Clickbait Headline Detection Using Supervised Learning Method

Vincent
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
vincent053@binus.ac.id

Sharlene Regina
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
sharlene.regina@binus.ac.id

Kartika Purwandari
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
kartika.purwandari@binus.ac.id

Felix Indra Kurniadi
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
felix.indra@binus.ac.id

***Abstract*— With the increasing use of the Internet of Things (IoT) as a means of communication in the 21st century, many news media now rely on the internet as an online news publication platform. News headlines are often made as attractive as possible to entice the reader's curiosity and thus increase the views of a news article. One of the many tactics employed is the use of clickbait. This research involves creating a model to detect headlines that contain clickbait. The model can act as a classifier between real news and clickbait-filled headlines with a Natural Language Processing (NLP) method approach. Bidirectional Long Short-Term Memory (Bi-LSTM), Decision Tree, and K-Nearest Neighbor (KNN) are all methods that can be used to distinguish actual news headlines from clickbait-laden headlines. This work is preliminary as research in this field is still being conducted, and improvements to the accuracy of these systems are still improving.**

Keywords*— *clickbait, Natural Language Processing, classifier, supervised learning

I. INTRODUCTION

The rise of clickbait cases on online news platforms often worries the public [1]. The use of clickbait by news journalists is often a source of doubt about the credibility of online news [2]. Judging from the number of journalists currently focusing on the number of views and clicks on their articles, they are competing to create sensational headlines [3]. Figure 1 shows an example of a headline that contains clickbait.


Politik

Gile, Luhut The Real President Ngamuk!

LAPORAN: WIDIAN VEBRIYANTO
Senin, 19 Maret 2018, 22:52 WIB

REPUBLIKMERDEKA Menteri Koordinator Bidang Kemaritiman Luhur Binsar Pandjaitan meluapkan kemarahan kepada Ketua Dewan Kehormatan PAN Amien Rais saat berpidato di Gedung BPK, Jakarta, Senin (19/3).

Fig. 1. Headlines that use clickbait formatting

Clickbait is a way of presenting news headlines intended to bind readers' emotions by creating a curiosity gap [4]. This gap is what causes the interest of the reader to click on the hyperlink and visit the publisher's website. Clickbait is usually characterized by excessive title writing style (Exaggeration), provocative titles (Inflammatory), and the use of capital letters and excessive punctuation (Formatting) [5].

Blocking Clickbait from users' social media feeds starts with identifying them. The social media websites Facebook, Twitter, and others have faced criticism for not up-ranking clickbait from users' feeds and failing to identify them [6], [7].

In this study, we rely on a classification determination of what constitutes a provocative title. Examples from English include the use of words like “fantastic” and “terrible”. The outcome of our results is dependent on this categorization. A model is trained on a labeled dataset containing some social media posts (Clickbait and non-Clickbait) to solve the problem. To detect Clickbait, most researchers employ supervised machine learning [9]–[11].

II. LITERATURE REVIEW

Clickbait on news headlines has been troubling the public for a long time [12]. There has been a lot of research on different methods and approaches to build engines that can detect clickbait news. The following are some related studies to address the clickbait problem. Based on research from [12], with a dataset of 3000 data, obtained an F1 Score of 76%. This research uses a regression model approach and a bag of word algorithm for the model. A similar study also came from [13], with a dataset of 6632 headlines, and an F1 score of 91% was obtained. This research uses the M-BERT language model approach in the model.

The methods that have been applied for modeling the classification of clickbait news also vary in complexity. All these research topics aim to find out which language method is the most effective for implementing clickbait news detection applications. The purpose of this study is to further examine the effectiveness of the application of the Long Short-Term Memory method in the application of the clickbait news classifier. The purpose of this study is to further examine the effectiveness of the application of the method to be tested between the Bidirectional Long Short-Term Memory, Decision Tree, and KNN methods in the application of the clickbait news classifier.

III. METHODOLOGY

Using a clickbait detection model requires the following architecture.

