FAKE NEWS DETECTION USING MACHINE

LEARNING AND PYTHON

ABSTRACT

The proliferation of misleading information in everyday access media outlets such as social media feeds, news blogs, and online newspapers have made it challenging to identify trustworthy news sources, thus increasing the need for computational tools able to provide insights into the reliability of online content. In this paper, we focus on the automatic identification of fake content in online news. Our contribution is twofold.

\*\* First, we introduce two novel datasets for the task of fake news detection, covering seven different news domains. We describe the collection, annotation, and validation process in detail and present several exploratory analysis on the identification of linguistic differences in fake and legitimate news content.

\*\* Second, we conduct a set of learning experiments to build accurate fake news detectors. In addition, we provide comparative analyses of the automatic and manual identification of fake news.

This is how the misleading information and the fake news is detected with the help of machine learning. The primary source is the creation of the dataset with a finite amount of data present in it.

Fake news detection in social media is a process of detecting false information that is intentionally created to mislead readers. The spread of fake news may cause social, economic, and political turmoil if their proliferation is not prevented.

However, fake news detection using machine learning faces many challenges. Datasets of fake news are usually unstructured and noisy. Fake news often mimics true news. In this study, a data pre-processing method is proposed for mitigating missing values in the datasets to enhance fake news detection accuracy. The experimental results show that Multilayer Perceptron (MLP) classifier combined with the proposed data pre-processing method outperforms the state-of-the-art methods.

Furthermore, to improve the early detection of rumours in social media, a timese1ies model is proposed for fake news detection in social media using Twitter data. With the proposed model, computational complexity has been reduced significantly in terms of machine learning models training and testing times while achieving similar results as state-of-the-art in the literature. Besides, the proposed method has a simplified feature extraction process, because only the temporal features of the Twitter data are used. Moreover, deep leaming techniques are also applied to fake news detection. Experimental results demonstrate that deep learning methods outperformed traditional machine learning models. Specifically, the Naïve Bayes classification model achieved top performance.