Journal of Applied Intelligent System (e-ISSN : 2502-9401 | p-ISSN : 2503-0493)
Vol. 7 No. 2, August 2022, pp. 193 – 205

# Predicting News Article Popularity with Multi Layer Perceptron Algorithm

**Arie Rachmad Syulistyo***[1], **Vira Meliana Agustin**[2]
*State Polytechnic Of Malang, Jl. Soekarno Hatta No. 9 Jatimulyo Kec. Lowokwaru Kota Malang Jawa Timur, (0341) 404424*
*E-mail : arie.rachmad.s@polinema.ac.id*[1], virameliana18@gmail.com[2]*

**Dwi Puspitasari**[3]
*State Polytechnic Of Malang, Jl. Soekarno Hatta No. 9 Jatimulyo Kec. Lowokwaru Kota Malang Jawa Timur, (0341) 404424*
*E-mail : dwi.puspitasari@polinema.ac[3]*

**Abstract –**Nowadays, news media seems to have been digitized. One of them is printed news which has now turned into online news. The increasing use of social media has made people interested in reading news online. News needs to attract readers with their headlines. Various online news media businesses want to know the future demand of readers, as well as whether the released news can reach more readers so that the news becomes popular. Therefore, with the increasing interest in online news today, this paper will analyze the performance of the Neural Network Algorithm and other artificial intelligence techniques in predicting the popularity of news articles that can help the media to know whether their news will become popular. The news article popularity prediction system can increase its revenue if there are advertisements in the news. The test results show that the accuracy of the Multi Layer Perceptron is 76% and Random Forest gives an accuracy of 70%.

**Keywords -** News, Popularity, Neural Network, Multi Layer Perceptron, Random Forest

## 1. INTRODUCTION

In the current era of information and technology, news media seems to have been digitized[1]. One of the misnews in print media which has now turned into online news. Reading, writing, and sharing information have become a part of life for people's entertainment[2], [3]. The emergence of online news makes people very interested in discussing all information for public consumption. This is supported by the development of social media such as YouTube, Instagram, Twitter, and Facebook so as to make people's interest in reading online news become increasing. It is undeniable, that the advancement of social media seems to increase the distribution of online news media. That’s why, information in online news flows so fast, so news becomes more dynamic with low cost but a relatively short life span[4].

The fast flow of online news certainly creates new problems for writers to continue to innovate and present news that is always up to date. More and more, online news enthusiasts are booming. News needs to attract readers with headlines or news titles, to anticipate the preference of readers. In another hand, the reader is able to anticipate the content of a news article before a headline, because knowledge inside the content of the news is certainly in accordance with the content of the news[5]. Various online news media businesses want to know the future demand of readers, as well as whether the released news can reach more readers. If they can find out if the news can reach more readers, of course, they will be better prepared to make decisions immediately in implementing news on their online platform [6].

Therefore, with the increasing interest in online news today, this paper will analyze the performance of the Neural Network Algorithm and other techniques in predicting news article popularity can help the media to know whether their news will become popular. The predicting news article popularity system can increase their income if there are advertisements in the news. This system is widely used in various types of applications such as media advertising, traffic management, and economic trend forecasting. This research uses an online news data set from Tribun News Articles which contains almost 1000 news titles, news_articles descriptions, publish time, publish date, and a number of views from February 01, 2021, to April 08, 2021, to be processed in a model so that it can be classified to predict the popularity.

Artificial Neural networks are models with a high enough level of accuracy to perform tasks such as classification and prediction. This model is considered a Multi Layer network of logistic regression units. This model also has more layers and a complex structure, so this research assume that neural networks are stronger for prediction systems than one-layer parametric logistic regression. Artificial Neural Network has widely used as one for predictive modeling. This method has a good ability in analyzing data patterns, that’s why this algorithm is good in prediction. One Artificial Neural Network that is often used as a predictive model is Multi Layer Perceptron.

The algorithm can help news media in classifying whether the news is worthy / can be published with an accurate prediction of the popularity. The media can do the classification first before the news is in the hands of the public so that the media can provide interesting content and headlines to achieve popularity.

## 2. RESEARCH METHOD

### *2.1. Research Position*

Previous research conducted by Priyanka Rathord (2019) under the title A Comprehensive Review on Online News Popularity Prediction using Machine Learning Approach conducted research with Comparative analysis of various popularity prediction methods, namely Random Forest, SVM, Ada Boost, KNN, Naive Bayes, Linear Regression, Logistic Regression and Genetic Algorithm[6], [7]. This research results in the accuracy of each algorithm in predicting the popularity of news, where Random Forest occupies the highest accuracy position. However, this research will still be improved by using the Neural Network algorithm and will be compared with the previous algorithm.

