With the development of the Internet, social networks and different communication channels, people can get information quickly and easily. However, in addition to real and useful news, we also receive false and unreal information. The problem of fake news has become a difficult and unresolved issue. For languages with few users, such as Vietnamese, the research on fake news detection is still very limited and has not received much attention.

In this paper, we present research results on building a tool to support fake news detection for Vietnamese. Our idea is to apply text classification techniques to fake news detection. We have built a database of 4 groups of 2 topics about politics (fake news and real news) and about Covid-19 (fake news and real news). Then use deep learning techniques CNN (Convolutional Neural Network) and RNN (Recurrent Neural Network) to create the corresponding models. When there is new news that needs to be verified, we just need to apply the classification to see which of the four groups they label into to decide whether it is fake news or not. The tool was able to detect fake news quickly and easily with a correct rate of about 85 %. This result will be improved when getting a larger training data set and adjusting the parameters for the machine learning model. These results make an important contribution to the research on detecting fake news for Vietnamese and can be applied to other languages. In the future, besides using classification techniques (based on content analysis), we can combine many other methods such as checking the source, verifying the author's information, checking the distribution process to improve the quality of fake news detection

Keywords: fake news detection, natural language processing, deep learning, CNN, RNN

UDC 82
DOI: 10.15587/1729-4061.2022.265317

DEVELOPMENT OF A FAKE NEWS DETECTION TOOL FOR VIETNAMESE BASED ON DEEP LEARNING TECHNIQUES

Trung Hung Vo
Corresponding author
Doctor of Computer Science, Professor, Vice-Rector
Department of Digital Technology
University of Technology and Education - The University of Danang
Cao Thang str., 48, Danang, Vietnam, 50209
E-mail: vthung@ute.udn.vn

Thi Le Thuyen Phan
Doctor of Computer Science, IT Lecturer
Department of Information Technology
FPT University
An Phu Thinh, 01 Quy Nhon, Binh Dinh, Vietnam, 55116

Khanh Chi Ninh
PhD Student, Lecturer of VKU
Department of Information Technology
Vietnam-Korea University of Information and Communication Technology (VKU)
The University of Danang
Tran Dai Nghia, 470, Khu do thi, Ngu Hanh Son, Danang, Vietnam, 50508

Received date 15.07.2022
Accepted date 22.09.2022
Published date 30.10.2022

***How to Cite**: Vo, T. H., Phan, T. L. T., Ninh, K. C. (2022). Development of a fake news detection tool for vietnamese based on deep learning techniques. Eastern-European Journal of Enterprise Technologies, 5 (2 (119)), 14–20. doi: https://doi.org/10.15587/1729-4061.2022.265317*

1. Introduction

On the Internet, especially social networking sites, there are more and more fake accounts, posting unverified information related to many topics about politics, epidemics, natural disasters, hydrometeorology, literature, superstition, false advertising, etc. This causes confusion, disturbance, and greatly affects people's daily life.

Along with the increase of fake news, the level of trust of users towards information on communication channels is decreasing. A proof of that is the survey data on user trust of Statista in 2021 in some countries as shown in Fig. 1.

Not a new concept, "fake news" refers to purported and largely verifiable false stories that are disseminated through formal or informal media. The prevalence of fake news has increased with the recent rise of social media, especially Facebook, and this misinformation is gradually seeping into the mainstream media. Several factors have been implicated in the spread of fake news, such as political polarization, post-truth politics, motivated reasoning, confirmation bias, and social media algorithms. It can be very persuasive and therefore it is necessary to develop strategies to identify and critically assess the news we read on social media. Around the world, a particularly large wave of fake news has attracted millions of people's attention, such as news related to the US presidential election in 2020, the Covid-19 epidemic and recently, news about the war in Ukraine.

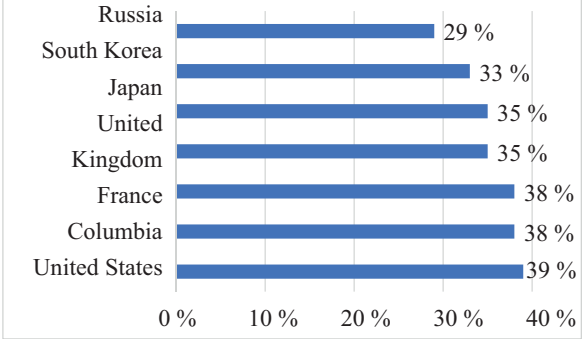

Fig. 1. Level of trust in media of some countries [1]

Up to this point, there is no tool that can help users know what is fake news, what is real news or how reliable a news is. The topic of automatic detection of fake news is attracting many scientists and businesses engaged in research and development.

