International Journal of Multidisciplinary Research and Publications
ISSN (Online): 2581-6187

Implementation of Conditional Random Field for Named Entity Recognition in Indonesian Traditional Arts Digital Article

Cyntia Fanny, Alexander Waworuntu*, Julio Christian Young
Department of Informatics, Universitas Multimedia Nusantara, Tangerang, Banten, Indonesia
*Corresponding author, Email address: alex.wawo@umn.ac.id

Abstract— *Digital articles are one type of media used as a reference when looking for information, one of which is information related to traditional arts in Indonesia. In traditional art articles, there are various artistic entities, such as dance, performing arts, music, artistic figures, and musical instruments. There are so many types of art in Indonesia, especially on the island of Java, making it more difficult to identify the entities in the article. This study aims to design a system that can label each entity in traditional Indonesian art articles through Named Entity Recognition (NER) using the Conditional Random Field (CRF) algorithm. The implementation of CRF in NER is carried out using the Python programming language, and is evaluated through the f1-score. Based on several test scenarios, the best performance was found in the distribution of training data and test data with a ratio of 4:1, which resulted in an average f1-score of 88.2%, recall 90.3%, and precision 88.2%.*

Keywords— *Conditional Random Field, Entity, Indonesian Arts, Named Entity Recognition, Traditional Arts.*

I. Introduction

Indonesia is a country with variety ethnics and cultures. Culture is the whole system of ideas, actions, and outcomes human work in the context of social life which is made one's own human by learning. In general, there are seven elements of culture include language, knowledge systems, social systems or social organization, living equipment systems, technology, livelihood, religious system, and art [1]. Art is an expression of the experience of beauty or aesthetic experience [2]. On the other hand, traditional art itself is a used to express a sense of beauty from within the human soul with the conventional background or cultural system of the community that owns the arts [3].

As a part of Indonesian culture, traditional arts very many in number. Overall, traditional arts developed in Indonesia are divided into music, dance, theatre, fine arts, and literary arts [4]. From 2013 to 2019, a total of 345 arts (in this case performing arts) as Intangible Cultural Heritage Indonesia and has recognized, and the highest number is on the island of Java with a total of 72 cultures [5]. The performing arts are in music, dance, and the performing arts itself (including theatre arts in it). In apart from the performing arts, many other traditional arts have not been noted. This matter shows that the information circulating about traditional arts will also be very diverse.

Along with the times, information in digital articles can be obtained easily through internet technology, including traditional art articles. With so many kinds of information about traditional art and the lack of public knowledge of traditional arts, the emergence of difficulties in processing the current information is very likely found [6][7]. Moreover, the culture found on the island of Java is so diverse and has similarities to each other. Therefore, Named Entity Recognition (NER) implementation on digital art articles is expected to help the reader identify the terms (entities) in the article. Then the meaning can then be searched based on a dictionary or particular reference.

Many NER kinds of research in Indonesia have been carried out, such as [8] Suwarningsih et al. carried out the application of NER in the field of health. The method applied in this research is Conditional Random Fields (CRF). In this study, medical entities are classified into several categories: location, facility, diagnosis, definition, and person. Next, the research is implemented as a medical question system answering in Indonesian. As training data, 1000 sentences are used for every feature. The result obtained the highest accuracy value of 90% with using 3000 sentences as testing data.

Another study entitled [9] Named Entity Recognition (NER) Language Indonesia Uses Conditional Random Field and Post-Tagging. The study applied the CRF method for NER in Indonesian text facilitates the acquisition of important information in a text, such as the names of people, organizations, and places. Data that used is divided into two; unlabelled data obtained from Wikipedia and labelled data obtained from DBpedia. The resulting precision and recall values are above 75%, with an F1 value of 77% from this research.

