ISSN: 2633-4828 Vol. 4, No.1, June 2022

International Journal of Applied Engineering & Technology

Fintech Lending in Indonesia: A Sentiment Analysis, Topic Modelling, and Social Network Analysis using Twitter Data

Sri Handika Utami, Anton Ade Purnama, and Achmad Nizar Hidayanto
Universitas Indonesia, sri.handika01@ui.ac.id, anton.ade@ui.ac.id, nizar@cs.ui.ac.id

Date of Submission: 12th April 2022 Revised: 18th May 2022 Accepted: 02nd June 2022

How to Cite: **S. H. Utami, A. A. Purnama, Achmad N. Hidayanto (2022). Fintech Lending in Indonesia: A Sentiment Analysis, Topic Modelling, and Social Network Analysis using Twitter Data. International Journal of Applied Engineering & Technology 4(1)**

***Abstract* - Digitalization in financing industries, has brought us into Financial Technology (Fintech). Fintech proved to be much more accessible to people with low financial literacy. Fintech Lending is one kind of business in fintech. Also known as Pinjol, Fintech Lending become one of the trending topics on Indonesian Twitter, which unfortunately is more about illegal Pinjol, which brings so many insecurities for government, financial regulations, and pinjol company itself. With this research, the sentiment and central character were analyzed. Using several algorithms for data model creation, Naïve Bayes is the most performing one, giving us the fact that negative sentiment still takes the lead on the earlier day of observation, but it decreased over time. We also discovered the most influential character on the issue, to help interested parties to deal with correction, confirmation, and endorsement to support the regulation and survival of the business. Using the pic model, the study concludes that despite its negative sentiment, pinjol still has a good opportunity to grow with the support of government, regulator, and law enforcement team. The policy, regulation, and government resources had to comply with the development speed of pinjol.**

Index Terms - Fintech Lending, Pinjol, Indonesian Twitter, sentiment analysis, social network analysis, topic model.

INTRODUCTION

The euphoria of digitalization uniquely hit the financing industries. Digitalization has changed the old and traditional way of doing finance into a new technology-centric business. Some said that it brought many more benefits for the lender and debtor, such as opening larger access to the crowd, as a debtor and as an investor [1]. Referring to [1], financial technology (fintech) utilized technology to give them the ability to do financial business in a new way, including the way to calculate interest rates, both for investors and debtors. The popularity of fintech also became proof of the inefficiency of the traditional channel [2].

In contrast with all the benefits, fintech also brought disadvantages with it. Chinese was one of the countries where the popularity of fintech business was at most [3]. The story of fintech lending in China brought many investors into losing all their lifesaving money. Some studies stated that the failure of fintech is sourced from the unethical way of business, such as the political money that derived regulation [3], the false motive from the owner, etc. Some worried that this type of financial industry will pull the country towards the next bubble burst.

In Indonesia, fintech lending, also known as peer-to-peer lending, is the most popular type of fintech. From here onwards, we will use the "Pinjol" term, as they are called in Indonesia. In Indonesia, unregulated pinjol or illegal pinjol grew rapidly. Since it was unregulated, either operation models, business models, or credit analysis is not standardized like the regulated pinjol. Therefore, the issues, such as higher interest rates, unethical way of debt collecting, unethical way of data using, and more, anger the citizen. This also degrades the image of pinjol in general.

The study on previous research has been held as part of Literature Study. Several have already studied the implementation of social media analysis on Pinjol. While some perform sentiment analysis [4], the other does the topic mining, or network analysis[5]. Some others combine sentiment analysis with topic mining to grab a deeper understanding of the topics[6]. In Indonesia, sentiment analysis for IGrow, a fintech lending platform, has already taken using Google Play Store reviews[7]. But a complete study of sentiment analysis, topic modeling, and social network analysis using Twitter data has not been done yet for Indonesian Pinjol.