2. Literature review and problem statement

The paper [2, 3] presents methods that people use to detect fake news. There are many ways for us to detect fake news and this depends on many factors such as knowledge, experience, analytical skills, judgment ability, critical thinking, etc. For humans, depending on skills and analytical ability, we can combine all the above factors to decide whether a news report is trustworthy or not. However, for computers to do all of the above, it still takes a long time to research and develop tools that can completely replace humans.

Manually detecting fake news usually involves all the techniques and processes that one person can use to verify news. However, the amount of online data generated daily is overwhelming. Furthermore, information spreads so quickly online that manual checking quickly becomes inefficient and impractical. Manual testing has the greatest difficulty when scaling the verification due to the large volume of data generated. Therefore, the task of automatic detection of fake news is an urgent and important need.

An automatic fake news detection system will help verify whether a news is fake or real without direct human intervention. There are different techniques and approaches used in fake news detection research. These techniques and approaches depend on the developer's point of view and tracking goals.

For Vietnamese, there are very few published studies on detecting fake news. The paper [4] presents an approach to detect fake news on social networking sites (SNS – Social Network Sites), by synthesizing linguistic features used by PhoBERT. The paper [5] presents the tasks shared on ReINTEL including three phases: initiation, public testing, private testing related to fake news. These results are very preliminary and not yet applicable. Moreover, they have not built a data warehouse for detecting fake news for Vietnamese, but only proposed solutions.

The paper [6] shows that current research often connects fake news with terms and concepts such as scam news. The challenges of fake news research begin with identifying what is fake news. Until now, no general definition has been provided for fake news, where it is considered "an article that is intentionally false and difficult to verify".

The paper [7] focuses on the fundamental theories of human cognition and behavior using to develop across various fields, such as the social and economic sciences, provide invaluable insights for the analysis of fake news. These theories may introduce new opportunities for qualitative and quantitative studies of fake news data. These theories can also facilitate the construction of plausible and interpretable models for detecting and intervening in the dispersal of fake news, which has been rarely available. News-related theories reveal possible characteristics of fake news content compared to real news content. The paper [8] makes the hypothesis that fake news is likely to differ from the truth in certain points such as writing style or statistics and emotional expressions. However, these theories are difficult to implement in real-world software/systems that automatically detect fake news.

The papers [9, 10] have proposed the application of neural network models and deep learning for fake news detection and have obtained some results. Furthermore, researchers have combined many models and methods to improve the quality of fake news detection. Combinations of CNN-RNN have been shown to be successful in a number of classification and regression tasks, as they are capable of capturing both local and sequential characterization of the input data. The papers [11, 12] proposed combining models to detect emotions or to extract features in the field of fake news detection but it is difficult to apply in practice and depends a lot on the nature of the texts. These models [9–12] have been applied to the processing of English documents and show that the construction of the model depends on language characteristics. The biggest limitation of the above studies is that they have not yet proposed a general model for fake news detection as well as how to adjust the parameters to achieve the best linguistic model for fake news detection. The language model is dependent on the features of the language (morphology, vocabulary, corpus, grammar, etc.) but has not shown the general rules. There are no studies that apply these techniques to Vietnamese.

In summary, compared with Vietnam, the world has conducted research on detecting fake news in recent years and initially proposed some solutions and models with good results. However, the identification and detection of fake news is still a big mystery to be discovered, a new research direction of artificial intelligence.

3. The aim and objectives of the study

The aim of this study is to develop a tool to assist in detecting fake news and first experiment on Vietnamese news.

To achieve this aim, the following objectives are accomplished:

– to propose a general solution for fake news detection based on research, analysis and evaluation of methods currently being applied to other languages;

– to create a dataset to serve on detecting fake news for Vietnamese. This dataset includes real and fake news that are classified by humans and pre-labeled for feature model creation and testing;

– to perform fake news detection based on deep learning with a Vietnamese dataset.