Based on the two studies above, as well as several reference studies other supports, including research entitled [10] Named Entity Recognition Model for Indonesian Tweet using CRF Classifier with 86% precision for composite tweets (Munarko et al., 2018); Medical Entity Recognition using Conditional Random Field (CRF) and [11] Implementation Comparison Conditional Random Fields and Random Forests for Named Entity Recognition on News Article; then the implementation of NER using the CRF method on traditional Indonesian arts. As for the data processed, namely in the form of digital articles obtained from the internet. Besides CRF method, Hidden Markov Model (HMM) and Maximum Entropy Markov Model (MEMM) are also often used in NER, but CRF was chosen because it is a refinement of HMM and MEMM. In addition, on

Cyntia Fanny, Alexander Waworuntu, and Julio Christian Young, "Implementation of Conditional Random Field for Named Entity Recognition in Indonesian Traditional Arts Digital Article," *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 5, Issue 2, pp. 51-55, 2022.

several studies whose input data is also in Indonesian, CRF can achieve higher accuracy scores.

II. Methodology

A. Named Entity Recognition (NER)

Named Entity Recognition is a type of technique from Natural Language Processing (NLP) helps extract entities from a file text and classify them into predefined categories before [12]. NER is widely applied to the question answering system, information retrieval, co-reference resolution, topic modelling, and others [13].

There are two groups of methods commonly applied to NER such as rule-based and machine learning. The rule-based system runs on a rules basis that has been determined by experts, while the machine learning system using several data to study the pattern of each data so that it can predict the label from the following data [14].

B. Pre-processing

Pre-processing is the initial stage applied to the input data, in this case in the form of digital articles. Pre-processing consists of several stages, including [11][15]:

1. Case Folding

Case Folding aims to make the information received only contain the letters a-z, while other characters are changed or omitted. Case Folding includes converting the entire text to lowercase, removing numbers, and remove punctuation in the input text.

2. Stopword Removal

This stage aims to eliminate conjunctions or words filler for each sentence. Words to be omitted, for example, “dan”, “atau”, ”misalnya”, ”adalah”, ”ialah”, ”yaitu”, “tetapi”, ”di”, and so on, so that only words that have meaning are produced to understand easier. Table I is an example of the implementation of stopword removal.

TABLE I. Example of Stopword Removal Implementation

Before Stopword Removal	[Masyarakat] [diwajibkan] [untuk] [menggunakan] [masker] [saat] [berada] [di] [luar] [rumah]
After Stopword Removal	[Masyarakat] [diwajibkan] [menggunakan] [masker] [saat] [berada] [luar] [rumah]

3. Stemming

Stemming is a process to remove word inflection; returns the word to its basic form. For example, “ditulis”, “tertulis”, “menulis”, “tulisan” will be transformed to “tulis”. Table II is an example of the implementation of stemming.

TABLE II. Example of Stemming Implementation

Before Stemming	Masyarakat] [diwajibkan] [menggunakan] [masker] [saat] [berada] [luar] [rumah]
After Stemming	[Masyarakat] [wajib] [guna] [masker] [saat] [ada] [luar] [rumah]

4. Tokenization

Tokenization is the process of separating the text into pieces called tokens. The token generated can be in the form of words or sentences, but in this study, the tokens generated in the form of words. Table III is an example of implementation of tokenization.

TABLE III. Example of Tokenization Implementation

Before Tokenization	Budi has children named Sinta and Santi.
After Tokenization	[Budi] [memiliki] [anak] [bernama] [Sinta] [dan] [Santi] [.]

C. K-Fold Cross Validation

K-Fold Cross Validation is a data resampling method, where the dataset is divided into k subsets (Figure 1). Partitioning is done without happening conversion. The model will be trained using k-1 subsets, while one remaining is used as testing data. This is done as many as k iterations (folds), with the selection of different testing subsets. The position of the subset used testing data according to the current fold value. For example, on fold first, then the first subset will be testing data, the rest will be data training. In the second fold, the second subset becomes testing data, and the rest becomes training data. And so on until the whole folds. This results in there are no repeated testing data for each fold [16].

