JURNAL EMACS
e-ISSN: 2686-2573
[Engineering, MAthematics and Computer Science] Vol.4 No.3 September 2022: 103-107 DOI: 10.21512/emacsjournal.v4i3.8670

Comparing SVM and Naïve Bayes Classifier for Fake News Detection

Nurhasanah[1*], Daniel Emerald Sumarly[2], Jason Pratama[3], Ibrahim Tan Kah Heng[4], Edy Irwansyah[5]

[1] Statistics Department, School of Computer Science,
[2,3,4,5] Computer Science Department, School of Computer Science,
Bina Nusantara University,
Jakarta, Indonesia 11480
nurhasanah001@binus.ac.id; daniel.sumarly@binus.ac.id; jason.pratama001@binus.ac.id;
ibrahim.heng@binus.ac.id; eirwansyah@binus.edu

*Correspondence: nurhasanah001@binus.ac.id

Abstract – *Fake news has been evolving into a problem that is getting even more challenging. Technology has been misused to spread false information about many things,such as war, pandemics, and the stock market. Unfortunately, this issue is not a big deal for some people without conscious consumption of that news. Hence, being part takes a role in combating the spread of false information using the advancement of technology. This study proposed two methods of machine learning model, Support Vector Machine (SVM) and Naïve Bayes, to classify fake news. Furthermore, to assert the applicability of models by examining news articles dataset which contain two labels, reliable and unreliable news. The higher accuracy is 96% using the SVM model.*

Keywords: *Fake news; Machine learning; SVM, Naïve Bayes, Classification*

I. INTRODUCTION

Because of the rapid development of technology, information and news can be obtained quickly and easily. Surfing websites, blogs, and social media can access the message in minutes. However, sophisticated information technology is a double-edged sword. On the one hand, helping people to consume updated news conveniently, and on the other hand, much news consumed by the public is fake news that is not yet known. Misinformation in the news has caused harm to many parties. The most frequent fake news stories include pandemic news, stock exchanges, and especially news about the recent war between Russia and Ukraine. In a short time, much fake news has been circulating, which can trigger a bigger impact on the war. An analysis has found that fake news continues to increase over time (Zhou etal., 2019). Therefore, detection of fake news is paramount. While technological sophistication exacerbates the problem, technologies such as artificial intelligence (AI) can be used to resolve hoaxes (Cassauwers, 2019).

Machine learning is part of artificial intelligence that can help create systems to learn and perform several actions (Ahmed et al., 2021). Generally, machine learning is used for various predictions or for detecting fraud. Machine learning algorithms are used to vary and must be trained with a dataset. The model from the training result can be used to classify or detect fake news. To detect fake news, several researchers created algorithms or systems to detect fake news based on the content, text, and language style contained in news articles, blogs, and social media. Identifying and classifying fake news by how the author or writers use language. (Torabi Asr & Taboada, 2019) found that fake news often uses words related to scandal, death, and terror. In addition, many language styles in misleading news are intentionally exaggerated or overly dramatic, and the use of second-person pronouns is directly related to fake news (Hancock et al., 2007; Rashkin et al., 2017). Using AI technology to overcome the frequent and rapid emergence of fake news.

Some studies already consent about the challenge to detect fake news detection using machine learning model. We use two algorithms: Support Vector Machine (SVM) and the Naïve Bayes algorithm. According to research, SVMs outperform avariety of supervised machine learning algorithms for fake news detection. In another side employing Naïve Bayes which using probability to deal with the classification task and provide lower accuracy. We compare both method using open-source dataset (site Kaggle) to define which algorithm give higher accuracy using the fake news dataset.

1.1 Literature Review

Fake news has been a never-ending problem. Almost every major incident or problem has its fake news spread across different platforms, either websites or social media platforms. Giant social media companies such as Google, Facebook, and Twitter have made various efforts to prevent fake news propagation. Nevertheless, determining news veracity becomes increasingly difficult with the ease and speed of fake news propagation.

