Indonesian Journal of Electrical Engineering and Computer Science
Vol. 30, No. 2, May 2023, pp. 1232~1241
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v30.i2.pp1232-1241

Deep learning for classifying thai deceptive messages

Panida Songram[1], Suchart Kummanee[1], Phatthanaphong Chomphuwiset[1], Chatklaw Jareanpon[1], Laor Boongasame[2], Khanabhorn Kawattikul[3]
[1]Department of Computer Science, Faculty of Informatics, Mahasarakham University, Bangkok, Thailand
[2]Department of Mathematics, Faculty of Science, King Mongkut's Institute of Technology, Bangkok, Thailand
[3]Department of Information Technology, Faculty of Social Technology, Rajamangala University of Technology Tawan-Ok, Chanthaburi, Thailand

Article Info

Article history:

Received Aug 30, 2022
Revised Jan 18, 2023
Accepted Jan 26, 2023

Keywords:

Deceptive messages
Deep learning
Online deception
Social networks
Text classification
Thai deception

ABSTRACT

Online deception has become a major problem affecting people, society, the economy, and national security. It is mostly done by spreading deceptive messages because message are quickly spread on social networks and are easily accessed by anyone. Detecting deceptive messages is challenging as the messages are unstructured, informal, and complex; this extends into Thai language messages. In this paper, various deep learning models are proposed to detect deceptive messages under two feature extraction trials. A balanced two-class dataset of deceptive and truthful Thai messages (n=2378) is collected from Facebook pages. Instance features are encoded using word embeddings (Thai2Fit) and one-hot encoding techniques. Five classification models, convolutional neural network (CNN), bidirectional long short-term memory (BiLSTM), bidirectional gated recurrent units (BiGRU), CNN-BiLSTM, and CNN-BiGRU, are proposed and evaluated upon the dataset with each feature extraction technique. The experimental results show that all the proposed models had excellent accuracy (95.59% to 98.74%) and BiLSTM with one-hot encoding gave the best performance, achieving 98.74% accuracy.

Corresponding Author:

Laor Boongasame
Department of Mathematics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand
Email: laor.bo@kmitl.ac.th

1. INTRODUCTION

Social networks connect us with other people, sharing aspects of our life. Communication via online social networks has become a part of one's life in this digital age. Social networks, especially Facebook, are widely used in the daily lives of people and have been growing rapidly. In Thailand, Facebook is the most popular social media network and reached around 50.75 million users in 2022, despite slower growth forecasted to reach 45 million by 2026 [1]. Due to the numerous users and sharing nature of social networks, deceivers may easily spread messages with malicious intentions to other users (victims). Therefore, detecting and acting upon deceptive messages within social networks is of rising importance to mitigate the number of victims of deception and related online fraud. Various approaches have been proposed in the literature to detect online deception in messages. They can be grouped into two categories. Machine learning methods are the first category exploited to detect deception in social media communications. Briscoe *et al.* [2] classified deception by machine learning methods–random forest, gradient boosting, support vector machine (SVM), and Perceptron–on features, including sentence length, sentence complexity, sentiment, emoticon usage, and informality. Appling *et al.* [3] proposed random forest and SVM to evaluate deception strategies by using textual cues. Ott *et al.* [4] investigated deceptive opinion spam on hotel reviews with several extracted features i.e., psychological deception, n-gram-based text categorization, and a combination of both n-gram and psychological deception. Naïve Bayes and SVM then classified the spam opinion inputs. Deceptive opinion spam detection was also investigated in [5] and [6]. Rayana and Akoglu [5] used the behavior and content of messages as features for classifying opinion spam with the SPEAGLE approach. The mislabeled instances were corrected to find deceptive opinion spam from message reviews [6]. Features including lexicon, part of speech, deep syntactic information of text, and psycholinguistic features, were extracted from the reviews and classified by Naïve Bayes, SVM, and logistic regression classifiers. Moreover, machine learning methods were applied to detect spam messages which led to online deception [7], [8]. Bhat *et al.* [9] compared single models and ensemble models for spammer classification, including decision tree, Naïve Bayes and K-nearest neighbors (KNN), Bagging, Boosting, and Stacking ensemble classifiers. Zheng *et al.* [10] and Gupta and Kaushal [11] detected spammers and non-spammers on social networks from message content and user behavior, using SVM, Naïve Bayes, and decision tree classifiers. Abdulqader *et al.* [12] proposed ten theories and nine relevant constructs to create a model for detecting fake online reviews. The ten theories include self-presentational theory, four-factor theory, interpersonal deception theory, leakage theory, truth-default theory, reality monitoring theory, criteria-based content analysis, scientific content analysis, verifiability approach, and information manipulation theory. The constructs include specificity, quantity, non-immediacy, affect, uncertainty, informality, consistency, source credibility, and deviation in behavior. Verbal and non-verbal features were investigated to validate the proposed model and found that non-verbal features are more important than verbal features.

