Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 483−489
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1956

Organization Entity Extraction Telkom University Affiliated using Recurrent Neural Network (RNN)

Muhammad Daffa Regenta Sutrisno[1,*], Donni Richasdy[2], Aditya Firman Ihsan[3]

School of Computing, Informatics Study Program, Telkom Univeristy, Bandung, Indonesia
Email: [1,*]daffaregenta@student.telkomuniversity.ac.id , [2]donnir@telkomuniversity.ac.id, [3]adityaihsan@telkomuniversity.ac.id
Email Penulis Korespondensi: daffaregenta@student.telkomuniversity.ac.id
Submitted:**25/07/2022**; Accepted:**18/08/2022**; Published**: 30/09/2022**

Abstract−In the news portal text, there is a lot of important information such as the name of the person, the name of the organization, or the name of the place. To obtain information in text documents manually, humans must read the contents of the entire news text. To overcome this issue, information extraction, commonly called Named Entity Recognition (NER) was used. The extraction of information expressly for the NER field is used to make it easier to process word or sentence data. It helps search engines and also helps to classify places, times, and organizations. There is a limited number of NER in Indonesian texts using only the Recurrent Neural Network (RNN) method. Similar previous studies only employed other versions of RNN such as LSTM (Long Short Term Memory), BiLSTM (Bidirectional Long Short Term Memory), and CNN (Convolutional Neural Network). NER is one of the answers to the problems that exist in a large number of news portal texts to obtain information effectively and efficiently. The results of this study indicate that the NER system using the RNN method in Indonesian news texts has an F1 -Score of 80%.

Keywords: Entity Recognition; NER; Recurrent Neural Network; RNN; Natural Language Processing; NLP

1. INTRODUCTION

In the news text, there is a lot of important information such as the name of the person, the name of the organization, or the name of the place. To obtain information in text documents manually, humans must read the contents of the entire news text. If the text document is very long, it takes a long time for humans to obtain the information contained in the text. One way to simplify the case is to use NER.

The Text Mining and Natural Language Processing procedure includes Named Entity Recognition (NER), which is highly helpful in the information extraction process. NER was first proposed in 1995 at the MUC-6 Conference [1]. Finding information from a document or natural language as input from the outcomes of valuable information in the form of structured information following a particular format is known as information extraction. Identification and classification of names in the text into predetermined classes is the primary function of NER [2]. The extraction of information expressly for the NER field is used to make it easier to process word or sentence data. It helps search engines and also helps to classify places, times, and organizations. NER is essential in NLP and is the most fundamental stage in extracting information. in its application, there are three extraction methods NER based on rules and templates, NER based on machine learning, NER based on deep learning[3].

Based on this description, this research implemented the Recurrent Neural Network (RNN) method in the case of Extraction of Organizational Entities Affiliated to Telkom University. The RNN method was chosen because similar studies used other versions of RNN such as LSTM (Long Short Term Memory), BiLSTM (Bidirectional Long Short Term Memory), CNN (Convolutional Neural Network), and others. Therefore, in this study, only the RNN method was used on Indonesian-language texts to see the performance and accuracy of the use of the RNN method on Indonesian-language text organizations and to evaluate performance using NER with the RNN method on Indonesian news portals. The extraction process takes place using NER and using methods or models from RNN. RNN is also called a feedback network, a network on a neural network where there is a loop as a feedback connection in the network. RNN network is a network that accommodates network output to be input to the network which is then used to generate new output [4]. Recently, RNN has demonstrated great success in various Natural Language Processing tasks such as speech recognition, machine translation, and language modeling[5]. The impact of this study is due to the lack of other studies that use Indonesian-language datasets using the RNN method to make the results of this study as an illustration for further research

In this study, the limitation of the study is that the dataset used was a collection of Indonesian news that discusses Telkom University. The dataset used was 20,061 words, on an Indonesian language news portal. The labeling process in Indonesian sentences was done manually. Aspect labels were only divided into 4 (four) categories, namely LOC (location), ORG (organization), PER (person), TIM (time), O (other), and the entity taken focuses on the ORG (organization) label. The performance of the RNN model was measured using a Confusion matrix such as F1-Score, recall, precision, and accuracy.