4. Materials and methods

When conducting research activities within the scope of this paper, we used the following methods:

– document analysis: we have studied the literature related to the concept of fake news, manual methods to detect fake news, solutions that researchers have used to detect fake news automatically. On that basis, we propose a possible research direction for the automatic detection of fake news in Vietnamese;

– apply text classification methods to fake news detection: in this case, we named each text class, including subject and type of news (fake or real), for example, covid_19_ fake_news, covid_19_real_news; political_fake_news, political_real_ news. Then apply text classification techniques to identify real or fake news. The testing is implemented at VAFC to support fake news detection on Vietnamese news;

– using the deep learning method in text classification: specifically, we use 2 deep learning models CNN and RNN to perform training and classification. Convolution Neural Network (CNN) is a class of feed-forward (where connections between nodes do not form a cycle) artificial neural networks and uses a variant of multilayer perceptrons designed to require minimal preprocessing. CNN is often used in the field of image processing, computer vision. However, it has recently been applied to natural language processing problems and has given very good results. In addition to the two layers input and output, CNN network architecture is composed of several types of layers including [13, 14]. Recurrent Neural Network (RNN) is an algorithm that has received a lot of attention recently because of its good results obtained in the field of natural language processing [15]:

– using the software reuse method: we reuse some available software in Vietnamese document processing such as data cleaning, word separation, vectorization in the process of building a tool to detect fake speech freeforms.

When developing program modules, we use the Python programming language because it is the most commonly used language in deep learning. We also choose to use the Keras library to perform natural language processing functions.

5. Results of studying fake news detection for Vietnamese

5. 1. Results of proposing solutions for detecting fake news in Vietnamese

We have proposed a solution to detect fake news for Vietnamese based on deep learning through CNN and RNN. We have proposed a CNN network consisting of 5 layers: input, convolution, pooling, full connection and output. The proposed RNN model consists of 3 layers: input, hidden and output. In it, the calculation of values on hidden layers is recorded information about transitions between hidden layers.

The solution includes:

a) Proposed model.

The general model of this approach is as Fig. 2.

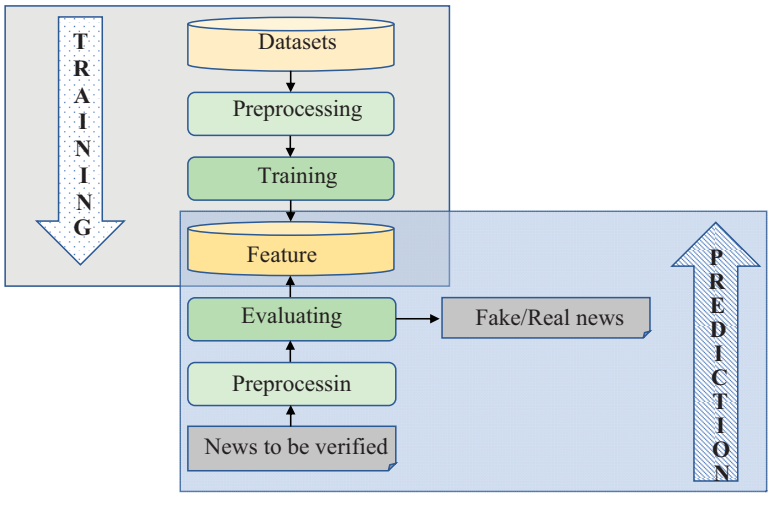

Fig. 2. Fake news detection process model

The specific functions of the main steps are as follows:

– the first step in this model is the data collection phase to build the training dataset. This dataset includes news that have been labeled as fake or real. In the case of supervised learning, all the data used for training must be labeled, in the case of semi-supervised learning, both labeled and unlabeled data;

– the preprocessing stage allows natural language processing techniques to be used to clean the data, remove unhelpful information, and represent the data;

– the training stage allows extracting the necessary linguistic features to create models for content classification and identification. On the basis of the extracted features, perform training according to the selection algorithms to build feature models. This model will be used for predicting whether a news report is fake or real;

– the prediction phase has the function of comparing the features of the news to be verified with the feature model created in the training phase to decide whether the news is fake or real.

b) Classification by using Convolution Neural Network.

