



1. INTRODUCTION

Many famous figures like to write articles or novels in Indonesia. The author is a term for people who have written works, both articles, and novels [1]. Many well-known Indonesian writers, such as KH. Abdurrahman Wahid, Dahlan Iskan, Pramoedyana Ananta Toer, and others. Written works can be in the form of novels, articles, or other writing-related things.

Every writer has a different style of language, and this is what makes each written work unique. Writing can be a fact or a work of fiction. A fact is a statement in which the statement can be proven false [2]. Fact is also the opposite of opinion [3]. The opinion is an expression of someone who cannot be proven [2].

A work of fiction is a work of imagination, not based on facts, although it may be based on a real story or situation [4]. An example of this work of fiction is a short story or a novel. A work of fiction usually has communicative language and an exciting storyline. In writing, it usually uses language that is not standard or language used in everyday life.

It is essential to identify the writings of well-known figures in Indonesia because it can help us understand the thoughts, ideas, and ideas that they convey. In addition, recognizing the writings of well-known figures in Indonesia can broaden our horizons and knowledge about Indonesian history, culture, and culture. Identifying these papers can also help us understand the development of thoughts and ideas in Indonesia and how these works influence Indonesian society and civilization.

For example, in 2020, there was research to find out changes in writing styles carried out by Iyer and Vosoughi with the title "Style Change Detection Using BERT." From the research that has been done, it is found that the accuracy of the Logistic Regression method is higher than Naïve Bayes, namely 91%, while Naïve Bayes is 90% [5]. However, this research still does not use explainable AI in its process, so the results obtained are still in the form of a black box which is still tricky for humans to understand. Explainable AI itself is a technique or method for understanding the output of a machine learning algorithm so that the output results are understandable to humans and easy to make decisions.

Based on research conducted by Aborisade and Anwar with the title "Classification for authorship of tweets by comparing logistic regression and naive Bayes classifiers," they researched the classification of authorship attribution or authorship ownership from Twitter data obtained through Twitter's RESTful API. The methods used alone are Logistic Regression and Naïve Bayes. From the research that has been done, it is found that the accuracy of the Logistic Regression method is higher than Naïve Bayes, namely 91%, while Naïve Bayes is 90%. The performance obtained depends on the corpora that have been trained, the more corpora that are trained, the higher the performance [6].

Based on research conducted by Zheng et al. with the title "The email author identification system based on support vector machine (SVM) and analytic hierarchy process (AHP)," they conducted research related to author attribution, author characteristics, and camouflage detection in an electronic mail (email) using Support Vector Machine and Analytic Hierarchy methods. The researcher collects a dataset from several emails sent, with 80% as train data and 20% as test data. From the data that has been collected, it will be identified who is the owner of the email by analyzing the writing procedures, choosing words, and so on. The resulting accuracy from the Support Vector Machine method is more than 95% [7].

Based on research conducted by Mathews with the title "Explainable artificial intelligence applications in NLP, biomedical, and malware classification: a literature review," they conducted research related to Natural Language Processing on tweets, which was then followed by cancer detection in biomedical signal classification, then researchers also conducted research related to malware detection on Windows PCs. The method used for biomedical classification uses the Deep neural network with an accuracy of 94%, while the classification of tweet data uses the XGB method with an accuracy of 84%. For malware detection, Microsoft's Windows PC malware classification is used, which has almost perfect accuracy, namely 99.83%. Of the three classification methods, LIME will be used as explainable [8].

Based on research conducted by Aslam et al. entitled "Interpretable Machine Learning Models for Malicious Domains Detection Using Explainable Artificial Intelligence (XAI)," they conducted research related to domains that were indicated to be dangerous using the Decision Tree, Naïve Bayes, Random Forest, Ada Boost, XGB, and Cat Boost methods. The data used alone is 45,000 samples from each of the dangerous domains and non-hazardous domains. Researchers use Explainable AI to assist in the models' explanations. The results obtained from the models used have high accuracy, but the XGB model has the highest accuracy, 98.18%. The framework used in Explainable AI is LIME [9].

Based on this background, a system was built to find out the style of language of several well-known writers in Indonesia with explainable AI. This research has never been done before, so it is still interesting to do further research. Explainable AI is later used to understand the output generated from a machine learning



algorithm. The algorithm used in this study is BERT and also Naïve Bayes. The algorithm is used to classify a paper from a well-known figure in Indonesia.

By using the BERT and Naïve Bayes algorithms, a comparison can be made of the two algorithms' performance in classifying a paper from a well-known figure in Indonesia. Then, by building this explainable AI system, it can be seen how the style of language used by well-known figures in Indonesia in their writings.

For the limitations of this study, the authors identified only KH. Abdurrahman Wahid, Dahlan Iskan, Emha Ainun Nadjib, and Pramodya Ananta Toer. Then the dataset used only Indonesian language datasets with no more than 2000 data.

This final project aims to compare the performance of the BERT and Naïve Bayes algorithms in classifying papers written by well-known figures in Indonesia, namely, KH. Abdurrahman Wahid, Dahlan Iskan, Emha Ainun Nadjib, and Pramodya Ananta Toer. In addition, this research was also conducted to find out the style of language used by the character in writing his essay.