A previous study related to NER in Indonesian was obtained from Wikipedia using a method that focuses on Bidirectional LSTM-CNNs[6]. using the label structure "BILOU" which means Beginning, Inside, and the Last token from a multi-token which means Outside and Unit for a single token. By using a corpus of 4139 sentences that had been labeled, the accuracy rate on the BiLSTM model obtained an F-Score of 76.47%. Then in the BiLSTM-CNNs model, the F-Score was 79.43%. meanwhile, an F-Score of 77.47% in the BiLSTM-LSTM model and an F-Score of

BITS is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 483−489
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1956

77.47% in the BiLSTM-CNNs-LSTM model. Hence, the best results were obtained using the Bidirectional LSTM-CNNs model.

Another study on NER by Ramadhyni [2], using the Multinomial Naïve Bayes Classifier on Indonesian-language tweets with BIO notation which means Beginning, Inside, Outside, obtained PER of 74%, ORG of 71%, LOC of 62%.

Then, a study by Andrej Zukov[7] using the CoNLL 2003 dataset using the parallel RNN method compared to the single component LSTM model concluded that they prefer to use smaller LSTM Units, obtaining an accuracy rate of 91.48%.

2. RESEARCH METODOLOGY

2.1 Research Stages

The following is the architecture of the Named Entity Recognition system using the Recurrent Neural Network method, which can be seen in Fig 1.

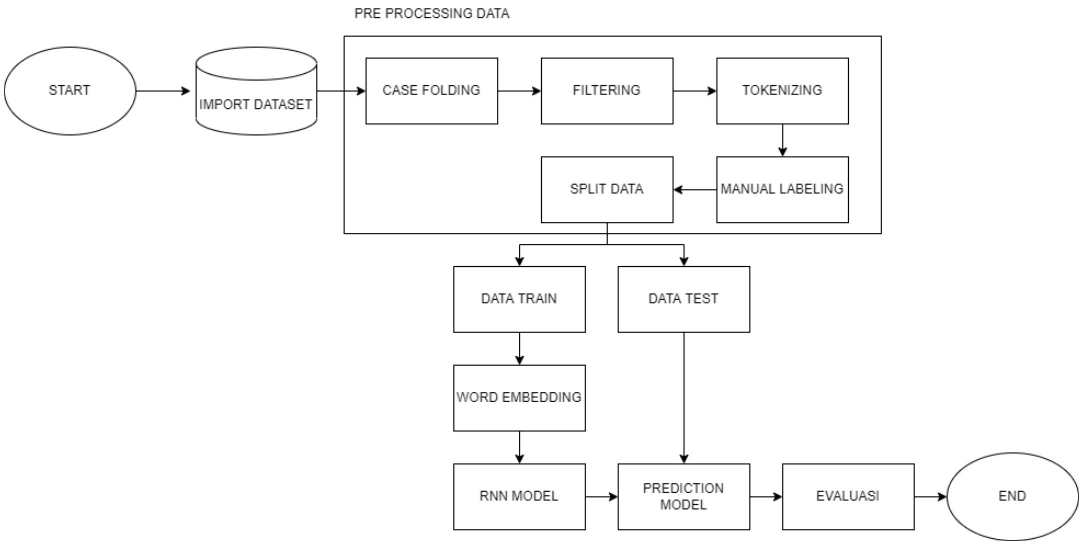

Figure 1. System Architecture

2.2 Dataset

The dataset used in this study was a collection of news portal texts related to Telkom University in Indonesian. The dataset was obtained by using a scrapy library with related news portal sources. In the obtained data of 424 articles, pre-processing was done by manually labeling. Therefore, the dataset used in this study was 51 articles, 20,061 words, and 1286 sentences. The following are some limitations in performing manual labeling on datasets.