A. Dataset

The dataset is taken from DATA_INDONESIA.xlsx, with the topic discussing the classification between clickbait and non-clickbait news from several news sites on the internet. There are 15000 labeled samples with a distribution ratio of about 1:1 for clickbait and non-clickbait. Sources of news headlines taken from 12 different sites; Liputan6, detikNews, Okezone, Fimela, Kapanlagi, Posmetro-Medan, Tempo, Kompas, Republika, Tribunnews, Wowkeren, and Sindonews. The labeling results can be seen in the "label" string column and the "label-score" binary column, which states whether the headline is clickbait or non-clickbait. If the headline containing clickbait is given a binary label score of "1" and the string label "clickbait." Meanwhile, non-clickbait headlines are given a binary label score of "0" and a string label of "non-clickbait."

The variable that will be used as the target of this modeling prediction is the label column. So far, there is no further information on how to classify headlines labeled clickbait and non-clickbait in the dataset.

B. Text Processing

Perform text processing of news headlines with case folding, stemming, and tokenizing. Case folding is a function that converts the target string to lowercase. The case folding method uses the string. Case fold function. Skimming is a method of converting both infix and prefix words into basic words. The skimming function used in this model is specifically for Indonesian. Tokenization changes the text of news headlines into words, symbols, punctuation marks, or numbers in the form of tokens. The tokenization method uses the Tokenizer function.

C. Text Embedding

Embedding text using the Word2Vec neural network. Each token will be converted into vector form.

D. Trainable dan untrainable parameter

Knowing the number of parameters that can be used (trainable and trainable parameters). In this model, out of a total of 987,101 parameters, all of them are trainable parameters.

E. Experimental Setting

For the feature model, the test training data process is carried out with a comparison between training data, test data, and validation data of 3:1:1. For details on the data proposition for the benefit of the train test model, as shown in table I.

TABLE I. DATA SPLITTING

Training Data		**Testing Data**
Data training	***Data validation***	
9000 data	3000 data	3000 data

F. Implementation Bidirectional LSTM

Bidirectional LSTM or bi-LSTM is a recurrent neural network that is commonly used in natural language processing-based research [14], [15]. When compared with the application of ordinary LSTM, the application of the bi-LSTM method has different data flow characteristics. The input flow data on the Bi-LSTM can come from 2 directions, making it possible to receive and perform processing from both opposite directions. The use of this model is also commonly applied to sequential dependencies, which pay attention to two words that squeeze a word/phrase from both directions.

The use of bi-LSTM is suitable when a word phrase can have a different meaning depending on the word that is before or after the intended word phrase. The following is an example of the use of bi-LSTM in table II.

TABLE II. IMPLEMENTATION BI-LSTM

Example	**Detect people's names**
"Teddy Bear in the sale"	The word 'Teddy' here does not indicate a person's name because there is the word 'Bear' after it.
"James Teddy gave his speech yesterday."	The word 'Teddy' here indicates a person's name due to the previous word 'James'.

G. Implementation Decision Tree

Decision tree is a decision-making tool shaped like a tree with roots, branches, and leaves in the algorithm model. A decision tree is included in supervised learning, which produces decisions based on previous decisions. This model will undergo training and testing with a selected set of data sets. Decision tree works similarly to a human's way of thinking, so it is easy to apply and understand. The keys to the decision tree [16], [17] are the base node in the tree. In addition, the decision node is where the sub-node will be divided into additional sub-nodes. Therefore, the leaf is the last node, where an additional sub-node does not have a branch, so it is considered output. The branch is a sub-section containing several nodes.

Decision trees can provide a simple visualization of complex relationships between two or more nodes. Decision trees show the cause and effect of non-linear relationships. This method makes it easier to identify risks, benefits, and outputs following the previously selected decision. There are three advantages of using a decision tree, namely easy to understand, more efficient because it only requires less data cleaning than data modeling, and displays many complex relationships.

H. Implementation KNN

K-Nearest Neighbor or KNN is one of the supervised learning algorithms which is also included as a non-parametric algorithm because it has a working procedure that is to store several data that will be used as a race to carry out the classification or regression process. The KNN algorithm is often considered instant-based learning because the algorithm does not make predictions by studying training data but instead classifies new data by finding matches with datasets that have been stored by the algorithm. KNN has the following basic working procedures [18], [19]:

- Take the K number of nearest neighbors from the new point you want to classify
- Calculates the Euclidean distance on the specified K neighbors.
- Take the smallest Euclidean distance value, then count the number of data points in each category class.
- Label the new data categories according to the maximum number of neighbors.

