

THE COMBINATION OF CONVOLUTION NEURAL NETWORKS AND DEEP NEURAL NETWORKS FOR FAKE NEWS DETECTION

ZAINAB A. JAWAD*, AHMED J. OBAID

Department of computer science, University of Kufa
College of Mathematical and Computer Science, Najaf 54001, Iraq
*Corresponding Author: Zainaba.alhasani@student.uokufa.edu.iq.

Abstract

Nowadays, People prefer to follow the latest news on social media, as it is cheap, easily accessible, and quickly disseminated. However, it can spread fake or unreliable, low-quality news that intentionally contains false information. The spread of fake news can have a negative effect on people and society. Given the seriousness of such a problem, researchers did their best to identify patterns and characteristics that fake news may exhibit to design a model that can detect fake news before publishing. This paper proposed a novel fake news detection system using the Fake News Challenge stage #1 (FNC-1) dataset. The procedure implies processing the text using different natural language techniques and encodes the text using the Term Frequency-Inverse Document Frequency (TF-IDF) technique. Then, the extracted features are reduced using the elbow truncated method. Then, the similarity between each pair of headline and body text is found using the soft cosine similarity method. The similarity feature is entered into the deep learning algorithms. Thus, the paper proposed a novel model using the elbow truncated method with the combination of CNN and DNN deep learning techniques. The proposed system detects all categories with high accuracy except the disagree category. As a result, it achieves up to 84.6 % accuracy, classifying it as the second-ranking when ranking the competitive studies regarding this dataset.

Keywords: Convolutional neural networks, Deep neural networks, FNC-1 dataset, Misleading content, Natural language processing, Soft cosine similarity, TF-IDF.

1. Introduction

The detection of fake news has recently stimulated the attention of the general audience and scientists as the spread of misleading information online grows [1], especially on the internet, such as newsfeeds, news articles, and newspaper articles. For example, according to a recent Jump-shot Tech Blog, on Facebook, 50% of the daily traffic represents fake news sites, while the remaining 20% present normal websites [2].

Identifying fake content in online sources is crucial since 62% of U.S. adults read news on social media [3]. Until now, computational methods for detecting fake news have depended heavily on humorous news sources like "The Onion" and fact-checking websites like "Factcheck.org" and "Snopes." Moreover, using these publications presents several difficulties and possible consequences.

For example, using humorous content as a source for misinformation can introduce underlying contributing effects into the analysis, such as jokes and illogicality. This is certainly relevant for humorous reports from "The Onion," which has previously been used to investigate other sentence characteristics such as jokes [4] and irony [5].

However, fact-checking websites are typically limited to a single area of interest, like politics. They necessitate human experience, making it challenging to find datasets that offer some level of generalization across several disciplines [6].

The fake news detection system can be evaluated from its ability to detect fake news accurately. A set of datasets has been collected for this purpose, which can be used to evaluate fake news detection systems. Consequently, recent work has explored how machine learning algorithms can extract linguistic features from textual articles [7]. Then, different learning algorithms can be used to verify that the headline and body text match, identify mismatched sentences in the body text, and measure the spread of fake news in social media [8].

This work proposed a novel approach for building a fake news detection system using the FNC-1 dataset. In contrast to the previous systems, the new system used the elbow truncated method with the combination of CNN and DNN deep learning techniques. As a result, the suggested system detects all the categories with high accuracy except the disagree category.

The organization of the paper is as follows: Section 2 describes the previous attempts for building fake news detection using the FNC-1 datasets, while Section 3 briefly describes the main components of the proposed system. Moreover, Section 4 presents and discusses the main results of the suggested system. Finally, Section 5 concludes the research.

2. Relate work

This section reviews existing attempts to use the FNC-1 dataset to build a misinformation news detection system. We especially focus on the approaches that use Deep Learning (DL) techniques, Machine Learning (ML) algorithms, and Natural Language Processing (NLP) techniques.

As an example, Riedel et al. [9] suggested using the UCL Machine Reading (UCLMR) method. The suggested method involves combining lexical and similarity

information with DL approaches. This system offers simplicity in contrast to the most advanced Neural Network algorithms. Despite the simplicity of the implementation, the authors acquired an 81.72% FNC-1 score. The researchers believed their work would be a foundation for competing attempts on the FNC-1 dataset. However, the system's accuracy is low and still needs much improvement.

