Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 2, April 2023
DOI : https://doi.org/10.33395/sinkron.v8i2.12090

e-ISSN : 2541-2019
p-ISSN : 2541-044X

Detect Fake Reviews Using Random Forest and Support Vector Machine

Zulpan Hadi[1)*], **Ema Utami**[2)], **Dhani Ariatmanto**[3)]
[1,2,3)]Universitas AMIKOM Yogyakarta, Indonesia
[1)]zlpnhadi@students.amikom.ac.id, [2)]ema.u@amikom.ac.id, [3)]dhaniari@amikom.ac.id

Submitted : Jan 2, 2023 | **Accepted** : Jan 26, 2023 | **Published** : Apr 1, 2023

Abstract: With the rapid development of e-commerce, which makes it possible to buy and sell products and services online, customers are increasingly using these online shop sites to fulfill their needs. After purchase, customers write reviews about their personal experiences, feelings and emotions. Reviews of a product are the main source of information for customers to make decisions to buy or not a product. However, reviews that should be one piece of information that can be trusted by customers can actually be manipulated by the owner of the seller. Where sellers can spam reviews to increase their product ratings or bring down their competitors. Therefore this study discusses detecting fake reviews on product reviews on Tokopedia. Where the method used is the distribution post tagging feature to perform detection. By using the post tagging feature method the distribution got 856 fake reviews and 4478 genuine reviews. In the fake reviews, there were 628 reviews written with the aim of increasing product sales or brand names from store owners, while there were 228 reviews aimed at dropping their competitors or competitors. Furthermore, the classification is carried out using the random forest algorithm model and the support vector machine. By dividing the dataset for training data by 80% while 20% for data testing. Here it is known that the support vector machine gets much higher accuracy than the random forest. The support vector machine gets an accuracy of 98% while the random forest gets an accuracy of 60%.

Keywords: Tokopedia**,** Fake Review, Pos Tagging, Support Vector Machine, Random Forest.

INTRODUCTION

With the rapid development of e-commerce, which makes it possible to buy and sell products and services online, customers are increasingly using these online shopping sites to fulfill their needs. After purchase, customers write reviews about their personal experiences, feelings, and emotions (Alsubari et al., 2021). Reviews are the main source of information for customers to decide to buy or not. For example, when a customer takes the initiative to buy a product, they will read reviews about what other customers think of the product. Depending on the feedback from the reviews, they decide to buy or not. If they get positive feedback from the reviews, they might proceed to buy. As such, past reviews become a very credible source of information for most people on several online stores. However, reviews are considered a form of sharing authentic feedback on positive or negative reviews, any attempts to manipulate those reviews by writing misleading or unauthentic content are considered fraudulent acts and such reviews are labeled as fake (Barbado et al., 2019). There are organizations or groups that are paid to provide reviews that decrease or increase the value of a product (Algotar & Bansal, n.d.). Spotting fake reviews and spammers is becoming more and more important as spam behavior becomes more destructive (Li et al., 2016).

Research conducted by (Le, 2020) it aims to improve the performance of fake review classifiers by integrating various techniques into the classifier model. More specifically, we analyze similarities between reviews and use an EM (Expectation Maximization) clustering algorithm to recognize review patterns. Researchers also apply sentiment analysis to analyze reviews. Using the results from the clustering model, sentiment analysis, and non-textual features of reviews and reviewers, we build a machine-learning model to classify fake reviews. Researchers compared three supervised machine learning algorithms: Support Vector Machine, Artificial Neural Network, and Random Forest. The empirical results from our experiments show that the Random Forest algorithm outperforms other algorithms.

*name of corresponding author

Here the researcher will try to use a textual feature, namely using pos tag distribution to detect fake reviews using the Random Forest algorithm and Support Vector Machine.