Over the years, scientists have tried to create a system that can find out news veracity and detect fake news automatically. Although, the impacts of fake news have pushed them to create a system with many different ways to achieve the best result. Twitter, one of the biggest social media, has become a primary source of information/ news nowadays. Unfortunately, become a huge amount of fake and manipulative news. Pavlyshenko from Ukraine has made a model capable of analyzing informal trends and detecting fake news on Twitter posts about an invasion of Russia to Ukraine (Pavlyshenko, 2022). A group of researchers (Chen et al., 2021) made a tool to detect fake news related to the COVID- 19 pandemic. Another method focused on preventing and minimalizing cyber terrorism on social media platforms (Divya Tiwari & Surbhi Thorat, 2021).

Automation systems play a significant role in identifying fake news. The proposed method varies and depends on the datasets used. One of the common methods is content-based detection. Content-based fake news detection uses information and text features in news articles and social media posts such as title, headline, image, and video. (Shu et al., 2017) have described and explained social and psychological theories related to fake news and the patterns in social media platforms. Many text characteristics and features such as lexical, sentence segmentation, and tokenization have helped researchers to create methods to classify fake news easier. For example, creating a model that detects potential clickbait information based on title and social media posts on Facebook and Reddit (Aldwairi & Alwahedi, 2018). Extracting news article text features and processing them through a model consisting of machine learning and natural language processing (Jain et al., 2019). (Lai et al., 2022) used word vectorization to convert words into numeric values to be understood by their neural network model. The result of the models made can always be improved with many more methods and techniques (Aldwairi & Alwahedi, 2018;ArunKumar et al., 2020; Farokhian et al., 2022; Humayoun, 2022) helping to improvise the accuracy of information appearing on the internet.

The growth and improvement of research methods and models on fake news detection have brought satisfying results. A fake news detection model made by (Aldwairi & Alwahedi, 2018) has an accuracy of up to 99.4%. Chauhan and Palivela (Chauhan & Palivela, 2021) improved the models and gained ground- breaking 99.88% accuracy. Tough, behind the experiment acquired almost-perfect results, the models may not always perform with the same accuracy. It turns out that the datasets used by researchers have a huge impact on their model performances. (Farokhian et al., 2022)stated that the dataset they used, FakeNewsNet, is a difficult dataset to be checked. While other models achieved 99% accuracy in some datasets, they barely achieved 85% when tested with FakeNewsNet.

Some challenges try to be faced in making fake news detection models. One of those challenges is non-Latin characters in the news. (Humayoun, 2022) studied Urdu fake news detection. The Urdu vocabulary consists of Persian, Arabic, and South Asian native languages. Some of the computing challenges are lack of capitalization, diacritic mark uses, and space not being a reliable word boundary marker. Because of these, the words and sentences have to be pre-processed through some steps in natural language processing such as diacritic removal, lemmatization, and more. Another challenge is detecting fake news in its early phase. Most existing models detect fake news 12 hours after its propagation, not being able to prevent it, remembering how fast information spreads.

Most fake news detection systems are based on texts in news articles and social media posts. Even though those models have achieved high accuracy, the news contains text and some images and videos that bring spurious and false information. The text content of news brings accurate and real news, yet the existence of fake images or videos makes question the veracity of the news. A model has been made to detect whether an image in a news article or social media post is real or edited using Convolutional Neural Network (CNN) (AlShariah & Khader, 2019), and the model achieved 97% accuracy. However, the research is only one of few that focus on fake news detection, not based on text. Many more fake news detection methods are still waiting to be explored and developed.

II. METHODS

In the following, we propose methods consisting of machine learning algorithms, a dataset to train the algorithms, and performance evaluation metrics to predict the veracity of news in the dataset.

2.1 Dataset

In order to distinguish between real and fake news, the dataset was taken from

Kaggle (Kaggle, 2018). The dataset has about 20,000 articles which included differentiated real and fake news. The dataset consists of 5 columns, namely id, title, author, and a label that indicates real or fake news.

