Fake News Detection by Using RecurrentNeural Network

Trung Hung Vo[1], Khanh Chi Ninh[2], Khanh Duy Ninh[3]

[1] The University of Danang – University of Technology and Education, vthung@ute.udn.vn

[2]The University of Danang – Vietnam Korea University of Information and Communication Technology, nkchi@vku.udn.vn

[3]The University of Danang – University of Science and Technology, nkduy@dut.udn.vn

Abstract: Along with the development of the Internet, social networks and different communication channels, people can get information quickly and easily. However, in addition to true and useful news, we must also receive false and untrue information. The problem of fake news has become a difficult and unresolved problem. In this paper, we present research results on building a tool to support the fake news detection by using RNN. 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 of RNN to create the corresponding models. When there is a 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 make a decision whether it is fake news or not. 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, check the distribution process,... to improve the quality of fake news detection.

Keywords: Fake News Detection, Text Classification, Machine Learning, Deep Learning, RNN

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.

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 the 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 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...

Similarly, in Vietnam, fake news appears more and more and has become an unresolved social problem. To correct the above situation, Vietnam's Ministry of Information and Communications has opened a portal to receive reports, publish fake news, and launch reception channels to receive fake news. Accordingly, the portal to receive fake news notices has the domain name www.tingia.gov.vn, operated by the Vietnam Fake News Processing Center (VAFC) under the Department of Radio, Television and Electronic Information and managed by the Ministry of Information and Communications. The tasks and functions of the VAFC include: Coordinating with relevant authorities to appraise and publish fake news; Assess the trend of information sharing, large interactions to label fake news warnings; Receiving, detecting, appraising and labeling fake news; Disclosure of authentic information; Instructions on how to recognize, prevent and deal with fake news. The Center focuses on the following areas of information: policy, law; economic, financial; medical field, health products related to human health; natural disasters, epidemics; national security, social order and safety; fake accounts; phishing links; other fields (https://tingia.gov.vn).

However, the detection of fake news in Vietnam is currently done entirely manually by humans, as the VAFC's published philosophy of action "Fake news is created by humans, so only humans can recognize and processing fake news".

In this paper, we present the results obtained when researching and applying fake news automatic detection techniques and testing at VAFC to support fake news detection on Vietnamese news. The method we choose to use is based on the text classification technique by deep learning. For each topic, for example news about politics or Covid-19, we build 2 datasets, one containing fake news and the other

containing real news. Next, we train to build a feature model, and when there is a news that needs to be evaluated we will perform a classification module to lablel into the real or fake news category. This is a method of detecting fake news based on content analysis and initially applied has yielded quite good results.

The paper is organized into the following main sections: Section 2 presents a summary of some problems related to fake news detection and the reasons for this research, especially for Vietnamese language; Section 3 introduces the aim and

objectives of the study, which is towards a tool to support fake news detection and implementing for Vietnamese language; Section 4 is devoted to presenting the background related to this research including the concept of fake news detection, fake news classification, and how to detect fake news by humans or automatically; Section 5 is the most important content for introducing the solution that we propose that is using deep learning classification technique based on RNN to support fake news detection; Section 6 presents the test implementation steps, results and comments; The paper ends with a conclusion, which presents the analysis of the significance of the obtained results and proposes new research directions in the future.

2 Literature review and problem statement

For Vietnamese, there are very few published studies on detecting fake news. The paper [1] presents an approach to detect fake news on social networking sites (SNS - Social Network Sites), by synthesizing linguistic features used by PhoBERT. The paper [2] 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.

Around the world, 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" [3].

The fundamental theories of human cognition and behavior developed 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 [4]. News-related theories reveal possible characteristics of fake news content compared to real news content. For example, theories imply that fake news is likely to differ from the truth in certain points such as writing style or statistics and emotional expressions [5]. However, these theories are difficult to implement into real-world software/systems that automatically detect fake news.

Currently, some scientists have proposed the application of neural network models and deep learning (Deep Learning) for fake news detection and have obtained some results [6][7]. 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. For example, they have been used to detect emotions [8] or to extract features by combining models [9].

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, is 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 for Vietnamese news.