Secondly, deep learning methods have been proposed to detect online deception. Jain *et al.* [13] evaluated multiple deep neural network (DNN) based approaches to detect deceptive reviews. The performances of these models were compared on multiple benchmark datasets. In addition, a multi-instance learning and hierarchical architecture handling variable length review texts were reported to have outperformed other machine learning methods. Anass *et al.* [14] reported the comparison between different neural network architectures and their effectiveness in the detection of deceptive opinion spam. Their results showed that convolutional neural network (CNN) performed better than recurrent neural network (RNN), long short-term memory (LSTM), bidirectional long short-term memory (BiLSTM), gated recurrent units (GRU), and bidirectional gated recurrent units (BiGRU). Zhang *et al.* [15] proposed a deep learning approach, called deep context representation by word vectors (DCWord), for text representation to deceptive review identification. The basic idea is that contextual information of words of deceptive reviews and truthful reviews should have different characteristics. The average-pooling technique is applied to the word vector-encoded data. Experimental results reported that the DCWord-M representation with logistic regression gave the highest accuracy for detecting deceptive reviews. Zhang *et al.* [16] proposed a deep learning method, called deceptive review identification by recurrent convolutional neural network (DRI-RCNN) to identify deceptive reviews. It used word contexts and a deep learning technique to detect deceptive reviews. Their experiment found that the DRI-RCNN outperformed SVM in deceptive review identification. Qureshi *et al.* [17] applied a feature set combining the connectivity patterns of news propagators with their profile features to detect COVID-19 Fake News. Various machine learning and deep learning models are investigated for the detection and found that CATBoost and RNN are the most effective.

These previous works show deep learning outperforms machine learning in deceptive message identification, where the ultimate cause is that hand-engineered feature extraction (in machine learning techniques) does not provide the necessary semantic information from the text data to discriminate the deceptive indicators [14]. Therefore, we apply deep learning models to detect deceptive Thai-language messages from Facebook sources. In addition, Thai language messages are complex in a different manner to English-language messages, and the techniques in this paper contribute to the much-needed work in this very challenging and under-explored specialization of deception detection.

The rest of this paper is organized as follows: in section 2, the proposed method is explained. In section 3, we show the experimental results and give a discussion. Finally, the conclusion and direction of future work are given in section 4.

2. METHOD

The overall process of the method is shown in Figure 1 and consists of 3 main parts; data preparation, the proposed models, and evaluation. For the proposed models, the CNN, BiLSTM, and BiGRU models are proposed to classify deceptive messages. We also proposed a set of hybrid models i.e, CNN-BiLSTM, and CNN-BiGRU, for detecting deceptive messages.

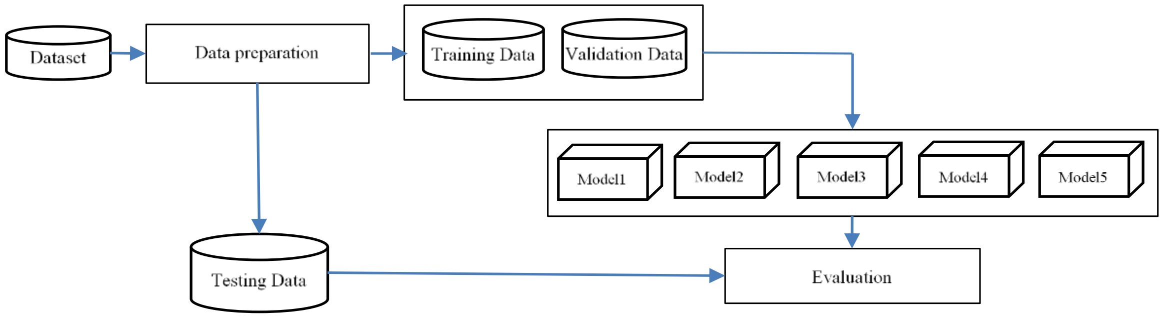

Figure 1. The proposed method

2.1. Data preparation

The dataset was collected by extracting textual messages from Facebook pages relevant to job applications written in Thai-language. Truthful messages were collected from reliable pages. The deceptive messages were collected from unreliable pages. The detail of the dataset collection process can be found in [18]. 1,189 truthful and 1,189 deceptive messages are in the dataset. The dataset's messages have a mean average length of 926 characters. Since deep learning models cannot learn directly from raw text data, a procedure to transform input messages into feature vectors is required. First, the messages are cleaned by removing numbers, signs, and stop words. Then the remaining character sequence (message) is segmented into words using OSKut [19]. These vectors of words undergo sequence padding by prepending 0's to ensure equal lengths at 534 words (the maximum length of a processed message vector within the dataset). Each message instance is thus a feature (or word) vector. In our experiments, two feature vector encodings are tested. The first is the one-hot encoding technique, the second is performed using Thai2fit [20]. Thai2fit is a pre-trained Thai word embedding technique that has been trained with Thai Wikipedia data by an ULMFit method–each word is represented by a 300-dimension vector.