I. Evaluation Method

Evaluation is needed to find out how well deep learning modeling can distinguish headlines that fall into the clickbait category and non-clickbait. In this model, the evaluation method can be obtained from the loss model performance curve, accuracy model performance curve, confusion matrix, and classification report.

IV. Result and Analysis

A. Training and Validation Performance Curve

Here a purely binary classification method has been employed through future work that includes a graduated scale that is monotonic [20] between 0 and 1, where weights are determined independently. One of the problems that are often experienced in Bidirectional LSTM learning modeling is an overfitting model. The overfitting condition causes the results of the two ratios, namely the accuracy ratio and loss ratio, to not intersect in the training and data validation phases.

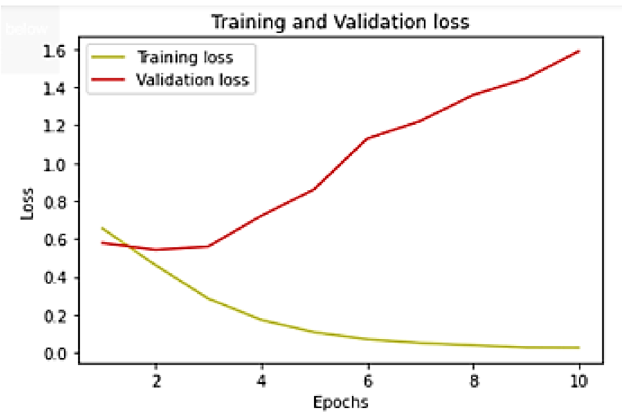

Fig. 2. Training and validation loss performance curves.

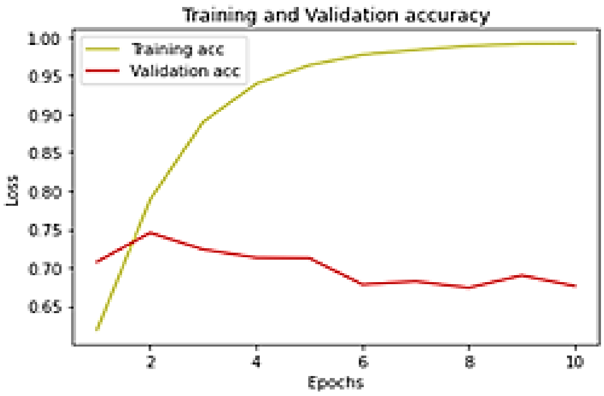

Fig. 3. Training and validation accuracy performance curves

Use the dropout function of the neuron layer to overcome data overfitting. The observation result of the two performance curve models in Figure 2 and Figure 3 shows the ratio for the loss level of 0.59 and the ratio for the accuracy level of 0.71.

B. Confusion Matrix

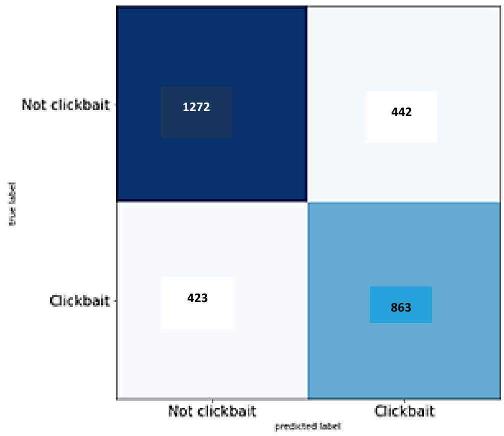

Fig. 4. Confusion matrix

There are four things needed to be considered from the confusion matrix to check the prediction results, namely

- TN / True Negative: true label 'non-clickbait' and predicted label 'non-clickbait'.
- TP / True Positive: true label 'clickbait' and predicted label 'clickbait.'
- FP/False Positive: true label 'non-clickbait' and predicted label 'clickbait.'
- FN/False Negative: true label 'clickbait' and predicted label 'non-clickbait.'