To achieve this aim, the following objectives are accomplished:

- Proposing a solution for detecting fake news for Vietnamese based on research, analysis and evaluation of methods currently being applied to other languages.
- Creating a dataset to serve research on detecting fake news for Vietnamese. This dataset includes real and fake news that are classified by humans and pre-labeled for modeling and testing.
- Building a tool to support fake news detection for Vietnamese based on deep learning.
- Proposing new research directions in the field of fake news detection in general and for detecting fake news in Vietnamese in particular

4 Background

4.1 The concept of fake news

Building a data warehouse for testing, we identify fake news based on the Oxford dictionary definition as follows: "False information that is broadcast or published as news for fraudulent or politically motivated purposes" (https://www.lexico.com).

4.2 Classification of fake news

Typical cases of fake news include deceptive advertising (in business and politics), government propaganda, original edited or misused images, fake documents, fake maps , Internet fraud, fake websites and untrue Wikipedia entries, etc. Fake news

can cause significant harm if people let it lie. To address this threat to information quality, we first need to understand exactly the types of fake news.

There are many studies on fake news and fake news classification, one of the most referenced and cited reports on fake news classification is by Claire Wardle and Hossein Derakhshan [10]. They therefore introduce a new conceptual framework for examining information disorder, identifying the three different types: mis-, dis- and mal-information. Using the dimensions of harm and falseness, they describe the differences between these three types of information:

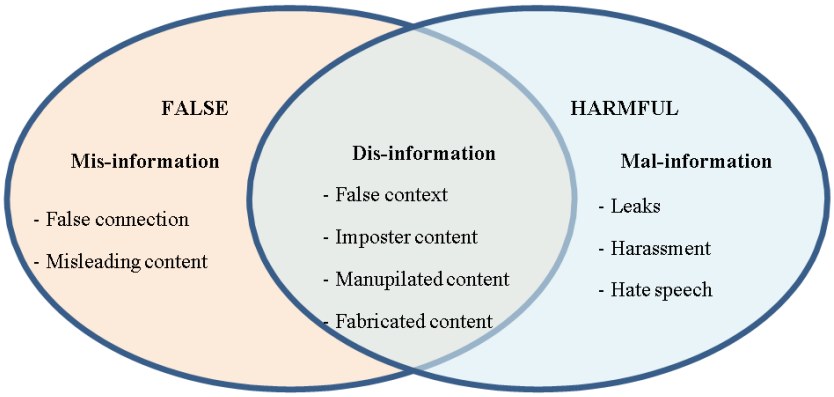

Figure 1

Classification of fake news

- Mis-information is when false information is shared, but no harm is meant.
- Dis-information is when false information is knowingly shared to cause harm.
- Mal-information is when genuine information is shared to cause harm, often by moving information designed to stay private into the public sphere.

4.3 Detecting fake news by humans

Before researching and developing an automatic fake news detection system, we can take a look at the 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,... [11].

Below, we summarize some common ways that people use to detect fake news online:

1) Check the source: Check the web address for the page you are viewing or where the content is distributed. Sometimes, fake news sites may have typos in their URLs or use less common domain name extensions like ".infonet" or ".offer".

Determining the source of the distribution of content will help assess the credibility of the content.

2) Check the author: Research on authors to see if they are trustworthy, e.g. if the author is real, does the author have a good reputation, does the author write about a particular area of expertise of distributed content? In particular, consider what the writer's motives might be.

3) Check other sources: Did the media or other reputable organizations cover this story? Check the reliable sources cited in the story? Professional news agencies around the globe have editorial guidelines and plenty of resources for fact-checking, so if they're also reporting the story, that's a good sign.

4) Maintain critical thinking: A lot of fake news is cleverly written to provoke strong emotional reactions like fear or anger to manipulate readers. Maintaining critical thinking by asking yourself: Why was this story written? Is it promoting a particular cause or agenda? Is it trying to make us visit another website?

5) Check the fact: Credible news stories will include lots of facts like data, statistics, quotes from experts, etc. If these are missing, ask why? Reports with false information often contain incorrect dates or altered dates, so it's a good idea to check when the article was published and the logic of the content.

6) Check the comments: Even if the article or video is legitimate, the comments below can help us find the truth. Note, links or comments posted in response to content may be automatically generated by network robots or people hired to deliver misleading information.

