simultaneously test which one is better used for Naïve Bayes. Next, split the train test data with several scenarios to get the most optimal model such as 70:30, 75:25, and 80:20.
5. Evaluation and Validation
   At this stage, measurements are made to evaluate the ability of the model using a confusion matrix to get accuracy, precision, recall, and F-score values. In addition, validation is done with k-fold cross validation.

4 Result and Discussion

In the research, the dataset used contains OVO application reviews on the Google Play Store with a total of 33,120 reviews. The dataset is divided into several scenarios; from several scenarios, the most optimal for this research is the 70:30 ratio. Furthermore, the aspects that have been produced by LDA will be used for grouping reviews, after which each group of reviews will be made a Naïve Bayes model and then evaluated and validated.

4.1 Topic Modeling

In the formation of this model, the gensim.models.LdaMulticore library is used. To train the LDA model in this study, using text data that has been transformed with BoW. In this study the model will do 100 times and with 10 number of topics, giving the highest coherence value reaching 0.396449 on 5 topics.

Fig. 4. LDA Visualization

Figure 4 shows the visualization of LDA. Of the 5 topics displayed, there are 2 topic clusters that intersect, namely topics 3 and 4. The intersection identifies that the two clusters can be put together. In each cluster there are top words or also called keywords in the topic.

Table 1. Aspect Interpertation

Aspect	Reviews	Topic Interpretation
1	*saya suka dengan semua* ***fitur*** *yang ada di aplikasi ovo ini*	Feature
2	*bagus dan mudah* ***layanan*** *nya*	Services
3	*mantap cepat* ***aksesn****ya dipercaya*	Access
4	*sudah lama pakai ovo mohon perbaikannya bagian bug* ***keamanan*** *aplikasi*	Security

Table 1 displays example sentences that address specific topics. The bolded words play an important role in the interpretation of the main topics, namely features, services, access, and security. These four topics are the aspects that will be scrutinized in the later stages.

4.2 Sentiment Classification

In sentiment classification, reviews that have been categorized based on their aspects will be labeled with the sentiment using scent strength. The results of data labeling are displayed in Table 2.

Table 2. Result of Aspect Extraction and Sentiment Identification

Text	Feature	Access	Services	Security	Sentiment
sesal download rugi					Negative
cepat transfer	Positive				Positive
server kendala keluh server pusat		Negative			Negative
daftar susah blok nomor sms verifikasi aneh		Negative	Negative		Negative

After labeling is done, the results show that reviews with neutral sentiment have a more dominant number, totaling 21,398 reviews, compared to reviews with positive sentiment with 9,578 reviews and negative sentiment totaling 12,476 reviews. This will have an impact on the imbalance of data, not only that reviews with neutral sentiments do not provide insights relevant to this research. Based on the things that have been explained, in this study, reviews with neutral sentiments were removed.

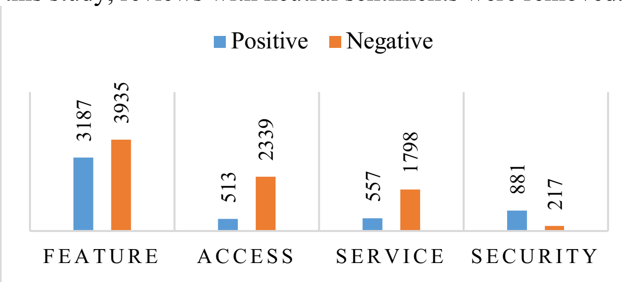

Fig. 5. The Number of Sentiments On Each Aspect

Figure 5 presents a division of positive and negative sentiments on each aspect. With the removal of reviews with "neutral" sentiment, it certainly reduces the overall number of reviews. The total number of reviews after removing reviews with neutral sentiment is 22,054 reviews. Furthermore, in the overall sentiment classification using Naïve Bayes and by aspects.

4.3 Evaluation

In all reviews, Naïve Bayes modeling was carried out using several split data scenarios and two methods of transformation text to get the most optimal model. The data is shown in Table 3.