2.2 Algorithms

2.2.1 Support Vector Machine (SVM)

The major purpose of this SVM is to determine the cut plane in an N-Dimensional space (Ahmad et al., 2020). For example, if the space is a two- dimensional plane, the intersection is a one-dimensional line. This cutting plane is used to determine data point classification. The mathematically displayed formula for SVM can be seen in equation 1.

$$J(\theta) = \frac{1}{2}\sum_{j=1}^{n}\theta_j^2 \quad (1)$$

To fit data points that are not easily separated or multidimensional data points by utilizing linear kernel function in Equation 2.

$$\theta^t x^{(i)} \geq 1, y^{(i)} = 1, \quad (2)$$

$$\theta^T x^{(i)} \leq -1, y^{(i)} = 0 \quad (3)$$

2.2.2 Naïve Bayes Classifiers

Naive Bayes applies the Bayes rule, assuming that the attributes are conditionally independent (Webb, 2016). The Naïve Bayes classifier provides a mechanism to use the information in the dataset to estimate the probability of class with a given object. Choose the highestprobability and use it for classification. This classifier is well-known for its computational efficiency and faster processing time compared to other classification algorithms. The mathematical formula for Nave Bayes is defined in Equation 4.

$$P(v|x) = \frac{P(y)\,P(x|y)}{P(x)} \quad (4)$$

Our detail to compare the SVM and Naïve Bayes as shown in Figure 1. Starting by input the dataset from news articles, a DataFrame was created. The DataFrame spreads the data from the dataset to a 2- dimensional data structure with rows and columns consisting of the news text features, making it easier to perform operations on each data. Meanwhile, the articles' text features had stop words such as 'is', 'an', 'we', and 'themselves' that needed to be removed. Therefore, the algorithm focused more on the important (1i)nformation in the text features and defined vectorizer parameters using TF-IDF Vectorizer. The vectorizer helps determine the stop words to be removed and create predictions based on words. For classification, exploit two machine learning methods that evaluate by a confusion matrix.

2.3 Performance Evaluation Metrics

We used different metrics to evaluate the performance of the algorithms on the datasets. The metrics consist of accuracy, precision, recall, and F1-score. These metrics are based on the confusion matrix, which is a performance measurement where the output can be two or more classes (Pandey et al., 2022) – in this case, the output will be fake and real news.

The confusion matrix is a table with 4 different classifications of predictedand actual values, which are:

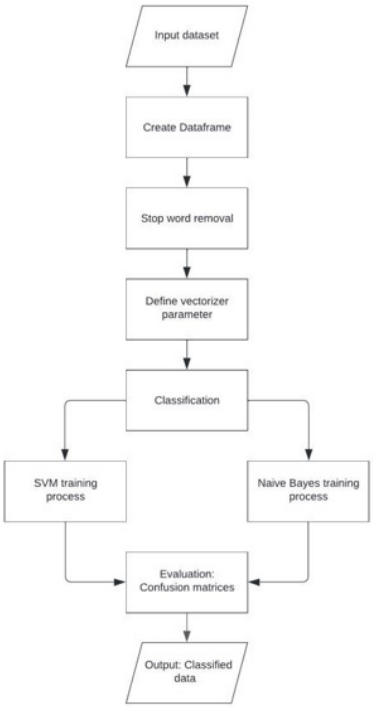

Figure 1. Flowchart of the training process with SVM and Naïve Bayes algorithm

- TP (true positive) determines the articles that have been correctly classified as real news.
- FN (false negative) determines the articles is true and predicts as fake news.
- FP (false positive) determines the articles is fake and predicts as real news.
- TN (true negative) determines the articles that have been correctly classified as fake news. The detail of confusion metric can be seen in Table 1

Table 1. Confusion Metric of Two Clasess

		Predicted Class: Positive	Predicted Class: Negative
True Class	**Positive**	TP	FN
	Negative	FP	TN

The value of confusion metric is used to calculate performance model: accuracy, precision, recall and F1-score.

2.3.1 Accuracy

Accuracy is the most intuitive and used metric, and it is a ratio orpercentage of correctly predicted observations. The Equation 5 is used tocalculate the accuracy:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

The algorithm has high accuracy, but keep in mind that a news article can be predicted true while it's false (false positive) or vice versa. This happens when a dataset is unbalanced. Therefore, accuracy alone is not enough to

determine the veracity of the article, hence the need for precision,recall, and F1-score metrics (Ahmad et al., 2020).