a. The first word in a sentence can only be labeled using the label ‘O’, or variations of the labels ‘B-PER’, ‘B-LOC’, ‘B-TIM’, ‘B-ORG’.
b. Labeling words that are not at the beginning of a sentence or words that are at the end of a sentence are labeled based on the word before it.
c. If the previous word is labeled 'O' then the next word can be labeled ‘O’ or variations of ‘B-PER’, ‘B-LOC’, ‘B-ORG’, ‘B-TIM’.
d. If the previous word is labeled ‘B-PER’, ‘B-LOC’, ‘B-ORG’, ‘B-TIM’, then the next word can be labeled ‘B-PER’, ‘B-LOC’, ‘B-ORG’, ‘B-TIM’ with different aspects, labeled ‘I-PER’, ‘I-LOC’, ‘I-ORG’, ‘I-TIM’ with the same aspect, or labeled ‘O’.
e. If the previous word is labeled ‘I-PER’, ‘I-LOC’, ‘I-ORG’, ‘I-TIM’ then the next word can be labeled by the labels ‘B-PER’, ‘B-LOC’, ‘B-ORG’, ‘B-TIM’ with different aspects. And the labels ‘I-PER’, ‘I-LOC’, ‘I-ORG’, and 'I-TIM’ with the same aspect as the label ‘O’.

The labeling process for Indonesian sentences was done manually, with the distribution of labels in Table 1 below:

Table 1. Number of Datasets

Label	Total
O	15170
B-LOC	1384
B-ORG	1099
I-LOC	977

BITS is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 483−489
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1956

I-ORG	692
B-TIM	370
B-PER	199
I-PER	152
I-TIM	18

2.3 Pre-Processing Data

Pre-Processing data is a process to produce words that can improve system performance[2] Pre-Processing is also the first stage in carrying out tasks in all fields of NLP, including Named Entity Recognition. This stage aims to prepare the data to be ready to be processed and improve performance in the Named Entity Recognition case. The following were the steps for Pre-Processing data, including:

a. Case Folding is the stage to change the uppercase character to lowercase. Case Folding is only useful for alphabetic characters.
b. Filtering is the stage to clean sentences from symbols, punctuation marks, and others that are non-alphabetic to facilitate data processing.
c. Tokenizing is to separate sentences word for word in the dataset.

2.4 Labeling

At this stage, entity labeling was carried out manually which served to train the system in predicting entities in its results. At this stage, IOB Notation marked each token with one of the other three tags. The three tags were I (Inside), O (Other), and B (Beginning) [2]. Labeling is carried out in 4 (four) categories in Table 2, namely:

Table 2. Label Category

Label Kategori	Label
Organization	B-ORG, I-ORG
Person	B-PER, I-PER
Time	B-TIM, I-TIM
Location	B-LOC, I-LOC
Other	O

2.5 Word Embedding

Word Embedding is also known as distributed representation, has recently been proposed to address some of the problems of NLP and has achieved great success. A word expressed by a distributed representation is a dense, low-dimensional, and real-valued vector. Vector words can capture many linguistic regularities, which are superior to one-hot-encode representations. For example, the biomedical terms 'gene', 'protein', and 'kinase' are close to each other, while far from another type of 'cat' noun, 'dog'. For one-hot encoding, all words are equally distant [8].

2.6 Recurrent Neural Network

Recurrent Neural Network (RNN) model is part of Deep Learning machine learning. Repetitive Neural Networks or RNNs are attractive options for sequence labeling they are flexible in the use of context information. Since they can learn what to store and what to ignore, they accept many different types and representations of data and they can recognize sequential patterns in the presence of distortion sequential[8]. The RNN model consists of an input layer ***x*** representing features at time ***t***, a hidden layer ***h***, and an output layer ***y*** representing the probability distribution of labels at time ***t*** [8]. The type of RNN in this study uses Many to Many which can accept many inputs with several output options that are determined by sequence. The following in Figure 2 is an overview of how RNN works.