LITERATURE REVIEW

Research conducted by (Elmurngi & Gherbi, 2017) it proposes a machine-learning approach to identify fake reviews. In addition to the review feature extraction process, this paper applies some feature engineering to extract various reviewer behaviors. This paper compares the performance of several experiments conducted on a dataset of real Yelp restaurant reviews with and without features derived from user behavior. In both cases, we compare the performance of several classifiers; KNN, Naive Bayes (NB), SVM, Logistic Regression, and Random forest. Also, different n-gram language models especially bi-gram and tri-gram are considered during the evaluation. The results show that KNN(K=7) outperforms the other classifiers in terms of f-scores achieving the best f-score of 82.40%. The results show that the f-score has increased by 3.80% when considering the extracted reviewer behavioral features.

In research conducted by (Alsubari et al., 2022) The proposed methodology adopted in this study uses a standardized fake hotel review dataset for the experiment and a data preprocessing method and a term frequency-Inverse document frequency (TF-IDF) approach to extract features and their representations. For detection and classification, n-gram review texts were entered into a model built to be classified as false or true. However, the experiment was conducted using four machine learning techniques that were supervised and trained, and tested on datasets collected from the Trip Advisor website. The classification results from this experiment show that naïve Bayes (NB), support vector machine (SVM), adaptive boosting (AB), and random forest (RF) receive 88%, 93%, 94%, and 95% respectively, based on test accuracy and F1 score. The results obtained were compared with existing work using the same dataset, and the proposed method outperforms comparable methods in terms of accuracy.

Research conducted by (Kashti & Prasad, 2019) The proposed methodology adopted in this study uses a standardized fake hotel review dataset for the experiment and a data preprocessing method and a term frequency-Inverse document frequency (TF-IDF) approach to extract features and their representations. For detection and classification, n-gram review texts were entered into a model built to be classified as false or true. However, the experiment was conducted using four machine learning techniques that were supervised and trained, and tested on datasets collected from the Trip Advisor website. The classification results from this experiment show that naïve Bayes (NB), support vector machine (SVM), adaptive boosting (AB), and random forest (RF) receive 88%, 93%, 94%, and 95% respectively, based on test accuracy and F1 score. The results obtained were compared with existing work using the same dataset, and the proposed method outperforms comparable methods in terms of accuracy.

METHOD

In every research process, there is always a research method or model used to explain the research flow to be used. Figure 1 illustrates the research stages starting from dataset collection, and data preprocessing, word extraction, at the classification stage using a predetermined algorithm model and the last is the testing or evaluation process.

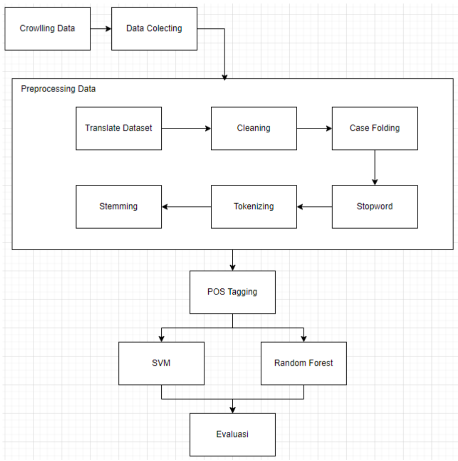

Figure 1. Research Flow

*name of corresponding author

Dataset

This study used a dataset taken from Kaggle with the keyword Tokopedia Product Reviews Category Food and Drink with a total of 5427 data. Table 1 is an example of the dataset that will be used.

Table 1. Sample Dataset

	Review	Rating
0	maaantap	5
1	Enak & segar	5
2	Maknyus	5
3	Dari berat 500 gram yg remuk 59 gram.\nEntah dari awal packing atau expedisi.\nUntuk masukan aja, tlg ada item dos packing, jadi barang tetap aman jika pengiriman jarak jauh.	4

Preprocessing

Preprocessing is the process of modifying data to make it suitable for use as input to an algorithm. Data cleansing is the process of replacing raw data with clean data (Kansal et al., 2016). In addition, the preprocessing process is also very important because it has a very large impact on the classification process. Because preprocessing can improve the accuracy of results in the classification process (Chandrasekar & Qian, 2016) (Prakoso, n.d.). In this preprocessing process, several stages will be carried out, such as data cleaning, translate, case folding, tokenizing, stopword removal, and stemming.