2.3.2 Precision

The precision value represents the ratio of correctly predicted positive observations to the total of predicted positive observations. In this case, the precision value shows the number of articles that are marked as true out of all the positively predicted articles. The precision value formula is:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (6)$$

2.3.3 Recall

The recall value represents the ratio of correctly predicted positive observations to the total number of the true positive class. Here is the equation to calculate the precision value:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (7)$$

2.3.4 F1-score

The F1-score combines precision and recall into a single metric and its value is an average of precision and recall.

$$F1 - \text{score} = 2\frac{\text{Precision x Recall}}{\text{Precision+Recall}} \quad (8)$$

III. RESULTS AND DISCUSSION

In Figure. 2 shows the confusion matrix for SVM algorithm testing on the dataset, whereas Figure. 3 shows the confusion matrix result on the Naïve Bayes algorithm.

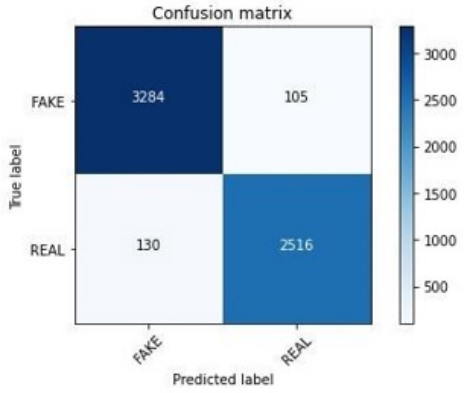

Figure 2. Confusion matrix result for SVM algorithm

From Figure 2. the total number of articles correctly predict as real news are 2516 and the number of fake news correcty predict (6i)s 3284 articles.

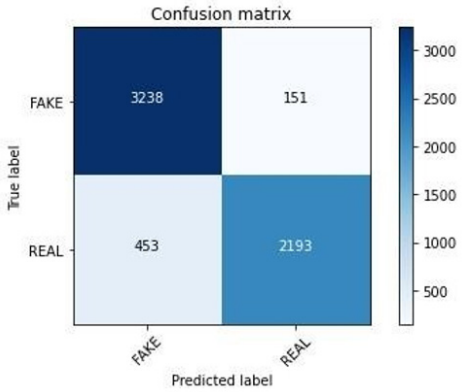

Figure 3. Confusion matrix result for Naïve Bayes algorithm

According to Figure 3, the total number of articles correctly predicted as real news is 2193, while the total number of articles correctly predicted as fake news is 3238 result of evaluation metrics for both algorithms can be seen in Table 2.

Table 2. SVM and Naïve Bayes Performance Results

Algorithm	Evaluation Metrics			
	Accuracy	Precision	Recall	F1-score
SVM	0.96	0.97	0.96	0.96
NB	0.89	0.95	0.87	0.90

Table 2. shows the performance metric evaluation of SVM and Naïve Bayes using The Fake News dataset. The results show that both algorithms achieved a high value on each evaluation metric. However, there are significant differences between the two results, especially in accuracy, recall, and F1-score.

Support Vector Machine achieved 96% accuracy, 97% precision, 96% recall, and an F1 score of 96%. Meanwhile, Naïve Bayes achieved 89% accuracy, 95% precision, 87% recall, and an F1 score of 90%. The results show that both algorithms nearly classified alldata correctly.

IV. CONCLUSION

Fake news as a long-standing problem has been evolving throughout the years, making it a difficult problem and requires in-depth knowledge and attention to detail in many aspects of the news beingresearched.

In this paper, the approach is implementing a simple method to detect the veracity of the news, with a dataset containing thousands of news articles, including their text features. The Support Vector Machine (SVM) and the Naïve Bayes algorithm are two machine learning algorithms used to detect fake news. SVM algorithms identify fake and real news with almost-perfect accuracy. However, the experiment operating one algorithm alone is not guaranteed to be effective nor applicable when it comes to fresh real-world news but gives a new insight into detecting fake news with a machine learning algorithm.

REFERENCES

Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). Fake News Detection Using Machine Learning Ensemble Methods. Complexity, 2020, 1–11. *https://doi.org/10.1155/2020/8885861*

Ahmed, A. A. A., Aljabouh, A., Donepudi,P. K., & Choi, M. S. (2021). Detecting Fake News Using Machine Learning: A Systematic Literature Review.