7) Check personal bias: We are all biased and should avoid letting bias overwhelm reason when evaluating content. Social media can create feedback chambers by recommending stories that match an individual's browsing habits, interests, and existing views. The more diverse sources and perspectives we read, the more likely we are to draw accurate conclusions.

8) Check if it's a joke: Satire sites are common, and it's not always clear whether a story is just a joke or parody. Check the website and author of the article to see if they are known for their satire or humorous stories to understand the true nature of the content.

9) Check the authenticity of the image: The illustrations we see with the content may have been edited or manipulated. Possible signs include warping where lines in the background now appear wavy, strange shadows, jagged edges, or skin tones that look too perfect. Also keep in mind that an image can be accurate but simply used in a misleading context. We can use tools like Google's Reverse Image Search to check where the image originated from and if the image has been altered.

For humans, depending on skills and analytical ability, we can combine all the above factors to come to a decision 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.

4.4 Automatic detection of fake news

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 and fast. 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.

5 Proposal solution

Within the scope of this paper, we present our proposed solution to detect fake news on news articles written in Vietnamese and based on text classification techniques by deep learning, specifically by RNN.

5.1 Text classification

The problem of text/document classification can be stated as follows [12]: give a set of documents D={ d1, d2, ..., dn }, di is the ith document and class set C= {c1, c2, ..., cm}, cj is the jth class. The purpose of the problem is to identify and assign the text di into the class cj.

The objective is to find the function

$f: D \times C \rightarrow \{True, False\}$

Inside:

– f(di,cj)=True, if the document di belongs to class cj

– f(di,cj)=False, if the document di does not belong to class cj

There are many cases of text classification such as binary classification (just need to determine a text belongs/does not belong to a given class), multi-layer classification (a text belonging to a certain class in the list of given classes),

multivalued classification (a text can belong to more than one class in the list of given classes).

If there are 3 classes c1, c2, c3 then we will have the combination (c1, c2), (c1, c3), (c2, c3). If there are n classes, there are n*(n–1)/2 combinations.

Example:

There is text d through a combination (c1, c2)→c2; d through a combination (c1, c3)→c3; d through a combination (c2, c3)→c2; → d belongs to class c2 because the result is that c2 which it appears most often.

5.2 Proposed model

The general model of this approach is as follows:

- The first step in this model is the data collection phase to build the training datasets. This datasets 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.
- Training stage allows to extract 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.

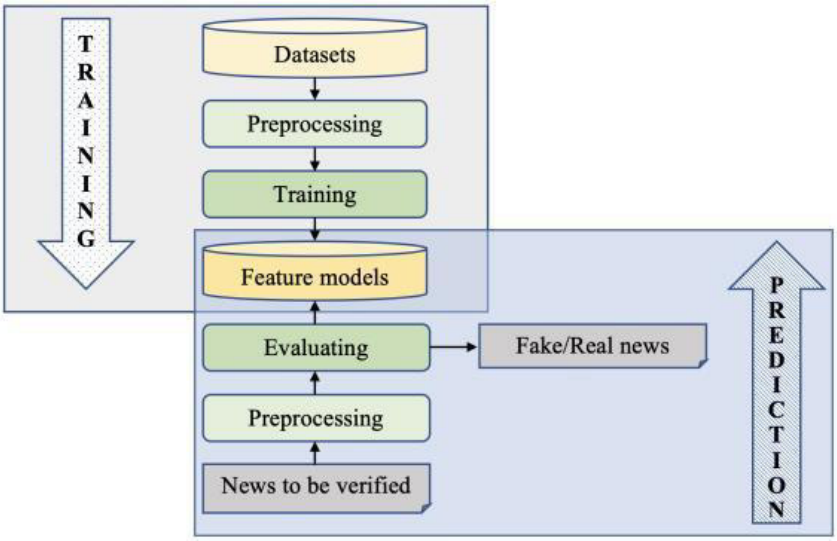

Figure 2

Fake news detection process model

5.3 Classification by using RNN

RNN (Recurrent Neural Network) is an algorithm that has received a lot of attention recently because of its good results obtained in the field of natural language processing [13].