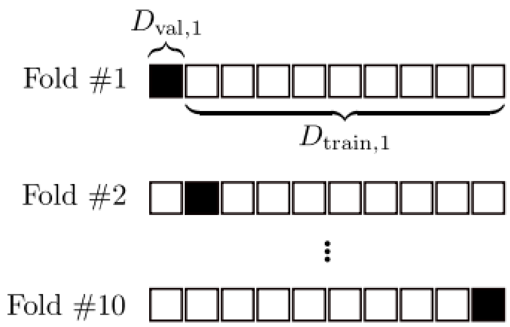

Fig. 1. Illustration of K-Fold application with 10 folds

D. Conditional Random Field (CRF)

Conditional Random Field is a discriminative and probabilistic model used in labeling or segmenting data sequences. CRF has many advantages compared to the model other probabilities, such as Hidden Markov Model (HMM) and Maximum Entropy Markov Model (MEMM), because in CRF, you can determine how many features you want, with free weights [17]. CRF has several forms, one of which is linear-chain CRF (Figure 2).

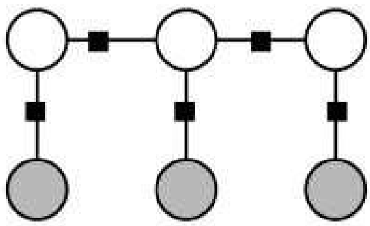

Fig. 2. Structure of Linear-chain CRF

Linear-chain CRF is used for NER because the result of NER is a sequence label. Linear-chain CRF meets the following definition [18]:

1. Given X, Y random vectors, λk is parameter vector, and fk(y, y0, xt) is real value of feature functions of k = 1 to k =

K, where K is number. So linear-chain CRF is distribution of p(y|x) in form of

$$(y \mid x) = \frac{1}{Z(x)} \exp\left(\sum_{k=1}^{K} \lambda_k f_k(y_t, y_{t-1}, x_t)\right) \quad (1)$$

where Z(x) is instance-specific normalization function

$$Z(x) = \sum_{k=1} \exp\left(\sum_{k=1}^{K} \lambda_k f_k(y_t, y_{t-1}, x_t)\right) \quad (2)$$

Main problem of linear-chain CRF is to find vector parameter $\lambda\vec{r} = \{\lambda 1, \ldots, \lambda k\}$ using maximum likelihood method during training process [9].

2. Suppose $\theta = \{\lambda_k\}$, given training data $\mathrm{D} = \{\mathrm{x(i)}, \mathrm{y(i)}\}^N$ where each $\mathbf{x}^{(i)} = \left\{\mathrm{x}_1^{(i)}, \mathrm{x}_2^{(i)}, \ldots, \mathrm{x}_T^{(i)}\right\}$ is sequences of input and each $\mathrm{y}^{(\mathrm{i})} = \left\{\mathrm{y}_1^{(i)}, \mathrm{y}_2^{(i)}, \ldots, \mathrm{y}_T^{(i)}\right\}$ is sequences of expected output, also each sequence assumed as independent instance. The estimation parameter is usually displayed in the form of maximum likelihood, but because this model is a conditional distribution, it will be displayed in the likelihood log, which fulfills

$$\mathrm{L}(\theta) = \sum_{i=1}^{N} \log\, p\left(y^{(i)} \mid x^{(i)}\right) \quad (3)$$

3. To obtain the conditional likelihood $\mathrm{p}(\mathrm{y} \mid \mathrm{x}; \theta)$ by eg the combination $\mathrm{p}(\mathrm{x}; \theta')$ which forms $p(y, x)$. And the resulting joint log-likelihood

$$\log\, \mathrm{p}(\mathrm{y}, \mathrm{x}) - \log\, \mathrm{p}(\mathrm{y} \mid \mathrm{x}; \theta) + \log\, \mathrm{p}(\mathrm{x}; \theta') \quad (4)$$

4. Substitution of the CRF model in point (1) to the likelihood in point (3), so that the resulting

$$\mathrm{L}(\theta) = \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{k=1}^{K} \lambda_k f_k\left(y_t^{(i)}, y_{t-1}^{(i)}, x_t^{(i)}\right) - \sum_{i=1}^{N} \log\, Z\left(x^{(i)}\right) \quad (5)$$

5. Regularization is used to avoid overfitting, which is a penalty for weight vectors whose norm calculation is too large. Penalty applied based on the Euclidean norm of and the regularization parameter 1/2σ2 as the strength of the penalty. Then the regularized log-likelihood is generated as

$$\mathrm{L}(\theta) = \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{k=1}^{K} \lambda_k f_k\left(y_t^{(i)}, y_{t-1}^{(i)}, x_t^{(i)}\right) - \sum_{i=1}^{N} \log\, Z\left(x^{(i)}\right) - \sum_{k=1}^{K} \frac{\lambda_k^2}{2\sigma^2} \quad (6)$$

NER using CRF is built on an undirected graphical model of independently trained probabilistic finite-state automata. CRF is used to calculate the conditional probability of each value that has been designed at the output nodes based on the value given to the input nodes [19].