Aldwairi, M., & Alwahedi, A. (2018). Detecting Fake News in Social Media Networks. Procedia Computer Science, 141, 215–222. *https://doi.org/10.1016/J.PROCS.2018.10. 171.*

AlShariah, N. M., & Khader, A. (2019). Detecting Fake Images on Social Media using Machine Learning. International Journal of Advanced Computer Science and Applications, 10(12). *https://doi.org/10.14569/IJACSA.2019.010 1224.*

ArunKumar, K., Preethi, G., & Vasanth, K. (2020). A Study of Fake News Detection using Machine Learning Algorithms. International Journal of Technology and Engineering System (IJTES), 11(1), 1–7.

Cassauwers, T. (2019, April 15). Can artificial intelligence help end fake news? Horizon Magazine.

Chauhan, T., & Palivela, H. (2021). Optimization and improvement of fake news detection using deep learning approaches for societal benefit. International Journal of Information Management Data Insights, 1(2), 100051. *https://doi.org/10.1016/J.JJIMEI.2021.100 051.*

Chen, M., Chu, X., & Subbalakshmi, K. P. (2021). MMCoVaR: Multimodal COVID- 19 Vaccine Focused Data Repository for Fake News Detection and a Baseline Architecture for Classification.

Divya Tiwari, & Surbhi Thorat. (2021). Terrorism And Fake News Detection. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 313–323. *https://doi.org/10.32628/CSEIT217670.*

Farokhian, M., Rafe, V., & Veisi, H. (2022, April 10). Fake news detection using parallel BERT deep neural networks.

Hancock, J. T., Curry, L. E., Goorha, S., & Woodworth, M. (2007). On Lying and Being Lied To: A Linguistic Analysis of Deception in Computer-Mediated Communication. Discourse Processes, 45(1), 1–23. *https://doi.org/10.1080/016385307017391 81.*

Humayoun, M. (2022). The 2021 Urdu Fake News Detection Task using Supervised Machine Learning and Feature Combinations.

Jain, A., Shakya, A., Khatter, H., & Gupta, A. K. (2019). A smart System for Fake News Detection Using Machine Learning. 2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), 1–4. *https://doi.org/10.1109/ICICT46931.2019. 8977659.*

Kaggle. (2018). Fake News. Kaggle. Available: *https://www.kaggle.com/competitions/fake-news/data.* [Accessed: 25-Jun-2022].

Lai, C.-M., Chen, M.-H., Kristiani, E.,Verma, V. K., & Yang, C.-T. (2022). Fake News Classification Based on Content Level Features. Applied Sciences, 12(3), 1116. *https://doi.org/10.3390/app12031116*

Pandey, S., Prabhakaran, S., Subba Reddy, N. v, & Acharya, D. (2022). Fake News Detection from Online media using Machine learning Classifiers. Journal of Physics: Conference Series, 2161(1), 012027. *https://doi.org/10.1088/1742-6596/2161/1/012027.*

Pavlyshenko, B. M. (2022). Methods of Informational Trends Analytics and Fake News Detection on Twitter.

Rashkin, H., Choi, E., Jang, J. Y., Volkova, S., & Choi, Y. (2017). Truth of Varying Shades: Analyzing Language in Fake News and Political Fact-Checking. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 2931– 2937. *https://doi.org/10.18653/v1/D17-1317.*

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake News Detection on Social Media: A Data Mining Perspective.

Torabi Asr, F., & Taboada, M. (2019). Big Data and quality data for fake news and misinformation detection. Big Data & Society, 6(1), 205395171984331. https://doi.org/10.1177/2053951719843310

Webb, G. I. (2016). Naïve Bayes. In Encyclopedia of Machine Learning and Data Mining (pp. 1–2). Springer US. *https://doi.org/10.1007/978-1-4899-7502- 7_581-1.*

Zhou, X., Zafarani, R., Shu, K., & Liu, H. (2019). Fake News. Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, 836–837. *https://doi.org/10.1145/3289600.32913.*