JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 1, Januari 2023, Page 170-177
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i1.5386

1. INTRODUCTION

News is information obtained through the form of newspapers, radio and online media [1]. Hoax is information or news that contains things whose truth has not been confirmed and is not following reality [2]. The impact of spreading hoax news is hazardous because it can cause divisions between communities, especially during the Covid-19 pandemic [3]. Hoax news in Indonesia, especially regarding the Covid-19 pandemic, is still a hot topic of conversation.

In contrast to the old days, people still find it difficult to find news, whereas, in today's digital era, it is effortless to get news or information. With increasingly sophisticated digital technology, a way has also been found to detect hoax news using algorithms that can determine whether the news is hoax news or not [4]. But until now, there is still no system that can see the reasons for the model to classify hoax news and non-hoax news. This is what needs to be done, considering that not everyone can understand what the hoax detection system is doing. Considering that not many detection systems have Explainable AI to explain the detection. Each hoax news has different words, so each piece has its point of view. By knowing the words in the news, it is expected to know whether the news leads to factual news or hoaxes.

Much research has been done on the detection of hoaxes related to news, such as research conducted by Ropikoh (2021), who examined the application of the Support Vector Machine algorithm for the classification of hoax news [5]. This research produces an accuracy value of 90.46%, a precision value of 66.86%, and a recall value of 64.53%. Furthermore, there is research from Ismayanti (2021), which examines the detection of Indonesian-language hoax content on Twitter using the Word2Vec expansion feature [6]. The research resulted in an accuracy value of 82.79% and an F1-Score of 0.8278.

Frista (2018), conducted research on the detection of hoax content using the Levenshtein distance method [4]. However, the test obtained a limit of 0.0014 in the results of scenario 2 due to the fact that more data was used than in scenario 1. The Tf-Idf calculation calculates the number of word occurrences in documents which causes the resulting performance to be poor. Alvanof (2020), conducted research on the detection of hoax content using 4 types of algorithms namely, multilayer perceptron or MLP, Naïve Bayes, SVM, and Random Forest [1]. The highest accuracy was obtained, namely, the Random Forest algorithm with a value of 75.37%. The random forest algorithm gets the highest accuracy value, so the precision and recall level is also high.

Palma (2021), conducted research on the text classification of COVID-19 hoax news articles using the K-Nearest Neighbor algorithm [3]. It was found that the test results obtained an accuracy value of 48% using a value of k = 5 and the percentage of data was 80% training data and 20% test data. However, the results obtained are not very good because there are several classes that have a small amount of data that affects the model's ability to classify these classes. Putra (2021), conducted research on the detection of the use of abusive sentences in Indonesian texts using the IndoBERT method [7]. Obtained from the results of testing the F1-Score value with the IndoBERT method obtained a value of 76.32%. The IndoBERT method has optimal results because it utilizes transformers which study contextual relationships between words in the text.

Pradana (2019), conducted research on the detection of hoaxes on Android-based social media [8]. In his research, good results were obtained with a minimalist and not confusing interface. However, the news analyzed can only be in Indonesian. Aldwairi (2018), conducted research on detecting fake news on social media networks [9]. The research conducted obtained a high classifier value when using a logistics classifier with a precision value of 99.4%.

Then, there is research on Explainable AI using the LIME framework by Ribeiro (2016), which examines LIME explanations [10]. Research conducted on experiments on datasets about religion with several sentences that will be predicted related to religion. The results are that the LIME framework's accuracy reaches 89%. Saini (2021), conducted research on optimizing the AI LIME Explainable method to BO-LIME or (Bayesian Optimation-Local Interpretability Model-agnostic Explainability) [11]. The BO-LIME method works significantly based on Bayesian optimization and provides good explanatory stability compared to the LIME model.

Based on this background, the detection of hoaxes about Covid-19 is still interesting for further research. The Explainable AI method using LIME is exciting to study because the framework can see the contents of the machine learning black box and then make it a white box or transparent, considering that it is rare to find research using Explainable AI. Then the SVM and logistic regression methods are used as models for the detection of hoax and non-hoax news. The two methods need to be compared to find the best classification report value so that explainable AI can see the black box correctly. The limitations of the problem in this research are datasets taken through the websites turnbackhoax.id, hoaxbuster, and Kompas. Then, this study uses a dataset with text in the form of news in Indonesian. The dataset used is 500 data. This final project will look at the performance of

Submitted: **22/12/2022**; Accepted: **12/01/2023**; Published: **28/01/2023**

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 1, Januari 2023, Page 170-177
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i1.5386
machine learning models from Support Vector Machine and Logistic Regression and what words are in hoax and non-hoax news related to Covid.

Submitted: **22/12/2022**; Accepted: **12/01/2023**; Published: **28/01/2023**