Table 3. Result of Model Evaluation

Trans Text	Ratio	*Split Data*	*True Positive*	*True Negative*	Number of Prediction Data	Actual Amount of Data	Accuracy
F	70:30	Train	6,308	8.675	14.983	15.437	97.05%

		Test	2,577	3.659	6.236	6.617	**94.24%**
	75:25	Train	6,751	9.284	16.035	16.540	96.94%
		Test	2,134	3.080	5.184	5.514	**94.01%**
	80:20	Train	7,169	9.919	17.088	17.643	96.85%
		Test	1.716	2.415	4.131	4.411	**93.65%**
	Mean Accuracy						**95%**
BoW	70:30	Train	6,340	8,664	15,004	15,437	97.19%
		Test	2,632	3,680	6,312	6,617	**95.39%**
	75:25	Train	8,791	9,275	16,066	16,540	94.01%
		Test	2,181	3,069	5,250	5,514	**95.21%**
	80:20	Train	7,214	9,910	17,124	17,643	97.05%
		Test	1,758	2,434	4,192	4,411	**95.03%**
	Mean Accuracy						**96%**

Table 3 shows that the ratio of 70:30 for both models with TF-IDF and BoW has the highest accuracy. Therefore, this ratio is used for the model in every aspect. In addition to high accuracy, other values such as precision (P), recall (R), and f-score (F) are also high, indicating that the model has good performance, where detailed data is shown in Table 4. Validation was carried out on the two models with k of 10. The result is that both models are stable with a standard deviation of only 0.01 which indicates that the model is not overfitting or underfitting.

Table VI. Result Model Performance

	P (TF-IDF)	P (BoW)	R (TF-IDF)	R (BoW)	F (TF-IDF)	F (BoW)
N	93	94	98	98	95	96
P	97	98	90	92	93	95

Furthermore, the models for each aspect, both the model using TF-IDF and the model using Bow, both model accuracies can be seen in Figure 6. It can be seen that BoW has better accuracy than TF-IDF in all aspects.

Fig. 6. Model Accuracy by Aspect and Text Transformation

5 Conclusion and Future Work

Based on the result of this study, implementation of topic modeling, especially using the Latent Dirichlet Allocation (LDA) algorithm, was able to identify the aspects contained in existing reviews from the Google Play Store written by OVO fintech users. However, when visualized with pyLDAvis, it was found that topic clusters 3 and 4 overlap, with the overlap indicating that the clusters can be merged into 1 topic. The topics generated and analyzed in this research are features, access, services, and security. The implementation of aspect-based sentiment analysis using the Naïve Bayes Multinominal algorithm produces a good model in classifying user sentiment. The accuracy obtained in this study is 97.05% for models using TF-IDF and 97.19% for models using BoW where each model uses a division of training data and test data with a ratio of 70:30, the ratio will be used for models on each aspect. The performance of the models for each aspect has high accuracy (above 80%) and no overfitting. However, in the aspect of access and service, it was found that the model using TF-IDF had difficulty predicting positive sentiment, which indicates that the naïve bayes model using BoW.

