Single Modal and Bimodal Fake News Detection: A Survey

Shubham Bhalerao[1], Jash Vyas[1], Meet Gogri[1], Kartiki Sanap[1], Prof. Seema Shrawne[2]

[1]B.Tech Student, Dept of Computer Engineering & IT, VJTI College, Mumbai, Maharashtra, India.
[2]Assistant Professor, Dept of Computer Engineering & IT, VJTI College, Mumbai, Maharashtra, India.

Abstract - *The spread of misinformation has skyrocketed with the rise of social media. Today, misinformation is more abundant and accessible than ever and it has shaken our trust in institutions, governments and even the COVID vaccine. Fake news also interferes with the ability of a user to discern useful information from the internet. Thus, an automated approach to detecting fake news has become a need of the hour. As an effort in this direction, this paper aims to perform a comprehensive survey on different machine learning and deep learning techniques used for the task of fake news detection. We explore various research papers using a single modal as well as bimodal approach for detecting fake news. Thus, we summarize various models and techniques used along with the datasets and corresponding accuracies for different models.*

Key Words: Fake News Detection, Machine Learning, Deep Learning, Natural Language Processing

1. INTRODUCTION

In today's technological world, a huge amount of data gets generated online every day. However most of the data which is flooded on the internet is fake and it is generated to attract the audience, to misguide people, to influence beliefs and decisions of people. Obtaining and spreading information through social media platforms has become extremely easy, which makes it difficult and nontrivial to detect fake news based merely on the content of news.

Several tech companies like Google, Facebook and Twitter have attempted to address this particular concern of fake news spread. Majority of the countries all across the world are trying to combat this challenge of fake news spread. Users continue to deal with the sites containing false information and whose involvement tends to affect the reader's ability to engage with actual news.

Most of the sites which contain such information also include a sharing option that implores users to disseminate the contents of the web page further. Social networking sites allow for efficient and fast sharing of material and thus, users can share the misleading information within a short time. Thus, to identify fake news and impede the spread of fake news is of pivotal importance in today's age.

2. LITERATURE REVIEW

Fake news detection is a very well-established and well-researched task in NLP. It is defined as determining from the news article including the head, body, images,etc whether the news can be determined as real or fake. Previously many researches have been conducted to detect fake-news on the internet and social media platforms using NLP tools and machine learning techniques.

The Fake News Challenge Stage-1 (FNC-1) which was held in 2017 featured many novel solutions for this problem. Riedel et al. and other teams who won in FNC-1 achieved accuracy close to 82% in the stance detection stage. After the completion of this competition, many different solutions have emerged using various NLP Techniques. The datasets used in FNC-1 are available publicly and we are close to having standard benchmarks to compare all the newly proposed techniques. Currently there are many other datasets available for fake-news detection like in [5].

2.1 Detecting Fake News in Social Media Networks

Monther Aldwairi et al. [1] proposed a solution to detect and filter out sites containing false and misleading information on the internet, using important features of the title and post to accurately identify fake posts. Different features like titles starting with numbers, all caps words, presence of question and exclamation marks, whether the user left the page immediately, and content related to the title- were extracted from the content of the webpage whose URL is provided. Most of the websites considered in this paper were from social media like Facebook, Forex and Reddit. WEKA classifiers like Random Forests, Naive Bayes, Bayes network, Logistic Regression etc. were used to train the model. For comparison of models, the paper used various metrics like precision, recall, F-measure and area under roc curve. F-measure of 99.3% and roc of 99.5% was achieved with the help of a logistic classifier.

2.2 Fake News Stance Detection Using Deep Learning Architecture (Cnn-lstm)

M. Umer et al. [2] has proposed a hybrid Neural Network architecture that combines the capabilities of CNN and LSTM for the detection of fake-news. The network architecture is used with two different dimensionality reduction approaches, Principal Component Analysis (PCA) and Chi-Square. The dataset used for training this hybrid model was the FNC-1 dataset. Also a comparison was made between the metrics of the proposed solution and already existing solutions like TalosComb, TalosTree, stackLSTM, BERT, XLNet, etc. Some researchers used different variations of LSTMs like BiDirectional LSTMs[15], LSTM attention models[7], etc. CNN Model with Chi-Square achieved 95.2% accuracy and 91.49% F1 score. Best results were achieved using a hybrid CNN LSTM Model with PCA. The results were 97.8% accuracy and 97.8% F1 score.