Using this research, we intend to complete the study using Twitter data for Pinjol in Indonesia. This research aims to give recommendations for government, regulator, and pinjol entities from the public sentiment on Twitter, so they could understand why such sentiment emerged and arrange the strategy to handle the situation. To complete the study, the social network analysis was carried out, to understand how the information runs through the network and who is the most significant persona to help balance the sentiment or validate the false rumors.

LITERATURE STUDY

I. Fintech Lending

Fintech Lending is one of the varieties in the Fintech Industry. According to [2]fintech is divided into crowdfunding, fintech lending, and Initial Coin Offerings (ICO's). Both crowdfunding and fintech lending (Peer-to-Peer lending), are categorized into Peer-to-Peer financing [8]. Among the three, fintech lending was the most promising type of business. For example, LendingClub, one of the most famous P2P landings in the world, achieved a 100% increment of loans in two years period, with USD 4.4 billion in a debit balance in 2013, to USD 8.8 billion [9]. Even during the COVID-19 pandemic, the Fintech Lending business ran more successfully than before, one of the cases was researched by [8]. The research which observed US Fintech Lending as the subject proved that Fintech Lending helped Americans on fulfilling their immediate needs during the pandemic and supported the economy nationally.

On the other side, fintech lending brought an adverse impact on investors in China [3]. The company declared bankruptcy to run off with the funds. According to [3], there were financial, political, and other issues that played parts in the failure.

In Fintech, the interest rate is one of the most crucial factors to determine the growth speed of a business [10]. The article stated that a higher interest rate will lower the payout rate and increase credit risks or repayment risks. The other factor is investor motivation. The research stated that investors with good social orientation and high ethics will bring the company into a better business.

II. Social Media and Social Media Data

Social media has become one of the effective marketing channels [11]. Social media can be utilized for several purposes [12], such as expressing an opinion (e.g., Twitter, Tripadvisor, Amazon), connecting with their network (e.g., Facebook, Coachsurfing), sharing their professional or expertise (e.g., Linkedln, Upwork), or expressing their passion (e.g., Instagram, Pinterest, Facebook). On the other side, social media also brought an issue with the false opinion that damaged public opinion about important matters, with an unprecedented opportunity for the people to communicate freely [13]. In social media, Social Media Influencers (SMIs) played a significant role damaged product branding, leading public opinions, or spreading issues [11].

This research will utilize twitter data to get the opinion about an issue, as claimed as the most popular source of social media data [14]. Bringing the concept of Microblogging word of mouth (MWOM), Twitter was the most logical platform to fuel the process of answering research questions.

In processing social media data, big data analytics is the favorite technique to mine useful information [15]. The reason was that social media data matched with one of the characteristics of big data, which involved unstructured data behavior, whether it is textual or non-textual data [16]. When talking about big data, the three Vs (Volume, Velocity, Variety) plus the "Value" as the fourth V was described as main indicators [17]. Since Twitter is given the negative sentiment resulting in early adoption of new products [14], doing analytics on that is expected to give much more constructive suggestions for stakeholders.

III. Sentiment Analysis

Sentiment analysis is a method to analyze and extract knowledge using subjective statements as a source of data, which is published through the Internet [18]. The data used in analysis ideally need to express explicitly the opinion of the writer on the aspect of the subject, including brands, products, and services [19]. The classification tasks of the analysis mean to label textual data as "Positive", "Negative", or "Neutral". But, in this study, the aim is to separate the data into "Positive" and "Negative" only.

Vary methods and techniques have been applied in previous studies. The methods could be categorized into three categories:

- Simple techniques [20], such as K-Nearest Neighbor (KNN), Decision Tree, Logistic Regression, and Naïve Bayes.
- Advanced Techniques, usually created for a specific form of data, such as some Deep Learning Algorithm [19] [18] [20].
- Ensemble or combination of techniques [21], such as Random Forest, Light Gradient Boosting Algorithm (LGBM), eXtreme Gradient Boosting XGBoost, Adaptive Boosting (AdaBoost), and Voting Classifier (Vot).