Google offers various APIs (Application Programming Interface) which are very useful for developers when using various functions including translation. Automatic language translation uses the Google Translate API, which is a set of Python libraries. Python libraries are combinations of packages and modules that make it easy to build systems or applications. Python has several translation libraries namely Deep Translator, Textblob, Goslate, Googletrans, Py-Translate, etc (Kadek et al., 2021). The library used in this step is googletrans.

Table 2. Translate Dataset

Before	After
Dari berat 500 gram yg remuk 59 gram.\nEntah dari awal packing atau expedisi.\nUntuk masukan aja, tlg ada item dos packing, jadi barang tetap aman jika pengiriman jarak jauh.	From the weight of 500 grams which crumble 59 grams.\nEither from the beginning of packing or expedition.\nFor just input, there are dos packing items, so the items remain safe if long distance delivery.

The table above shows the differences after the data was translated and before the previous Indonesian language dataset was converted to English.

The recently collected dataset may contain many irrelevant or even missing pieces. Therefore, it is necessary to carry out a data cleaning process which is often referred to as data cleaning. The goal of the data cleansing process is to provide better data quality, which is very useful in preparing data for the analysis phase (Putu et al., n.d.). at this stage, there are several characters will be removed such as numbers, symbols, punctuation, emoji, etc. The results of cleaning data can be seen in table 3.

Table 3. Cleaning Data

Before	After
From the weight of 500 grams which crumble 59 grams.\nEither from the beginning of packing or expedition.\nFor just input, there are dos packing items, so the items remain safe if long distance delivery.	From the weight of grams which crumble grams Either from the beginning of packing or expedition For just input there are dos packing items so the items remain safe if long distance delivery

In the table above it can be seen that several characters were removed in the cleaning process such as numeric characters, and punctuation marks.

At the case folding stage, each letter is converted to lowercase, because the model built is case sensitive, while the use of capital letters may be inconsistent in all words. For more details, see table 4 below

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 2, April 2023
DOI : https://doi.org/10.33395/sinkron.v8i2.12090
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Table 4. Case folding

Before	After
From the weight of grams which crumble grams Either from the beginning of packing or expedition For just input there are dos packing items so the items remain safe if long distance delivery	from the weight of grams which crumble grams either from the beginning of packing or expedition for just input there are dos packing items so the items remain safe if long distance delivery

The table above shows the difference before and after the case where all text is justified to lower case.

The tokenization step is performed to divide sentences into words based on spaces. See Table 5 below for more details.

Table 5. Tokenizing

Before	After
from the weight of grams which crumble grams either from the beginning of packing or expedition for just input there are dos packing items so the items remain safe if long distance delivery	'from', 'the', 'weight', 'of', 'grams', 'which', 'crumble', 'grams', 'either', 'from', 'the', 'beginning', 'of', 'packing', 'or', 'expedition', 'for', 'just', 'input', 'there', 'are', 'dos', 'packing', 'items', 'so', 'the', 'items', 'remain', 'safe', 'if', 'long', 'distance', 'delivery'

After the tokenizing stage is done, all words are solved based on their spaces

Filtering removes meaningless words and stops words. For more details, see table 6.

Table 6. Stopword

Before	After
'from', 'the', 'weight', 'of', 'grams', 'which', 'crumble', 'grams', 'either', 'from', 'the', 'beginning', 'of', 'packing', 'or', 'expedition', 'for', 'just', 'input', 'there', 'are', 'dos', 'packing', 'items', 'so', 'the', 'items', 'remain', 'safe', 'if', 'long', 'distance', 'delivery'	'weight', 'grams', 'crumble', 'grams', 'either', 'beginning', 'packing', 'expedition', 'input', 'dos', 'packing', 'items', 'items', 'remain', 'safe', 'long', 'distance', 'delivery'

At this stage, some words are omitted such as the words \`the\`, \`of\`,\`or\`,\`for\`,\`just\`,\`are\`,\`so\` etc.