The main idea of RNN is to use sequence of information. In traditional neural networks all inputs and outputs are independent of each other, that is, they are not chained together. This model is not suitable for many problems. For example, if we want to guess the next word that may appear in a sentence, we need to know how the previous words appear one after the other (recurrent) because they perform the same task for all elements of a sequence whose output depends on previous calculations, in other words, RNNs are capable of remembering previously computed information. In theory, the RNN can use the information of a very long text, but in practice it can only remember the previous few steps.

The RNN is designed as Figure 7. As shown in the previous section, 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.

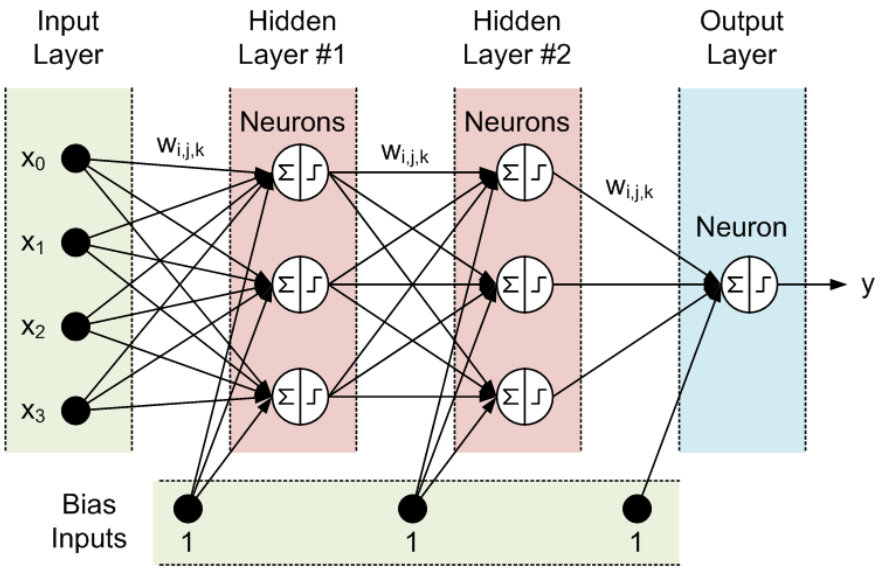

Figure 3

Architecture of RNN

Therefore, RNN was born with the main idea of using a 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.

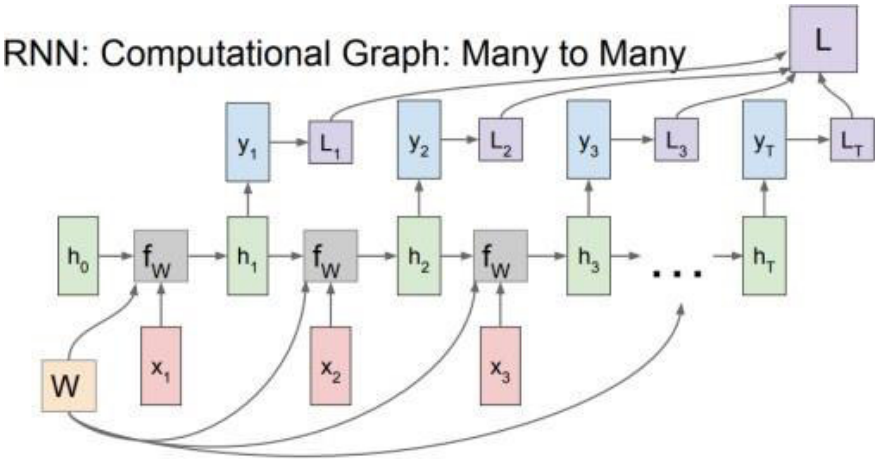

Figure 4

RNN 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 xt will be

combined with hidden layer ht-1 by fw function to calculate the current hidden layer ht and the output yt will be calculated from ht. W is the set of weights and it is obtained in all clusters, the L1, L2,..., Lt are the loss functions which will be explained later. Thus, the results from previous computations have been "remembered" by adding ht-1 to calculate ht to increase the accuracy of current predictions. Specifically, the calculation process is written in mathematical form as follows:

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 xt will be combined with hidden layer ht-1 by fw function to calculate the current hidden layer ht and the output yt will be calculated from ht. W is the set of weights and it is obtained in all clusters, the L1, L2,..., Lt are the loss functions which will be explained later. Thus, the results from previous computations have been "remembered" by adding ht-1 to calculate ht to increase the accuracy of current predictions. Specifically, the calculation process is written in mathematical form as follows:

$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t)$

$y_t = W_{hy}h_t$

At this point, 3 new things appear: W_{xh}, W_{hh}, W_{hy}. For NN using only a single weight matrix W, for RNN it 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" ht from there combined with Why to calculate y_t.