Another team, Gaurav Bhatt et al. [10], introduced a model that combines DL methods, statistical, neural, and feature engineering heuristics. The suggested model was straightforward. Furthermore, it fulfils higher accuracy in detecting all categories, except the disagree instances, where it achieved 83.08% accuracy.

On the other hand, Mohtarami et al. [11] used a novel method for stance identification, where he applied memory networks with Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) approaches. A stance filtering component and the similarity-based matrix further enhanced the model. The suggested extension significantly improves the performance gains and adds the benefit of competitiveness. Results proved the ability of the method to remember previous information, and it reached a score of 81.23%. However, the proposed model cannot deal with a set of instances; instead, it handles one instant at a time.

Another interesting attempt was made by Borges et al. [12], where they presented a novel approach that combines Deep Neural Network (DNN) approaches with string external similarity features. The suggested model significantly improves the results of earlier research, which scored 82.23%. However, the authors reported that advanced sentence modelling techniques are required. Furthermore, instead of feed-forward approaches, the Recurrent Neural Network (RNN) method must be used for text encoding.

In another attempt by Robiert et al. [13] used RoBERTa as a foundation to improve the classification. In this case, two dense layers are included to reduce dimensions and the SoftMax classification layer. In addition, they used soft cosine similarity. Furthermore, using summarizing methods for the text provided good results in this preliminary research. However, the accuracy in detecting the agree and disagree instances still needs more improvement.

From studying the existing works, it is concluded that the proposed approaches fail to build a system that can detect fake news efficiently, where in most cases the accuracy was not enough in detecting the fake news categories. The current work proposed a new system for detecting fake news. The new system used the elbow truncated method with a combination of CNN and DNN deep learning techniques. As a result, the proposed system achieves high accuracies in detecting most instances, where it outperforms most previous attempts.

3. The proposed system

The main components of our proposed system are presented in this section. Figure 1 illustrates these components which are described in detail.

We used the FNC-1 dataset in this search. The FNC-1 is a collective effort that includes over 100 volunteers and 71 teams from academia and industry worldwide. This dataset was designed for stance detection with 75,385 labeled headlines and articles. There are four label categories: agree, disagree, discuss, and unrelated. In the dataset, each headline is phrased as a statement. The goal of the FNC-1 is to

explore how AI technologies could be employed to tackle fake news problems. They organize an annual competition to encourage the development of new tools that could assist human fact-checkers using ML, AI, and NLP.

