WDPS Assignment 2: Utilizing LSTM Model and Ensemble Algorithm to Detect Fake News

Yilin Li (2737659)
Vrije Universiteit Amsterdam
y45.li@student.vu.nl

Ruijia Lei (2733510)
Vrije Universiteit Amsterdam
r.lei@student.vu.nl

Haohui Zhang (2722930)
Vrije Universiteit Amsterdam
h17.zhang@student.vu.nl

Tianhao Xu (2740825)
Vrije Universiteit Amsterdam
t4.xu@student.vu.nl

1 INTRODUCTION

In the past many years, the development of information exchange lead to the rapid spread of fake news on the internet. Fake news may bring disastrous consequences and may plunge society into a serious crisis. In order to detect and eliminate fake news in the early stages to avoid social hazard and help people obtain the true information. We intended to use machine learning models (e.g., LSTM, Bi-LSTM) to establish a fake news detection system. Simultaneously, some optimizations (e.g., ensemble learning methods, trained classifiers) were made as well. The result of our test showed that our approach improved the performance of fake news detection on specific dataset.

2 RELATED WORK AND DATASET

The project refers a scientific dataset for guaranteeing the strictness and availability of data. The ISOT Fake News Dataset contains both true and fake articles extracted from the World Wide Web. The true articles are extracted from reuters.com which is a renowned news website, while the fake articles were extracted from multiple sources, mostly websites which are flagged by politifact.com. The dataset contains a total of 44,898 articles, out of which 21,417 are truthful articles and 23,481 fake articles. Each article contains the title of articles, the text of articles, issuing time and label of fake news (i.e., fake=1, true=0). Dataset we have used can be downloaded from 'https://www.kaggle.com/clmentbisaillon/fake-and-real-news-dataset'

In 2017, N-Gram Analysis and machine learning models were used to detect fake news [1]. Ahmed's team evaluated the performance of N-Gram model and LSVM classifier with phenotype data on dataset which is almost the same as ours. In their test, N-Gram model with LSVM achieved an average of 92% for accuracy. Considering that our experiment uses the same data set as Ahmed, we can treat their experimental results as a control group to evaluate the classification effect of our model. Not only that, Ahmed's work also inspired us to use multi-layer machine learning models to extract features from input data. So we designed a multi-layer LSTM model for feature extraction, and then used the trained classifier to complete the classification task. Details are as follows.

3 OVERVIEW OF SOLUTION

(1) Pre-processing raw data and utilizing TF-IDF to vectorize text.
(2) Designing a four-layers LSTM model to analyze title and text data separately to extract features and achieve the combination of them.
(3) Training suitable classifier and using ensemble learning to enhance the performance of classification. And evaluation.

3.1 Data Pre-processing

Whether the data is processed correctly and well or not, it largely lays a good foundation for the subsequent training of the model, etc. There are 6 steps that process the raw data. (1) Fill the blank data which is null(since there are about 3% null values). Avoid some mistakes when iterating data. (2) Normalize all the text, convert all text to the lowercase. Because the machine can't distinguish the upper and lower case. (3) Remove non-alpha text. This step is to eliminate the influence of non-alphabetic text, since all text is based on English. (4) Erase stop words(the, of, and, etc.). Eliminate white spaces. (5) Remove unwanted words, we expect to maintain those words that have the length of more than 2 and contain only alphabets. (6) Lemmalization. Get help from nltk package and use the wordnet lemmatizer.

3.2 TF-IDF

Combining with supervised Machine Learning methods, using TF-IDF will generate better performance compared to normal term frequency statistics to extract features. We imported TfidfVectorizer from sklearn package to vectorize text and title. We selected top 600 words in each text and top 20 words in each title. Each word will be translate to an integer between 0 to 9999. Then, each text will be expressed as $text = (x_0, x_1, ..., x_{599}), (0 <= x_i < 10000)$

3.3 Four-Layers LSTM Model and Classification

Figure 1 shows the overview of structure and function of the Four-Layers LSTM model. This model is divided into three steps. In the first step, with the aiming of providing more features for extraction, we established an embedding layer to map the vectors of text and title to a high-dimensional space (i.e., (1, 600)->(30, 600)). Then, we imported a LSTM layer and train this layer independently to complete feature extraction. More specifically, we abstracted a vector with 128 nodes to express the feature of text from output of embedding layer. In the final step, we designed two full-connected (dense) layers to reduce the number of nodes in feature vectors in order to prepare for further analysis and classification. To sum up, we stacked these four models to a four-layers LSTM model. And we generated two independent stacked models for analysis and extracting features from text and title (e.g., the dimension of