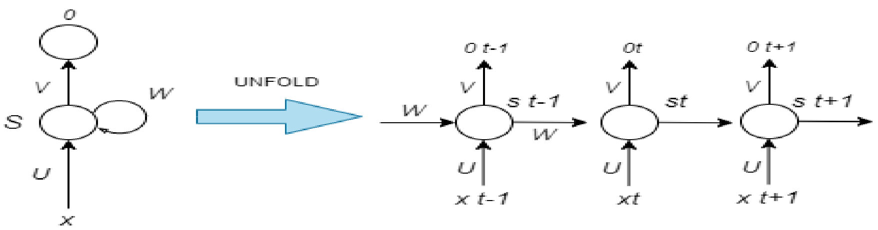

Figure 2. RNN Architecture

In the architectural drawing above in Figure 2, the image on the left is a circuit diagram, where the black box shows the time delay of one-time step. The diagram shows the RNN in the unrolled position to the full network. Meanwhile, the picture on the right shows the RNN that has been opened (unrolled/unfolded) into a full network. Therefore, the sequence becomes complete. For example, if the sequence obtained is 1 sentence with 5 words, then the network will be opened in a 5-Layer neural network, one layer for each word. The following is a description of the formula language from the image above.

Copyright © 2022 **Muhammad Daffa Regenta Sutrisno**, Page 485
BITS is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 483−489
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1956

a. $\boldsymbol{x_t}$ is the input in the time step. For example, $\boldsymbol{x_1}$ can be a one-hot vector corresponding to the second word of a sentence being processed.
b. $\boldsymbol{S_t}$ is the hidden state at each time step **t**. Hidden state can be referred to as the "memory" of a network that functions to store the results of calculations and records that have been carried out. **St** is calculated based on the previous hidden state and based on the input in the current state.

$$\boldsymbol{S_t = F(Ux_{t-1} + Ws_{t-1})} \quad (1)$$

c. The **F** function is usually non-linear like **tanh** or **ReLU**. $\boldsymbol{S_{t-1}}$is used to calculate the first hidden state, usually, initialization always starts with 0 (zero).
d. $\boldsymbol{0_t}$ is the output at step **t**. For example, if the "next word" is expected in a sentence, then 0_t is a probability vector across the vocabulary in our database.

$$\boldsymbol{0_t = softmax\ (Vs_t)} \quad (2)$$

2.7 Evaluation

To evaluate the performance of the RNN model on NER, the following evaluation metrics are used. Precision is the number of positive samples classified correctly divided by the total number of positive samples. In percent (%). Precision value is obtained by the following equation. [2]

$$P = \frac{TP}{TP+FP} \quad (3)$$

Accuracy is an evaluation parameter to measure the accuracy of a built classification system. In percent (%). Accuracy value is obtained by the following equation.[2]

$$A = \frac{TP+FN}{TP+FP+FN+TN} \quad (4)$$

Recall is the ratio of the total number of positive classifications divided by the total number of positives. In percent (%). The recall value is obtained by the following equation.[2]

$$R = \frac{TP}{TP+FN} \quad (6)$$

F1-Measure is the average harmonic value of Precision and Recall to get a balanced Precission and Recall value. In percent (%). The F1-Measure value is obtained by the following equation.[2]

$$F1 - Measure = \frac{2 \times Precission \times Recall}{Precission+Recall} \quad (7)$$

3. RESULT AND DISCUSSSION

To achieve the objectives of this study, 4 scenarios were carried out, by adjusting the data split ratio, the number of RNN layers, the number of word embedding dimensions, and the number of batch sizes in the model to achieve maximum accuracy and performance results from the model, using the structure of the default RNN model as follows in Table 3:

Table 3. Model RNN Parameter Default

Max_length	Embedding Dimension	Dense	Units Cell	*RNN Layer*	Batch Size	Epoch
213	213	9	213	1	30	50

3.1 First Scenario (Split Data Ratio)

In the first scenario, a comparison was made by changing the split data ratio in the dataset by dividing it into 3 ratios as follows in Table 4 and Table 5:

Table 4. Result Data Split Ratio B-ORG

B-ORG				
Training Data	**Test Data**	**Precision (%)**	**Recall (%)**	***F1-Score* (%)**
90%	10%	77	51	62
80%	20%	77	60	67
70%	30%	77	57	66

