# Intelligent Classifier for Identify Reliable On-Demand Messages

Jiann-Liang Chen, Yi-Wei Ma, Song-Yun Tsai

Department of Electrical Engineering, National Taiwan University of Science and Technology, Taiwan
lchen@mail.ntust.edu.tw, ywma@mail.ntust.edu.tw, su11221122su@gmail.com

## Abstract

Accurately extracting useful messages from bodies of information is important. This work proposes an intelligent system, called AI@nti-Fake system, to categories social news and determine whether it is true or false. The news is preprocessed using a Natural Language Processing technique. The text sentiment analysis in the on-demand message is analyzed to identify the fake news. A dataset from the International Workshop on Semantic Evaluation is used in this study. The on-demand message is related to the public's attention, and the analyzed text sentiment is identified as positive, neutral or negative. The accuracies of the proposed AI@nti-Fake system in the training stage and the real data test can reach 90% and 80%, respectively. The F1-Score of the proposed approach and two others methods are 78.50, 64.84 and 64.59, respectively. The results of the analysis reveal that the F1-Score of our approach can get better performance in classifying on-demand messages and detecting disinformation. The proposed AI@nti-Fake system, which is based on social media analysis and the judgment of sentiment may have applications in business.

**Keywords:** Deep learning, Long Short Term Memory (LSTM) algorithm, Natural language processing, Fake news, On-demand message

## 1 Introduction

In this era of mass information, accurately and quickly finding information that we want to know is important [1-2]. Information in TV reports, online news and social media concerns important event. Some people say that the present era is the post-truth era [3]. What is the post-truth? Like a disinformation news has been reported- a dog is crawling on the ground; some people might think that the dog's hind feet is broken, so the dog is in pain, some may imagine that the dog has been abused. In fact, the dog is lazy and does not want to walk and the ground is relatively cool, so it remains close to the floor. The different understandings generate different emotions. Prevent crime, identifying psychological inclinations, understanding public opinion and analyzing election polls can all benefit from an accurate and quick method for analyzing content on social media to extract valuable information. So, this work proposes an intelligent system, called AI@nti-Fake system, to categories social news and determine whether it is true or false.

Section 1 below explores the importance of accurately extracting the on-demand message and eliminating fake news. Section 2 explores related research. Section 3 introduces the proposed AI@nti-Fake system architecture for classifying social news and identifying fake news. Section 4 analyzes the performance of the system in terms of the variation of different parameters. The system is also compared with other methods. The conclusion is given in Section 5.

## 2 Related Works

A hung amount of texts is published daily on social media and in newspapers. To explore this text and rapidly to find reliable on-demand messages, text analysis techniques must be used. Text analysis includes text classification, abstract extraction, sentiment analysis, keyword extraction, etc. [4]. Traditional text analysis techniques have been unable to cope with the recent explosion of information, so new technologies, such as natural language processing and artificial intelligence are being increasingly used [5-6].

In Reference [7], authors designed a machine learning model to assess the risk of substance abuse from images, captions and comments on Instagram posts. They used CNN (Convolutional Neural Network) and Long Short-Term Memory (LSTM) models to determine the personal risk of using alcohol, prescription drugs, illegal drugs and tobacco from social media content. In Reference [8], authors propose a model for detecting real-time events, drawing public attention to phenomena, such that each acts as a sensor of social information. The approach tracks the abnormal phenomena in the social data stream, and its accuracy is about to 76.9%. In Reference [9], astroturfing was appearing in numerous contexts on social media– a matter of concern because of intended deception. The

*Corresponding Author: Yi-Wei Ma; E-mail: ywma@mail.ntust.edu.tw
DOI: 10.3966/160792642020122107013

analysis involved a binary n- gram technique and KNN (k-nearest neighbor) learning, which has been demonstrated to be effective at accurately identifying authors when a training set from the same authors. A system for detecting cyberbullying with an eye to warning mechanisms for the prediction of posted images vulnerable to attacks is designed [10]. That work identified the importance of advanced features in posted comments in detecting cyberbullying. The accuracy of the proposed system reached 95%.