IV. Topic Modelling

Topic modeling is a method for text mining, to analyze big data, such as social media data, in text format [22]. There are several techniques used in topic modeling. One of them is Latent Dirichlet Allocation (LDA). Some previous studies stated that LDA was the most popular probabilistic-based method used in developing topic models [19] and provided a helpful technique for machine-assisted interpretation of texts [20].

LDA could be divided into three parts of processes [23]:

- For each tweet, sample the topic distribution using Dirichlet Distribution.

- Then, every word in a tweet differentiates into topics, which are defined in the first step.
- Sampling each word from words polynomial distribution using the generated topics sampling on steps before.

V. Social Network Analysis

Social Network Analysis is a great technique to analyze how things are connected in social media. One of the roles of social network analysis is to visualize information spreading. How information spreads have the power to influence the behavior of people and influence the effectiveness of a program, strategy, or regulation [24].

Influential people analysis is one analysis of the social network. One of the methods for finding influential people is using betweenness centrality. One of the examples is using betweenness centrality on finding influential people behind the spreading of information about COVID-19 on Twitter [25]. This research also uses this analysis to find influential people for pinjol information spreading on Twitter.

METHODOLOGY

This section will present a methodology that is used to answer research questions on this paper. Overall, the research process can be seen in Figure 2.

I. Data Collection

The Data Collection process starts with collecting the training data from Oct 6th to Oct 13th with "pinjol" as the keyword. Regarding the two kinds of fintech lending, legal and illegal, we also tried to find another keyword that has an impact on balancing the sentiment in pinjol with a positive perspective. But the other keyword like "pijol", "fintech", and others had too much noise on the data. The data have so many either unrelated tweets or non-Indonesian tweets. So, "pinjol" was decided to be the only keyword that will be used in this research.

We used 2.000 tweets as training data. The training data are labeled into positive or negative sentiments with the help of three annotators. The annotations are then processed using a consensus mechanism, with the help of one annotation supervisor to guarantee annotation result quality. Then it will be trained using several algorithms to build a good model to be used for sentiment analysis. As one of the most important processes in data mining, data preprocessing is also carried out using tokenize, stop words removal, lemmatization, POS (part of speech) tagging, and a bag of words [26].

Then we collect the data for analysis from 27th October to 5th November 2021 using Twitter API (Application Programming Interface). From that, we got 112.381 tweets. Those tweets will be used for sentiment analysis, topic modeling, and social network analysis.

II. Sentiment Analysis

For sentiment analysis, we used several algorithms to build a model. They are K Nearest Neighbor (kNN), Stochastic Gradient Descent (SGD), Neural Network (NN), Logistic Regression (LR), Naïve Bayes (NB), Random Forest (RF), Adaptive Boosting (AdaBoost). We evaluated the models from each algorithm to find the best model for sentiment analysis. Model performance evaluation can be seen in Table 1.

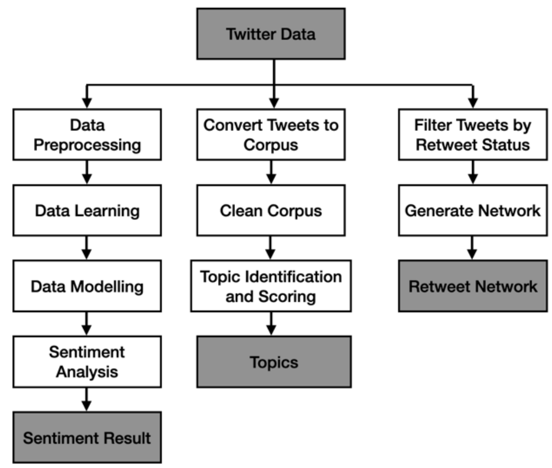

FIGURE 1
RESEARCH PROCESS

TABLE 1
MODEL PERFORMANCE EVALUATION

Model	AUC	CA	F1	Precision	Recall
kNN	0.606	0,565	0.525	0.608	0.565
SGD	0.667	0.666	0.667	0.667	0.666
Random Forest	0.754	0.688	0.686	0.696	0.688
Neural Network	0.712	0.656	0.656	0.657	0.656
Naïve Bayes	0.760	0.691	0.689	0.699	0.691
Logistic Regression	0.743	0.688	0.688	0.690	0.688
AdaBoost	0.677	0.663	0.664	0.666	0.663