The general architecture is as Fig. 3.

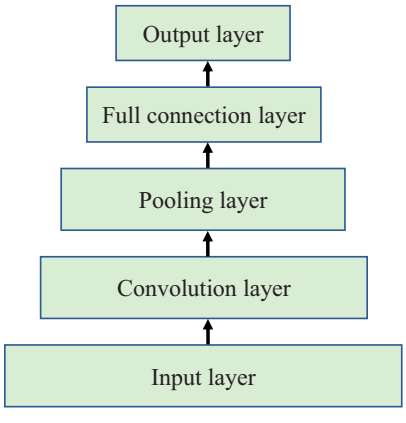

Fig. 3. Architecture of Convolution Neural Network

The general architecture:

1. Convolution layer. Convolutional neural networks are based on the math of convolution. We can simply imagine that the meaning of convolution is like an operation of mixing information together. Convolution is widely applied in many different branches of science and engineering.

In mathematics, the convolution between two functions f and g produces a third function that represents the transformation of one function with respect to the other. Considering two functions f and g, the convolution between these two functions is defined as follows:

$$h(x)=f\otimes g=\int_{-\infty}^{\infty} f(x-u)g(u)\mathrm{d}u=$$

$$=F^{-1}\left(\sqrt{2\pi}F[f]F[g]\right)=$$

$$=\sum_{y=0}^{columns}\left(\sum_{x=0}^{rows}\mathrm{input}(x-a,y-b)\mathrm{kernel}(x,y)\right)=$$

$$=F^{-1}(\sqrt{2\pi}F[\mathrm{input}]F[\mathrm{kernel}]). \quad (1)$$

In it, we note:

$$\text{featuremap} = \text{input} \otimes \text{kernet}. \quad (2)$$

We consider in one-dimensional space, the convolution between two functions f and g is described by the following equation:

$$(f * g)(x) = \sum_t f(t) g(x+t). \quad (3)$$

In fact, we perform the convolution by sliding the kernel/filter according to the input data. At each location, we perform matrix multiplication and sum the values to put into the feature map.

For 2-dimensional (2D) input data, such as text data (each document is represented as a vector – column of a matrix and has a set of n documents), we have two inputs for convolution. The first input is the weight matrix, the other is called the kernel or mask, which acts like a filter for the input data and produces a processed result for the output. We consider the following 2D convolution in detail:

$$(K * I)(i,j) = \sum_{m,n} K(m,n) I(i+n, j+m). \quad (4)$$

An example of a convolution is shown in Fig. 4.

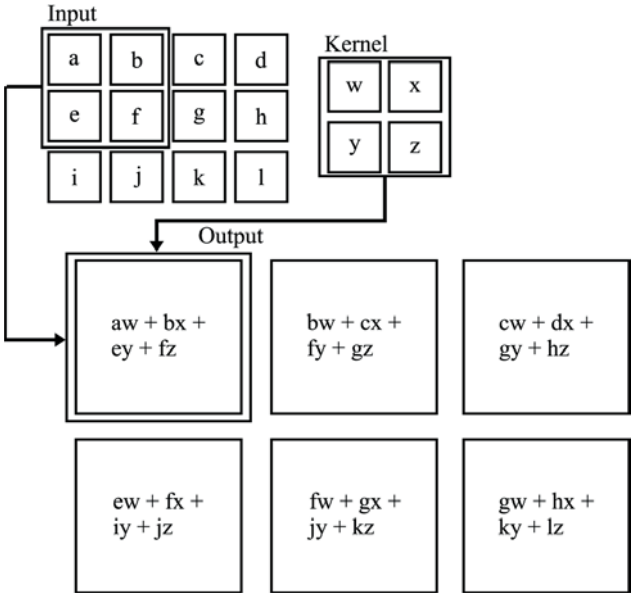

Fig. 4. Example of a convolution

2. Pooling layer. The general function of the Pooling layer is to reduce the spatial size of the feature with the main purpose of reducing the number of parameters and the amount of computation in the network. The Pooling layer operates on each Feature Map independently of each other. The most common approach used is Max Pooling.

An example of Max Pooling is shown in Fig. 5.

