

Abstract

A Hybrid Approach for Clickbait Detection

by:

Shaqina Yasmin
18/429297/PA/18688

To generate revenue, online news media employ a variety of attention-getting strategies. Clickbait is one of the techniques used. It is an enticing headline whose purpose is to attract readers' clicks. The headlines of clickbait articles frequently use catchy or exaggerated language. Clickbait articles mostly contain poor content. Due to the poor quality of the content, the promises made in the headlines cannot be met. The presence of clickbait is disruptive to readers.

A variety of approaches, such as traditional machine learning, have been used to overcome clickbait issues. Traditional machine learning methods, however, rely significantly on feature engineering. This research presents a hybrid technique that combines traditional machine learning with deep learning to overcome this issue. There are eight models compared against each other, from logistic regression, SVM, CNN1D, CNN2D, CNN1D with logistic regression, CNN1D with SVM, CNN2D with logistic regression, CNN2D with SVM.

Combining Convolutional Neural Network (CNN) with traditional machine learning models such as logistic regression or Support Vector Machine (SVM) demonstrates that the hybrid approach outperformed the stand-alone traditional machine learning and deep learning approaches in the Indonesian news headlines from twelve online news sources. The CNN and SVM combination achieves the highest performance in terms of accuracy of 91.2% compared to other models.

Keywords: *clickbait detection, machine learning, deep learning, hybrid approach, convolutional neural network, support vector machine, logistic regression*