In Reference [11], a deep learning approach to solving the problem of domain adaptation of sentiment classifiers. In Reference [12], the F1-Score of the proposed approach is about 65. Customers' opinions in user-generated content are valuable for market and trend analysis. Sentiment analysis helps retailers to extract these written opinions automatically. The polarized terms based on SentiWordNet and used the number of positive or negative polarity words were introduced in Reference [13]. The F1-Score of the approach was close to 68.2.

## 3 Proposed AI@nti-Fake System

In the proposed AI@nti-Fake system, a deep learning algorithm, is utilized to classify on-demand message. Since social news evolves content features over time, the classification algorithm used must be time-dependent. Therefore, this research includes LSTM as the research algorithm. Text sentiment analysis is used to identify fake news to investigate the difference in the emotion used in these articles. The system architecture is shown in Figure 1.

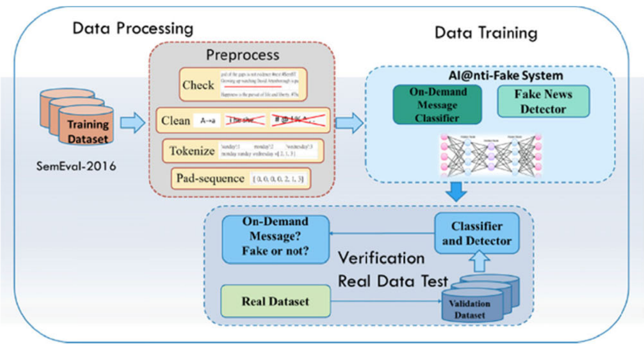



**Figure 1.** The proposed AI@nti-Fake system architecture

### 3.1 Data Processing Layer

A dataset from the International Workshop on Semantic Evaluation (http://alt.qcri.org/semeval2016/) which is concerned with research into text [14]. Based on the Natural Language Processing (NLP) techniques, data are processed in four steps. The first, "Check" excludes blank without content. The second step is "Clean". The original data are from Twitter. Many tweets identify Twitter users, so words must be deleted. The processed text will be displayed without punctuation, and punctuation marks are replaced by spaces. Converting all letters to lowercase letters, replacing the letters with lowercase and cleaning up the unimportant elements eliminates many steps, because it can avoid having to repeat them dozens of times during the cleanup. The end of each text sample is added "SemST", indicating that the material is from the International Workshop on Semantic Evaluation.

The third step is "tokenization". To convert a piece of text into a word vector to enable computing in the LSTM embedding layer, each word is assigned a number. Based on the word cloud analysis. A more frequently used work is assigned a smaller value.

The last step is "Pad-sequence". To carry out the training smoothly, the length of the input must be equalized. If the length of each piece of text is less than a defined length, it must add a "0" vector to the front of the piece of text. In the used dataset, 95% of the data samples have a length of 15, so the defined length is set to 15, eliminating the need for many "0" vectors, which would result in training errors and inaccurate outputs.

### 3.2 Data Training Layer

In the training stage, a deep-learning mechanism that is based on the LSTM algorithm is used. The algorithm is an evolution of the recurrent neural network. The symbols of C, X and h are indicated as memory cell, input and hidden (short-term memory), respectively. The operations of LSTM computing from time t-1 to time t is performed with three gates, the forget gate, the input gate and the output gate.

According to the LSTM training model and based on number of issues, the classifier is divided into six models. One issue is the subject to be predicted by a model, and the other five issues are used as other categories, using sigmoid activation function. The output of the sigmoid function is between 0 and 1, then closer the sigmoid value is to the label, the more relevant to the issue. The text sentiment analysis is also similar as that of on-demand message classifier. The classifier is divided to three models, using softmax activation function. The softmax output is the probability, and all categories get a probability value between 0 and 1, and the category with the highest probability and the final forecast category.