2.2. The proposed models

2.2.1. CNN

CNN has widely been used for text classification and had favorable results in various domains of text classification [21]. Underlying, it is a feed-forward neural network. It has a convolutional layer for generating feature maps. Then the size of the feature map is reduced by using the pooling layer. Finally, the softmax layer (one of several activation functions) is used as a classification output. The structure of the proposed CNN model for classifying deception messages is shown in Figure 2. An input feature vector is fed into the convolutional layer to learn information from words in sentences through the filters. In the proposed CNN model, the input feature vector is convolved by using 32 filters with sizes 3×300 each. The output from the convolution of each filter is a feature map with size h-(3-1) when striding = 1, where h is the number of words in the message. The feature map is pooled by using 1D-max pooling layer to generate a convolutional feature with a size of 32. The feature map is flattened into a single column using the flatten layer. A dropout layer is applied for regularization to reduce overfitting in the model. Finally, the output layer gives the predicted class.

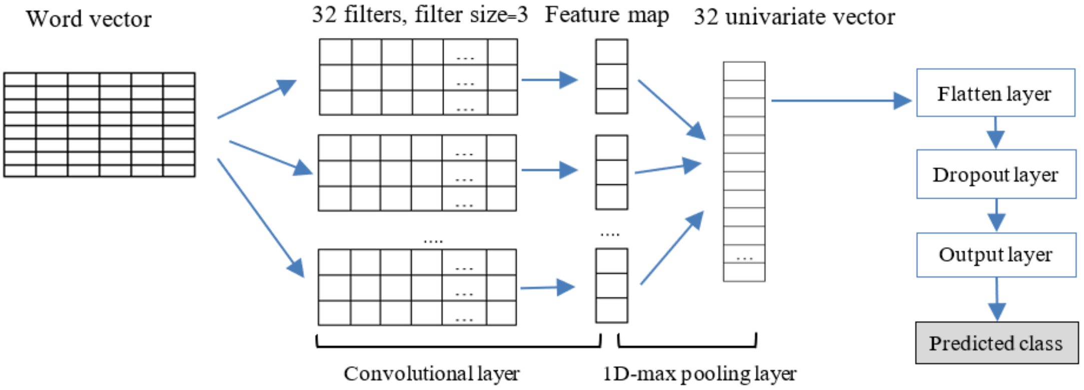

Figure 2. The structure of the proposed CNN model

2.2.2. Bi-LSTM

LSTM is a deep learning method involving sequential data [22]. It is an algorithm in the RNN family. LSTM processes data in a forward direction with the ability to remember and forget the information. LSTM consist of forget gate (f_t), input gate (i_t), input modulation gate ($\tilde{c}_t$), cell state (c_t), output gate (o_t), and hidden state (h_t) [23]. Forget gate is used to decide which information is kept to calculate the cell state and which information should be forgotten. The information of sample (x_t) and previous hidden state (h_{t-1}) are fed through a Sigmoid function in forget gate that can be expressed by (1), where *W* is the weight matrix and *b* is the bias vector. Input gate helps to find important information of a sample (x_t) with the previous hidden state (h_{t-1}). It helps to find out important information. The input gate can be expressed as (2). The input modulation gate is the candidate cell state. It learns both new information and the previous hidden state, as shown in (3). The cell state combines old information that is dropped by a forget gate and new information that is produced by the input gate and modulation gate, as shown in (4). The output gate gives the next hidden state, as shown in (5). The hidden state holds the information which is seen by LSTM, as shown in (6). LSTM's extension, called BiLSTM, was proposed to learn both past and future input data sequences so data is processed in the forward direction and backward direction in parallel [24]. For BiLSTM the hidden state of forward direction and backward direction are saved. For text classification, BiLSTM views text as a sequence of words. A sample (x_t) is a word vector of words in a sentence [25]. In this paper, the structure of the proposed BiLSTM model for classifying deceptive messages is shown in Figure 3.

$$f_t = sigmoid(W_f[h_{t-1}, x_t] + b_f) \tag{1}$$

$$i_t = sigmoid(W_i[h_{t-1}, x_t] + b_i) \tag{2}$$

$$\tilde{c}_t = tanh\ (W_c[h_{t-1}, x_t] + b_c) \tag{3}$$

$$c_t = f_t\, .\, \mathrm{c}_{t-1} + i_t\, .\, \tilde{c}_t \tag{4}$$

$$o_t = sigmoid(W_o[h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t\, .\, tanh\ (c_t) \tag{6}$$