From the clickbait analogy in Figure 4, the data obtained are 863 True Positives, 442 False Positives, 423 False Negatives, and 1272 True Negatives data.

C. Classification Report

To prove that the bi-LSTM approach is the most effective model for clickbait detection, we added two other supervised learning methods: the KNN method (n_neighbors =9) and the Decision Tree method. The following table III displays the comparison results of each learning method.

TABLE III. COMPARISON BI-LSTM, KNN, AND DECISION TREE.

	Bi-LSTM	KNN	Decision Tree
Precision	0.67	0.51	0.49
Recall	0.69	0.313	0.49
Accuracy	0.71	0.577	0.56
F1 Score	0.65	0.388	0.49

In table III, here are four things that are needed to build a classification report:

- Precision: Precision shows the classifier's ability to display the ratio of predicted positive labels. The precision ratio in Bi-LSTM shows a score of 67%, KNN scores of 51%, and Decision Tree scores of 49%. From the comparison value obtained, Bi-LSTM proved to show the best precision ratio, followed by KNN and Decision Tree.

- Recall: Recall shows the classifier's ability to display the ratio of predicted positive labels that have been identified correctly. The recall ratio in bi-LSTM shows a score of 69%, KNN shows a score of 31.3%, and Decision Tree scores 49%. From the comparison value obtained, Bi-LSTM proved to show the best recall ratio, followed by Decision Tree and KNN.

- F1 Score: The F1 Score shows the classifier's ability to display the ratio weighted harmonic mean between precision and recall ratio. F1 score on bi-LSTM shows a score of 65%, KNN shows a score of 38.8%, and Decision Tree shows a score of 49%. From the comparison value obtained, Bi-LSTM proved to show the best F1 score, followed by Decision Tree and KNN.

- Accuracy: Accuracy is the exact predictive ratio between the positive and negative of the entire dataset. The accuracy ratio in bi-LSTM shows a score of 71%, KNN scores of 57.7%, and Decision Tree scores of 56.2%. From the comparison value obtained, Bi-LSTM proved to show the best F1 score, followed by KNN and Decision Tree. Followed by KNN with an accuracy ratio of 57,7% and KNN with an accuracy ratio of 56,2%.

The accuracy value obtained proves that the Bidirectional LSTM learning model is the best model to identify and classify clickbait and non-clickbait headlines effectively.

D. System Testing

The following process is to test the model by manually entering news headlines as test data. Table IV shows the result of headline classification using the pre-trained model.

TABLE IV. EXPERIMENTAL RESULTS ENTERING THE HEADLINE TITLE

Headline Title	Label
Denny Siregar kritik pedas demo anak STM jadi trending, ejek jangan jadi banci	Clickbait
Gagal rencana 3.000 orang main petak umpet di IKEA	Clickbait
Wanita dengan tubuh berisi terbukti bikin pria lebih bahagia	Clickbait
Hujan asteroid di bulan September	Non-clickbait

E. Overfitting Handling

Three ways can be done to anticipate the occurrence of overfitting data: augmented data, dropout function, and early stopping. In this model, we use the network model's dropout function on the neuron layers. The dropout value used is 20% of the total neurons because a small value will have an almost invisible effect. Otherwise, the network will become under-learning.

V. CONCLUSION

The conclusion obtained from this study is that using the Bidirectional LSTM model equipped with a dropout function in anticipation of overfitting the data shows an accuracy of 71% and a loss of 59%. With the accuracy value obtained, compared to other supervised methods such as KNN and Decision Tree, the Bi-LSTM method seems to be the most effective method for the application of clickbait and non-clickbait headlines detection. With the importance of elections and the role of public opinion, there is no doubt that this kind of research will continue. We have presented one approach to identifying clickbait in headlines. However, there are many other ways that statistical learning [21] can be employed to classify sensationalized news presentations and perhaps warn readers.

REFERENCES

[1] M. Alimardani and M. Elswah, "Online Temptations: COVID-19 and Religious Misinformation in the MENA Region," *Social Media+ Society*, vol. 6, no. 3, p. 2056305120948251, 2020.