2.3 Supervised Learning for Fake News Detection

Julio C et al. [3] has proposed supervised machine learning techniques to classify fake news. A total of 141 textual features were evaluated in this paper. This paper roughly classifies its features into three types : 1) Features from news content, 2) Features from news source and 3) features from the environment. These features are then fed into various classifiers like random forest, XGBoost, Naive Bayes and Support Vector Classifier to classify fake news. Best results were obtained with random forest classifier and XGBoost Classifier. Random forest achieved an AUC (Area under curve) of 0.85 with F1 score of 0.81. XGBoost achieved an AUC of 0.86 with F1 score of 0.81.

2.4 Towards Automatic Identification of Fake News: Headline-Article Stance Detection with LSTM Attention Models

In [4], a solution has been proposed for fake-news stance detection using various neural architectures built on top of LSTM models to label the headline-body pair as agree, disagree, or discuss. The dataset used here was the FNC-1 Dataset. The paper first divides the problem further into two classification problems. The first classification is whether the headline and body of the news article are related or not and the second classification is a three-way stance detection. The first classification is done using a SVM classifier with a Radial basis function(RBF) Kernel trained on TF-IDF cosine similarity features. The second classification is done using different variations of the LSTM models after transforming the input words into 300 dimensional vectors using Glove Embedding. Different variations like Conditionally Encoded LSTMs, LSTM with Global Attention, LSTM with word-by-word attention, Bidirectional Global Attention, Bidirectional Conditional LSTM with Bidirectional Global Attention, etc were used for three-way classification. Conditionally Encoded Lstm with Bidirectional Global Attention and Bidirectional CE LSTM with Bidirectional Global Attention showed really good metrics on the FNC-1 Dataset.

2.5 r/Fakeddit: A New Multimodal Benchmark Dataset for Fine-grained Fake News Detection"

In [5], Fakeddit presents a multimodal approach to detecting fake news. It has a multimodal dataset which consists of more than 1 million samples through multiple categories of fake news. They processed the dataset through several stages of review and labelled all of the samples for a 2-way, 3-way or 6-way classification by following a distant supervision. They constructed hybrid models consisting of both text and images to perform various extensive experiments so as to form several different variations of classifications which in turn shows the importance for multimodality and fine-grained classifications which is the novel aspect absolutely unique to Fakedd. For textual data, accuracy of 0.8644 was achieved on 2-way classification using the BERT model. For 3-way classification and 6-way classification accuracy was 0.8580 and 0.7677 respectively. For multimodal classification, accuracy was 0.8909 for 2-way classification using the BERT + ResNet50 Model. 3-way classification and 6-way classification achieved 0.8890 and 0.8588 accuracy respectively.

2.6 MVAE: Multimodal Variational Autoencoder for Fake News Detection

For multimodal fake-news detection, numerous deep learning architectures have been proposed including MVAE(Multimodal Variational Autoencoder for Fake News Detection)[7], Hybrid LSTM-CNN architectures, EANN(Event-Adversarial Neural Networks)[6], etc. A bimodal variational autoencoder coupled with a binary classifier is used in MVAE for detecting fake news.The variational autoencoder is capable of learning probabilistic latent variable models by optimizing a bound on the marginal likelihood of the observed data. The multimodal representations obtained from the bimodal variational autoencoder are utilized by the fake news detector to classify posts as fake or not. Two standard fake news datasets were collected from popular microblogging websites: Weibo and Twitter. The results achieved show that across the two datasets, on average their model outperformed state-of-the-art methods by margins as large as ~ 6% in accuracy and ~ 5% in F1 scores.

2.7 EANN: Event Adversarial Neural Networks for Multi-Modal Fake News Detection

EANN is an end-to-end framework which can derive event-invariant features and thus benefit the detection of fake news on newly arrived events. In EANN, extracting the textual and visual features from posts is performed by the multi-modal feature extractor. It cooperates with the fake news detector to learn the discriminable representation for the detection of fake news. For the removal of event-specific features and to keep shared features among events, EANN uses the event discriminator. The dataset used was from twitter and weibo. On the twitter dataset, EANN Model achieved a F1 score of 0.719 and on weibo dataset it achieved a F1 score of 0.829.

2.8 Fake News Detection using Bi-directional LSTM-Recurrent Neural Network

Pritika Bahad et. al. [8] has proposed a solution for detecting fake news by determining whether the information is true or false on the basis of the relationship between the news headline and body, and the quantification of the bias of a news article. To find out the best deep learning approach for this problem, different models were trained on the same data like CNN, Vanilla-RNN, LSTM-RNN and Bidirectional LSTM-RNN. These models were used because they can detect complex patterns in textual data. Two datasets were used for this study which can be obtained from the open Machine Learning Repository at Kaggle. They used pre-trained word embeddings called Glove for converting words to vectors. The Bidirectional LSTM-RNN showed better accuracy than the other models on both the datasets which showed that this model is significantly more effective than the rest of the unidirectional models. Accuracy of 0.9108 was achieved on the first dataset whereas accuracy of 0.9875 was achieved on the second dataset using the Bidirectional LSTM-RNN Model.