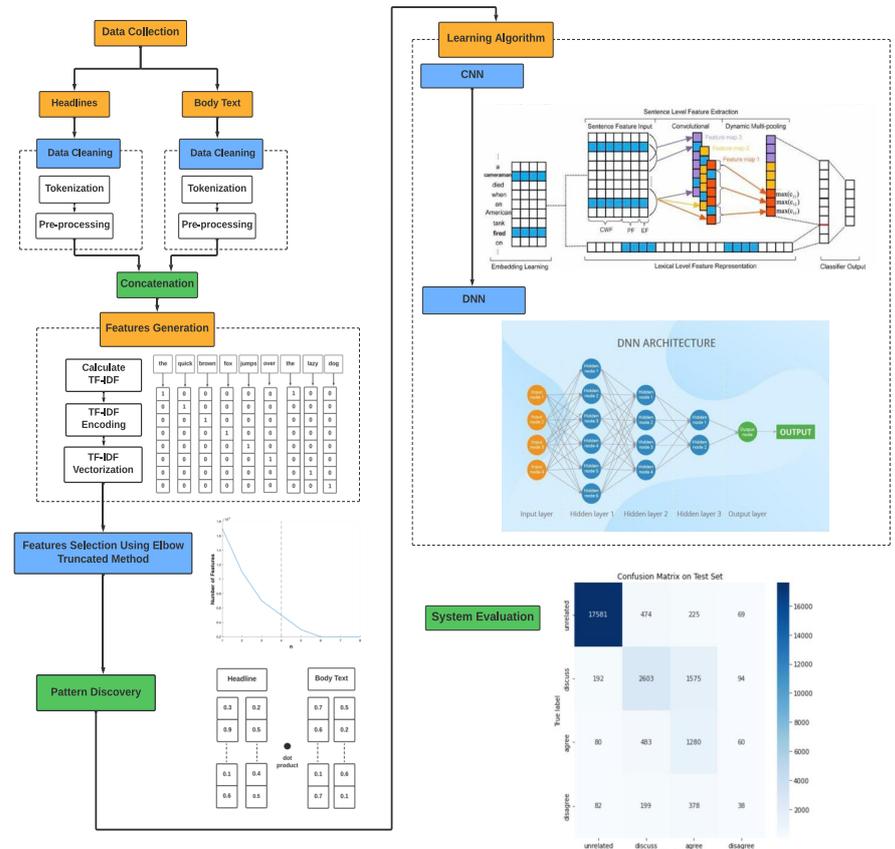


Fig. 1. The main components of the proposed system.

After that, we used data cleaning. Raw data must be transformed into a format that can be understood and used. A real-world dataset is usually incomplete and has inconsistent formatting. It is possible to resolve such issues by preprocessing data and making it more efficient and complete for analysis. A successful data mining or machine learning project depends heavily on this process. As a result, knowledge can be discovered more quickly from datasets, which eventually affects the performance of machine learning models—the main steps of preprocessing, as shown in Fig. 2.

Among NLP tasks, text generation involves generating text with constraints like starting characters or starting words. In addition, generating text appears indistinguishable from the human-written text. In the literature, this task is referred to as "natural language generation." Text Generation is a branch of NLP. It uses mathematical linguistics and AI knowledge to automatically generate natural language texts that can fulfil certain communicative requirements. Which include:

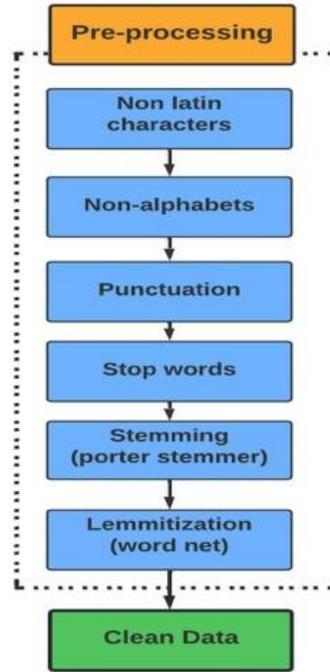


Fig. 2. The main components of the preprocessing.

Calculate and Encoding by the TF-IDF: It is a statistical measure that determines how relevant a text is to a document in a list of documents. It is accomplished by multiplying two metrics: the number of times a term makes it appears in a document and the phrase's inverse document frequency across a list of documents. In this step, we perform the TF-IDF for the headline and the body text, which produce two encoding lists: one for the headline and the other for the body text. as shown in Fig. 3.

the	quick	brown	fox	jumps	over	the	lazy	dog
1	0	0	0	0	0	1	0	0
0	1	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	1

Fig. 3. An example of TF-IDF vectorization.

Feature selection is crucial when building a fake news detection system for removing noisy features and keeping only those relevant to the system. After generating the TF-IDF vectors for both headlines in train stances and body text in train bodies, we

use the elbow truncated method to find the optimal number of features. The Elbow method a well suited method to estimate the truncated values than other methods, also gives a good perception about the correlation between data points and their related features [14]. Furthermore, we have 17000 concepts, while the optimal number of features after using this method is 5000. This step illustrates in Fig. 4.

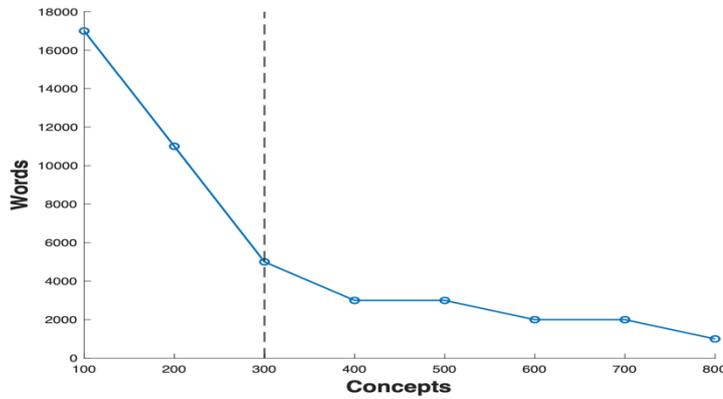


Fig. 4. Feature selection process.