[2] A. Chakraborty, B. Paranjape, S. Kakarla, and N. Ganguly, "Stop clickbait: Detecting and preventing clickbaits in online news media," in *2016 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)*, 2016, pp. 9–16.

[3] T.-L. Wang, "Presentation and impact of market-driven journalism on sensationalism in global TV news," *International Communication Gazette*, vol. 74, no. 8, pp. 711–727, 2012.K. Elissa, "Title of paper if known," unpublished.

[4] P. Mormol, "'I Urge You To See This...'. Clickbait as One of the Dominant Features of Contemporary Online Headlines," *Social Communication*, vol. 5, no. 2, pp. 1–10, 2019.

[5] P. Biyani, K. Tsioutsiouliklis, and J. Blackmer, "' 8 amazing secrets for getting more clicks': detecting clickbaits in news streams using article informality," 2016.

[6] N. Hurst, "To clickbait or not to clickbait?: an examination of clickbait headline effects on source credibility," University of Missouri–Columbia, 2016.

[7] J. Cable and G. Mottershead, "'Can I click it? Yes you can': Football journalism, Twitter, and clickbait," *Ethical Space*, vol. 15, no. 1/2, 2018.

[8] Newman, M.L.; Pennebaker, J.W.; Berry, D.S.; Richards, J.M. Lying words: Predicting deception from linguistic styles. Personal. Soc. Psychol. Bull. 2003, 29, 665–675

[9] X. Cao, T. Le, and others, "Machine learning based detection of clickbait posts in social media," *arXiv preprint arXiv:1710.01977*, 2017

[10] D. M. Daoud and M. El-Seoud, "An Effective Approach for Clickbait Detection Based on Supervised Machine Learning Technique.," *International Journal of Online & Biomedical Engineering*, vol. 15, no. 3, 2019.

[11] B. Siregar, I. Habibie, E. B. Nababan, and others, "Identification of Indonesian clickbait news headlines with long short-term memory

recurrent neural network algorithm," in *Journal of Physics: Conference Series*, 2021, vol. 1882, no. 1, p. 12129.

[12] M. Potthast, T. Gollub, M. Hagen, and B. Stein, "The clickbait challenge 2017: Towards a regression model for clickbait strength," *arXiv preprint arXiv:1812.10847*, 2018.

[13] M. N. Fakhruzzaman, S. Z. Jannah, R. A. Ningrum, and I. Fahmiyah, "Clickbait Headline Detection in Indonesian News Sites using Multilingual Bidirectional Encoder Representations from Transformers (M-BERT)," *arXiv preprint arXiv:2102.01497*, 2021.

[14] D. Wang, J. Su, and H. Yu, "Feature extraction and analysis of natural language processing for deep learning English language," *IEEE Access*, vol. 8, pp. 46335–46345, 2020.

[15] T. S. Heo *et al.*, "Prediction of stroke outcome using natural language processing-based machine learning of radiology report of brain MRI," *J Pers Med*, vol. 10, no. 4, p. 286, 2020.

[16] S. Pathak, I. Mishra, and A. Swetapadma, "An assessment of decision tree based classification and regression algorithms," in *2018 3rd International Conference on Inventive Computation Technologies (ICICT)*, 2018, pp. 92–95.

[17] J. R. Quinlan, "Decision trees and decision-making," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, no. 2, pp. 339–346, 1990.

[18] O. Kramer, "K-nearest neighbors," in *Dimensionality reduction with unsupervised nearest neighbors*, Springer, 2013, pp. 13–23.

[19] S. B. Imandoust, M. Bolandraftar, and others, "Application of k-nearest neighbor (knn) approach for predicting economic events: Theoretical background," *Int J Eng Res Appl*, vol. 3, no. 5, pp. 605–610, 2013.

[20] Kearsley, Anthony J. "Projections onto order simplexes and isotonic regression." Journal of Research of the National Institute of Standards and Technology 111, no. 2 (2006): 121-125.

[21] T. Hastie, R. Tibshirani, and J. Friedman. The Elements of Statistical Learning, Springer (2001), 468–469.