3. PROPOSED SYSTEM

3.1 Problem Statement

"To detect different forms of fake news media using neural networks."

3.2 Problem Elaboration

Fake News can be used to create unnecessary unrest in the society. It can be in different forms such as text or image or audio and video and can have various degrees of impact based on the severity of the content. In this paper we propose a single modal as well as a bimodal approach to fake news detection. In the single modal approach we use textual data of news while for bimodal approach we use text as well as images. Deep learning methods are used for textual and image processing tasks. Various preprocessing methodologies and neural network architectures are implemented for these cases. Thus, we have decided to implement neural networks for the task of fake news detection.

3.3 Proposed Methodology

Single Modal:
For single modal i.e textual data, we will first preprocess data to convert textual data to a suitable form for the purpose of model training. We can then use various machine learning as well as deep learning models to classify fake news. As the LSTM model helps in finding out longer range dependencies and also helps in mitigating vanishing gradients problem, we plan on using this model.

Bimodal:
For bimodal data i.e for text and images, we will first apply several preprocessing techniques to convert the image and text into a suitable form for model training. As CNNs give good results on images, we plan on using this model. Additionally, CNN is useful in extracting features from the images and textual data and also helps reduce the number of parameters saving both memory and processing time. All of the models can be implemented using Keras library.

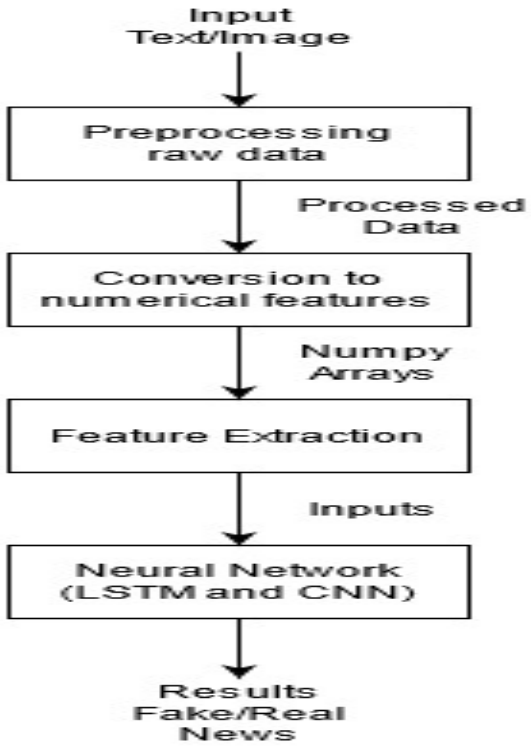

Fig -1: Proposed Methodology

4. CONCLUSION

To conclude, this paper summarizes the results of our survey on different approaches to Fake News Detection. We have seen that most of these systems use various complex deep learning architectures like LSTMs and CNNs. We have proposed a single modal as well as bimodal Fake News Detection System using either text (single modal) or with text and images both (bimodal).

5. REFERENCES

[1] Monther Aldwairi, Ali Alwahed. "Detecting Fake News in Social Media Networks",in the 9th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2018).

[2] Muhammad Umer, Zainab Imtiaz, Saleem Ullah ,Arif Mehmood, Gyu Sang Choi,Byung-won On "Fake News Stance Detection Using Deep Learning Architecture (Cnn-lstm)",Ieee Access,2020.

[3] Julio C. S. Reis, A. Correia, Fabricio Murai, Adriano Veloso, F. Benevenuto, E. Cambria "Supervised Learning for Fake News Detection",Published 2019,IEEE Intelligent Systems.

[4] Sahil Chopra, Saachi Jain, John Merriman Sholar. "Towards Automatic Identification of Fake News: Headline-Article Stance Detection with LSTM Attention Models".

[5] Kai Nakamura, Sharon Levy, William Yang Wang. "r/Fakeddit:A New Multimodal Benchmark Dataset for Fine-grained Fake News Detection".

[6] Yaqing Wang, Fenglong Ma, Zhiwei Jin, Ye Yuan, Guangxu Xun,Kishlay Jha, Lu Su,Jing Gao "EANN: Event Adversarial Neural Networks for Multi-Modal Fake News Detection".

[7] Dhruv Khattar, Jaipal Singh Goud, Manish Gupta, Vasudeva Varma. "MVAE: Multimodal Variational Autoencoder for Fake News Detection"

[8] Pritika Bahad, Preeti Saxena, Raj Kamal "Fake News Detection using Bi-directional LSTM-Recurrent Neural Network".