Stemming is used to change all words into basic words, one of which is removing affixes. The algorithm used in this process is Porter's Algorithm.

Table 7. Stemming

Before	After
'weight', 'grams', 'crumble', 'grams', 'either', 'beginning', 'packing', 'expedition', 'input', 'dos', 'packing', 'items', 'items', 'remain', 'safe', 'long', 'distance', 'delivery'	'weight', 'gram', 'crumbl', 'gram', 'either', 'begin', 'pack', 'expedit', 'input', 'do', 'pack', 'item', 'item', 'remain', 'safe', 'long', 'distanc', 'deliveri'

At this stage, several words change because at this stage the word 'grams' becomes 'gram' 'crumble' becomes 'crumbl' 'beginning' becomes 'begin' etc.

One of the fundamental processes involved in extracting information from Twitter data is POS (part-of-speech) coding. POS coding identifies word tags in tweets based on syntax analysis. The study of tagging POS data on Twitter was conducted using a rule-based and statistical approach (Suryawati et al., 2018). In part of speech (POS), the task is to assign each word in a sentence a predefined sentence identifier, which describes the grammatical role of that word in the sentence (Bast et al., 2016).

Table 8. pos tagging

Before	After
'weight', 'gram', 'crumbl', 'gram', 'either', 'begin', 'pack', 'expedit', 'input', 'do', 'pack', 'item', 'item', 'remain', 'safe', 'long', 'distanc', 'deliveri'	('weight', 'NN'), ('gram', 'NN'), ('crumbl', 'NN'), ('gram', 'NN'), ('either', 'DT'), ('begin', 'VB'), ('pack', 'NN'), ('expedit', 'NN'), ('input', 'NN'), ('do', 'VBP'), ('pack', 'VB'), ('item', 'NN'), ('item', 'NN'), ('remain', 'VBP'), ('safe', 'JJ'), ('long', 'JJ'), ('distanc', 'NN'), ('deliveri', 'NN')

In the table above it can be seen that the markers or tags of each word in the product review sentence are dominated by nouns (NN).

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 2, April 2023
DOI : https://doi.org/10.33395/sinkron.v8i2.12090
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Fake Review Detection

Opinion spam refers to fake opinions that attempt to deceive readers or automated systems by giving a target item an undue positive opinion to promote that item, or by giving a harmful negative opinion to tarnish its reputation. Detecting such spam is very important for the app. One type of opinion spam is a fake opinion (Jindal & Liu, 2008), intentionally misleads readers by giving a target property an unreasonably positive rating to promote that property, or by giving another property an unreasonably negative rating to tarnish its reputation. To detect fake reviews or opinion spam, textual features are used where the method used is posted tagging distribution. Where the tagging post is used to get the tag or marker of a word. Honest or real reviews tend to have lots of nouns (NN) and adjectives (ADJ) while fake reviews have lots of verbs (VERB) and adverbs (ADV) (Pasaribu et al., 2019), (Alsubari et al., 2022).

Support vector machine

SVM is a supervised learning technique with a high level of accuracy and quality, so it is very popular among other algorithms. However, its implementation requires a series of training stages and must go through a testing process (Laurensz & Eko Sediyono, 2021).

SVM (Support Vector Machine) can solve linear and nonlinear problems. The core approach developed in SVM can also be used to deal with the number of different classes and a large number of classes. SVM for prediction is called regression SVM and consists of linear and nonlinear functions. The SVM regression method is called a nonparametric method because it relies on kernel functions (Lumbanraja et al., 2021).