E. BIO-format

BIO Format is the format for each word. There are three types of tag pieces to label each existing token, including B (Beginning), I (Inside), and O (Outside). The B tag helps mark the first word of an entity. The I tag helps mark the second word and so on from an entity. The O tag helps mark words that are not an entity [20][21].

Next, the BIO tag is followed by the entity type. In this study, five types of entities were used, resulting in eleven tags, as shown in Table IV.

TABLE IV. Tag with BIO Format

Entity	Tag
Dance (Tarian)	B-Tarian and I-Tarian
Music (Musik)	B-Musik and I-Musik
Person (Tokoh)	B-Tokoh and I-Tokoh
Performing Arts (Pertunjukan)	B-Pertunjukan and I-Pertunjukan
Musical Instrument (Alat music)	B-Alat and I-Alat
Uncategorized	O

F. Limited-Memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS)

L-BFGS is a Quasi-Newton (QN) method used for parameter optimization. Newton's method contains a Jacobian matrix that requires large-scale time in its execution because it is a derivative computation. Therefore, there is a Quasi-Newton method that replaces the negative computation into a direct computation function [22].

L-BFGS itself is a development of BFGS, where L-BFGS requires less memory than BFGS. While BFGS stores a dense n*n approximation into the Hessian inverse, L-BFGS stores it as a vector representing the implicit conjecture. It causes the memory usage on L-BFGS to be much smaller [23].

G. L1 and L2 Regularization

Regularization is the most frequently used technique to penalize models in machine learning, thereby minimizing overfitting and making models perform better for new inputs. The regression model that uses L1 regularization is called Lasso Regression, while the regression model that uses L2 regularization is called Ridge Regression [24].

L1 regularization gives a penalty equal to the absolute value of the coefficient. On the other hand, L2 regularization penalizes the value of the square of the absolute value of the coefficient. Giving high weight to L1 and L2 can cause the model to become underfit [24].

III. RESULT AND DISCUSSION

A. Research Data Set

This study used two kinds of datasets—the first dataset contains 23,488 words derived from 66 Indonesian-traditional-art articles sourced from kompas.com which have gone through preprocessing. It will be split into training subsets and a testing subset. The second dataset is the content of an Indonesian-traditional-art article sourced from Wikipedia.com. it will be used to external testing. Both of them in the Indonesian language.

B. Testing Scenario

This study tested regularization with hyperparameters c1 and c2. c1 represents the weight of L1, while L2 represents the weight of c2. Both c1 and c2 were tested with values of 0.01, 0.1, 0, 1, 10, and 100, resulting in 36 combinations. According to K-Fold Cross Validation with K=5, for each combination, training and testing are carried out five times (five folds) with

Cyntia Fanny, Alexander Waworuntu, and Julio Christian Young, "Implementation of Conditional Random Field for Named Entity Recognition in Indonesian Traditional Arts Digital Article," *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 5, Issue 2, pp. 51-55, 2022.

the distribution of testing and training subsets, resulting in five models, five f1 score, precision, and recall values. F1 score, precision, and recall values for each fold are calculated by weighted average. This is done because the distribution of labels in the dataset is uneven.

The next step, calculated the average f1 score, precision, and recall for each combination. The average is calculated using a macro average because the position of each fold is equal. To find the most stable combination of c1 and c2, it is seen from the highest average f1 score. Then from this combination, the model is taken from the fold with the highest f1 score. The model is then used for external testing.

C. Testing Results

Based on the test results, Table V are some of the evaluation results that can be presented. The first is the average value resulting from each combination.