The max pool step here allows reducing a matrix 4×4 to a matrix 2×2.

3. Fully connected layer. After the input data has passed through many Convolutional Layers and Pooling Layers, the model relatively learned the characteristics of the training data, the output at the last layer is usually reconnected from the previous layer and converted to a vector (one-dimension). Then, we use Fully Connected Layers to combine the characteristics of the data to get the Output of the model.

An example of data flattening is shown in Fig. 6.

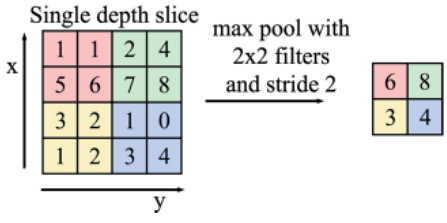

Fig. 5. Example of max pooling

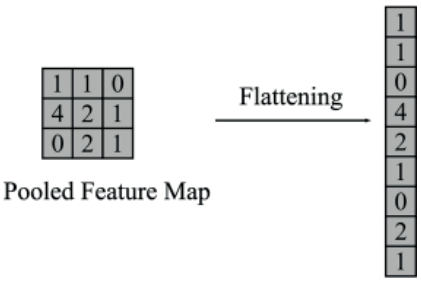

Fig. 6. Example of Full Connection

The flattening step here converts the elements of a matrix to be elements of a vector.

In summary, this section has detailed our proposal regarding the design of a CNN network to transform input data as training documents and output as a feature model.

c) Classification by using Recurrent Neural Networks.

The RNN is designed as Fig. 7.

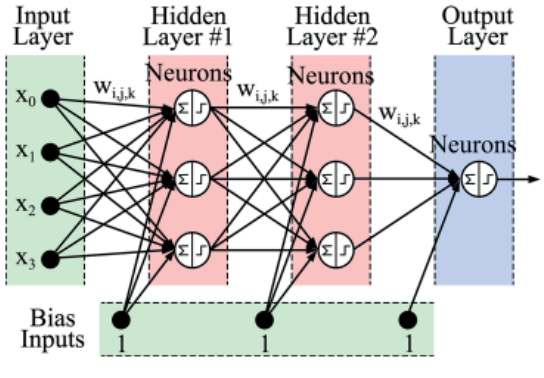

Fig. 7. Architecture of Recurrent Neural Network

As shown in the previous section, the neural network consists of 3 main parts: Input layer, Hidden layer and Output layer. We can see that the input and output of this neural network are independent of each other. Thus, this model is not suitable for string problems such as description, sentence completion, etc., because subsequent predictions such as the next word depend on its position in the sentence and the words before it.

Therefore, RNN was born with the main idea of using memory to store information from previous computation steps so that based on it, it can make the most accurate prediction for the current prediction step as Fig. 8.

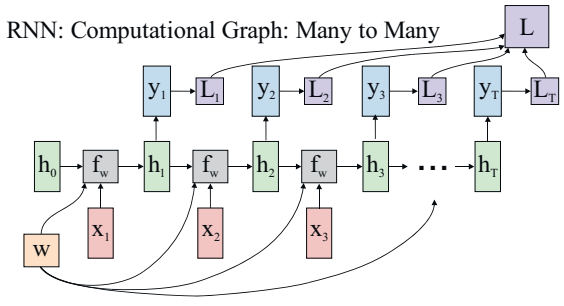

Fig. 8. Recurrent Neural Network for Many to Many

If the Neural Network is just input layer x going through hidden layer h and results in Output layer y with Full Connected between layers. In RNN, input x_t will be combined with hidden layer h_{t-1} by f_w function to calculate the current hidden layer h_t and the output y_t will be calculated from h_t. W is the set of weights and it is obtained in all clusters, L_1, L_2, ..., L_t are the loss functions, which will be explained later. Thus, the results from previous computations have been "remembered" by adding h_{t-1} to calculate h_t to increase the accuracy of current predictions. Specifically, the calculation process is written in mathematical form as follows:

$$h_t = \tanh\left(w_{hh}h_{t-1} + w_{xh}x_t\right), \quad (5)$$

$$y_t = w_{hy}h_t. \quad (6)$$

At this point, 3 new things appear: W_{xh}, W_{hh}, W_{hy}. For NN, only a single weight matrix W is used, RNN uses 3 weight matrices for 2 computations: W_{hh} combined with "pre-memory" h_{t-1} and W_{xh} combined with x_t to calculate "current step's memory" h_t from there combined with W_{hy} to calculate y_t.