Thus, RNN is a specialized neuron-based method that is effective in sequential information processing. An RNN recursively applies a 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. The following figure illustrates a simple RNN framework below.

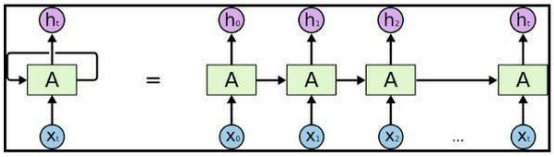

Figure 5

Frame of RNN

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 at natural language processing tasks.

6 Experiment

To build and test Vietnamese fake news detection software, we follow these steps:

- Data preparation (including training data and data for testing the evaluation);
- Data preprocessing;
- Model building (through training to model and refine);
- Build software and test to evaluate text as fake news or real.

1) 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).

We divide downloads into 8 categories:

Table 1.

Data used for testing fake news detection system

Nº	**Category**	**Number of files**
1	Real political news (for training)	500
2	Political fake 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	Political fake 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 1000-2000 words and the size of each file is about 5KB on average.

2) Data preprocessing

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

xã_hội... different from English and some other languages, the words are separated. We will import the necessary libraries to process the data:

```
from pyvi import ViTokenizer, ViPosTagger # Library NLP forVietnamese
from tqdm import tqdm
import numpy as np

import gensim # Library NLP of Python
```

The command `lines = gensim.utils.simple_preprocess(lines)` is used to process the deletion of special characters and the command `lines = ViTokenizer.tokenize(lines)` is used to separate words in Vietnamese documents. The results achieved at this stage are "clean" documents in the form of TXT.

3) Text vectorization

To convert text into a vector, we will 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). In which, 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. Here we will continue to use the *Gensim* library:

```
from gensim.models import KeyedVectors

dir_path      =      os.path.dirname(os.path.realpath(os.getcwd()))
word2vec_model_path   =   os.path.join(dir_path,   "Data/vi/vi.vec")
w2v     =     KeyedVectors.load_word2vec_format(word2vec_model_path)
vocab = w2v.wv.vocab

wv = w2v.wv

def get_word2vec_data(X):

      word2vec_data = []

      for x in X:

            sentence = []

            for word in x.split(" "):

                  if word in vocab:

                        sentence.append(wv[word])

            word2vec_data.append(sentence)

      return word2vec_data
X_data_w2v = get_word2vec_data(X_data)
X_test_w2v = get_word2vec_data(X_test)
```

Thus, we have converted all training and testing documents into vectors. From here, we can use them for Deep Learning models for text classification.

4) Model building and evaluation

To test the construction of RNN models as well as evaluate the classification on 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 [14].

For example, to create custom models, we only need to adjust the parameters when building the model:

```
model = Sequential()
model.add(layers.Embedding(input_dim=vocab_size,

output_dim=embedding_dim, input_length=maxlen))

model.add(layers.GlobalMaxPool1D())

model.add(layers.Dense(10, activation='relu'))

model.add(layers.Dense(1, activation='sigmoid'))
```

To test the classification on datasets, we just need:

```
history = model.fit(X_train, y_train,

epochs=50, verbose=False,

validation_data=(X_test, y_test), batch_size=10)
```

5) Results and comments

Based on the classification results of the messages used for testing in the comparison between CNN and RNN models, the correct classification results for the newsgroups are as follows:

Table 2.

Experimented results on CNN and RNN

Nº	Topics	Correct ratio	
		CNN	RNN
1	Real political news (for testing)	85 %	89 %
2	Political fake 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 %

Through the test results, we observed:

- The detection of fake news through content analysis for classification based on deep learning is positive, but the results are not high. In particular, the

classification is mainly based on word usage and writing style rather than on the intention of the writer. Using only content-based classification makes it difficult to tell if it is fake news or not.