Then we did k-fold cross-validation to estimate how accurately each predictive model will perform in practice. We set k as 5 and did cross-validation for each predictive model from Table 1, where they have an F1 score of more than 0.5. The result of the model comparison can be seen in Table 2.

Based on Table 1 and Table 2, the predictive model from Naïve Bayes has the biggest score on average F1-macro and bigger chance to give better performance than the others. It means the Naïve Bayes model will perform better than the other mode. Thus, we use the model to do sentiment analysis for this research.

TABLE 2
MODEL COMPARISON

	kNN	SGD	RF	NN	NB	LR	Ada Boost
kNN		0.025	0.000	0.008	0.000	0.000	0.002
SGD	0.975		0.013	0.013	0.002	0.003	0.232
RF	1.000	0.987		0.939	0.360	0.798	0.999
NN	0.992	0.987	0.061		0.012	0.033	0.896
NB	1.000	0.998	0.640	0.988		0.971	0.998
LR	1.000	0.997	0.202	0.967	0.029		0.999
AdaBoost	0.998	0.768	0.001	0.104	0.002	0.001	

III. Topic Modelling

To discover patterns across tweets data, we use LDA. LDA is used to retrieve information from sentiment analysis results. The result of sentiment analysis will be divided into 2 data which are negative sentiment and positive sentiment. Then we generated 4 topics for each sentiment using LDA.

IV. Social Network Analysis

To identify key prayers on information dissemination about pinjol on Twitter, we used SNA. We filtered 112.281 tweets from data collection by retweeted status. The total retweeted data that we got are 54.894 tweets. Then we generated retweeted network based on the tweets.

RESULT AND DISCUSSION

I. Fintech Lending Twitter Sentiment

Based on Table 2, the Naive Bayes model as the best performance model is used to analyze 112.381 tweets. The model found that negative sentiment for fintech lending on Twitter in Indonesia is bigger than positive sentiment. They are 21,24% positive tweets and 78.76% negative tweets. It can be seen in Figure 2.

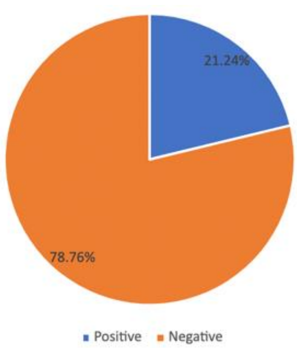

FIGURE 2
FINTECH LENDING SENTIMENT ANALYSIS CHART

We also produce the trends of sentiment on pinjol during the observed period in Figure 4. From the graphic, we can interpret that from 16th October to 10th November of 2021, the negative sentiment on pinjol decreased gradually. This trend is coordinated with the tweet trend of pinjol. This phenomenon is suited to [14] statement about the tendency of new products or issue sentiment to be negative.

Some events analyze to affect the sentiment trends. The peak during the early day of observation was suspected as the impact of the first arrest of illegal pinjol on the 14th. The government and regulator agreement on the moratorium of legitimizing new pinjol licenses, also contribute to the number. The arrest continued to 18th October in Yogyakarta for another illegal pinjol.

The peak of negativity also happened on 28th October MahfudMD, one of the presidential team, released the statement to not repay the debt from illegal pinjol. This issue continued with the statement on 23rd October that the illegal pinjol will not be able to charge the consumer who failed to pay the debt. On 23rd October, another arrest occurred in Tangerang. Finally, the Chinese leader for illegal debt collectors was also arrested and brought down the trends the day after.