Thus, RNN is a specialized neuron-based method that is effective in sequential information processing. An RNN recursively applies computation to every instance of the conditional input sequence based on previously computed results. These sequences are usually represented by a fixed-size vector that is supplied sequentially (one by one) for a periodic unit.

Fig. 9 illustrates a simple RNN framework below.

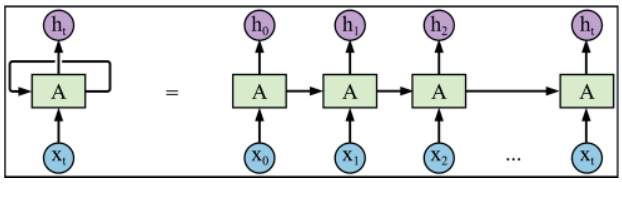

Fig. 9. Framework of Recurrent Neural Network

The main advantage of an RNN is its ability to remember the results of previous calculations and use that information in the current computation. This makes RNN models fit the context-dependent model in inputs of arbitrary length to produce an appropriate layout of the inputs. RNNs have been used to study various NLP tasks such as machine translation, image annotation, and language modeling, among others.

When compared with the CNN model, the RNN model can have similar or even better performance in natural language processing tasks.

5. 2. Results of building the dataset for Vietnamese fake news detection

To prepare training as well as testing data for fake news detection, we downloaded news from Vietnamese government websites (assuming they are real news) and news from anti-communist opposition websites (assuming these are fake news sites).

We divide downloads into 8 categories as in Table 1.

Table 1

Data used for testing the fake news detection system

No.	Category	Number of files
1	Real political news (for training)	500
2	Fake political news (for training)	500
3	Real news related to Covid-19 (for training)	500
4	Fake news related to Covid-19 (for training)	500
5	Real political news (for testing)	100
6	Fake political news (for testing)	100
7	Real news related to Covid-19 (for testing)	100
8	Fake news related to Covid-19 (for testing)	100

Each news contains about 1,000 2,000 words and the size of each file is about 5 KB on average. Test data and detailed processing steps are published on the Web: https://omomega.com/ombot/fake-news/database/.

5. 3. Results of building a tool to support fake news detection for Vietnamese

We developed a tool for Vietnamese fake news detection following these steps:

– data preparation (including training data and data for testing the evaluation);

– data preprocessing;

– model building (through training to model and refine);

– software building and testing to evaluate a text as fake news or real.

Data preparation. To prepare training as well as testing data for fake news detection, we downloaded news from Vietnamese government websites (assuming they are real news) and news from anti-communist opposition websites (assuming these are fake news sites).

Data preprocessing. First of all, we need to remove special characters in the original text such as periods, commas, opening and closing brackets, etc. by using the *Gensim* library. Then we use the *PyVi* library to separate Vietnamese words. A special feature in Vietnamese text is that a word can be combined with many different languages, for example: chính_trị, xã_hội... different from English and some other languages, the words are separated.

Text vectorization. To convert text into a vector, we convert each word in the dictionary into an element of an n-dimensional vector, using the Bag-of-words algorithm. In this model, each document will be represented by an n-dimensional vector (n is the number of words in the dictionary used). If the text contains the word corresponding to the dictionary, the vector element carries the frequency of occurrence of that word in the text, if it does not appear, the value is 0.

Model building and evaluation. To test the construction of CNN and RNN models as well as evaluate the classification of the evaluation datasets, we used Keras. Keras is a

model-level library, providing high-level building blocks for developing deep learning models. It doesn't handle low-level operations such as tensor manipulation and differentiation. Instead, it relies on a specialized, well-optimized tensor library to do so, serving as the backend engine of Keras [16].

Based on the classification results of the messages used for testing, the correct classification results for the newsgroups are as in Table 2.