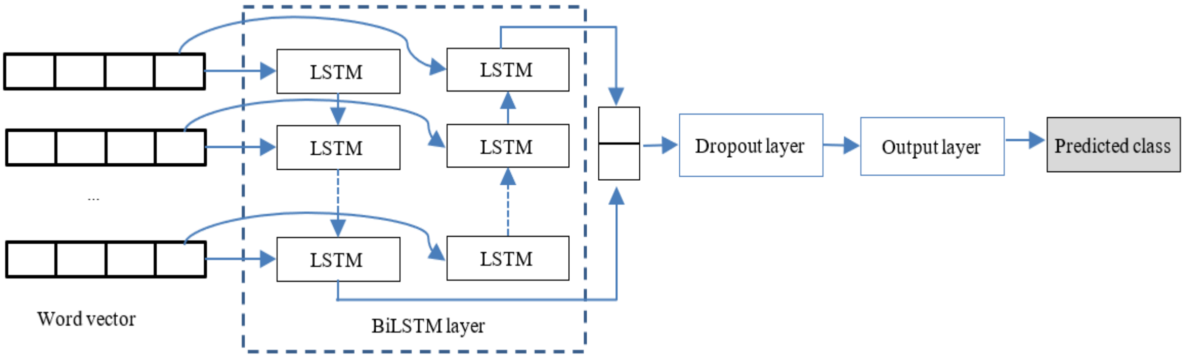

Figure 3. The structure of the proposed BiLSTM model

2.2.3. BiGRU

GRU is an algorithm of the RNN family [26]. It requires sequential data for learning. To decrease computation time, GRU reduces external signal gates from LSTM as shown in (7) and (8), where *U* is a weight matrix. In addition, it includes two gates, called an update gate z_t as shown in (9) and a reset gate r_t as shown in (10). The model parameters (*W*, *U*, *b*) are shared at all time steps and learned during the training stage. In this paper, we propose BiGRU which allows for the use of information from both previous time steps and later time steps to make predictions about the current state. The proposed BiGRU model in this paper is presented in Figure 4. The hidden state of forward and backward directions are generated in the BiGRU layer and combined to go through the dropout layer. Finally, the output from the dropout layer goes to the output layer and predicts the class.

$$h_t = (1 - z_t) \, . \, h_{t-1} + z_t \, . \, \tilde{h}_t \tag{7}$$

$$\tilde{h}_t = tanh(W_h x_t + U_h \, (r_t . h_{t-1}) + b_h \tag{8}$$

$$z_t = \sigma(W_z x_t + U_z \;\; h_{t-1} + b_z) \tag{9}$$

$$r_t = \sigma(W_r x_t + U_r \;\; h_{t-1} + b_r) \tag{10}$$

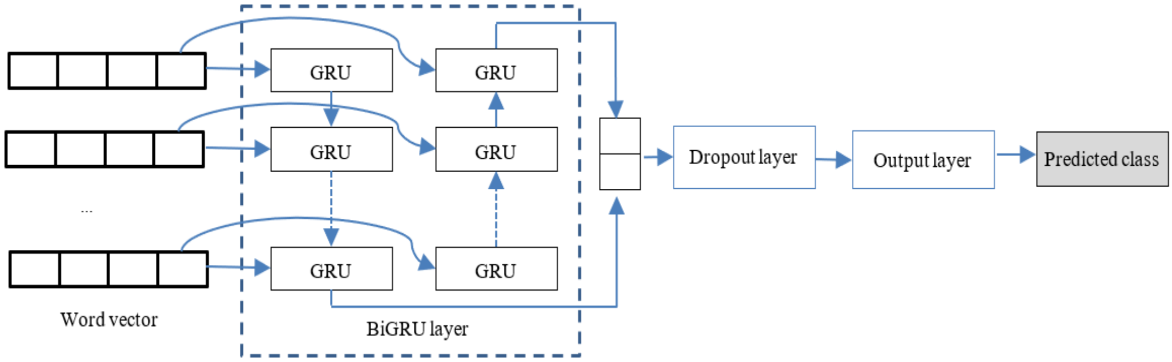

Figure 4. The structure of the proposed BiGRU model

2.2.4. CNN-BiLSTM

The proposed CNN-BiLSTM aims to learn the local features of the text input vector using CNN and then long-range dependency in the sequence of words is learned by BiLSTM. The proposed CNN-BiLSTM model is presented in Figure 5. The output from the convolutional layer is the feature map. Then the feature map is fed into the BiLSTM layer to learn the features in forward and backward directions. The hidden states of forward and backward directions are combined and fed through the dropout layer. Finally, the output from the dropout layer goes to the output layer and predicts the class.

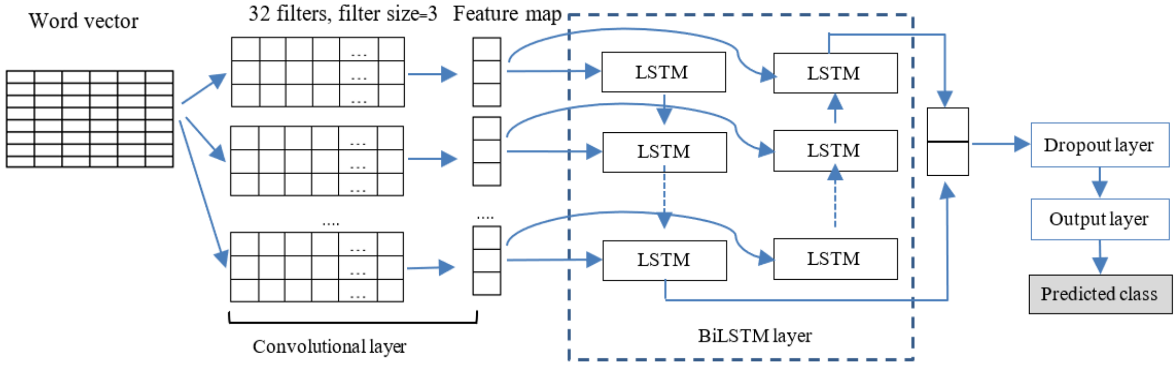

Figure 5. The structure of the proposed CNN-BiLSTM model

2.2.5. CNN-BiGRU

The proposed CNN-BiGRU model learns the local features of text via CNN and then long-range dependency between the sequences of words is learned by BiGRU. The proposed CNN-BiGRU model is presented in Figure 6. The output from the convolutional layer is fed into the BiGRU layer to learn the feature in forward and backward directions. The hidden states of forward and backward directions are combined and go through the dropout layer. Finally, the output from the dropout layer goes to the output layer and predicts the class.