In general, the sentiment analysis shows us that every issue and action was taken by the government, law enforcement, and regulator has brought the trend to its local peak and to be cooled down for the day after until another news release. Overall, we see that the stakeholder can cool down the heat along with the pinjol issue in the long term. This also becomes a good sign for pinjol entities to run their business. Using this data, we recommend government and pinjol to promote more and correct false issues or negative image of pinjol, both using mainstream media and social media. We also recommend regulators and the government to improve pinjol regulation and supervisory to increase the safety feeling of the business environment. The actions will become a tremendous help to speed up acceptance period of pinjol.

II. Topic Modelling

To get a visual representation of the most frequently used words, a bag of words (BoW) is used. BoW can show words based on specific criteria such as word frequency, by reducing and simplifying the representation of a text document. In this research, the result of sentiment analysis is divided into 2 BoW. First, we generated BoW of negative sentiment. It can be seen in Figure 3. Second, we generated BoW of positive sentiment which can be seen in Figure 5.

FIGURE 3
BOW OF NEGATIVE SENTIMENT

From both BoW, we can analyze the topics using LDA.

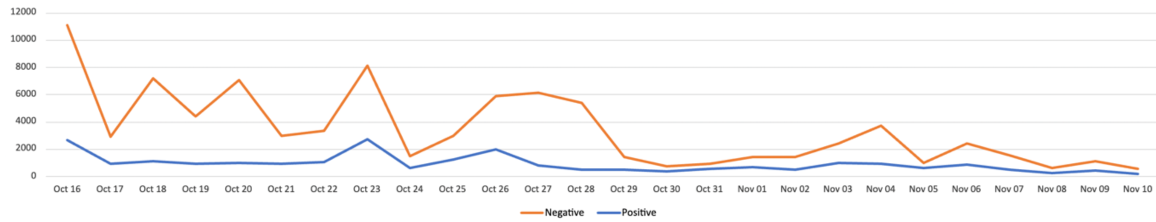

FIGURE 4
FINTECH LENDING SENTIMENT TREND

TABLE 3
TOPICS OF NEGATIVE SENTIMENT

Topic	Words of Topic				
Topic 1: Terror on pinjol	tagih	ilegal	bunuh	teror	chat
Topic 2: legality of pinjol	ilegal	legalitas	satgas	waspada	ojk
Topic 3: Pinjol crime reports	akun	hapus	polri	lapor	data
Topic 4: Higher Interest Rates	Habis	masyarakat	uang	ngeri	main

TABLE 4
TOPICS OF POSITIVE SENTIMENT

Topic	Words of Topic				
Topic 1: Easy money	pinjam	hidup	jasa	upaya	duit
Topic 2: Illegal pinjol office raid	kantor	gerebek	polisi	aman	orang
Topic 3: List of authorized pinjol	ojk	daftar	izin	otorisasi	usaha
Topic 4: Low-income people can be served by pinjol	economic	lowincome	mudah	keluarga	hutang

Table 3 shows the topics of negative sentiment. There are 4 topics on the negative sentiment which are terror on pinjol, the legality of pinjol, pinjol crime report, and higher interest rate on pinjol. Understanding of these topics will means so much on correcting false issue and negative image of pinjol. The topic about legality and pinjol crime report indicate that regulators and law enforcement need to be stricter about pinjol license and be more transparent about the processes or actions required on supervisory. So far, via Twitter, this transparency is already reflected from tweets from Polri and OJK. The topic about interest rate could be controlled by regulation, which is recommended on sentiment analysis result. Meanwhile, producing bright advertisement material and promoting customer centric services are additional steps to ease the topic of terror.

