

Fake News Detection using Temporal Features Extracted via Point Process

Problem Statement:

Many people use social networking services (SNSs) to easily access various news. There are numerous ways to obtain and share “fake news,” which are news carrying false information. To address fake news, several studies have been conducted for detecting fake news by using SNS-extracted features. There is a need to make an attempt to use temporal features generated from SNS posts by using a point process algorithm to identify fake news from real news. Temporal features in fake news detection have the advantage of robustness over existing features because it has minimal dependence on fake news propagators. Further, design a model which includes linguistic and user features alongside temporal features is the main need for effective detection of fake news from SNS posts.

Challenge: -To construct an efficient deep learning framework using linguistic and user features alongside temporal features of the news articles

Early studies attempted to detect fake news based on linguistic features extracted from texts. Recent studies used deep learning models to capture temporal–linguistic features using recurrent neural networks (RNNs), which capture temporal–linguistic features from a bag-of-words of user posts. Some researchers have used recursive neural networks based on the texts of a reply tree. Further examples include convolutional neural, hierarchical attention networks and neural-network models using discourse level structures. Temporal features are also useful for short term fake news detection. We have to design a model to handle/utilizes linguistic, user, and temporal features, which are easy to obtain, to capture the characteristics of fake news.

Experimental Design

Dataset: - Weibo , Twitter15, and Twitter16 released

Links:

https://www.dropbox.com/s/7ewzdrbelpmrnxu/rumdetect2017.zip?dl=0&file_subpath=%2Frumor_detection_acl2017

<https://github.com/yaqingwang/EANN-KDD18/tree/master/data>

Table 2: Summary of datasets

Dataset	Weibo	Twitter15	Twitter16
No. of true news	2351	371	204
No. of fake news	2313	363	205
No. of unverified news	-	373	205
No. of debunking	-	372	199
No. of training posts	2973	942	517
No. of validation posts	525	167	97
No. of test posts	1166	370	204

Evaluation Measures

Evaluation is measured in terms of Precision, Recall, F1-Score, Accuracy, false positive cases, and fake negative cases.

Software and Hardware Requirements

Python-based Deep Learning libraries will be exploited for development and experimentation of the project. Tools such as Anaconda Python, and libraries such as, Tensorflow, and Keras will be utilized for this process.

Previous benchmark Results: -

In previous benchmark results, there is a large scope to improve the results using deep learning techniques.

Dataset	Weibo			Twitter15					Twitter16							
	Acc.	T	F_1	F	Acc.	T	F	F_1	U	D	Acc.	T	F	F_1	U	D
SVM-TS	0.827	0.831	0.837		0.599	0.772	0.598	0.608	0.544		0.574	0.743	0.488	0.551	0.549	
CSI	0.780	0.750	0.803		0.556	0.601	0.631	0.550	0.530		0.507	0.552	0.511	0.475	0.443	
GRU-2	0.876	0.872	0.879		0.794	0.822	0.815	0.849	0.697		0.750	0.761	0.750	0.771	0.723	
PPC	0.914	0.912	0.917		0.806	0.748	0.840	0.807	0.730		0.778	0.803	0.760	0.711	0.767	

Mentor Name: Rohit Kumar Kaliyar