In Jalal Rezaeenour's (2018) research in a journal entitled Developing a New Hybrid Intelligent Approach for Prediction Online News Popularity, he conducted research on popular news prediction by utilizing the ELM (Extreme Learning Machine) Neural Network algorithm[4]. This research shows that the most important predictors of popularity are the time for publishing news (higher number of visitors on weekends) and news topics (lifestyle and social media are the most popular topics on the site).

In Feras Namous' (2018) research entitled Online News Popularity Prediction, he conducted a study to determine popular news predictions using data sets from the Mashable

News Website and compared algorithms for classification and prediction. The best algorithms with the highest level of accuracy for popular news prediction cases are Random Forest and Multi Layer Perception Neural Network.

Based on the current research, the method with the best accuracy is Multi Layer Perceptron and Random Forest, so the paper will compare the level of accuracy in the application of Multi Layer Perceptron with Random Forest.

### *2.2. Dataset*

The data used to conduct this research are articles from Tribun News https://www.kaggle.com/waseemakramkhan/the-tribune-news-articles. This data set contains almost 1000 news titles, news_articles descriptions, publish time, publish date, number of views and popularity from February 01, 2021, to April 08, 2021, collected from the tribune newspapers. On Tribun News Article Dataset, there is a column that defines the popularity, conducted of is_popularity. This research will apply the prediction of popularity by headlines / titles of the news articles. The data set used is balanced, the two classes have a number that is not much different, which is 432 for popular news data and 589 for non-popular news data.

The data set is also taken from https://www.kaggle.com/datasets/szymonjanowski/internet-articles-data-with-users-engagement?resource=download. In the data set there is a top_article attribute which indicates that the article is popular. Of the total 10436 data, only 3853 data were taken for research so that the overall data set was more balanced for both popularity classes. The data set with the top article class has unbalanced data, so the researcher only takes some and combines it with the first data set. The result of the two data sets is a balance.

From total data set will be split into three sets, consisting of train, validation, and test. From research by Islam, M. M., Karray, F., Alhajj, R., & Zeng, J. (2021) entitled “A Review on Deep Learning Techniques for the Diagnosis of Novel Coronavirus (COVID-19), the data partitioning step splits the data into training, validation, and testing set for the experiment. In the "Diagnosis Using Computer Tomography (CT) Images" scheme, this study uses a split data set train, testing, validation scheme of 80:10:10, 60:20:20, 60:25:15, 60:30:10, 50:30:20.

In this research, the data set will be divided into three sets namely training, testing, and validation for the experiment with 3 schemes, consisting of 80:10:10, 60:20:20, 60:25:15, 60:30:10, and 50:30:20 experiment patterns.

### *2.3. System Design*

Main design of Predicting News Article Popularity with Multi Layer Perceptron Algorithm research as follows :

This paper used the text mining method[9], [10] to preprocess the data which has been divided into training, validation, and testing stages. At the preprocessing stage, the data will be processed using the text mining method. In this method, the data will go through a case folding stage to convert letters to lowercase, tokenizing, stop words removal, and lemmatization.