FIGURE 5
BOW OF POSITIVE SENTIMENT

Table 4 shows the topics of positive sentiment. There are 4 topics on the positive sentiment which are easy money because people can loan money easily from their handphones, illegal pinjol office raid, a list of authorized pinjol from OJK, and low-income people can be served by pinjol without too many requirements to get a loan. These topics will help stakeholders to understand the pros points of pinjol from Indonesian society perspectives. In addition to some recommendations above, pinjol needs to focus on easy access topics to increase its appealing points. Improving accessibility while implementing safety credit analyze process will help pinjol to maintain its strong point without ignoring business risks. Risks analysis is another key point to maintain positive sentiment, due to its impact on building better quality of credit. In the end, with better credit quality, the customer in no need to meet debt collectors, which brings a significant contribution to creating negative sentiment of pinjol.

TABLE 5
BETWEENNESS CENTRALITY

Account	Betweeness Centrality
Partono_ADjem	0.007588879
YRadianto	0.006390355
didhopratama	0.00633519
MkomEliya	0.003831733
BossTemlen	0.003756136

III. Social Network Analysis

The retweeted network was created to identify key players in information dissemination on Twitter. From the visualization of the retweet network in Figure 6, we could see the influential player on the pinjol issue. Partono_Adjem is one of the most vocal influencers who often speaks up about his concern about the negative issue in the society, one of them about illegal pinjol. Unfortunately, from his tweet, even though he comments about illegal pinjol, the term he used did not differentiate it from regulated pinjol. It resulted in a bad sentiment effect on the whole business. The betweenness centrality score can be seen in Table 5.

TABLE 6
IN-DEGREE CENTRALITY

Account	Betweeness Centrality
BossTemlen	3871
Partono_ADjem	2490
DivHumas_Polri	1134
pinjollaknat	1052
WidasSatyo	939
VIVAcoid	799
txtdarionlshop	695
CNNIndonesia	664
CCICPolri	604
detikcom	603

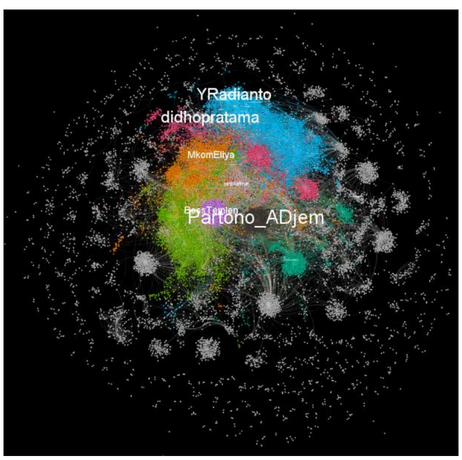

FIGURE 6
RETWEETED NETWORK BASED ON BETWEENNESS CENTRALITY

Table 7 presents the data of network overview on the generated networks. The average degree for the network is 1.565 and the network diameter is 19. There are also 476 connected components.

TABLE 7
NETWORK OVERVIEW

Property	Score
Average Degree	1.565
Network Diameter	19
Connected Components	476
Avg. Clustering Coefficient	0.024
Avg. Path Length	6.674

The other thing that we can analyze is in-degree centrality. Table 6 shows that the top influential persona based on in-degree centrality was also from government and law enforcement official accounts. So, the tasks on issue balancing and validation could be assigned to this account. It will be much easier to control the inner circle account for government and regulators than endorsed personal account who has their stands on the issue.

CONCLUSION

This research concludes that, the trouble pinjol that mixed up, in a hot messed situation, in Indonesia is pinjol who operates without a license. This kind of pinjol provokes negative sentiment in online communities, which brought implications into regulated legal pinjol who play within boundaries set by government and regulator. Therefore, strict regulation and fast action are needed to dismiss illegal pinjol and force them to play within regulation and ethical constraints. With all this strategy, hopefully, society will be able to enjoy the benefit of pinjol, such as low-interest rates and easy access.

Collaboration of work from the government, regulator and law enforcement will be needed. The government needs to publish a good law and policy to embrace pinjol correctly and appropriately. The regulator needs to prepare the regulation to regulate and supervise pinjol operation, procedure, and system. In general, the stakeholders need to keep the constraints and tools in supervising and regulating pinjol to be coordinated with the rapid development speed of the industry.