TABLE V. Test result

Combination c1	c2	f1-score	Precission	recall
0	0	0.87933111	0.87933111	0.900517344
0	0.01	0.8804444	0.8804444	0.901425027
0	0.1	0.87930517	0.87930517	0.901993996
0	1	0.85957087	0.85957087	0.89078641
0	10	0.81197619	0.81197619	0.867488627
0	100	0.79556527	0.79556527	0.860119676
0.01	0	0.87975451	0.87975451	0.900698595
0.01	0.01	0.88232634	0.88232634	0.903754777
0.01	0.1	0.88121704	0.88121704	0.903603785
0.01	100	0.79556527	0.79556527	0.860119676
0.1	0	0.88155543	0.88155543	0.903368001
0.1	0.01	0.87885391	0.87885391	0.901543448
0.1	0.1	0.87850497	0.87850497	0.90191956
0.1	1	0.85867136	0.85867136	0.890399172
0.1	10	0.81197619	0.81197619	0.867488627
0.1	100	0.79556527	0.79556527	0.860119676
1	0	0.87068501	0.87068501	0.896752179
1	0.01	0.86971682	0.86971682	0.895962892
1	0.1	0.86662662	0.86662662	0.894079318
1	1	0.85327937	0.85327937	0.887205263
1	10	0.81173593	0.81173593	0.867336061
1	100	0.79556527	0.79556527	0.860119676

From the test results, the highest average value of f1 is 0.88232634 when the value of c1 is 0.01 and c2 is 0.01. Therefore this combination is the most stable parameter combination. Furthermore, the results of the testing of each fold for the combination are displayed (Table VI).

TABLE VI. Test result with c1=0.01 and c2=0.01

Fold	f1-score	precision	recall
1	0.879586917	0.890362774	0.896151819
2	0.899857117	0.904132802	0.917658258
3	0.85832081	0.863236118	0.890974272
4	0.861243762	0.878675326	0.891057649
5	**0.912623114**	0.915332312	0.922931888
Avg	**0.882326344**	**0.882326344**	**0.903754777**

The fifth fold has the highest f1 score from all folds, so the model from that fold is saved. Furthermore, the model is used in external testing. As for the detailed report of the model, it is shown in the Figure 3.

```
c1:0.01-c2:0.01-fold:5
                precision    recall  f1-score   support

         B-Alat      0.91      0.65      0.76       112
        B-Musik      0.58      0.58      0.58        36
   B-Pertunjukan     0.84      0.25      0.39       106
       B-Tarian      0.36      0.20      0.26        49
        B-Tokoh      0.65      0.58      0.61        53
         I-Alat      0.91      0.47      0.62        43
        I-Musik      0.60      0.57      0.59        21
   I-Pertunjukan     0.43      0.12      0.19        25
       I-Tarian      0.28      0.35      0.31        23
        I-Tokoh      0.33      0.24      0.28        21
              O      0.94      0.99      0.97      3754

       accuracy                          0.92      4243
      macro avg      0.62      0.46      0.50      4243
   weighted avg      0.92      0.92      0.91      4243
```

Fig. 3. Detail Report of Best Model

When viewed from the prediction performance of each label, the best prediction occurred on the O label, while the worst prediction occurred on the I-Tarian label. Label O has the best prediction results because the support value is very high, indicating that non-entity words are more numerous in articles in the dataset. For the I-Tarian label, it gets a low predictive value because the writing of dance entities often changes its format, which can only be the name of the dance itself or begins with the word "tari" or "tarian". Therefore it is possible to make wrong predictions between B-Tarian and I-Tarian, this is also supported by the low predictive value of B-Tarian.

For all combinations, if one or both c1 and c2 are filled with 100, then the lowest f1 value is 0.79557. When c1 is 0 and 0.01, the average value of f1 increases when there is a change from c2 = 0 to c2 = 0.01, but then the average f1 decreases from c2 = 0.1 and so on. For c1 with values of 0.1, 1, 10, and 100, the higher the value of c2, the lower the average value of f1.

The same thing happens the other way around. When c2 is 0 and 0.01, the average value of f1 increases when there is a change from c2 = 0 to c2 = 0.01 (specifically for c2 = 0, the average increase in f1 occurs until c1 = 0.1), then the average f1 experiences decreases as the value of c1 increases. For c2 values of 0.1, 1, 10, and 100, the higher the c1 value, the lower the f1 average value.