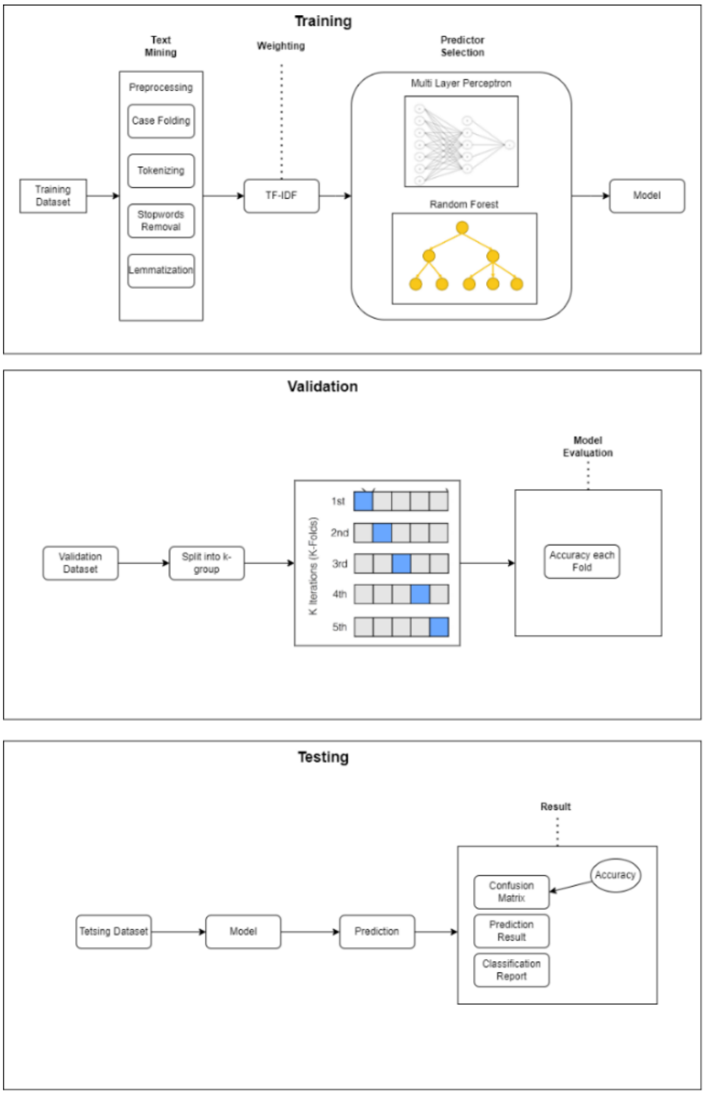


Figure 1. System Design

The next training data will be a prediction process using the Multi Layer Perceptron method and Random Forest with input in the form of news article titles. The first method is MLP. For example, this method has 3 layers, that is input layers, hidden layers, and output layers. After the modeling is saved, the next step is the validation process.

After the training stage for the MLP algorithm, the data is then trained with the Random Forest algorithm which is also stored in the model. The preprocessing stage for this method is the same as previously described, the difference lies in the classification process of news popularity.

Validation will be applied by 5-fold cross-validation, which is split into a K number (on this stage K = 5) of section or fold where each fold is used as a testing set at some point. This step also gives an output of the prediction and accuracy of each k-subset value. The testing step will use testing data and use the model to process it. At the testing stage, the system will apply a confusion matrix to find out the percentage of the classification accuracy.

Both models will be compared between Random Forest and Multi Layer Perceptron, which means which method has the best accuracy rate for news popularity prediction. In today's prejudice, MLP has a better level of accuracy and performance. However, the results will still be determined based on the accuracy results at the testing stage.

The system needs one parameter which is an input news article title to show the prediction. The algorithm will predict the popularity. The entire process is built in the python programming language using the Scikit-learn, NLTK, Numpy, and Pandas to Regular Expression libraries and is integrated on the website using the Flask RESTful API so that users can use the system more easily.

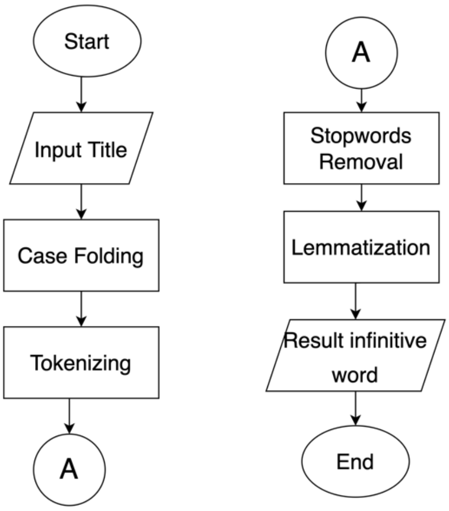


Figure 2. Flowchart Preprocessing

## 2.4. *Preprocessing*

This stage is the stage for processing the data set. The first begins with inputting data on article news titles which will be processed using Stopwordss, tokenizing and filtering all words. At the preprocessing stage on figure 2 will produce clean data. This means that the data is in lowercase format, does not contain meaningless data, and consists of infinitive words with a valid meaning from the lemmatization step. In figure 3, the heading data will be processed through casefolding. All letters will be returned to lowercase and remove punctuation marks. From this sentence, it goes to the tokenizing stage, which is dividing it into several word tokens. And each of these words will go through a stop word removal process that filters the words listed in the stop word list. Furthermore, the clean words enter the lemmatization stage, which is changed to basic words in English.