LIMITATION AND FUTURE RESEARCH

When doing research, there are limitations applied to the subject. One of them is the timeframe of the data source as stated before, which also limits the result of analysis into that specific window of time. Sarcasm interprets with its denotative meaning or explicit expression. The other limitation is passing over the tweets in the local language.

For future research, we suggest using advanced techniques, such as deep learning algorithms. The longer period will bring a better analysis result and produce a better recommendation for the purpose.

REFERENCES

[1] G. Buchak, G. Matvos, T. Piskorski, and A. Seru, "Fintech, regulatory arbitrage, and the rise of shadow banks," *J. financ. econ.*, vol. 130, no. 3, pp. 453–483, 2018, doi: 10.1016/j.jfineco.2018.03.011.

[2] H. Bollaert, F. Lopez-de-Silanes, and A. Schwienbacher, "Fintech and access to finance," *J. Corp. Financ.*, vol. 68, no. December 2020, p. 101941, 2021, doi: 10.1016/j.jcorpfin.2021.101941.

[3] Q. He and X. Li, "The failure of Chinese peer-to-peer lending platforms: Finance and politics," *J. Corp. Financ.*, vol. 66, no. December 2020, p. 101852, 2021, doi: 10.1016/j.jcorpfin.2020.101852.

[4] Q. Zhang, X. Zhu, J. L. Zhao, and L. Liang, "Discovering signals of platform failure risks from customer sentiment: the case of online P2P lending," *Ind. Manag. Data Syst.*, vol. 122, no. 3, pp. 666–681, 2022, doi: 10.1108/IMDS-05-2021-0308.

[5] E. M. Dermineur, "The evolution of credit networks in pre-industrial Finland," *Scand. Econ. Hist. Rev.*, vol. 70, no. 1, pp. 57–86, 2022, doi: 10.1080/03585522.2021.1884594.

[6] B. Niu, J. Ren, A. Zhao, and X. Li, "Lender trust on the P2P lending: Analysis based on sentiment analysis of comment text," *Sustain.*, vol. 12, no. 8, 2020, doi: 10.3390/SU12083293.

[7] Mussalimun, E. H. Khasby, G. I. Dzikrillah, and Muljono, "Comparison of K-N earest Neighbor (K-NN) and Naïve Bayes Algorithm for Sentiment Analysis on Google Play Store Textual Reviews," *2021 8th Int. Conf. Inf. Technol. Comput. Electr. Eng. ICITACEE 2021*, pp. 180–184, 2021, doi: 10.1109/ICITACEE53184.2021.9617217.

[8] K. Najaf, R. K. Subramaniam, and O. F. Atayah, "Understanding the implications of FinTech Peer-to-Peer (P2P) lending during the COVID-19 pandemic," *J. Sustain. Financ. Invest.*, vol. 0, no. 0, pp. 1–16, 2021, doi: 10.1080/20430795.2021.1917225.

[9] J. G. Wang and H. Xu, *Financing the Underfinanced*. 2015.

[10] G. Dorfleitner, E. M. Oswald, and R. Zhang, "From Credit Risk to Social Impact: On the Funding Determinants in Interest-Free Peer-to-Peer Lending," *J. Bus. Ethics*, vol. 170, no. 2, pp. 375–400, 2021, doi: 10.1007/s10551-019-04311-8.

[11] A. Audrezet, G. de Kerviler, and J. Guidry Moulard, "Authenticity under threat: When social media influencers need to go beyond self-presentation," *J. Bus. Res.*, vol. 117, no. November 2017, pp. 557–569, 2020, doi: 10.1016/j.jbusres.2018.07.008.

[12] A. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of Social Media," *Bus. Horiz.*, vol. 53, no. 1, pp. 59–68, 2010, doi: 10.1016/j.bushor.2009.09.003.