References

1. Dr. C. Muthurasu and Dr. M. Suganthi, "an Overview on Digital Library," *Glob J Res Anal*, no. October 2017, pp. 1–2, 2019, doi: 10.36106/gjra/8906567.
2. Ivan, N. R. Arkaan, K. R. Venir, A. S. Ananda, and B. K. A., "Fintech Indonesia 101," *Digits Insight 3*. pp. 63–67, 2018.
3. Ernest & Young, "Global FinTech Adoption Index 2019, https://fintechauscensus.ey.com/2019/Documents/ey-global-fintech-adoption-index-2019.pdf," pp. 1–44, 2019.
4. AFTECH, "Annual Members Survey 2021," *Fintech Indonesia*, pp. 1–59, 2021.
5. Y. K. Lee, "Impacts of digital technostress and digital technology self-efficacy on fintech usage intention of Chinese gen Z consumers," *Sustainability (Switzerland)*, vol. 13, no. 9, 2021, doi: 10.3390/su13095077.
6. U. N. U. Ahmad, S. M. Amin, and W. K. W. Ismail, "The Impact of Technostress on Organisational Commitment among Malaysian Academic Librarians.," *Singapore Journal of Library & Information Management*, vol. 38, pp. 103–123, 2009, [Online]. Available: http://search.ebscohost.com/login.aspx?direct=true&db=lih&AN=47506385&site=ehost-live
7. J. Tacy, "Technostress : A concept analysis," no. September, 2021.
8. D. H. Granello and J. E. Wheaton, "Online data collection: Strategies for research," *Journal of Counseling and Development*, vol. 82, no. 4, pp. 387–393, 2004, doi: 10.1002/j.1556-6678.2004.tb00325.x.
9. A. S. A. Alwabel and X. J. Zeng, "Data-driven modeling of technology acceptance: A machine learning perspective," *Expert Syst Appl*, vol. 185, no. July, p. 115584, 2021, doi: 10.1016/j.eswa.2021.115584.
10. M. Axhiu and A. Aliu, "Aspect-term Extraction from Albanian Reviews with Topic Modeling Techniques," *12th International Conference, ICT Innovations 2020*, pp. 26–35, 2020.
11. I. Vayansky and S. A. P. Kumar, "A review of topic modeling methods," *Inf Syst*, vol. 94, p. 101582, 2020, doi: 10.1016/j.is.2020.101582.
12. R. Albalawi, T. H. Yeap, and M. Benyoucef, "Using Topic Modeling Methods for Short-Text Data: A Comparative Analysis," *Front Artif Intell*, vol. 3, no. July, pp. 1–14, 2020, doi: 10.3389/frai.2020.00042.
13. A. Rahman and M. S. Hossen, "Sentiment Analysis on Movie Review Data Using Machine Learning Approach," *2019 International Conference on Bangla Speech and Language Processing, ICBSLP 2019*, pp. 27–28, 2019, doi: 10.1109/ICBSLP47725.2019.201470.

14. Sarwin Kiko Napitupulu, A. F. Rubini, K. Khasanah, and A. D. Rachmawati, "Kajian Perlindungan Konsumen Sektor Jasa Keuangan: Perlindungan Konsumen Pada Fintech," *Departemen Perlindungan Konsumen, Otoritas Jasa Keuangan*, vol. 1, pp. 1–86, 2017.
15. Y. Kalepalli, S. Tasneem, P. D. P. Teja, and S. Manne, "Effective Comparison of LDA with LSA for Topic Modelling," *Proceedings of the International Conference on Intelligent Computing and Control Systems, ICICCS 2020*, no. Iciccs, pp. 1245–1250, 2020, doi: 10.1109/ICICCS48265.2020.9120888.
16. D. M. Blei, A. Y. Ng, and M. T. Jordan, "Latent dirichlet allocation," *Adv Neural Inf Process Syst*, vol. 3, pp. 993–1022, 2002.
17. D. H. Wahid, "Peringkasan Sentimen Ekstraktid di Twitter Menggunakan Hybrid TFIDF dan Consine Similarity.pdf," *Circular Técnica EMBRAPA-CNPF n. 20. EMBRAPA, Brazil*, vol. 10, no. 2, pp. 207–218, 2016.
18. Mike Thelwall, "Sentiment Strength Detection Short Informasl Text.pdf," *Journal of the American Society for Information Science and Technology*, vol. 64, no. July, pp. 1852–1863, 2010, doi: 10.1002/asi.
19. A. Khan *et al.*, "Journal of Advances in Information Technology," vol. 1, no. 1, 2010.
20. H. Hasanli and S. Rustamov, "Sentiment Analysis of Azerbaijani twits Using Logistic Regression, Naive Bayes and SVM," *13th IEEE International Conference on Application of Information and Communication Technologies, AICT 2019 - Proceedings*, 2019, doi: 10.1109/AICT47866.2019.8981793.
21. D. A. Permatasari, H. Fakhrurroja, and C. Machbub, "Human-Robot Interaction Based on Dialog Management Using Sentence Similarity Comparison Method," *Int J Adv Sci Eng Inf Technol*, vol. 10, no. 5, pp. 1881–1888, 2020, doi: 10.18517/ijaseit.10.5.7606.
22. C.-F. Tsai, "Bag-of-Words Representation in Image Annotation: A Review," *ISRN Artificial Intelligence*, vol. 2012, pp. 1–19, 2012, doi: 10.5402/2012/376804.