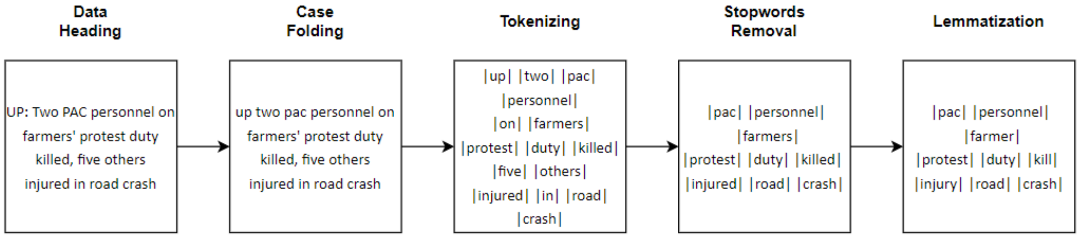

Data Heading
Case Folding
Tokenizing
Stopwords Removal
Lemmatization
UP: Two PAC personnel on farmers' protest duty killed, five others injured in road crash
up two pac personnel on farmers' protest duty killed, five others injured in road crash
|up| |two| |pac| |personnel| |on| |farmers| |protest| |duty| |killed| |five| |others| |injured| |in| |road| |crash|
|pac| |personnel| |farmers| |protest| |duty| |killed| |injured| |road| |crash|
|pac| |personnel| |farmer| |protest| |duty| |kill| |injury| |road| |crash|

Figure 3. Preprocessing Process Example

### 2.5. TF-IDF Process

In the TF-IDF weighting, the process for each word has gone through the preprocessing stage. In the stage of giving weights to words, it is necessary to use the TF-IDF method. This weighting aims to assign a value to a word that will be used as input in the implementation of the model.

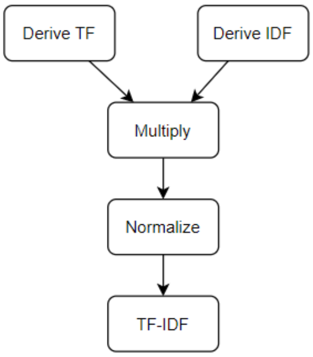


Figure 4. TF-IDF Process

### 2.6. Data Processing

At this stage, the data processing will be carried out. To predict the popularity of news articles, it is necessary to apply several techniques. After collecting data and through all preprocessing steps, the processed data will be classified using the Multi Layer Perceptron Algorithm method, which will classify predictions of popularity from news articles.

### 2.7. Prediction using Multi Layer Perceptron

From the results of the words that have been processed by TF-IDF, then the prediction process will be carried out using the Multi Layer Perceptron method. The steps that need to be done in the Multi Layer Perceptron method are as follows:

- Determine the number of input inputs, hidden layers, and outputs as training targets.
- Randomly assigns initial values to all weights between the input-hidden and hidden-output layers.
- Doing Feed forward.
- Processing back propagation.

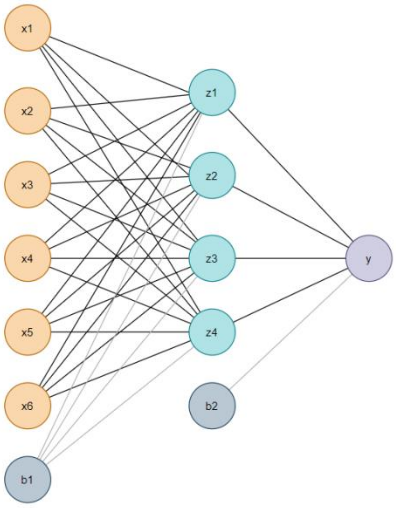


Figure 5. The Illustration of MLP Structure

In figure 5, Example Multi Layer Perceptron's structure consists of 3 layers, namely 1 input layer, 1 hidden layer, and 1 output layer. For input layer, consists of 6 inputs neuron and 1 bias neuron. The hidden layer consists of 4 hidden neurons and 1 bias neuron. And the last one is the output layer consists of 1 output neuron.