Table 5. Result Data Split Ratio I-ORG

I-ORG				
Training Data	**Test Data**	**Precision (%)**	**Recall (%)**	***F1-Score* (%)**
90%	10%	79	65	71

Copyright © 2022 **Muhammad Daffa Regenta Sutrisno**, Page 486
BITS is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 483−489
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1956

80%	20%	75	86	80
70%	30%	69	67	68

In the Table 4 and Table 5 above, it can be seen that the difference in the split data ratio can affect the results of the test. The test uses the F1-Score to measure the performance of the model. In the B-ORG entity, the highest result was obtained at a ratio of 80:20 obtained with a mark of 67%, but the results obtained were not satisfactory because the training data for the B-ORG entity was relatively small so the machine could not train the data properly. Reasonably good results were obtained on the I-ORG entity with the highest yield at a ratio of 80: 20 with an output of 80%. This is because I-ORG entities have sufficient data to conduct training on the machine to provide the best results. low results were obtained for the I-ORG entity at a ratio of 70:30 with a mark of 68% due to insufficient training data because there were relatively few for the dataset used.

3.2 Second Scenario (RNN Layer)

In the first scenario, in the previous split data, a good ratio was 80:20. Meanwhile, in the second scenario, a data split ratio of 80:20 was used to test the number of RNN layers used. Can be seen in the Table 6 and Table 7.

Table 6. B-ORG Result RNN Layer

B-ORG			
RNN Layer	**Precision (%)**	**Recall (%)**	***F1-Score* (%)**
1	77	60	67
2	83	50	63
3	72	54	62

Table 7. I-ORG Result RNN Layer

I-ORG			
RNN Layer	**Precision (%)**	**Recall (%)**	***F1-Score* (%)**
1	75	86	80
2	95	32	48
3	84	39	53

In Tables 6 and Table 7 With the addition of layers to the model, the results obtained were reduced for B-ORG entities with RNN layer 1 with the F1-Score of 67%. Then, the addition of RNN layer 2 results in a decrease of 63%, then the addition of RNN layer 3 gets the same results, decreased to 62%. Furthermore, in the I-ORG entity, the F1-Score result for RNN layer 1 is 80%, then at RNN layer 2 the results decreased quite drastically to 48%, then at RNN Layer 3 the results decreased to 53%. This is because adding layers to the model causes the model to find more complicated patterns in the given dataset and is at risk of overfitting. Hence, the deeper the layer the model will learn more complicated patterns and affect the results of prediction accuracy.

3.3 Third Scenario *(Embedding Dimension)*

Embedding Dimension is adding dimensions to word embedding to classify words into similar words, and be able to understand the context of a word. Accordingly, similar words have detailed word embedding. Because in the previous test, the best number of RNN Layers was 1. Accordingly, the number of RNN Layer 1 with a change in the number of Embedding Dimensions was used to test this scenario. The followings are the results of the tests carried out in Table 8 and Table 9.

Table 8. B-ORG Dimension Embedding Result

B-ORG			
Embedding Dimension	**Precision (%)**	**Recall (%)**	***F1-Score* (%)**
100	85	57	68
213	77	60	67
300	77	61	68
400	81	66	73

Table 9. I-ORG Dimension Embedding Result

I-ORG			
Embedding Dimension	**Precision (%)**	**Recall (%)**	***F1-Score* (%)**
100	81	58	68
213	75	86	80
300	81	72	76
400	85	65	73

BITS is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 483−489
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1956

In Table 8 and Table 9 above test results obtained by reducing the number of Embedding Dimensions affect the performance of the RNN. By reducing the number of embedding dimensions to 100, the classification of words with dimensions is limited. Therefore, it affects the results obtained, causing the engine places words that do not match the number of dimensions that are reduced, and the F1-Score results are very low for entities I-ORG at 68% respectively. The results are quite good for the I-ORG entities for the number 213 on the embedding dimension, 80%, respectively because the machine can adjust similar words in the dimensions that match the sentence length in table 3, namely 213. Furthermore, on the number of embedding dimensions 300 and 400 for the B-ORG entity, the results decreased again by 68% and 73% because the amount of performance used by the machine was getting bigger. Therefore, the results decreased again. In the I-ORG entity where the results are increased on the number of embedding dimensions of 300 and 400 with both 76% and 73%.