The distance of the change in the average f1 from one combination to another varies. The smaller the value of c1 or c2, the distance of the change in the average value of f1 is also getting smaller. There are even some points where there is no change in the average value of f1.

D. External Testing Result

In external testing, the processing is carried out on the dataset that has been provided previously. The article content is preprocessed to be word-by-word, then it will be predicted by the model. Predict is done twice, namely on content that has been stemming and content without stemming. The goal is to recognize entities even though they are affixed. However, if a word, the predictive result without stemming has been identified as an entity, then the result without stemming is used so that the meaning does not change.

Figure 4 is the result of external testing in the form of every word of the article followed by the style and label according to the type of entity. Styles are displayed in the form of background colors, including blue for dance; purple color for show; yellow color for musical instruments; pink color for music; and green for the characters. If the word is labeled "O" it will be displayed without a background color or label. The label will be printed following the last word of an entity.

Fig. 4. External Testing Result

IV. Conclusion and Future Works

Based on the results of the implementation and the tests' results, it can be concluded that Named Entity Recognition using Conditional Random Field has been successfully implemented and get with the average of f1 score in range 79.5% until 88.2%. The highest average f1 value was obtained when the values of c1 = 0.01 and c2 = 0.01, which was 88.2%. It also shows that the combination of parameters c1 = 0.01 and c2 = 0.01 produces the most stable f1 value for each fold compared to other combinations. For future works, adding a variety of articles that can include more new entities as training data will increase the model's performance. In addition, the addition of processes using POS-Tagging can also make the model more reliable in recognizing context.

Acknowledgement

Thank you to Universitas Multimedia Nusantara, Indonesia for supporting this research. Hopefully this research can contribute to the advancement of Information Technology.

References

[1] K. Koentjaraningrat, *Manusia dan kebudayaan di Indonesia.[Humans and culture in Indonesia]*. Jakarta: Djambatan, 2010.

[2] P. Soeriadiredja, *Fenomena kesenian dalam studi antropologi*. Denpasar: Fakultas Ilmu dan Budaya, Universitas Udayana, Denpasar, 2016.

[3] L. A. RM, "Seni pertunjukan tradisional di persimpangan zaman: studi kasus kesenian menak koncer sumowono semarang," *Humanika*, vol. 23, no. 2, pp. 25–31, 2016.

[4] Leonard, "Macam Macam Kesenian Daerah di Indonesia," 2018. https://ilmuseni.com/dasar-seni/macam-macam-kesenian-daerah.

[5] _, *Statistik Kebudayaan 2020*. Tangerang Selatan: Pusat Data dan Teknologi Informasi Kementerian Pendidikan dan Kebudayaan, 2020.

[6] Dhammajoti, J. C. Young, and A. Rusli, "A Comparison of Supervised Text Classification and Resampling Techniques for User Feedback in Bahasa Indonesia," in *2020 Fifth International Conference on Informatics and Computing (ICIC)*, 2020, pp. 1–6, doi: 10.1109/ICIC50835.2020.9288588.

[7] A. Rusli, A. Suryadibrata, S. B. Nusantara, and J. C. Young, "A Comparison of Traditional Machine Learning Approaches for Supervised Feedback Classification in Bahasa Indonesia," *Int. J. New Media Technol.*, vol. 7, no. 1, pp. 28–32, 2020, doi: 10.31937/ijnmt.v1i1.1485.

[8] W. Suwarningsih, I. Supriana, and A. Purwarianti, "ImNER Indonesian medical named entity recognition," in *2014 2nd International Conference on Technology, Informatics, Management, Engineering & Environment*, 2014, pp. 184–188.

[9] A. Willyawan, "Named entity recognition (NER) bahasa indonesia menggunakan conditional random field dan pos-tagging," Universitas Sumatera Utara, 2018.

[10] H. Raditya, M. A. Jiwanggi, and M. Adriani, "Medical entity recognition using conditional random field (CRF)," in *2017 International Workshop on Big Data and Information Security (IWBIS)*, 2017, pp. 57–62.