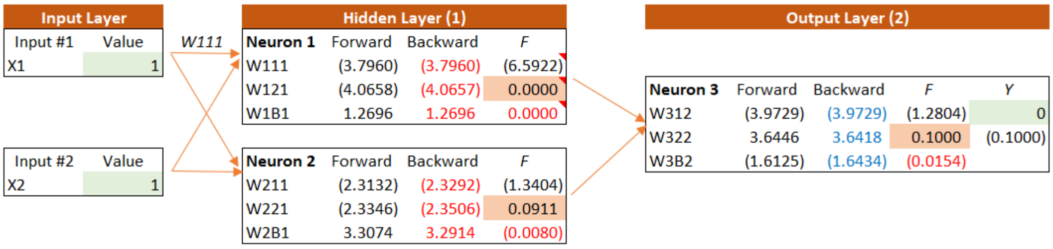
Table 1 below, is a sample calculation of the Multi Layer Perceptron using 2 input layer neurons, 2 hidden layer neurons, and 1 output layer.

Table 1. Table Sample of input value

| Prediction | Y | X1 | X2 |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 1 | 1 | 1 | 0 |
| 1 | 1 | 0 | 1 |
| 0 | 0 | 1 | 1 |

From the sample above, X1 and X2 is the value of the input layer and Y is actual result or popularity of the news article. The prediction result needs to be the same as the Y value. So first thing to do is define Network Parameters:

- Epoch = 150
- Bias = 1



Input Layer

| Input #1 | Value |
|---|---|
| X1 | 1 |

| Input #2 | Value |
|---|---|
| X2 | 1 |

Hidden Layer (1)

| **Neuron 1** | Forward | Backward | *F* |
|---|---|---|---|
| W111 | (3.7960) | (3.7960) | (6.5922) |
| W121 | (4.0658) | (4.0657) | 0.0000 |
| W1B1 | 1.2696 | 1.2696 | 0.0000 |

| **Neuron 2** | Forward | Backward | *F* |
|---|---|---|---|
| W211 | (2.3132) | (2.3292) | (1.3404) |
| W221 | (2.3346) | (2.3506) | 0.0911 |
| W2B1 | 3.3074 | 3.2914 | (0.0080) |

Output Layer (2)

| **Neuron 3** | Forward | Backward | *F* | *Y* |
|---|---|---|---|---|
| W312 | (3.9729) | (3.9729) | (1.2804) | 0 |
| W322 | 3.6446 | 3.6418 | 0.1000 | (0.1000) |
| W3B2 | (1.6125) | (1.6434) | (0.0154) | |

Figure 6. MLP Process Sample

The input layer neurons are forwarded to the Hidden Layer in Neurons 1 and 2. The first step is to calculate the Weight. W111 means Weight of Hidden Neuron Layer 1, Input Layer 1, first Weight while B means Bias. The value at the input layer will be calculated using the formula described in the previous chapter.

The higher the epoch value, the more accurate the results will be. Giving an epoch value that is too high also does not have a good effect on training performance, so it is necessary to determine the right epoch value.

After calculating the weight for forward, the next updates weight for backward. From the updated weight, next is defined the induced field and neuron output, which is 0.0 and 0.9 in both neurons of the hidden layer.

### *2.8. System Testing*

Testing will be carried out after the implementation phase is complete. Testing is very helpful for research to find out whether the system is running properly and appropriately. Testing of the Predicting News Article Popularity system with Multi Layer Perceptron Algorithm can be done by:

a. Perform User Acceptance Testing to run website-based applications that have implemented algorithms.
b. Testing the accuracy of all implemented methods and comparing the accuracy results between Multi Layer Perceptron and Random Forest.

Calculation of accuracy can be done with the Confusion Matrix table. From the table, the calculation of accuracy, recall, and precision can be displayed.

## 3. RESULTS AND DISCUSSION

After building the system, it is necessary to do blackbox testing according to the scenario prepared in the previous chapter. The results of blackbox testing are represented in the following table:

Table 2. Table Black Box Testing

| No | Scenario | Hoped Result | Result | Status |
|---|---|---|---|---|
| 1 | Submit form news popularity prediction on home page | Prediction result will show in the bottom of form with keyword | Pass | Succeed |
| 2 | Register new account for user | User success create new account | Pass | Succeed |
| 3 | Submit form login account | If the form validations is true, user redirect to dashboard page. If not, user still on login page | Pass | Succeed |
| 4 | Submit form news popularity prediction on dashboard page | Prediction result will save on database and redirect to history prediction | Pass | Succeed |
| 5 | Export data prediction | Export success with PDF, CSV, Word format | Pass | Succeed |
| 6 | Edit prediction data in history page and resubmit | The system will generate new prediction result and edit the last one | Pass | Succeed |
| 7 | Delete data prediction | The system will delete the selected data | Pass | Succeed |
| 8 | Display dashboard of classification and accuracy report | Display the dashboard with classification report | Pass | Succeed |

Based on the testing scenarios, the results are appropriate and can be concluded as successful. All features have been tested with the expected results.