According to (Prasetyo, 2012) Kernel functions include:

1. Kernel linier
   $K(u,v) = uv^T$
2. Kernel polynomial
   $K(u,v) = (1+uv^T)^d, d \geq 2$
3. Kernel RBF (Radial Basis Function):
   $K(u,v) = \exp(-y||u-v||^2), y > 0$
4. Kernel Gaussian:
   $$K(u,v) = \exp\left[-\frac{||u-v||^2}{2\sigma^2}\right]$$

Random forest

Random Forest is an ensemble learning method by Breiman in 2001. Random forest relies on random vector values that have the same distribution in all trees, where each decision tree has maximum depth. Random forest is a classifier consisting of classifiers in the form of a tree {h(x, θ k), k = 1, . . .} where θk is an independently distributed random vector and each unit tree selects the most popular class from the input x. The following are random forest accuracy characteristics:.

1. Centralize the random forest
   There are classifiers h1(x), h2(x), . . . , hk (x) and with the training set of the random vector distribution Y,X, the following functions are formed
   $mg(X,Y) = au_k I(h_k(X) = Y) - \max\ au_k I(h_k(X) = j)$
   The error function used
   $PE* = Px, y(mg(X,Y) < 0)$
   The result of combining functions
   $P_{X,Y}(P_\Theta(h(X,\Theta) = Y) - \max P_\Theta(h(X,\Theta) = j) < 0)$
   This result explains why the random forest does not overfit when the tree is added but produces a limited error value.

2. Power and correlation
   The resulting function is
   $$PE* \leq \sum var\left(P_\Theta(h(X,\Theta) = \mathrm{Y}) - P_\Theta(h(X,\Theta) = \mathrm{j})\right)s_j^2$$
   In this function, the strength does not depend on the forest.

3. Random Forest uses random input selection
   Bagging is used for random feature selection. Each training set is replaced with the original training set. Then trees are grown in the training set using random feature selection. There are two reasons for using

*name of corresponding author

DOI : https://doi.org/10.33395/sinkron.v8i2.12090
e-ISSN : 2541-2019
p-ISSN : 2541-044X

bagging, the first is the use of bagging to improve accuracy when using the random property. The second is bagging, which is used to provide estimates of the generalized error (PE*) of tree connections to estimate strength and correlation. The simplest random forest with the random property is constructed by randomly selecting a small set of common input variables at each node. Grow the tree to its maximum size using the CART methodology.

4. Random Forest uses a linear combination of inputs
Assuming there are many inputs, M, F takes a fraction of M, resulting in increased efficiency but with high correlation. Another approach is formed by defining more features by taking random linear combinations of several input variables. The characteristic variable L is the combined number of variables. Variable L is chosen randomly and added to the coefficients with random numbers [-1,1]. The resulting linear combination F This procedure is called Forest-RC.

RESULT

In the process of detecting fake reviews, researchers use the distribution tagging post method. Where in the reviews of food and beverage products on Tokopedia there are more real reviews or genuine reviews. More details can be seen in Figure 2.

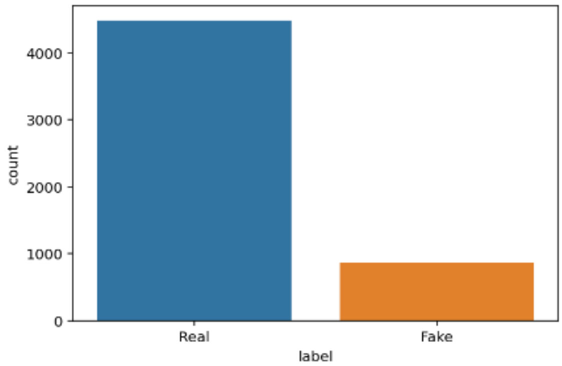

Figure 2. Fake review detection

After detecting fake reviews, a sentiment analysis process is then carried out using the python library, namely textblob. The results of the sentiment analysis can be seen in Figure 3.

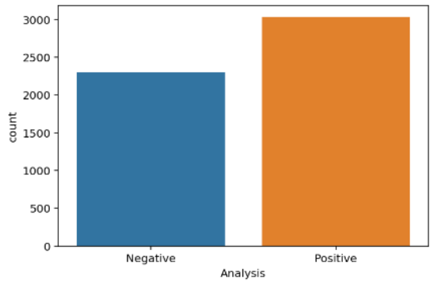

Figure 3. Sentiment analysis

Then the two processes are combined to know how many fake or fake reviews are positive or negative and whether genuine reviews are positive or negative. For more details, it can be seen in Figure 4.