[11] F. Wong, "Perbandingan Implementasi Conditional Random Fields dan Random Forest untuk Named Entity Recognition pada Artikel Berita," Universitas Multimedia Nusantara, 2020.

[12] I. Roldós, "Named Entity Recognition: Concept, Tools and Tutorial," 2020. https://monkeylearn.com/blog/named-entity-recognition/.

[13] V. Yadav and S. Bethard, "A Survey on Recent Advances in Named Entity Recognition from Deep Learning models," Oct. 2019, [Online]. Available: https://doi.org/10.48550/arXiv.1910.11470%0A.

[14] M. Konkol, "Rozpoznávání pojmenovaných entit [Named Entity Recognition]," University of West Bohemia, 2016.

[15] S. Vijayarani, M. J. Ilamathi, and M. Nithya, "Preprocessing techniques for text mining-an overview," *Int. J. Comput. Sci. Commun. Networks*, vol. 5, no. 1, pp. 7–16, 2015, [Online]. Available: https://www.researchgate.net/profile/Vijayarani-Mohan/publication/339529230_Preprocessing_Techniques_for_Text_Mining_-_An_Overview/links/5e57a0f7299bf1bdb83e7505/Preprocessing-Techniques-for-Text-Mining-An-Overview.pdf.

[16] D. Berrar, "Cross-Validation," *Encycl. Bioinforma. Comput. Biol.*, vol. 1, pp. 542–545, 2018, [Online]. Available: https://www.researchgate.net/profile/Daniel-Berrar/publication/324701535_Cross-Validation/links/5cb4209c92851c8d22ec4349/Cross-Validation.pdf.

[17] N. Jaariyah and E. Rainarli, "Conditional Random Fields Untuk Pengenalan Entitas Bernama Pada Teks Bahasa Indonesia," *Komputa J. Ilm. Komput. dan Inform.*, vol. 6, no. 1, pp. 29–34, 2017, [Online]. Available: https://ojs.unikom.ac.id/index.php/komputa/article/view/2474.

[18] C. Sutton and A. McCallum, "An introduction to conditional random fields," *Found. Trends® Mach. Learn.*, vol. 4, no. 4, pp. 267–373, 2012, [Online]. Available: https://www.nowpublishers.com/article/Details/MAL-013.

[19] N. Patil, A. Patil, and B. V. Pawar, "Named entity recognition using conditional random fields," *Procedia Comput. Sci.*, vol. 167, pp. 1181–1188, 2020, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050920308978.

[20] B. Tang, H. Cao, Y. Wu, M. Jiang, and H. Xu, "Recognizing clinical entities in hospital discharge summaries using Structural Support Vector Machines with word representation features," in *BMC medical informatics and decision making*, 2013, vol. 13, no. 1, pp. 1–10, [Online]. Available: https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/1472-6947-13-S1-S1.

[21] F. Y. Azalia, M. A. Bijaksana, and A. F. Huda, "Name Indexing in Indonesian Translation of Hadith using Named Entity Recognition with Naïve Bayes Classifier," *Procedia Comput. Sci.*, vol. 157, pp. 142–149, 2019, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050919310695.

[22] C. Susmitha and D. Haritha, "An Ensemble Technique for Named Entity Recognition using Conditional Random Fields and L-BFGS Optimization Algorithm," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 3, pp. 3752–3756, 2020, [Online]. Available: https://d1wqtxts1xzle7.cloudfront.net/63915607/ijatcse19093202020200714-26536-1bbuz8r-with-cover-page-v2.pdf?Expires=1655000048&Signature=TLFoExc0SXXmjhdJlGdgvgPup4vHP4hFeVZqeLQJ2V1pBiBl7AiKwMSwK8PzjwYZK90ZyJSqgDAz6GJGf-l1hMnn55zdPdXgvddIjd64DDkPyH3ju-Nue.

[23] D. R. S. Saputro and P. Widyaningsih, "Limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method for the parameter estimation on geographically weighted ordinal logistic regression model (GWOLR)," 2017, [Online]. Available: https://doi.org/10.1063/1.4995124.

[24] N. Tyagi, "L2 and L1 Regularization in Machine Learning," *Anal. Steps*, vol. 28, no. 2, 2021.