Table 2

Experimented results on CNN and RNN

No.	Topics	Correct ratio	
		CNN	RNN
1	Real political news (for testing)	85 %	89 %
2	Fake political news (for testing)	86 %	88 %
3	Real news related to Covid-19 (for testing)	80 %	82 %
4	Fake news related to Covid-19 (for testing)	81 %	82 %

The correct identification of fake news is over 80 % and the correct result for the political news group is higher than for the news about Covid-19.

6. Discussion of the results of research and development of a fake news detection tool for Vietnamese

Through the research results, we observed:

The solution for detecting fake news for Vietnamese based on text classification:

– our solution for the detection of fake news through content analysis and text classification based on deep learning is positive. The model we propose in Fig. 2 is divided into two stages: training and fake news detection. This is a convenient model for practical implementation when we have a large volume of labeled documents, however, the training data must be constantly updated to ensure the quality of fake news detection;

– the selection of deep learning models CNN and RNN has many advantages over traditional machine learning models such as support vector machine, decision tree and Bayesian network. In particular, the design of the layers as in Fig. 3 and Fig. 7, and the way of processing in the layers as we propose is relatively easy to install the neural network and suitable for Vietnamese word processing. The calculation formulas to handle data transformation in formulas (1) to (6) are language-independent, however, we need to try to adjust the input parameters and kernel to match each language. Furthermore, with the experimental data being too small, the advantages of these models have not been clearly demonstrated;

– using only content-based classification makes it difficult to tell if it is fake news or not. The rate of accurate detection of fake news based on content is not high and in the future, we should combine multiple evaluation forms such as checking the source, verifying the author's information, checking the distribution process, etc. besides just analyzing the content;

– deep learning applications with big data always require high-speed computers and systems, making it difficult to deploy on single computers or ordinary servers. In order to operate the fake news detection system in practice, high-performance computing systems or deployment in the cloud are required.

Building a dataset for Vietnamese fake news detection:

– the dataset we built and presented in Table 1 is quite small and it can affect the results when applied in practice to identify fake news for new news. The coverage of training data is essential because it will allow the creation of a model that contains all the features needed to analyze the content of the news;

– labeling often has to be done manually because there are currently no tools to automatically label fake and real news. Currently, we are only developing support tools for collection and pre-processing, but we can't assign automatic labels yet. We need a lot of time and expense to create a correctly labeled training dataset. For a number of popular languages such as English, Chinese, French, etc., the data sets are usually very large and tools for collection and pre-processing are available;

– in particular, the training dataset must be continuously added to improve the quality of the model. The limitation of the current method is that every time more training data is added, it must be recalculated to create the corresponding language model;

– the biggest disadvantage of this method is manual labeling and in the future, it is necessary to research methods to be able to assign labels automatically based on artificial intelligence application;

– in the future, we will continue research to be able to automatically assign labels and find a solution to update the language model when adding new training data without having to recalculate from scratch.

Development of a tool to support fake news detection for Vietnamese:

– initially, a tool to support fake news detection for the two fields of politics and Covid-19 has been built with good accuracy as shown in Table 2 but requires human participation to make a final decision;

– the outstanding feature of the tool is content analysis based on the multilayer design of CNN and RNN models. Compared with existing methods such as traditional machine learning (vector support machines, decision trees), this method can handle more optimally in the case of large data;

– the limitation of the current tool is that the accuracy is still low and it is necessary to continue experimenting to adjust the parameters of the neural network to create the best possible language model;

– the tool has some disadvantages such as the user interface is not convenient, the model has not been automatically updated when new data is available, and only runs experimentally on Google Colab. These disadvantages can be overcome if we deploy in the cloud with high-performance computing and continue to improve the user interface;

– we can further improve the results if we combine some other deep learning models such as Deep Belief Network (DBN), Long Short Term Memory (LSTM), etc. or continue experimenting to adjust the network parameters to optimize the calculation process;

– we can easily apply this tool to other languages by changing the dataset used and adjusting the input parameters to suit the characteristics of the language used (the adjustment must be through the experimental results).

7. Conclusions

1. The solution based on text classification and deep learning suitable for fake news detection for Vietnamese news with a content analysis approach is proposed.

2. A dataset for fake news detection for Vietnamese with two topics about politics and Covid-19 was built. This dataset includes 2,400 news and each news contains about 1,000–2,000 words and the size of each file is about 5 KB on average. When building datasets, the solution is to use available support tools for data collection, preprocessing, and vectorization, but the labeling is done manually. This dataset has high reliability and can be used for future studies on fake news detection for Vietnamese.