### 3.3 Verification Layer

In the data training stage, the classifiers are designed based on the LSTM model with better performance. The 10-Fold Cross Validation is used to validate the trained model through generating different combinations of the data the study already has. In this study, we have 4973 sampled data, the model is trained on the first 4476, and test on the last 497. Then we train on data 1-3997 & 4476-4973, and test on samples 3998-4475. Then repeat. We get different combinations of train/test data to validate the trained model.

## 4 Performance Analysis

Considering the computational complexity in this study, the performance is first analyzed with various LSTM sizes {100,200}, multiple fully connected {True, False} and fully connected unit {50,100}. The accuracies of each issue and sentiment are shown in Figure 2. Multiple fully connected means the first input layer must have the number of complete dataset size. Fully connected unit means the neurons between the input layer and the hidden layer are completely connected. This study conducted 100 experiments to obtain analysis results. This experiment confidential level is 95%.

Based on the variation in the number of LSTM layer (hidden unit), the accuracies of the results of analyses are shown in Figure 3. The results reveal that a one-LSTM layer delivers the best average performance, because the characteristics of the dataset that is used in this study are not complicated.

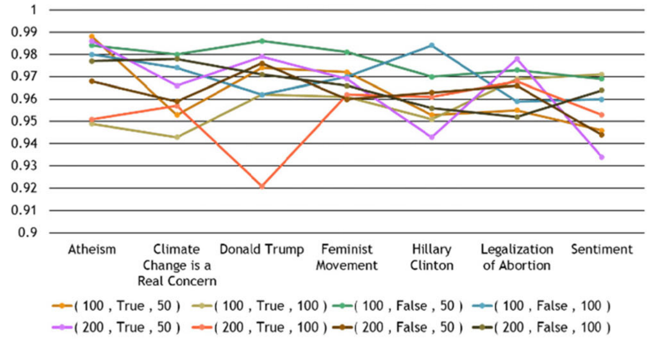



**Figure 2.** Accuracies with variation of parameters (LSTM size, multiple fully connected and fully connected unit)

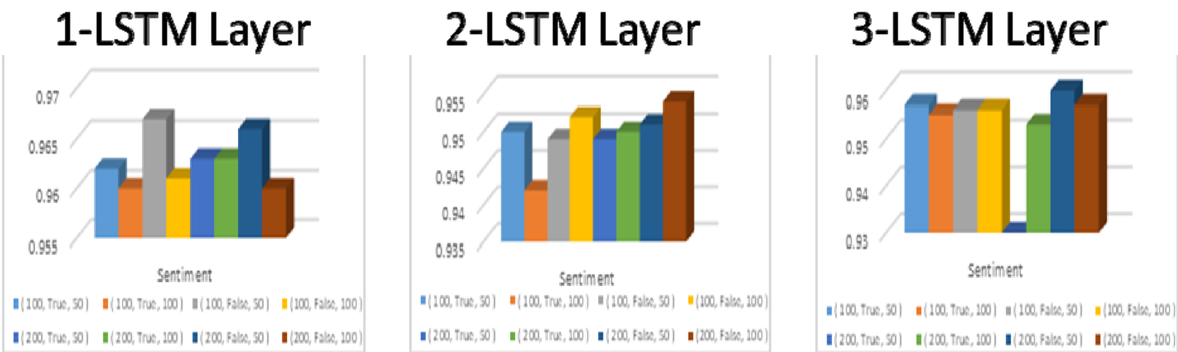



**Figure 3.** The analysis with the variable number of LSTM layer

Figure 4 shows a performance analysis with various sizes of the embedded layer. The size is set to 100, 150, and 200. The performance is the best at a size of 200. The higher the dimension yield better results of training with the dataset, because a high dimensional vector space better captures of the information in the training document.