Model 9. Label Analysis

Label	Analysis	Total
Fake	Positive	628
	Negative	228
Real	Positive	2402
	Negative	2076

The application of the random forest algorithm and the SVM in the classification of fake reviews is carried out by dividing the dataset by 80% for training data and 20% for testing data. After the classification process is

*name of corresponding author

complete, then the process of evaluating the performance of the model used is carried out. The performance level of the model can be seen in the table below.

Model 10. Testing Table

	Random Forest	SVM (karnel linier)
Accurasi	0.60	0.98
Precision	0.30	0.98
Recall	0.5	0.98
F1-Score	0.37	0.98

After calculating the performance level of each algorithm used. Furthermore, the level of performance of each algorithm can be described using the confusion matrix below.

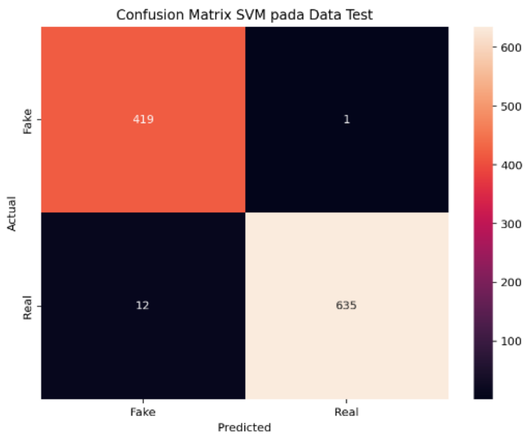

Figure 4. Confusion Matrix SVM

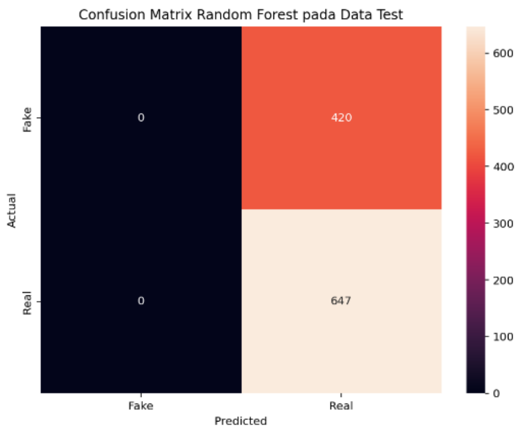

Figure 5. Confusion matrix Random forest

DISCUSSIONS

In this study, it is known that there are more genuine reviews than fake or fake reviews where there are 4478 genuine reviews and 856 fake reviews. Of the 4478 original reviews, it is known that 2402 reviews were positive and 2076 that were negative. As for the fake reviews, there were 628 positive reviews and 228 negative reviews.

For this classification process, the model used is a random forest algorithm and a SVM that uses a linear kernel. Where in this process SVM far outperforms random forest by getting 98% accuracy and random forest getting an accuracy score of 60%.

In addition, the obstacle found in this study is that reviews using foreign languages cannot be detected using this method. The way to overcome this is to generalize the language used by using the googletrans API python library

*name of corresponding author

CONCLUSION

In the classification process for detecting fake reviews using the pos tagging method in the SVM algorithm model, it gets the highest accuracy score with a score of 98%, while random forest gets a score of 60%.

REFERENCES

Algotar, K., & Bansal, A. (n.d.). *Detecting Truthful and Useful Consumer Reviews for Products using Opinion Mining*.