3.4 Fourth Scenario (Batch Size)

Because in the third scenario (Embedding Dimension) the best results were obtained with the number of embedding dimensions of 213. The same number of embedding dimensions was used in this fourth scenario by adjusting the number of batch sizes. can be seen in the Table 10 and Table 11.

Table 10. B-ORG Batch Size Result

B-ORG			
Batch Size	**Precision (%)**	**Recall (%)**	***F1-Score* (%)**
30	77	60	67
50	82	59	69
100	84	58	68

Table 11. I-ORG Batch Size Result

I-ORG			
Batch Size	**Precision (%)**	**Recall (%)**	***F1-Score* (%)**
30	75	86	80
50	75	73	74
100	84	62	71

In the change in batch size, it can be seen that in Table 10 and Table 11 the batch size affects the performance of the machine itself. From the results obtained for the number of batch sizes 30 and 50 for B-ORG, the F1-Score is 67% and 69%, but for the number of 100, the F1-Score decreases to 68%. For I-ORG, the F1-Score results for the number of 30 is 80%, the number of 50 is 74%, the number of 100 is 71%. For the best results in batch size 30, for B-ORG and I-ORG, the results were 67% and 80%, respectively. Setting the batch size depends on the training data used, in this test 80% of the training data is used or equal to 1014 sentences and the words and tags. Therefore, if the larger batch size used in the data train affects the machine when processing, which is the number of data samples that will be propagated through the neural network. The number of batch size 100 experienced a decrease in accuracy because the number of batch sizes with the number of train data was not enough to conduct training on the data which made the training short.

4. CONCLUSION

The problem found in this study is that the labeling of words or entities is still done manually. Thus, it becomes difficult to give the right entity naming to match the actual meaning of the word. And the Recurrent Neural Network method used is still minimally used for Indonesian-language texts. Hence, there is a lack of references for the method used. Tests have been carried out with 4 scenarios, namely by setting the split data ratio, the number of RNN layers, the number of word embedding dimensions, and the number of batch sizes in the model. From the above test, it can be concluded that with a data split ratio of 80:20, the number of RNN Layer 1, the number of Embedding Dimensions 213, and the batch size of 30 are the best results with F1-Score B-ORG 67% and I-ORG 80%. Moreover, the low F1-Score results obtained at the split data ratio with 70:30 for B-ORG and I-ORG entities are 66% and 68%, the number of RNN Layer 3 obtained the lowest results are F1-Score 62% and 53%, the number of Embedding dimension 100, the lowest results are F1-Score 65% and 51%, and batch size 100 with the lowest F1-Score results in B-ORG and I-ORG entities are 68% and 68%, respectively. It can be seen that the addition of the RNN layer greatly affects the machine in processing. Thus, the results obtained are greatly reduced from the addition of the RNN layer because the machine learns deeper and more complex patterns. Besides tuning the parameters on the model, the accuracy and performance of the model are also affected by the number of datasets and the labeling of entities in the dataset, because there are words that have more than one entity. Accordingly, the machine makes errors in predicting entities. More datasets and more entity categories are expected for further research. Moreover, adding model parameters is suggested to get better results.

Copyright © 2022 **Muhammad Daffa Regenta Sutrisno**, Page 488
BITS is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 483−489
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1956

REFERENCES

[1] Q. Guo, S. Wang, and F. Wan, "Research on named entity recognition for information extraction," *Proc. - 2020 2nd Int. Conf. Artif. Intell. Adv. Manuf. AIAM 2020*, pp. 121–124, 2020, doi: 10.1109/AIAM50918.2020.00030.