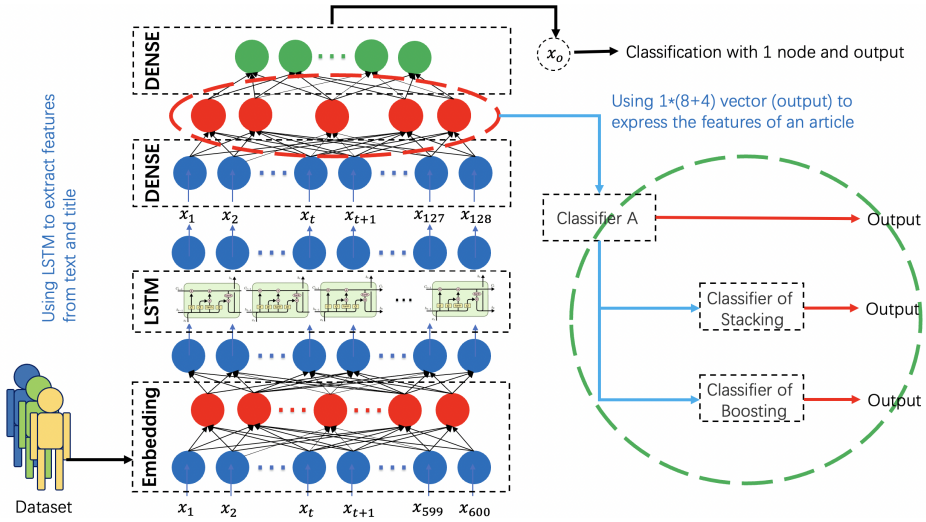

Figure 1: The structure detail of model

Table 1: Overview of model's performance. *: Ahmed et al., 2017

Models	Accuracy	Specificity	Precision	Sensitivity	F-1 Score
N-Gram, LSVM *	0.92	-	-	-	-
4-LSTM, title only	0.9316	-	-	-	-
4-LSTM, text only	0.9037	-	-	-	-
4-LSTM, both, LR	0.9572	0.9567	0.9605	0.9576	0.9591
4-LSTM, both, Stacking	0.9746	0.9712	0.9739	0.9776	0.9757

each layer's output for title is (30, 20)->(1, 8)->(1, 4)->(1, 1); the dimension of each layer's output for text is (30, 600)->(1, 128)->(1, 8)->(1, 1)). After processing inputs by LSTM model, we directly splice the feature vector of text (1,8) with the feature vector of title (1,4), then use single classifier or ensemble learning algorithm to realize the classification of fake news (**Figure 1**, green panel).

4 TECHNICAL CHALLENGES

1. How to design the pipeline and process the raw data in order to abstract possible features to recognize a news is fake or not?

- Considering that each article in the dataset records the whole context of text and a label for each article to represent the news is fake or not, we intended to use supervised machine learning methods (i.e., LSTM) to extract features from text and title to use for detecting fake news.

2. What is the innovation of the project?

- Instead of using LSTM for classification, we trained a classifier and spliced two different feature vectors as input, which were extracted from text and title through using machine learning model, and achieve classification by analyzing the comprehensive feature vectors. We believed this method is robustness because it can correctly classify input with fake title and real text.

3. How to improve the performance of the model and how to ensure the accuracy of the results?

- Using ensemble learning methods and K-folds cross validation.

5 RESULT

Table 1 list the performance of several different models in comprehensive test. We conducted four comprehensive tests are: (1) using 4-LSTM model with title data, (2) using 4-LSTM model with title data, (3) using 4-LSTM model with both data and (4) using 4-LSTM model with ensemble learning algorithm (i.e., Stacking). And the control experiment is summarized from Ahmed's work. In control test, N-Gram model with LSVM achieved an average of 92% for accuracy. Compare with control experiment, the accuracy of 4-LSTM model with both data increased by 3.7%, which reached to 95.72%. Besides, not only the accuracy of classification, but the values of other evaluation indicators have also increased to a certain extent. More specifically, the fake news classification with an average of 95.67%, 96.05%, 95.76% and 95.91% for specificity, precision, sensitivity and F-1 score. Moreover, compare with test (4), the ensemble learning algorithm used by comprehensive test can also enhance the classification of fake news. After detailed experiments, we believe that our model can better improve the accuracy of fake news recognition and has the value of further research.

REFERENCES

[1] Hadeer Ahmed, Issa Traore, and Sherif Saad. 2017. Detection of online fake news using n-gram analysis and machine learning techniques. In *International conference on intelligent, secure, and dependable systems in distributed and cloud environments*. Springer, 127–138.