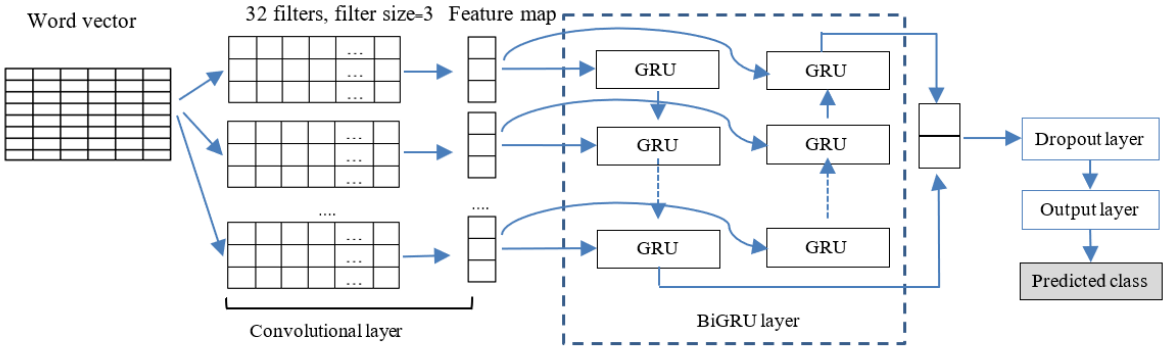

Figure 6. The structure of the proposed CNN-BiGRU model

2.3. Performance evaluation

This section explains the detail of the performance evaluation. All the proposed models are evaluated by using accuracy, precision, recall, and F-measure. Accuracy (Acc) measures the overall predictions that were correct (and is a reliable measure on our balanced dataset). The precision is the number of correct predictions from the predictions per class. The recall is the number of correct predictions from all true data for each class. F-measure is the harmonic mean between precision and recall. Precision, recall, and f-measure for the deceptive class are represented by P_D, R_D, F_D, respectively. Precision, recall, and f-measure for the truthful class are represented by P_T, R_T, F_T, respectively.

3. RESULTS AND DISCUSSION

3.1. Experimental settting

The dataset is split into three subsets, training, validation, and test sets, at a ratio of 60:20:20. All of the subsets inherited the same characteristics of the original dataset, including class distribution and sentence length distribution. Adam optimizer with a learning rate of 0.001 is set for all models. All models are trained for 10 epochs. Dropout is used as a regularization technique in all models. The dropout value is set to 0.3. For the proposed CNN model, it applies 32 filters with a size 3×300 of each. All models are implemented in Python.

3.2. Experimental result

Table 1 reports the performance of the proposed models using the one-hot vector encoding technique. In Table 1, we can see that all the proposed models gave high classification performance. The proposed CNN model gave 98.33% accuracy. The proposed BiLSTM model had the best accuracy at 98.74%, and the highest recall for detecting deceptive messages at 99.15%. Moreover, BiLSTM provided the highest precision for predicting truthful messages at 99.16%. Conversely, the proposed BiGRU model resulted in the lowest accuracy of all the models at 97.27%.

Among the hybrid models, CNN-BiLSTM had the highest accuracy at 98.53%. In addition, we found that the combination of CNN and BiGRU (CNN-BiGRU) improved the accuracy when compared to only BiGRU, where CNN-BiGRU gave 98.32% accuracy. In conclusion, Table 1 results show that the proposed BiLSTM outperforms CNN, BiGRU, CNN-BiLSTM, and CNN-BiGRU when using one-hot vector encoding. Figure 7 shows the convergence curve of the loss function for BiLSTM. We can see that at the final epoch, the training and validation losses reached a point of the minimal gap with stability between the two loss values, indicating a qualified model-to-data fit.

Table 1. The performance of the proposed models with one-hot vector encoding (%)

		Deceptive			Truthful		
Model	Accuracy	P_D	R_D	F_D	P_T	R_T	F_T
CNN	98.33	97.89	98.72	98.31	98.76	97.94	98.35
BiLSTM	**98.74**	**98.31**	**99.15**	**98.73**	**99.16**	**98.34**	**98.75**
BiGRU	97.27	96.20	98.28	97.23	98.33	96.31	97.31
CNN-BiLSTM	98.53	98.31	98.73	98.52	98.74	98.33	98.54
CNN-BiGRU	98.32	98.31	98.31	98.31	98.33	98.33	98.33