3. Fake news detection for Vietnamese news was tested through two models CNN and RNN. Through testing, these tools correctly detected news as fake or real in about 85 %. The test results are quite good but we can improve them if we increase the dataset and adjust the parameters of the network.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorship or otherwise, that could affect the research and its results presented in this paper.

Acknowledgments

This research was funded by the Ministry of Education and Training (Vietnam) through the project code B2022-DNA-17.

References

1. Watson, A. (2022). Trust in media worldwide 2021. Statista. Available at: https://www.statista.com/statistics/683336/media-trust-worldwide/
2. Fallis, D. (2015). What Is Disinformation? Library Trends, 63 (3), 401–426. doi: https://doi.org/10.1353/lib.2015.0014
3. Wardle, C., Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policy making. Council of Europe, 109.
4. Nguyen, D. Q., Tuan Nguyen, A. (2020). PhoBERT: Pre-trained language models for Vietnamese. Findings of the Association for Computational Linguistics: EMNLP 2020. doi: https://doi.org/10.18653/v1/2020.findings-emnlp.92
5. Le, D.-T., Vu, X.-S., To, N.-D., Nguyen, H.-Q., Nguyen, T.-T., Le, L. et. al. (2020). ReINTEL: A multimodal data challenge for responsible information identification on social network sites. arXiv. doi: https://doi.org/10.48550/arXiv.2012.08895
6. Molina, M. D., Sundar, S. S., Le, T., Lee, D. (2019). "Fake News" Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content. American Behavioral Scientist, 65 (2), 180–212. doi: https://doi.org/10.1177/0002764219878224
7. Miller, T., Howe, P., Sonenberg, L. (2017). Explainable AI: Beware of inmates running the asylum or: How I learnt to stop worrying and love the social and behavioural sciences. arXiv. doi: https://doi.org/10.48550/arXiv.1712.00547
8. Chadwick, A., Stanyer, J. (2021). Deception as a Bridging Concept in the Study of Disinformation, Misinformation, and Misperceptions: Toward a Holistic Framework. Communication Theory, 32 (1), 1–24. doi: https://doi.org/10.1093/ct/qtab019
9. Zhou, X., Wu, J., Zafarani, R. (2020). SAFE: Similarity-Aware Multi-modal Fake News Detection. Lecture Notes in Computer Science, 354–367. doi: https://doi.org/10.1007/978-3-030-47436-2_27
10. Zhou, X., Zafarani, R. (2019). Network-based Fake News Detection. ACM SIGKDD Explorations Newsletter, 21 (2), 48–60. doi: https://doi.org/10.1145/3373464.3373473
11. Kollias, D., Zafeiriou, S. (2021). Exploiting Multi-CNN Features in CNN-RNN Based Dimensional Emotion Recognition on the OMG in-the-Wild Dataset. IEEE Transactions on Affective Computing, 12 (3), 595–606. doi: https://doi.org/10.1109/taffc.2020.3014171
12. Elhadad, M. K., Li, K. F., Gebali, F. (2019). A Novel Approach for Selecting Hybrid Features from Online News Textual Metadata for Fake News Detection. Lecture Notes in Networks and Systems, 914–925. doi: https://doi.org/10.1007/978-3-030-33509-0_86
13. Keeling, R., Chhatwal, R., Huber-Fliflet, N., Zhang, J., Wei, F., Zhao, H. et. al. (2019). Empirical Comparisons of CNN with Other Learning Algorithms for Text Classification in Legal Document Review. 2019 IEEE International Conference on Big Data (Big Data). doi: https://doi.org/10.1109/bigdata47090.2019.9006248
14. Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). doi: https://doi.org/10.3115/v1/d14-1181
15. Yu, Y., Si, X., Hu, C., Zhang, J. (2019). A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. Neural Computation, 31 (7), 1235–1270. doi: https://doi.org/10.1162/neco_a_01199
16. Ketkar, N. (2017). Introduction to Keras. Deep Learning with Python, 97–111. doi: https://doi.org/10.1007/978-1-4842-2766-4_7