*3.1. User Acceptance Testing Result*

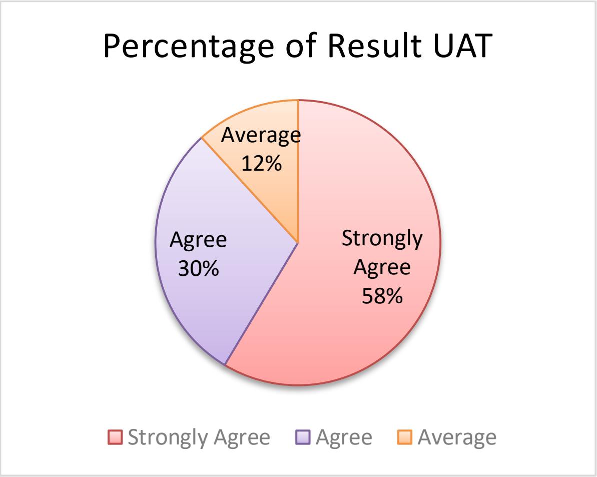


Figure 7. UAT Result

Testing needs to be done to find out the Design and Build of a News Popularity Prediction Website as needed and has been running correctly. Tests are carried out using User Acceptance Testing (UAT). The following are the results of the questionnaire testing the UAT method which is implemented in the News Popularity Prediction system. This stage aims to obtain information on whether the system that has been built is in accordance with user needs. Testing is intended to test the extent to which the application can function and be useful according to needs. From each percentage of respondents as many as 15 users taken on August 2, 2022, then the highest and lowest score can be calculated as follows:

Table 3. Table Highest and Lowest Score

| | |
|---|---|
| Highest Score | 15 x 8 x 5 = 600 (if all respondent answer strongly agree) |
| Lowest Score | 15 x 8 x 1 = 120 (if all respondent answer strongly disagree) |

From the calculation which states the highest value is 4800 so that the results of the percentage of UAT tests can be found as follows:

$$\text{Percentage} = 486 \times 600 \times 100\% = 81\%$$

From the results of the percentage above, it can be concluded that the level of usability of the system is strong, which is 81% from 100%.

### *3.2. Accuracy Testing Result*

Accuracy testing aims to determine the level of success of the system in predicting the popularity of news by using several testing samples that have been split in the system.

### *3.2.1 Multilayer Perceptron Testing*

Before entering the testing stage, researchers need to calculate the validation value using K Fold Cross Validation. From the model parameters used for training and data validation, there are 5 ratio split data set model to validate.

Table 4. Table MLP K-Fold Cross Validation

| Ratio | Fold | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| 80:10:10 | 0.56 | 0.67 | 0.61 | 0.49 | 0.57 |
| 60:20:20 | 0.57 | 0.6 | 0.61 | 0.55 | 0.58 |
| 60:25:15 | 0.65 | 0.57 | 0.6 | 0.64 | 0.63 |
| 60:30:10 | 0.64 | 0.65 | 0.6 | 0.61 | 0.61 |
| 50:30:20 | 0.59 | 0.63 | 0.65 | 0.6 | 0.61 |

The validation test result shows the highest average value lies in the ratio of 50:30:20 which is equal to 61%. Values from k-fold 1 to 5 have insignificant differences. Testing the accuracy of the Multi Layer Perceptron method using the MLP Classifier with 5 data set split data represent by confusion matrix on table as follows:

Table 5. MLP Confusion Matrix for Ratio 80:10:10

| Actual/Prediction | Not Popular’ | Popular’ |
|---|---|---|
| Not Popular | 181 | 59 |
| Popular | 58 | 190 |

From the table 5 above, there are 190 popular classes and 181 not popular class data which are predicted to be correct. Meanwhile, the other 118 data were predicted to be incorrect. From the table above, the results of the classification report calculation are as follows:

Table 6. Table MLP Classification Report

| Ratio | Precision | Recall | F1 | Acc |
|---|---|---|---|---|
| 80:10:10 | 1 = 0.76<br>0 = 0.76 | 1 = 0.77<br>0 = 0.75 | 1 = 0.76<br>0 = 0.76 | **0.76** |
| 60:20:20 | 1 = 0.76<br>0 = 0.68 | 1 = 0.63<br>0 = 0.79 | 1 = 0.69<br>0 = 0.73 | 0.71 |
| 60:25:15 | 1 = 0.73<br>0 = 0.67 | 1 = 0.63<br>0 = 0.77 | 1 = 0.67<br>0 = 0.71 | 0.69 |
| 60:30:10 | 1 = 0.74<br>0 = 0.69 | 1 = 0.63<br>0 = 0.79 | 1 = 0.68<br>0 = 0.73 | 0.71 |
| 50:30:20 | 1 = 0.69<br>0 = 0.68 | 1 = 0.66<br>0 = 0.72 | 1 = 0.67<br>0 = 0.70 | 0.68 |

The table 6 is a classification report. From the report, it is stated that the MLP Classifier model gives a highest accuracy at ratio 80:10:10 split data set which is 76% with the precision value for not popular predictions is 76%, recall is 77%, and f1-score is 76%. Meanwhile, for popular predictions, precision is 76%, recall is 75% and f1-score is 76%.

### *3.2.2 Random Forest Testing*

In this study, the Random Forest Algorithm was used as a comparison method with the Multilayer Perceptron. This method uses the same split data set with Multi Layer Perceptron. The stage before testing is to run the validation function using K Fold Cross Validation with a total of 5 Folds.

Table 7. Table Random Forest K-Fold Cross Validation

| Ratio | Fold | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| 80:10:10 | 0.55 | 0.54 | 0.52 | 0.50 | 0.59 |

| | | | | | |
|---|---|---|---|---|---|
| 60:20:20 | 0.62 | 0.62 | 0.59 | 0.56 | 0.62 |
| 60:25:15 | 0.63 | 0.58 | 0.64 | 0.62 | 0.62 |
| 60:30:10 | 0.61 | 0.61 | 0.62 | 0.57 | 0.57 |
| 50:30:20 | 0.63 | 0.64 | 0.67 | 0.61 | 0.62 |

From the validation test, the highest average value lies in ratio 50:30:10 which is equal to 63%. Values from k-fold 1 to 5 have insignificant differences. Testing the accuracy of the Random Forest method using the Random Forest Classifier with 5 data set split data represent by confusion matrix on table as follows:

Table 8. Table Random Forest Confusion Matrix Ratio 80:10:10

| Actual/Prediction | Not Popular’ | Popular’ |
|---|---|---|
| Not Popular | 193 | 47 |
| Popular | 98 | 150 |

From the confusion matrix table, there are 193 not popular data and 150 popular data which are predicted as correct, while the other 155 data are predicted as wrong. Table 9 is also given a classification report that describes the value of precision, recall, and f1-score in each class.

Table 9. Table Random Forest Classification Report

| Ratio | Precision | Recall | F1 | Acc |
|---|---|---|---|---|
| **80:10:10** | 1 = 0.76<br>0 = 0.66 | 1 = 0.60<br>0 = 0.80 | 1 = 0.67<br>0 = 0.73 | **0.70** |
| 60:20:20 | 1 = 0.72<br>0 = 0.61 | 1 = 0.48<br>0 = 0.81 | 1 = 0.58<br>0 = 0.69 | 0.64 |
| 60:25:15 | 1 = 0.73<br>0 = 0.60 | 1 = 0.46<br>0 = 0.83 | 1 = 0.57<br>0 = 0.70 | 0.64 |
| 60:30:10 | 1 = 0.73<br>0 = 0.62 | 1 = 0.48<br>0 = 0.83 | 1 = 0.58<br>0 = 0.71 | 0.65 |
| 50:30:20 | 1 = 0.71<br>0 = 0.62 | 1 = 0.50<br>0 = 0.80 | 1 = 0.58<br>0 = 0.70 | 0.65 |

From the table 9, the highest accuracy is at ratio 80:10:10 split data with 70% accurate. The classification report for the not popular class has a precision of 76%, recall of 60% and an f1-score of 67%. Meanwhile, the popular class has a precision of 66%, a recall of 80% and an f1-score of 73%.