[2] R. Rifani, M. A. Bijaksana, and I. Asror, "Named Entity Recognition for an Indonesian Based Language Tweet using Multinomial Naive Bayes Classifier," *Indones. J. Comput.*, vol. 4, no. 2, pp. 119–126, 2019, doi: 10.21108/indojc.2019.4.2.330.

[3] X. Liu, H. Chen, and W. Xia, "Overview of Named Entity Recognition," *J. Contemp. Educ. Res.*, vol. 6, no. 5, pp. 65–68, 2022, doi: 10.26689/jcer.v6i5.3958.

[4] D. Tarkus, S. R. U. A. Sompie, and A. Jacobus, "Implementasi Metode Recurrent Neural Network pada Pengklasifikasian Kualitas Telur Puyuh," *J. Tek. Inform.*, vol. 15, no. 2, pp. 137–144, 2020.

[5] J. P. C. Chiu and E. Nichols, "Named Entity Recognition with Bidirectional LSTM-CNNs," *Trans. Assoc. Comput. Linguist.*, vol. 4, no. 2003, pp. 357–370, 2016, doi: 10.1162/tacl_a_00104.

[6] W. Gunawan, D. Suhartono, F. Purnomo, and A. Ongko, "Named-Entity Recognition for Indonesian Language using Bidirectional LSTM-CNNs," *Procedia Comput. Sci.*, vol. 135, pp. 425–432, 2018, doi: 10.1016/j.procs.2018.08.193.

[7] A. Žukov-Gregorič, Y. Bachrach, and S. Coope, "Named entity recognition with parallel recurrent neural networks," *ACL 2018 - 56th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. (Long Pap.*, vol. 2, pp. 69–74, 2018, doi: 10.18653/v1/p18-2012.

[8] D. Song, L. Li, L. Jin, and D. Huang, "Biomedical named entity recognition based on recurrent neural networks with different extended methods," *Int. J. Data Min. Bioinform.*, vol. 16, no. 1, pp. 17–31, 2016, doi: 10.1504/IJDMB.2016.079799.

[9] H. Ceovic, A. S. Kurdija, G. Delac, and M. Silic, "Named Entity Recognition for Addresses: An Empirical Study," *IEEE Access*, vol. 10, pp. 42094–42106, 2022, doi: 10.1109/ACCESS.2022.3167418.

[10] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, and C. Dyer, "Neural architectures for named entity recognition," *2016 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. NAACL HLT 2016 - Proc. Conf.*, pp. 260–270, 2016, doi: 10.18653/v1/n16-1030.

[11] A. Vlachos, "Evaluating and combining biomedical named entity recognition systems," *ACL 2007 - Proc. Work. BioNLP 2007 Biol. Transl. Clin. Lang. Process.*, pp. 199–206, 2007, doi: 10.3115/1572392.1572430.

[12] K. Sintoris and K. Vergidis, "Extracting business process models using natural language processing (NLP) techniques," *Proc. - 2017 IEEE 19th Conf. Bus. Informatics, CBI 2017*, vol. 1, pp. 135–139, 2017, doi: 10.1109/CBI.2017.41.

[13] H. S. Al-Ash and W. C. Wibowo, "Fake news identification characteristics using named entity recognition and phrase detection," *Proc. 2018 10th Int. Conf. Inf. Technol. Electr. Eng. Smart Technol. Better Soc. ICITEE 2018*, pp. 12–17, 2018, doi: 10.1109/ICITEED.2018.8534898.

[14] N. M. Rezk, M. Purnaprajna, T. Nordstrom, and Z. Ul-Abdin, "Recurrent Neural Networks: An Embedded Computing Perspective," *IEEE Access*, vol. 8, pp. 57967–57996, 2020, doi: 10.1109/ACCESS.2020.2982416.

[15] V. Yadav and S. Bethard, "A survey on recent advances in named entity recognition from deep learning models," *COLING 2018 - 27th Int. Conf. Comput. Linguist. Proc.*, pp. 2145–2158, 2018.

Copyright © 2022 **Muhammad Daffa Regenta Sutrisno**, Page 489
BITS is licensed under a Creative Commons Attribution 4.0 International License