After selecting the important feature, we contacted the headline in train stances and text bodies in train bodies in one CSV. This step is shown in Fig. 5.

	Headline	Body_Text
0	soldier shot parliament locked gunfire erupts ...	A small meteorite crashed into a wooded area i...
1	tourist dubbed 'spider man' spider burrows ski...	A small meteorite crashed into a wooded area i...
2	luke somers killed failed rescue attempt yemen	A small meteorite crashed into a wooded area i...
3	breaking soldier shot war memorial ottawa	A small meteorite crashed into a wooded area i...
4	giant 8ft 9in catfish weighing 19 stone caught...	A small meteorite crashed into a wooded area i...

Fig. 5. An example of the merge process.

Furthermore, it seeks to extract similar patterns from temporal data, such as periodic patterns. Discovering similar patterns is considered among the most crucial data-mining processes and may be used in various fields. Soft Cosine Similarity (SCM) allows us to assess the similarity between two documents in a meaningful way. The sentences have no words in common, but SCM can measure their similarity accurately by modelling synonymy. As part of the method, the documents are also represented as TF-IDF vectors (in other words, their frequency in the documents). It outperforms many state-of-the-art methods when applied to a semantic text similarity task within a community question-answering context. Thus, we use SCM in this study.

Finally, learning algorithms define the last level in building a fake news system. This procedure is essential for building a fake news system that can identify misleading news in real time. Different strategies, including ML classification, and DL approaches, are available for this purpose. As our main technique in this study, we used multiple deep learning classification algorithms, which include:

- Convolutional Neural Networks: This is a part of artificial neural networks commonly used in text mining. Compared to other text classification algorithms, CNNs require relatively little pre-processing. Unlike traditional algorithms, the network learns to optimize the filters (or kernels) through automated learning, whereas, in traditional algorithms, these filters are designed by hand. This approach is advantageous because it is independent of prior knowledge and human intervention when obtaining features.
- Deep Neural Networks: In DNNs, data flows from the input layer to the output layer without looping back. A DNN begins by creating a map of virtual neurons and assigning random numerical values to connections between them. A combination of weights and inputs results in an output between 0 and 1. An algorithm would adjust the weights if the network failed to recognize a particular pattern accurately. As a result, the algorithm can gradually increase the importance of certain parameters until it finds the correct mathematical manipulation to process the data fully.

After finding the SCM between headlines and body text, we add the stance column with four labels (agree, disagree, discuss, and unrelated), which will represent our dependent variable. In addition, we enter these columns into CNN learning models with some properties, as shown in Table 1.

Table 1. Properties of CNN learning model.

CNN layer	Dense	Activation
Input layer	1024	relu
Hidden layer	128	relu
Output layer	4	SoftMax

On the training CSV, with the batch size being 512 and the number of epochs being 80, the best result with a categorial accuracy (average accuracy of stances) of 0.997 and loss function of 0.659 because the first ten epochs were over-fitting. We compute the confusion matrix as shown in Fig. 6.



Fig. 6. The confusion matrix of CNN algorithm.

That is challenging on the competition's CSV, which has real-time news and does not exist in training or testing CSV.

After we learn the training CSV, we do the same Learning on the competition CSV. The best result was validation-categorical accuracy (average accuracy of stances) of 0.805 and validation-loss function of 0.938 because the first ten epochs were over-fitting. Furthermore, we compute the confusion matrix as shown in Fig.7. Finally, we loaded the weighted CNN to the DNN model to enhance our results, which increased the accuracy to 84.6%.

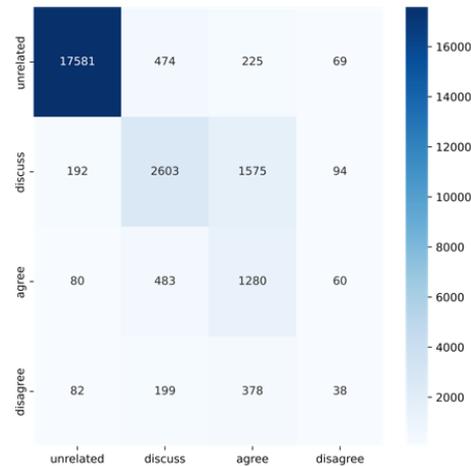


Fig. 7. The confusion matrix of DNN algorithm.