- RNN is a well-suited model for natural language processing because its iterative structure is well-suited for variable-length text processing. The RNN-based text classification method improves the accuracy of text classification to a certain extent, however, it still has some shortcomings and can be improved. During neural network training, how to efficiently tune the parameters in each algorithm is a matter that needs to be studied further.

Conclusion

Fake news is a global problem, not specific to any one country. Identifying fake news is very difficult and so far, scientists are focusing on research to have better solutions. In particular, research on detecting fake news for Vietnamese is an urgent issue, but the research results are still very limited.

In this paper, we present a solution to apply deep learning technique of RNN to support fake news detection for 2 fields of politics and news related to Covid-19 in Vietnamese. For testing, we built a dataset with 2000 news for training and 400 for testing. Fake news detection simply uses text classification to see if a news is classified as fake news or true news. The detection/classification rate is over to 80%. Classification rerult depends on the topic, for example political news results are often higher than Covid-19 news.

In the future, issues that need to be researched to improve the quality of fake news detection may include the following main directions: 1) Combining multiple evaluation forms such as checking the source, verifying the author's information , check the distribution process, ... besides just analyzing the content; 2) Study how to update the model through continuously updating the training data warehouse as more content is generated every day; 3) Continue to propose new methods and improve the existing ones towards the application of deep learning and Big Data techniques.

Acknowledgement

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

References

[1] Dat Quoc Nguyen and Anh Tuan Nguyen. (2020): “PhoBERT: Pre-trained language models for Vietnamese”, Proceedings EMNLP 2020 (The 2020 Conference on Empirical Methods in Natural Language Processing), pages 1037–1042.

[2] Duc-Trong Le et al. (2020), "ReINTEL: A multimodal data challenge for responsible information identification on social network sites", 7th Annual Workshop on Vietnamese Language and Speech Processing - VLSP 2020.

[3] M.D. Molina, S.S. Sundar, T. Le, D. Lee. (2021): “Fake News Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content”, American Behavioral Scientist, Volume: 65, Issue: 2, pages 180-212.
[4] T. Miller, P. Howe, and L. Sonenberg. (2017): "Explainable AI: Beware of inmates running the asylum or: How I learnt to stop worrying and love the social and behavioural sciences", IJCAI, Workshop on Explainable Artificial Intelligence.
[5] A. Chadwick, J. Stanyer. (2022): " Deception as a Bridging Concept in the Study of Disinformation, Misinformation, and Misperceptions: Toward a Holistic Framework", Communication Theory, Volume 32, Issue 1, Pages 1–24.
[6] X. Zhou, J. Wu, and R. Zafarani. (2020): "SAFE: Similarity-Aware Multi-modal Fake News Detection", In Advances in Knowledge Discovery and Data Mining, Springer International Publishing, p.p. 354–367.
[7] X. Zhou and R. Zafarani. (2019): "Network-based Fake News Detection: A Pattern-driven Approach", arXiv e-prints.
[8] D. Kollias, S.P. Zafeiriou. (2020): "Exploiting multi-CNN features in CNN-RNN based dimensional emotion recognition on the OMG in-the-wild dataset", IEEE Transactions on Affective Computing.
[9] M.K. Elhadad, K.F. Li, F. Gebali. (2019): "A novel approach for selecting hybrid features from online news textual metadata for fake news detection", International conference on p2p, parallel, grid, cloud and internet computing, Springer, p.p. 914-925.
[10] C. Wardle, H. Derakhshan. (2017): "Information disorder: Toward an interdisciplinary framework for research and policy making", Report of DGI (Directorate General Human Rights and Rule of Law), Published by the Council of Europe.
[11] D. Fallis. (2015): “What Is Disinformation?", Library Trends, Johns Hopkins University Press, Volume 63, Number 3, pp. 401-426.
[12] Vo Trung Hung. (2020): " Development of a Document Classification Method by Using Geodesic Distance To Calculate Similarity of Documents". Eastern-European Journal of Enterprise Technologies, Volume 4, Issue 2 (106), pages 25-32.
[13] Y. Yu. (2019): "A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures", Neural Computation, Vols. 31, 1235–1270.
[14] N. Ketkar. (2017): "Introduction to Keras", In: Deep Learning with Python. Apress, Berkeley, CA.