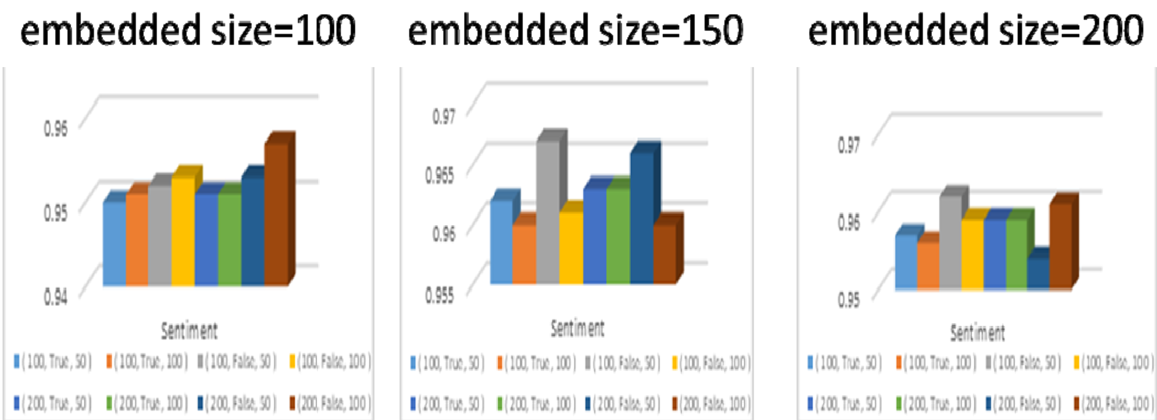



**Figure 4.** The analysis with different size of embed ayer

### 4.1 Comparison with Other Existed Methods

This section compares the proposed with others relevant studies. Table 1 shows the F1-Score of the method herein and two other methods. The F1-Score of our approach exceeds those of the other methods.

### 4.2 Real Data Test

To investigate the feasibility of our approach, this study crawls on the tweets for relevant tweets. Table 2 presents the accuracy of test results. The accuracy of self-collected data sets is similar to these obtained using test sets. However, Table 2 only shows the

**Table 1.** Comparison with other existed methods

| Method | AI@nti-Fake system -LSTM | Webis: An Ensemble for Twitter Sentiment Detection [14] -SVM | UNITN: Training Deep Convolutional Neural Network for Twitter Sentiment Classification [12] -CNN |
|---|---|---|---|
| F1-Score | 78.5 | 64.84 | 64.59 |

accuracy of the model related to the issues, and no test results of a test of the sentiment model are provided. Even though this study includes sentiment tests, no standard label can be used as a basis for evaluation. The labels in the original data set, SemEval-2016, are all marked by multiple people and are subject to review.

**Table 2.** Real data test with self-collected data set

| Issue | Accuracy |
|---|---|
| Atheism | 0.85 |
| Climate Change is a Real Concern | 0.875 |
| Donald Trump | 0.887 |
| Hillary Clinton | 0.862 |
| Legalization of Abortion | 0.85 |
| Feminist Movement | 0.885 |

## 5 Conclusion

In this work, an intelligent classifier, called AI@nti-Fake system, is designed to extract the on-demand message from social media and identify fake news by sentiment detection. In the proposed AI@nti-Fake system, a deep learning algorithm, Long Short-Term Memory (LSTM), classifies the on-demand message. Text sentiment analysis identifies fake news. The social media news is preprocessed using a Natural Language Processing technique. The accuracies of the proposed AI@nti-Fake system on training data and real test data reach 90% and 80%, respectively.

## Acknowledgments

This work is a partial result of project No. MOST 108-2221-E-011-068-MY2 conducted by National Taiwan University of Science and Technology under the sponsorship of Ministry of Science and Technology, Taiwan.