We used this metric for evaluating classification models that specifically measures the similarity between predicted and actual fake news. Furthermore, it is possible to calculate using an equation. The number of documents detected by the system is denoted by True Positive (TP), while the number of correctly predicted fake documents is represented by True Negative (TN). Furthermore, False Positive (FP) represents the number of real documents classified as fake by the system. Finally, False Negative (FN) refers to the number of fake documents classified as real by the system.

The F-score is an important metric that reflects the success of detecting fake news rather than real news. Precision represents the percentage of correctly classified observations as a fake document. Sensitivity, in contrast, is defined as the ratio of real news classified to actual real news [15]:

- Precision = $\frac{|TP|}{|TP| + |FP|}$ (1)

- Recall = $\frac{|TP|}{|TP| + |FN|}$ (2)

- F1-Score = $\frac{2 * precision * Recall}{precision + Recall}$ (3)

- Accuracy = $\frac{|TP| + |TN|}{|TP| + |TN| + |FN| + |FP|}$ (4)

These metrics are widely used in the ML and DL community and allow to evaluate a classifier's performance from various angles.

4. Results and Discussion

System implementation was performed on a Macintosh platform with a dual core Intel Core i5 processor and 16 GB of memory, illustrated in Table 2.

Both Microsoft Visual Studio 2019 and MATLAB R2020b were used. As shown in Table 3, pre-processing takes 12 seconds on average, converting the data into a TF-IDF vectorizer takes 16 minutes, and calculating SCM between headlines and body text takes 24 minutes. The test and the training of CNN take 5.55 minutes, respectively. On DNN, the system can be applied in real-time in 2.034 seconds, demonstrating its ability to be applied in real-time.

The proposed system is evaluated via two classifiers based on the features obtained using Elbow truncated algorithms. CNN and DNN are these classifiers. Based on the CNN classifier, which is most commonly used to analyse or understand natural language processing, the two classifiers perform well in detecting fake news.

Table 2 demonstrates the combination of CNN and DNN output of learning algorithms. Furthermore, we calculate the results of each stance in the dataset, as shown in Table 4.

Table 2. Summary of the experiment results.

Classifier	Accuracy	Precision	Recall
Proposed Model(CNN, DNN)	84.6%	84.6%	84.6%

Table 3. Summary of the execution time.

Model	Execution time
Pre-processing	12 s
TF-IDF	16 min
SCM	24 min
CNN	5.55 min
DNN	2
Totally	45.69 min

Table 4. Comparison between the stances.

Stance	Accuracy	Precision	Recall	F-score
Unrelated	95.04	98.03	95.81	96.91
Discuss	87.70	69.25	58.31	63.31
Agree	88.47	37.02	67.26	47.75
Disagree	96.00	5.45	14.56	7.93

Across all stances except disagree, we achieved higher scores and class-wise accuracy, as shown in Fig. 8. It is a problem since the importance of disagree is similar to the importance of agree and discuss. Not only are there a relatively small number of news pairs in the disagree category, but they also contain divergent news articles. Deep models, including top teams, do not perform well when classifying disagree labels.

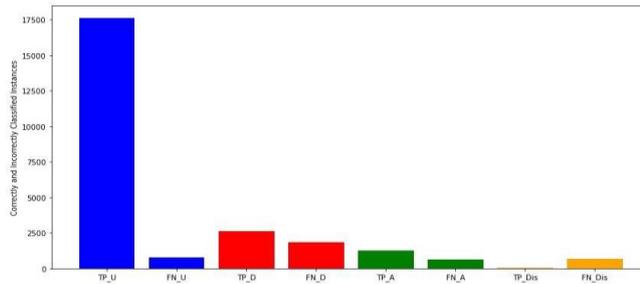


Fig. 8. The results of the proposed system.

Moreover, we compared the proposed system's performance to previous works based on metrics of accuracy and methods used. Table (5) shows that our proposed approach outperforms most previous approaches. For example, Riedel et al. [9] used the ULCMR method to achieve 81.72% accuracy, while [10] used statistical, neural, and feature engineering heuristics to achieve 83.08% accuracy. In another example, the authors in [1] used Novel memory network enhancements with CNN and LSTM to achieve 81.23% accuracy. Although in [13], the researcher achieved 90% accuracy. On the other hand, the proposed system achieved 84.6% accuracy and needs to combine CNN and DNN with SCM. CNN and DNN are used to demonstrate the ability of the system to detect fake news, as shown in Fig. 9.