### *3.2.3 Comparison Method Between MLP and Random Forest*

After getting the accuracy results from the 2 algorithms, the next step is to compare the accuracy comparisons. Both methods use the same amount of data and balance in each separate data set. Comparison of Multi Layer Perceptron and Random Forest is presented in the following table:

Table 10. Table Comparison Accuracy

| | MLP | Random Forest |
|---|---|---|
| Accuracy | 76% | 70% |

From the results of the comparison of the method models, it can be seen that the highest accuracy is obtained by the Multi Layer Perceptron algorithm with MLP Classifier with an accuracy value of 76%. The difference in accuracy with the Random Forest algorithm is quite large, namely 6% because this algorithm reaches 70% accuracy for news headline data.

## 4. CONCLUSION

Based on the research that has been done, it can be concluded as follows, from the test results that have been described in detail in chapter V, the accuracy of the Multi Layer Perceptron algorithm is 76%. Comparison of accuracy is done with Random Forest which gives an accuracy of 70%. The difference in accuracy is 6%, so it can be concluded that for the Predicting News Popularity with Multi Layer Perceptron research, the algorithm with the best classification and accuracy results is achieved by Multi Layer Perceptron. Second, the News Popularity Prediction system is built with a website-based front-end so that it can be accessed flexibly and can assist writers in managing the right headlines for news content development. Third, predicting the popularity of news in the system is given the Setting feature, which is a feature that can choose the preferred method for prediction in the system. Prediction can be processed using Multi Layer Perceptron Method or Random Forest.

## *REFERENCES*

[1] J. Boumans, D. Trilling, R. Vliegenthart, and H. Boomgaarden, “The Agency Makes the (Online) News World Go Round: The Impact of News Agency Content on Print and Online News,” *Int. J. Commun.*, vol. 12, pp. 1768–1789, 2018.

[2] H. Ren and Q. Yang, “Predicting and Evaluating the Popularity of Online News,” *Conf. Proc.*, 2015, [Online]. Available: https://pdfs.semanticscholar.org/9e91/6a3469e9e2fc5f0c8f927d7d1d05f5575729.pdf%0Ahttp://cs229.stanford.edu/proj2015/328_report.pdf.

[3] P. Meesad, “Thai Fake News Detection Based on Information Retrieval, Natural Language Processing and Machine Learning,” *SN Comput. Sci.*, vol. 2, no. 6, pp. 1–17, 2021, doi: 10.1007/s42979-021-00775-6.

[4] J. Rezaeenour, M. Y. Eili, E. Hadavandi, and M. H. Roozbahani, “Developing a new hybrid intelligent approach for prediction online news popularity,” *Int. J. Inf. Sci. Manag.*, vol. 16, no. 1, pp. 71–87, 2018.

[5] D. Hardt, D. Hovy, and S. Lamprinidis, “Predicting news headline popularity with syntactic and semantic knowledge using multi-task learning,” *Proc. 2018 Conf. Empir. Methods Nat. Lang. Process. EMNLP 2018*, pp. 659–664, 2018, doi: 10.18653/v1/d18-1068.

[6] P. Rathord, D. A. Jain, and C. Agrawal, “A Comprehensive Review on Online News Popularity Prediction using Machine Learning Approach,” *Smart Moves J. Ijoscience*, vol. 5, no. 1, p. 7, 2019, doi: 10.24113/ijoscience.v5i1.181.

[7] H. A. Khoirunissa, A. R. Widyaningrum, and A. P. A. Maharani, “Comparison of Random Forest, Logistic Regression, and Multilayer Perceptron Methods on Classification of Bank Customer Account Closure,” *Indones. J. Appl. Stat.*, vol. 4, no. 1, p. 14, 2021, doi: 10.13057/ijas.v4i1.41461.

[8] F. Namous, A. Rodan, and Y. Javed, “Online News Popularity Prediction,” *ITT 2018 - Inf. Technol. Trends Emerg. Technol. Artif. Intell.*, no. Itt, pp. 180–184, 2019, doi: 10.1109/CTIT.2018.8649529.

[9] S. A. Salloum, M. Al-Emran, A. A. Monem, and K. Shaalan, “Using text mining techniques for extracting information from research articles,” *Stud. Comput. Intell.*, vol. 740, pp. 373–397, 2018, doi: 10.1007/978-3-319-67056-0_18.

[10] M. A. Rosid, A. S. Fitrani, I. R. I. Astutik, N. I. Mulloh, and H. A. Gozali, “Improving Text Preprocessing for Student Complaint Document Classification Using Sastrawi,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 874, no. 1, 2020, doi: 10.1088/1757-899X/874/1/012017.