## References

[1] S.I. Manzoor, J. Singla, Nikita, Fake News Detection Using Machine Learning Approaches: A Systematic Review, *3rd International Conference on Trends in Electronics and Informatics*, Tirunelveli, India, April 2019, pp. 230-234.

[2] T. Traylor, J. Straub, Gurmeet, N. Snell, Classifying Fake News Articles Using Natural Language Processing to Identify In-Article Attribution as a Supervised Learning Estimator, *13th IEEE International Conference on Semantic Computing*, Newport Beach, CA, USA, January 2019, pp. 445-449.

[3] R. Keyes, *The Post-truth Era: Dishonesty and Deception in Contemporary Life*, St. Martin's Press, 2004.

[4] C. D. Manning, H. Schutze, *Foundations of Statistical Natural Language Processing*, MIT Press, 1999.

[5] R. Feldman, J. Sanger, *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*, Cambridge University Press, 2007.

[6] S. J. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach*, Pearson Education Limited, 2016.

[7] S. Hassanpour, N. Tomita, T. DeLise, B. Crosier, L. A. Marsch, Identifying Substance Use Risk Based on Deep Neural Networks and Instagram Social Media Data, *Neuropsychopharmacology*, Vol. 44, No. 3, pp. 487-494, February, 2019.

[8] D. T. Nguyen, J. J. Jung, Real-time Event Detection on Social Data Stream, *Mobile Networks and Applications*, Vol. 20, No. 4, pp. 475-486, August, 2015.

[9] J. Peng, S. Detchon, K. K. R. Choo, H. Ashman, Astroturfing Detection in Social Media: A Binary n-gram-based Approach, *Concurrency and Computation: Practice and Experience*, Vol. 29, No. 17, pp. 1-14, September, 2017.

[10] H. Zhong, H. Li, A. C. Squicciarini, S. M. Rajtmajer, C. Griffin, D. J. Miller, C. Caragea, Content-Driven Detection of Cyberbullying on the Instagram Social Network, *25th International Joint Conference on Artificial Intelligence*, New York, New York, USA, 2016, pp. 3952-3958.

[11] X. Glorot, A. Bordes, Y. Bengio, Domain Adaptation for Large-scale Sentiment Classification: A Deep Learning Approach, *28th International Conference on Machine Learning*, Bellevue, WA, USA, 2011, pp. 513-520.

[12] A. Severyn, A. Moschitti, Unitn: Training Deep Convolutional Neural Network for Twitter Sentiment Classification, *9th International Workshop on Semantic Evaluation*, Denver, Colorado, USA, 2015, pp. 464-469.

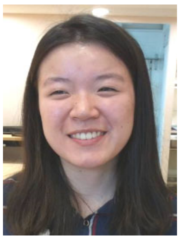
[13] E. Dovdon, J. Saias, ej-sa-2017 at Semeval-2017 Task 4: Experiments for Target Oriented Sentiment Analysis in Twitter, *11th International Workshop on Semantic Evaluation*, Vancouver, Canada, 2017, pp. 644-647.

[14] M. Hagen, M. Potthast, M. Büchner, B. Stein, Webis: An Ensemble for Twitter Sentiment Detection, *9th International Workshop on Semantic Evaluation*, Denver, Colorado, USA, 2015, pp. 582-589.

## Biographies

**Jiann-Liang Chen** was born in Taiwan on December 15, 1963. He received the Ph.D. degree in Electrical Engineering from National Taiwan University, Taipei, Taiwan in 1989. Since August 2008, he has been with the Department of Electrical Engineering of National Taiwan University of Science and Technology, where he is a professor now. His current research interests are directed at cellular mobility management and personal communication systems.

**Yi-Wei Ma** is an assistant professor in National Taiwan University of Science and Technology. His research interests include internet of things, cloud computing, future network and ubiquitous computing.



**Song-Yun Tsai** received the M.S. degree in Electrical Engineering of National Taiwan University of Science and Technology, Taipei, Taiwan. Her research interests include deep learning, and natural language processing.