Alsubari, S. N., Deshmukh, S. N., Alqarni, A. A., Alsharif, N., Aldhyani, T. H. H., Alsaade, F. W., & Khalaf, O. I. (2022). Data analytics for the identification of fake reviews using supervised learning. *Computers, Materials and Continua*, *70*(2), 3189–3204. https://doi.org/10.32604/cmc.2022.019625

Barbado, R., Araque, O., & Iglesias, C. A. (2019). A framework for fake review detection in online consumer electronics retailers. *Information Processing and Management*, *56*(4), 1234–1244. https://doi.org/10.1016/j.ipm.2019.03.002

Bast, H., Buchhold, B., & Haussmann, E. (2016). Semantic search on text and knowledge bases. *Foundations and Trends in Information Retrieval*, *10*(2–3), 119–271. https://doi.org/10.1561/1500000032

Chandrasekar, P., & Qian, K. (2016). The Impact of Data Preprocessing on the Performance of a Naïve Bayes Classifier. *Proceedings - International Computer Software and Applications Conference*, *2*, 618–619. https://doi.org/10.1109/COMPSAC.2016.205

Elmurngi, E., & Gherbi, A. (2017). Detecting Fake Reviews through Sentiment Analysis Using Machine Learning Techniques. *DATA ANALYTICS 2017 : The Sixth International Conference on Data Analytics Detecting*, *c*, 65–72.

Jindal, N., & Liu, B. (2008). Opinion spam and analysis. *WSDM'08 - Proceedings of the 2008 International Conference on Web Search and Data Mining*, 219–229. https://doi.org/10.1145/1341531.1341560

Kadek, N., Sari, R., Made, I., Suarjaya, A. D., & Buana, W. (2021). *Perbandingan Translation Library Pada Python (Studi Kasus: Analisis Sentimen Penyakit Menular Di Indonesia)* (Vol. 2, Issue 3).

Kansal, A., Singh, Y., Kumar, N., & Mohindru, V. (2016). Detection of forest fires using machine learning technique: A perspective. *Proceedings of 2015 3rd International Conference on Image Information Processing, ICIIP 2015*, 241–245. https://doi.org/10.1109/ICIIP.2015.7414773

Kashti, R. P., & Prasad, P. S. (2019). ANALYSIS OF CLASSIFIERS FOR FAKE REVIEW DETECTION. In *International Journal For Technological Research In Engineering* (Vol. 6, Issue 9). www.ijtre.com

Laurensz, B., & Eko Sediyono. (2021). Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, *10*(2), 118–123. https://doi.org/10.22146/jnteti.v10i2.1421

Le, H. (2020). Detection of Fake Reviews on Social Media Using Machine Learning Algorithms. *Issues In Information Systems*, *21*(1), 185–194. https://doi.org/10.48009/1_iis_2020_185-194

Li, Y., Feng, X., & Zhang, S. (2016). Detecting Fake Reviews Utilizing Semantic and Emotion Model. *Proceedings - 2016 3rd International Conference on Information Science and Control Engineering, ICISCE 2016*, 317–320. https://doi.org/10.1109/ICISCE.2016.77

Lumbanraja, F. R., Saputra, R. A., Muludi, K., Hijriani, A., & Junaidi, A. (2021). Implementasi Support Vector Machine Dalam Memprediksi Harga Rumah Pada Perumahan Di Kota Bandar Lampung. *Jurnal Pepadun*, *2*(3), 327–335. https://doi.org/10.23960/pepadun.v2i3.90

Pasaribu, B. E., Herdiani, A., & Astuti, W. (2019). *Deteksi Fake Reviews Menggunakan Support Vector Machine*. *6*(2), 8788–8797.

Prakoso, B. H. (n.d.). *Pengaruh Preprocessing Data pada Metode SVR dalam Memprediksi Permintaan Obat*.

Putu, N., Widiari, A., Agus, M., Suarjaya, D., & Putra Githa, D. (n.d.). *Teknik Data Cleaning Menggunakan Snowflake untuk Studi Kasus Objek Pariwisata di Bali*.

Suryawati, E., Munandar, D., Riswantini, D., Abka, A. F., & Arisal, A. (2018). POS-Tagging for informal language (study in Indonesian tweets). *Journal of Physics: Conference Series*, *971*(1). https://doi.org/10.1088/1742-6596/971/1/012055

*name of corresponding author