Improving Fake News Detection Using K-means and Support Vector Machine Approaches

Authors : Kasra Majbouri Yazdi, Adel Majbouri Yazdi, Saeid Khodayi, Jingyu Hou, Wanlei Zhou, Saeed Saedy

Abstract : Fake news and false information are big challenges of all types of media, especially social media. There is a lot of false information, fake likes, views and duplicated accounts as big social networks such as Facebook and Twitter admitted. Most information appearing on social media is doubtful and in some cases misleading. They need to be detected as soon as possible to avoid a negative impact on society. The dimensions of the fake news datasets are growing rapidly, so to obtain a better result of detecting false information with less computation time and complexity, the dimensions need to be reduced. One of the best techniques of reducing data size is using feature selection method. The aim of this technique is to choose a feature subset from the original set to improve the classification performance. In this paper, a feature selection method is proposed with the integration of K-means clustering and Support Vector Machine (SVM) approaches which work in four steps. First, the similarities between all features are calculated. Then, features are divided into several clusters. Next, the final feature set is selected from all clusters, and finally, fake news is classified based on the final feature subset using the SVM method. The proposed method was evaluated by comparing its performance with other state-of-the-art methods on several specific benchmark datasets and the outcome showed a better classification of false information for our work. The detection performance was improved in two aspects. On the one hand, the detection runtime process decreased, and on the other hand, the classification accuracy increased because of the elimination of redundant features and the reduction of datasets dimensions.

Keywords : clustering, fake news detection, feature selection, machine learning, social media, support vector machine

Conference Title : ICCNSIP 2020 : International Conference on Communications, Networking, Signal and Image Processing

Conference Location : Melbourne, Australia

Conference Dates : February 03-04, 2020