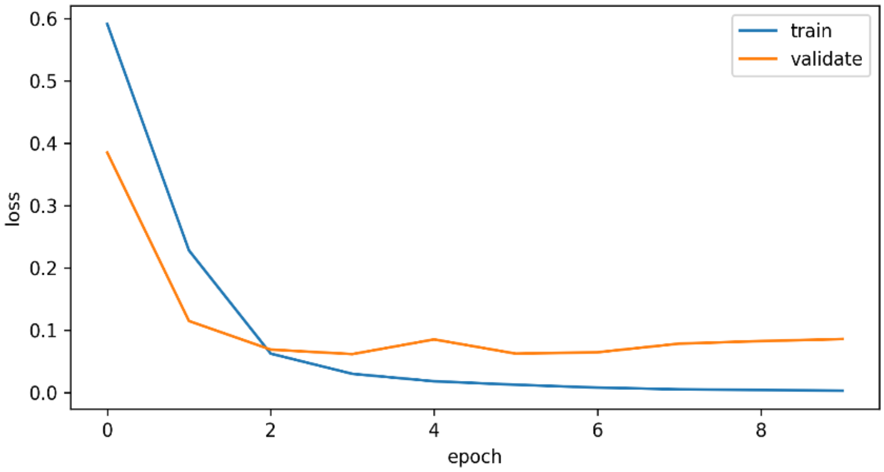

Figure 7. Convergence curve of the loss function for BiLSTM with one-hot vector encoding

Table 2 shows the performance of the proposed models tested with the word embedding vector feature encoding technique. From Table 2, we can see that all the proposed models gave high classification accuracy, especially CNN, BiLSTM, BiGRU, and CNN-BiLSTM. The proposed CNN-BiLSTM model gave the highest accuracy, precision, and f-measure of deceptive messages which are 97.90%, 97.15%, and 97.95%, respectively. The CNN and BiGRU models gave the highest recall of deceptive messages. While the proposed CNN-BiGRU model gave the lowest accuracy at 95.59%. In conclusion, the proposed CNN-BiLSTM had the highest accuracy performance in the word embedding vector trial. Figure 8 shows the convergence curve of the loss function for CNN-BiLSTM. We can see that training loss and validation loss decrease to a point of stability and share a small gap, indicating a good fit.

Table 2. The performance of the proposed models with word embedding vector encoding (%)

Model	Accuracy	Deceptive			Truthful		
		P_D	R_D	F_D	P_T	R_T	F_T
CNN	97.17	95.24	**99.17**	97.17	99.11	94.87	96.94
BiLSTM	97.06	95.97	98.35	97.14	98.25	95.73	96.97
BiGRU	97.69	96.39	**99.17**	97.76	**99.12**	96.15	97.61
CNN-BiLSTM	**97.90**	**97.15**	98.76	**97.95**	98.70	**97.01**	**97.84**
CNN-BiGRU	95.59	95.85	95.45	95.65	95.32	95.73	95.52

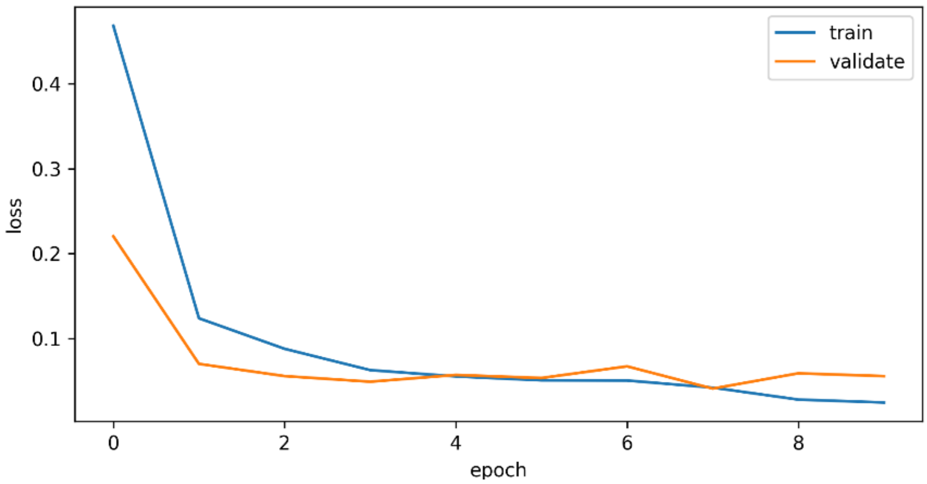

Figure 8. Convergence curve of the loss function for CNN-BiLSTM with word embedding vector encoding

From Tables 1 and 2, we can conclude that the proposed models gave excellent accuracy performance (95.59% to 98.74%) for detecting deceptive messages. The combination of CNN and BiLSTM with word embedding data encoding gave the highest accuracy performance of the word embedding trial (97.90%). Ultimately, the proposed BiLSTM with one-hot encoding technique gave the best overall classification accuracy performance (98.74%) on the dataset when compared to CNN, BiGRU, CNN-BiLSTM, and CNN-BiGRU, under all trail conditions.

