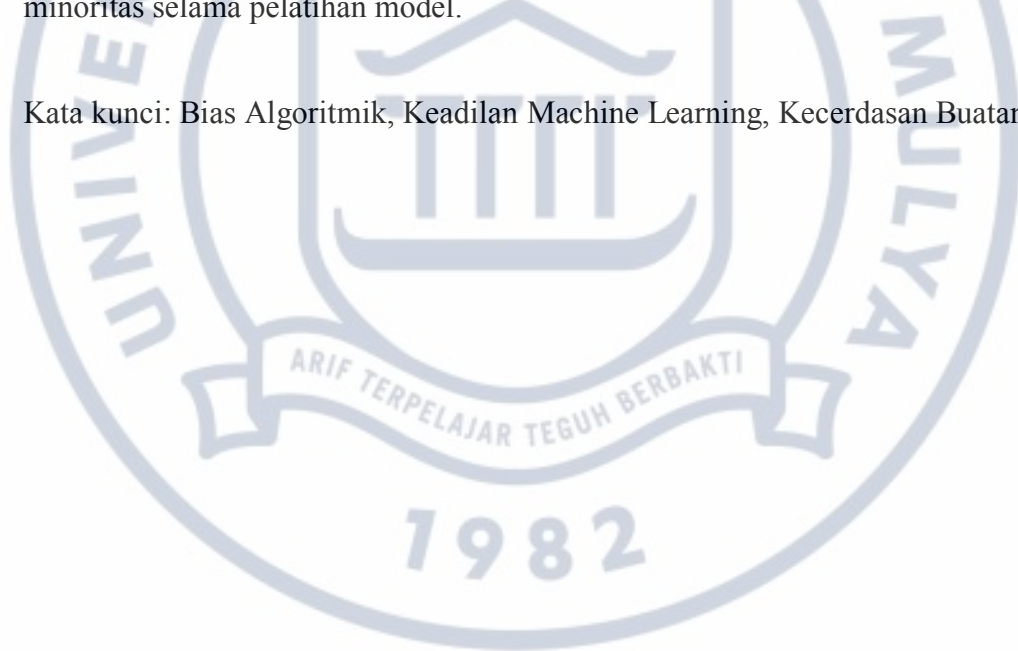


ABSTRAK

Machine Learning adalah bidang studi yang berfokus pada algoritma komputer yang mampu beradaptasi dan meningkatkan diri secara otomatis melalui pelatihan berulang dengan dataset pelatihan yang diberikan. Pada umumnya algoritma tersebut dianggap netral, adil, dan tidak memihak, serta merupakan sistem yang lebih objektif dan absolut dalam membuat keputusan otomatis. Tetapi, ketika algoritma tersebut memiliki bias terhadap suatu kelas tertentu, maka algoritma tersebut sudah tidak tergolong sebagai algoritma yang adil. Dalam studi ini, penulis melakukan penelitian dan percobaan untuk mencari seberapa penting keadilan algoritma Machine Learning dan apabila proporsi dataset pelatihan dapat mempengaruhi keadilan sebuah algoritma Machine Learning. Percobaan tersebut dilakukan dengan melatih beberapa model algoritma dengan dataset pelatihan yang sudah dimodifikasi agar penulis dapat melihat korelasi dari proporsi data pelatihan dengan kinerja dari algoritma Machine Learning tersebut. Melalui penelitian ini penulis menemukan bahwa untuk mengurangi potensi adanya bias dan ketidakadilan dalam proses pelatihan model Machine Learning, pengembang model tersebut tidak harus mengurangi jumlah data mayoritas yang terdapat dalam dataset pelatihan. Penulis juga mendapatkan bahwa faktor yang mempengaruhi keadilan dari sebuah model merupakan perbedaan frekuensi ditemuinya data mayoritas dan minoritas selama pelatihan model.

Kata kunci: Bias Algoritmik, Keadilan Machine Learning, Kecerdasan Buatan



ABSTRACT

Machine Learning is a field of study that focuses on computer algorithms that are capable of adapting and improving themselves automatically through repeated training with a given training dataset. In general, these algorithms are considered neutral, fair, and impartial, as well as being a more objective and absolute system in making automated decisions. However, when the algorithm has a bias towards a certain class, then the algorithm is no longer classified as a fair algorithm. In this study, the author conducts research and experiment to find out the importance of Machine Learning algorithm fairness and if training dataset proportions can affect the fairness of a Machine Learning algorithm. The experiment is performed by training several models with modified training dataset in order for the author to observe the correlation of training data proportions with the performance of said Machine Learning algorithm. Through this research the author discovered that in order to reduce the potential of bias and unfairness existing in the process of Machine Learning model training, the developer of said model doesn't have to reduce the amount of major data in the training dataset. The author also found that the factor which affects the fairness of a model is the difference of the frequency of the model coming across major and minor data during the model's training.

Keywords: Algorithmic Bias, Machine Learning Fairness, Artificial Intelligence