Table 5. Comparison between the proposed approach with previous works.

Reference	Year	Technique	Accuracy
[9]	2017	UCL Machine Reading's (UCLMR) system.	0.8172
[10]	2017	Statistical, neural and feature engineering heuristics.	0.8308
[11]	2018	Novel memory network enhances with CNN and LSTM.	0.8123
[16]	2018	Rich stacked LSTM approach.	0.8020
[12]	2018	Combining deep neural network approaches with string similarity features.	0.8223
[17]	2020	Based on the core concept of Credibility Reviews (CRs).	0.7160
[13]	2021	DL (SoftMax activation) with SCM and RoBERTa.	0.9073
Proposed model	2022	Combine CNN and DNN with SCM.	0.8460

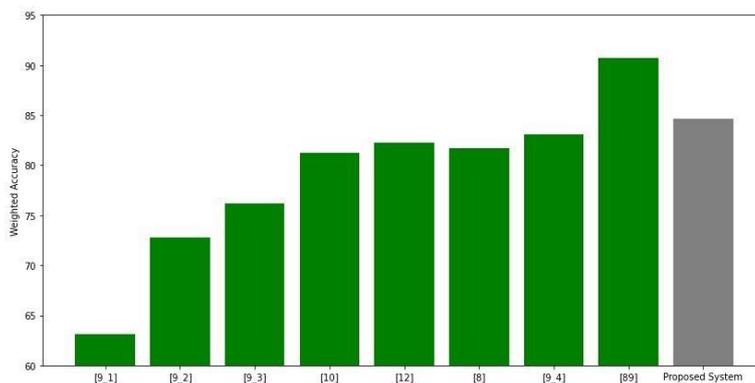


Fig. 9. Comparison between the proposed approach with previous works.

Figure 10 illustrates the purpose system with the other previous works in state of the art, which is higher than most of this model.

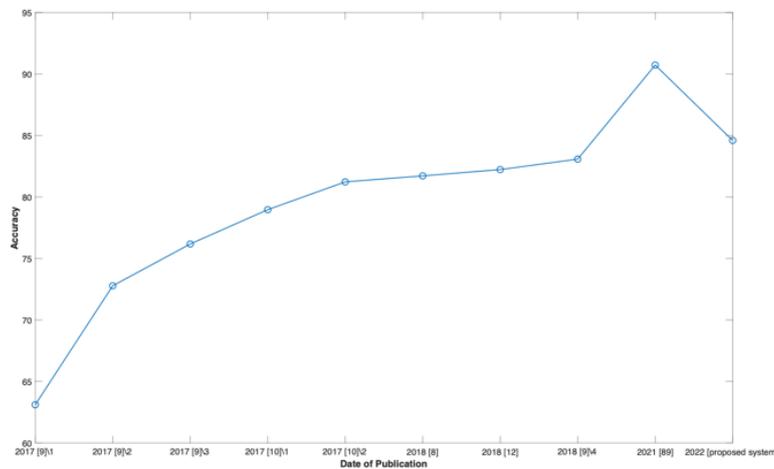


Fig. 10. Comparison between the proposed approach with state the art.

5. Conclusions and Future Work

Due to social media's increasing popularity, more and more people are consuming news via social media instead of traditional media. In addition, social media has been used to spread fake news, which can negatively impact individual users and society. Researchers attempted different approaches to build fake news detection systems using machine learning and deep learning approaches with different natural language processing techniques.

This work focused on the competitive studies that used the FNC-1 dataset. From studying the existing works, it is concluded that the proposed approaches fail to build a system that can detect fake news efficiently.

The current work proposed a new system for detecting fake news efficiently. The procedure of the proposed system implies cleaning the data, extracting new features using different natural language processing techniques, and reducing the number of features using the elbow truncated method. Furthermore, it includes finding the similarity between each pair of headlines and body text, and the resulting feature is entered into CNN and DNN learning approaches.

As a result, the suggested system can recognize if the headline agrees with the body text, discuss it, disagree with it, or they are unrelated. The systems give up to 84.6% accuracy, where it detects all these cases with high accuracy except the disagree case, and the main reason for that is that the training pairs of this category are not frequent.