4. CONCLUSION

This paper proposed deep learning models to classify deceptive messages written in Thai language. Five classification models–CNN, BiLSTM, BiGRU, CNN-BiLSTM, and CNN-BiGRU–were proposed and evaluated upon two different feature encoding techniques, one-hot encoding, and word embedding. From the experimentation, we found that all the proposed models gave excellent accuracy performance (95.59% to 98.74%) upon the Thai deceptive messages dataset collected from Facebook pages. We interpret that each of the deep learning models' high performance is due to their provision of semantic information extraction and their self-adaptability to extract highly discriminant features without intervention. Each of the proposed models gave high accuracy, recall, precision, and f-measure in detecting truthful and deceptive messages. The proposed BiLSTM model gave the best accuracy performance (98.74%) when features were encoded using the one-hot encoding technique. In future work, we will consider applying the proposed models to further datasets and run trials using part-of-speech tagging and semantic tagging to extract features.

ACKNOWLEDGEMENTS

This work was supported by faculty of Informatics, Mahasarakham University.

REFERENCES

[1] Statistics. "Number of Facebook users in Thailand from 2017 to 2020 with a forecast through 2026," Statistics.com http://www.statista.com/statistics/490467/number-of-thailand-facebook-users/ (accessed Aug. 8, 2022).
[2] E. J. Briscoe, D. S. Appling, and H. Hayes, "Cues to deception in social media communications," in *Proc 47th Hawaii International Conference on System Sciences*, Waikoloa, HI, USA, 2014, pp. 1435-1443, doi: 10.1109/hicss.2014.186.
[3] D. S. Appling, E. J. Briscoe, and C. J. Hutto, "Discriminative models for predicting deception strategies," in *Proceeding 24th International Conference on World Wide Web*, Florence, Italy, 2015, pp. 947–952, doi: 10.1145/2740908.2742575.
[4] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination," in *Proceeding 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Portland, Oregon, USA 2011, pp. 309–319, doi: 10.48550/arXiv.1107.4557.
[5] S. Rayana and L. Akoglu, "Collective opinion spam detection: bridging review networks and metadata," in P*roceeding 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Sydney, NSW, Australia, 2015, pp. 985–994, doi: https://doi.org/10.1145/2783258.2783370.
[6] Y. F. Ren, D. H. Ji, L. Yin, and H. B. Zhang, "Finding deceptive opinion spam by correcting the mislabeled instances," *Chinese. Journal Electronics*, vol. 24, no. 1, pp. 52-57, Mar. 2015, doi: 10.1049/cje.2015.01.009.
[7] S. More and R. Kalkundri, "Evaluation of deceptive mails using filtering and WEKA," in *Proc. 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS),* Coimbatore, India, 2015, pp. 1-4, doi: 10.1109/iciiecs.2015.7193262.
[8] S. More and S. A. Kulkarni, "Data mining with machine learning applied for email deception," in *Proceeding 2013 International Conference on Optical Imaging Sensor and Security (ICOSS),* Indonesia, 2013, pp. 1-4, doi: 10.1109/icoiss.2013.6678403.
[9] S. Y. Bhat, M. Abulaish, and A. A. Mirza, "Spammer classification using ensemble methods over structural social network features," in *Proc. 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT),* Warsaw, Poland, 2014, vol. 2, pp. 454-458, doi: 10.1109/wi-iat.2014.133.
[10] X. Zheng, Z. Zeng, Z. Chen, Y. Yu, and C. Rong, "Detecting spammers on social networks," *Neurocomputing,* vol. 159, pp. 27-34, Jul. 2015, doi: 10.1016/j.neucom.2015.02.047.
[11] A. Gupta and R. Kaushal, "Improving spam detection in online social networks," in *Proceeding 2015 International Conference on Cognitive Computing and Information Processing (CCIP),* Noida, India, 2015, pp. 1-6, doi: 10.1109/ccip.2015.7100738.
[12] M. Abdulqader, A. Namoun, and Y. Alsaawy, "Fake online reviews: a unified detection model using deception theories," *IEEE Access,* vol. 10, pp. 128622-128655, Dec. 2022, doi: 10.1109/ACCESS.2022.3227631.
[13] N. Jain, A. Kumar, S. Singh, C. Singh, and S. Tripathi, "Deceptive reviews detection using deep learning techniques," in *Proceeding International Conference on Natural Language Processing and Information Systems*, 2019, pp. 79-91, doi: 10.1007/978-3-030-23281-8_7.
[14] F. Anass, R. Jamal, M. A. Mahraz, Y. Ali, and H. Tairi, "Deceptive opinion spam based on deep learning," in *Proc. 2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS),* Oct. 2020, pp. 1-5, doi: 10.1109/ICDS50568.2020.9268772.
[15] W. Zhang, Q. Wang, X. Li, T. Yoshida, and J. Li, "DCWord: A novel deep learning approach to deceptive review identification by word vectors," *Journal of Systems Science and Systems Engineering,* vol. 28, no. 6, pp. 731-746, Dec. 2019, doi: 10.1007/s11518-019-5438-4.
[16] W. Zhang, Y. Du, T. Yoshida, and Q. Wang, "DRI-RCNN: An approach to deceptive review identification using recurrent convolutional neural network," *Information Processing and Management,* vol. 54, no. 4, pp. 576-592, Jul. 2018, doi: 10.1016/j.ipm.2018.03.007.