[13] S. L. Wilson and C. Wiysonge, "Social media and vaccine hesitancy," *BMJ Glob. Heal.*, vol. 5, no. 10, pp. 1–7, 2020, doi: 10.1136/bmjgh-2020-004206.

[14] T. Hennig-Thurau, C. Wiertz, and F. Feldhaus, "Does Twitter matter? The impact of microblogging word of mouth on consumers' adoption of new movies," *J. Acad. Mark. Sci.*, vol. 43, no. 3, pp. 375–394, 2015, doi: 10.1007/s11747-014-0388-3.

[15] X. Liu, H. Shin, and A. C. Burns, "Examining the impact of luxury brand's social media marketing on customer engagement: Using big data analytics and natural language processing," *J. Bus. Res.*, vol. 125, no. January 2018, pp. 815–826, 2021, doi: 10.1016/j.jbusres.2019.04.042.

[16] S. Erevelles, N. Fukawa, and L. Swayne, "Big Data consumer analytics and the transformation of marketing," *J. Bus. Res.*, vol. 69, no. 2, pp. 897–904, 2016, doi: 10.1016/j.jbusres.2015.07.001.

[17] M. Lycett, "'Datafication': Making sense of (big) data in a complex world," *Eur. J. Inf. Syst.*, vol. 22, no. 4, pp. 381–386, 2013, doi: 10.1057/ejis.2013.10.

[18] M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya, "ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis," *Futur. Gener. Comput. Syst.*, vol. 115, pp. 279–294, 2021, doi: 10.1016/j.future.2020.08.005.

[19] X. Liu, A. C. Burns, and Y. Hou, "An Investigation of Brand-Related User-Generated Content on Twitter," *J. Advert.*, vol. 46, no. 2, pp. 236–247, 2017, doi: 10.1080/00913367.2017.1297273.

[20] B. Zhong, X. Pan, P. E. D. Love, L. Ding, and W. Fang, "Deep learning and network analysis: Classifying and visualizing accident narratives in construction," *Autom. Constr.*, vol. 113, no. January, p. 103089, 2020, doi: 10.1016/j.autcon.2020.103089.

[21] A. Onan, "Sentiment analysis on massive open online course evaluations: A text mining and deep learning approach," *Comput. Appl. Eng. Educ.*, vol. 29, no. 3, pp. 572–589, 2021, doi: 10.1002/cae.22253.

[22] S. Tirunillai and G. J. Tellis, "Tirunillai-Mining Meaning," vol. 2437, 2014.

[23] W. E. I. Wang, Z. Gong, J. Ren, F. Xia, Z. Lv, and W. E. I. Wei, "Venue Topic Model – enhanced Joint Graph Modelling for," vol. 20, no. 1, pp. 1–15, 2021.

[24] M. Cinelli *et al.*, "The COVID-19 social media infodemic," *Sci. Rep.*, vol. 10, no. 1, pp. 1–10, 2020, doi: 10.1038/s41598-020-73510-5.

[25] W. Ahmed, J. Vidal-Alaball, J. Downing, and F. L. Seguí, "COVID-19 and the 5G conspiracy theory: Social network analysis of twitter data," *J. Med. Internet Res.*, vol. 22, no. 5, pp. 1–9, 2020, doi: 10.2196/19458.

[26] U. A. Chauhan, M. T. Afzal, A. Shahid, M. Abdar, M. E. Basiri, and X. Zhou, "A comprehensive analysis of adverb types for mining user sentiments on amazon product reviews," *World Wide Web*, vol. 23, no. 3, pp. 1811–1829, 2020, doi: 10.1007/s11280-020-00785-z.

.

AUTHOR INFORMATION

Sri Handika Utami, Graduate Student of Information Technology Management, Universitas Indonesia, and Practitioner

Anton Ade Purnama, Graduate Student of Information Technology Management, Universitas Indonesia.

Achmad Nizar Hidayanto, Professor, Faculty of Computer Science, Universitas Indonesia.