

DEBUNKING FAKE NEWS BY LEVERAGING SPEAKER CREDIBILITY AND BERT BASED MODEL

Thoudam Doren Singh, Divyansha, Apoorva Vikram Singh, Anubhav Sachan, Abdullah Khilji

IEEE/WIC/ACM International Joint Conference on
Web Intelligence and Intelligent Agent Technology (WI-IAT 2020)

TABLE OF CONTENTS

Introduction

Motivation

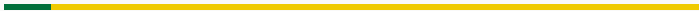
Feature

Extraction

Architecture

Results

INTRODUCTION



Fake news is a type of propaganda that consists of deliberate misinformation or hoaxes spread via traditional print and broadcast news media or online social media.

A new study kills the notion that fake news swung the US election to Trump

A team of Howard University researchers wants to know how disinformation impacts Black people

Information Overload Helps Fake News Spread, and Social Media Knows It

Trump may owe his 2016 victory to 'fake news,' new study suggests ■

Facebook, Google extend political ad ban amid misinformation rise

Facebook, Google extend political ad ban amid misinformation rise

Singapore: Study suggests that exposure to misinformation adversely impacts public confidence in COVID-19 vaccines

MOTIVATION



- Most of the fake news content is created by fabricating true news so it is more reasonable to perform multi-class classification that can help us assess the level of falsification of the fake news.
- Adding additional details about the subject can make the prediction model more efficient by assisting it in making more informed decisions.

FEATURE EXTRACTION



LIAR	Speaker2Credit
Statement	
Topic	
Speaker Profile	Global Credit History
Credibility Count	
Six different labels	

The choice of a political news dataset was due to the fact that false political news travels nearly three times faster and appeals to a larger audience than news of any other category.

Speaker Statement

Speaker Profile

Attributes

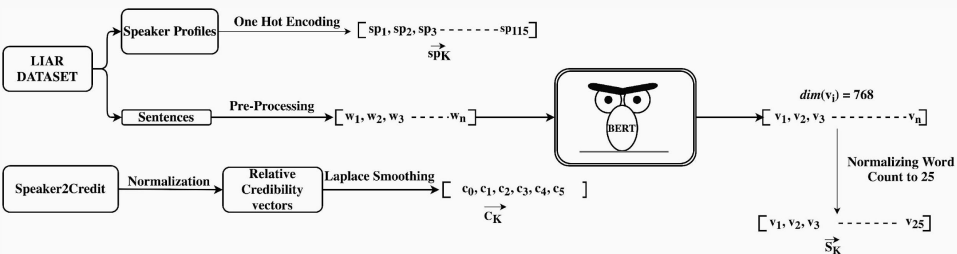
job of the speaker,
political party affiliation of the speaker,
state of residency of the speaker,
topic on which statement is made, and
the location of the speech

Speaker Credibility

Speaker2Credit

$$\text{credit}_{i\alpha} = \left(\frac{\text{count}(\mu_0|i) + \alpha}{\sum_{\mu} \text{count}(i) + 5\alpha} \cdots \frac{\text{count}(\mu_5|i) + \alpha}{\sum_{\mu} \text{count}(i) + 5\alpha} \right)$$

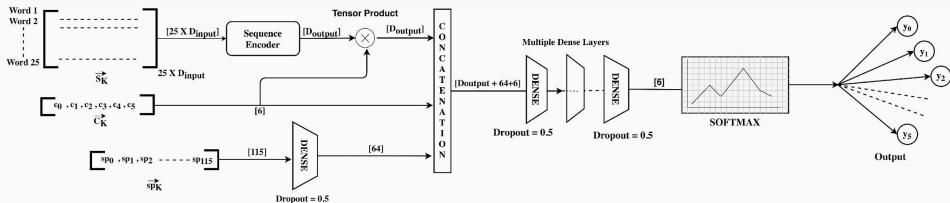
FEATURE GENERATION



ARCHITECTURE



PROPOSED ARCHITECTURE



RESULTS



- Dataset: LIAR (labeled articles from Politifact with speaker credentials)
- Baselines:
 - Hybrid-CNN (Proposed in LIAR dataset paper) [5]
 - Multi-source multi-class fake news detection (MMFD) [2]
 - LSTM-Attention [4]
 - Bi-LSTM [1]
 - CreditLSTM [3]
- Evaluation Metric: Accuracy (Fraction of dataset correctly labeled)

RESULTS

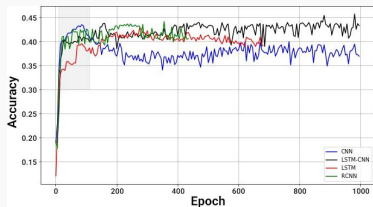


Figure 1: Accuracy Curves for Different Sequence Models

Model	Accuracy	Batch size
LSTM	0.427827	64
CNN	0.442835	32
LSTM-CNN	0.468040	128
RCNN	0.443182	128

Table 1: Accuracy comparison among different sequence encoding models.

Features	Accuracy
Word2Vec	0.2636
BERT	0.2718
Word2Vec + MetaData	0.2741
BERT + MetaData	0.3099
BERT + MetaData + Credit	0.4680

Table 2: Performance analysis of different feature models

Model	Accuracy
Hybrid-CNN [5]	0.274
MMFD [2]	0.348
LSTM-Attention [4]	0.415
Bi-LSTM [1]	0.415
CreditLSTM [3]	0.457
BERT + MetaData + Credit	0.4680

Table 3: Performance of previous works on the LIAR dataset



M. K. Balwant.

Bidirectional lstm based on pos tags and cnn architecture for fake news detection.

In 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pages 1–6, 2019.



H. Karimi, P. Roy, S. Saba-Sadiya, and J. Tang.

Multi-source multi-class fake news detection.

In Proceedings of the 27th International Conference on Computational Linguistics, pages 1546–1557, Santa Fe, New Mexico, USA, Aug. 2018. Association for Computational Linguistics.

-  A. Kirilin and M. Strube.
Exploiting a speaker's credibility to detect fake news.
In *Proceedings of Data Science, Journalism & Media workshop at KDD (DSJM'18)*, 2018.
-  Y. Long, Q. Lu, R. Xiang, M. Li, and C.-R. Huang.
Fake news detection through multi-perspective speaker profiles.
In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 252–256, Taipei, Taiwan, Nov. 2017. Asian Federation of Natural Language Processing.



W. Y. Wang.

“liar, liar pants on fire”: A new benchmark dataset for fake news detection.

In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pages 422–426, 2017.

QUESTIONS?

Reach out: anubhav_ug@ece.nits.ac.in
twitter.com/anubhav4sachan