[17] K. A. Qureshi, R. A. S. Malick, M. Sabih, and H. Cherifi, "Complex network and source inspired COVID-19 fake news classification on Twitter," *IEEE Access*, vol. 9, pp. 139636-139656, Oct. 2021, doi: 10.1109/ACCESS.2021.3119404.
[18] P. Songram, A. Choompol, P. Thipsanthia, and V. Boonjing, "Detecting thai messages leading to deception on Facebook," in *Proc. International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making*, pp. 293-304, Nov. 2016, doi: 10.1007/978-3-319-49046-5_25.
[19] P. Limkonchotiwat, W. Phatthiyaphaibun, R. Sarwar, E. Chuangsuwanich, and S. Nutanong, "Handling cross and out-of-domain samples in Thai word segmentation," *Association for Computational Linguistics*, Aug. 2021, pp. 1003-1016, doi: 10.18653/v1/2021.findings-acl.86.
[20] C. Polpanumas and W. Phatthiyaphaibun, "ULMFit Language modeling, text feature extraction and text classification in Thai Language," Github.com. https://github.com/cstorm125/thai2fit (accessed Jul. 5, 2022).
[21] S. Minaee, N. Kalchbrenner, E. Cambria, N. N. Khasmakhi, M. Asgari-Chenaghlu, and J. Gao, "Deep learning--based text classification: a comprehensive review," *ACM Computing Surveys*, vol. 54, pp. 1-40, Apr. 2021, doi: 10.1145/3439726.
[22] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
[23] K. Pasupa and T. S. N. Ayutthaya, "Hybrid deep learning models for Thai sentiment analysis," *Cognitive Computation*, vol. 14, no. 1, pp. 167-193, Jan. 2022, doi: 10.1007/s12559-020-09770-0.
[24] G.-H. Kwak, C.-W. Park, H.-Y. Ahn, S.-I. Na, K.-D. Lee, and N.-W. Park, "Potential of bidirectional long short-term memory networks for crop classification with multitemporal remote sensing images," *Korean Journal of Remote Sensing*, vol. 36, no. 4, pp. 515-525, Aug. 2020, doi: 10.7780/KJRS.2020.36.4.2.
[25] Y. Li, L. N. Harfiya, K. Purwandari, and Y.-D. Lin, "Real-time cuffless continuous blood pressure estimation using deep learning model," *Sensors*, vol. 20, Sep. 2020, doi: 10.3390/s20195606.
[26] R. Dey and F. M. Salem, "Gate-variants of gated recurrent unit (GRU) neural networks," in *Proceeding 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, Medford, MA, United States, 6-9 Aug. 2017, pp. 1597-1600, doi: 10.1109/MWSCAS.2017.8053243.

BIOGRAPHIES OF AUTHORS

Panida Songram is currently an Associate Professor at the Faculty of Informatics, Mahasarakham University (MSU), Thailand. She received the M.Sc. and Ph.D. degrees in Computer Science at King Mongkut's Institute of Technology Ladkrabang. Her research areas are in association rule mining, sequential pattern mining, text mining, and data mining. She can be contacted at email: panida.s@msu.ac.th.

Suchart Khummanee is a full lecturer of Computer Science at the Mahasarakham University, Thailand. He received a B.Eng. degree in Computer Engineering from King Mongkut's Institute of Technology Ladkrabang, M.Sc. degree in Computer Science from Khon Kaen University, Thailand, where he was awarded a Ph.D. degree in Computer Engineering. His research interests are in network security, computer networks, agricultural robotics, and internet of things (IoT). He can be contacted at email: suchart.k@msu.ac.th.

Phatthanaphong Chomphuwiset is a full-time lecturer at the Department of Computer Science, Mahasarakham University, Thailand. His main research interests fall into a variety areas, including medical image processing, computer vision applications and machine learning. He can be contacted at email: phatthanaphong.c@msu.ac.th.

Chatklaw Jareanpon is an assistant professor at the Department of Computer Science, Faculty of Informatics, Mahasarakham University, Thailand. Her interest areas are in data science, artificial intelligence, image processing and robotics. She can be contacted at email: chatklaw.j@msu.ac.th.

Laor Boongasame is a lecturer in the School of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand. She attained her Ph.D. in Computer Engineering from King Mongkut's University of Technology Thonburi, Thailand. Her research interests involve buyer coalitions, n-person game theory, and investment. She has published several research papers in internationally refereed journals and has presented papers at several international conferences. She can be contacted at email: laor.bo@kmitl.ac.th.

Khanabhorn Kawattikul is a full time lecturer at Department of Information Technology, Faculty of Social Technology, Rajamangala University of Technology Tawan-Ok (Chanthaburi Campus), Thailand. Khanabhorn received a master's degree in information technology from Mahasarakhsm University, Thailand. The main research of interests falls into variety areas, including human computer interaction, computer vision and machine learning. She can be contacted at email: khanabhorn_ka@rmutto.ac.th.