As a result, the proposed system can be considered the second-ranking among the competitive studies with the FNC-1 dataset, where it outperforms most of the previous works in terms of accuracy. The accuracy can be further increased by enhancing the system's performance in detecting disagreeing records. This can be

done by initiating a dictionary that contains the words that discriminate the disagreeing pairs.

Furthermore, different summarizing methods can be used for the text in the headline and the body text columns before any pre-processing, and this step can significantly improve the accuracy.

Finally, the Glove method can be used instead of the TF-IDF method to generate the new features through encoding and vectorizing the text, as this method is considered more efficient.

References

1. Vishwakarma, D.K.; and Jain, C. (2020). Recent state-of-the-art of fake news detection: A review. *Proceedings of the 2020 International Conference for Emerging Technology (INCET)*, Belgaum, India, 1-6.
2. Zhang, X.; and Ghorbani, A.A. (2020). An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, 57(2), 102025.
3. Gottfried, J.; and Shearer, E. (2016). *News use across social media platforms 2016*. Pew Research Center.
4. Mihalcea, R.; and Strapparava, C. (2005). Making computers laugh: Investigations in automatic humor recognition. *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*. Vancouver, British Columbia, Canada, 531-538.
5. Wallace, B.C. (2015). Computational irony: A survey and new perspectives. *Artificial Intelligence Review*, 43(4), 467-483.
6. Pérez-Rosas, V.; Kleinberg, B.; Lefevre, A.; Mihalcea, R. (2017). Automatic detection of fake news. arXiv preprint arXiv:1708.07104, 2017.
7. Demilie, W.B.; and Salau, A.O. (2022). Detection of fake news and hate speech for Ethiopian languages: a systematic review of the approaches. *Journal of Big Data*, 9:66, 1-17.
8. Zhou, X. and R. Zafarani, *A survey of fake news: Fundamental theories, detection methods, and opportunities*. ACM Computing Surveys (CSUR), 2020. 53(5): p. 1-40.
9. Riedel, B.; Augenstein, I.; Spithourakis, G.P.; and Riedel, S. (2017). A simple but tough-to-beat baseline for the fake news challenge stance detection task. arXiv preprint arXiv:1707.03264.
10. Bhatt, G.; Sharma, A.; Sharma, S.; Nagpal, A.; Raman, B.; and Mittal, A. (2017). On the benefit of combining neural, statistical and external features for fake news identification. arXiv preprint arXiv:1712.03935.
11. Mohtarami, M.; Baly, R.; Glass, J.; Nakov, P.; Marquez, L.; and Moschitti, A. (2018). Automatic stance detection using end-to-end memory networks. arXiv preprint arXiv:1804.07581.
12. Borges, L.; Martins, B.; and Calado, P. (2019). Combining similarity features and deep representation learning for stance detection in the context of checking fake news. *Journal of Data and Information Quality (JDIQ)*, 1(3), 1-26.

13. Sepúlveda-Torres, R.; Vicente, M.; Saquete, E.; Lloret, E.; and Palomar, M.S. (2021). Exploring summarization to enhance headline stance detection. *Proceedings of the 26th International Conference on Applications of Natural Language to Information Systems, NLDB 2021*, Saarbrücken, Germany, 43-254.
14. Al-Asadi, T.; Obaid, A.J.; Alkhayat, A.A. (2017). Proposed method for web pages clustering using latent semantic analysis. *Journal of Engineering and Applied Science*, 12(8), 8270-8277.
15. Shu, K., Sliva, A.; Wang, S.; Tang, J.; and Liu, H. (2017). Fake news detection on social media: A data mining perspective. arXiv preprint, arXiv:1708.01967.
16. Hanselowski, A.; Avinesh, P.V.S.; Schiller, B.; Caspelherr, F.; Chaudhuri, D.; Meyer, C.M.; and Gurevych, I. (2018). A retrospective analysis of the fake news challenge stance detection task. arXiv preprint arXiv:1806.05180.
17. Denaux, R.; and Gomez-Perez, J.M. (2020). Linked credibility reviews for explainable misinformation detection. *The Semantic Web – ISWC 2020: 19th International Semantic Web Conference*, Athens, Greece, 147-163.