

IMPLEMENTATION OF ARTIFICIAL NEURAL NETWORK WITH BACKPROPAGATION ALGORITHM FOR RATING CLASSIFICATION ON SALES OF BLACKMORES IN TOKOPEDIA

Dalfa Habibah Nurul Aini¹, Dian Kurniasari*², Aang Nuryaman³, Mustofa Usman⁴

^{1,2,3,4} Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung, Indonesia

Email: dalfahabibah@gmail.com, dian.kurniasari@fmipa.unila.ac.id, aang.nuryaman@fmipa.unila.ac.id, usman.alfha@yahoo.com

(Naskah masuk: 6 Agustus 2022, Revisi: 26 Agustus 2022, Diterbitkan: 23 Maret 2023)

Abstract

The rating assessment classification contains feedback from consumers, which is given in the form of stars which aims to assess a product. However, the amount of data in the classification process often have differences in each class or is called an imbalanced dataset. These problems can affect the classification results. An imbalanced dataset can be overcome by applying random oversampling. To classify the rating assessment, this study proposes the Neural network method, which has a good accuracy level with the backpropagation algorithm and applies random oversampling to overcome the unbalanced amount of data. The results indicate that the neural network method with the backpropagation algorithm can classify the available data with an accuracy level of 85%. The application of resampling data using random oversampling and determining the amount of distribution of training data, testing data, number of epochs and the correct number of batch sizes affect the results obtained.

Keywords: Artificial Neural Network, Backpropagation, Classification.

IMPLEMENTASI ARTIFICIAL NEURAL NETWORK DENGAN ALGORITMA BACKPROPAGATION UNTUK KLASIFIKASI PENILAIAN RATING PADA PENJUALAN BLACKMORES DI TOKOPEDIA

Abstrak

Klasifikasi penilaian rating berisi feedback dari konsumen yang diberikan dalam bentuk bintang yang bertujuan untuk menilai suatu produk. Akan tetapi, jumlah data pada proses klasifikasi seringkali terdapat perbedaan pada setiap kelasnya atau disebut *imbalanced dataset*. Permasalahan tersebut dapat mempengaruhi hasil klasifikasi. *Imbalanced dataset* dapat diatasi dengan menerapkan *random oversampling*. Untuk melakukan klasifikasi penilaian rating, penelitian ini mengusulkan metode *Neural network* yang merupakan salah satu metode klasifikasi yang memiliki hasil tingkat akurasi yang baik dengan algoritma *backpropagation* dan menerapkan *random oversampling* untuk mengatasi jumlah data yang tidak seimbang. Hasil yang didapatkan menunjukkan bahwa metode *neural network* dengan algoritma *backpropagation* mampu melakukan klasifikasi data yang tersedia dengan tingkat *accuracy* sebesar 85%. Penerapan *resampling* data menggunakan *random oversampling* serta penentuan jumlah pembagian data *training*, data *testing*, jumlah *epoch* dan jumlah *batch size* yang tepat mempengaruhi hasil yang diperoleh.

Kata kunci: Artificial Neural Network, Backpropagation, Klasifikasi.

1. INTRODUCTION

The existence of a pandemic in Indonesia from the beginning of 2020 encouraged an increase in buying and selling transactions for health products such as hand sanitisers, masks, face shields, and supplements containing vitamins. However, these activities were hampered due to the implementation of social distancing to minimise the spread of covid,

so many human activities are currently carried out digitally, such as buying and selling transactions offered through marketplaces or e-commerce. One of the e-commerce marketplaces is tokopedia.

Tokopedia is a company that offers its products via the internet online. Tokopedia provides several features that can provide information about the goods that consumers will buy. One of the features

that can be used as a reference for consumers to buy a product is rating.

Rating is feedback from consumers given in the form of stars which aims to assess a product. The higher the rating on a product, the higher the buying interest of consumers. Therefore, it is necessary to conduct research that can help business actors increase the sales rating on products to increase consumer buying interest later. One way that can be used to overcome these problems is by classifying the sales rating of a product.

Classification is a process of finding a model or function that can describe and distinguish data classes or concepts, with the aim that the model can be used to predict an unknown class of an object of observation [1]. The amount of data in the classification process often experiences differences in each class or is commonly called an imbalanced dataset. These problems can affect the results of the classification obtained. Imbalanced datasets can be overcome by applying random oversampling and increasing the number in the minority class until it is equal to the majority class, so the accuracy results obtained in the classification process will be better.

A neural network is one method with a prediction classification algorithm with reasonably high accuracy. The method is implemented using a computer program so that it can complete several calculation processes [2]. A neural network works like a human brain neural network where the method has interconnected neurons to process data. In the neural network algorithm, backpropagation is included in the multi-layer perceptron, where the algorithm has a hidden layer between the input and output layers. Algorithm backpropagation has three stages in the process, namely the forward stage, backward stage, and weight modification, where the three stages are carried out repeatedly until the loss with a good level of accuracy.

Research on classification using neural networks has been carried out previously, such as the classification of diabetes [2], the classification of bank customer loans [3], and the classification of bitcoin sentiment analysis [4]. The classification of rating ratings on Blackmores sales at tokopedia using the neural network algorithm backpropagation has never been done. Therefore, a rating assessment classification will be carried out on Blackmores at tokopedia using the neural network algorithm backpropagation, which aims to get a good level of accuracy by applying random oversampling.

2. RESEARCH METHODS

The data used in this study is rating assessment data on Blackmores products obtained from <https://www.kaggle.com/>. There are five variables obtained from the website, namely product prices, product stock, official stores (stores that sell original products) in the form of categorical data (true and false), power badges (stores that subscribe to power

merchants on tokopedia) in the form of categorical data. (true and false). The following variable is rating. The rating consists of six parts, namely rating 0 (the product has no rating) with a total of 11768 data, rating 1 with a total data of 2, rating 2 with a total of 5 data, rating 3 with a total data of 7, rating 4 with a total number of data. as many as 47, and the last is a rating of 5 with a total data of 844. Initial data in this study amounted to 12673 data, and after resampling the data using random oversampling, the data would amount to 70608 data, with each rating amounting to 11768 data. The steps in this study are:

- Perform data input into google collab and select the data to be used in the study, with the variable X_1 = price on the product, X_2 = product stock, X_3 = official store, X_4 = store gets power badge and Y = rating on the product.
- Perform labelling to convert categoric data into numeric data using encoder labels.
- Perform the data transformation process using z-score normalisation on the variables X_1 and X_2 .
- We are applying random oversampling to overcome classes that have an unbalanced amount of data.
- Distribute training data and testing.
- Build the model and determine the number of epochs and batch size.
- Test and evaluate the model to see the level of accuracy obtained in the study.

Some of the theories that will be used are as follows:

2.1 Machine learning

Machine learning technology is the latest learning developed to channel human knowledge and reasoning in forming a model that can automatically build machines and system engineering (can learn by itself without direction from other users) [5].

2.2 Knowledge Discovery in Database

The KDD process can be broadly explained as follows [6]:

1. Data Selection

It is the process of selecting data to ensure the processing will be better by achieving the research targets. The research target is divided into X data input and Y data output. The input data includes five variables: prices on products, product stock, products from official stores, and products from power badge stores, while the output data is product ratings.

2. Data Pre-processing

Pre-processing data includes, among others, removing duplicate data, checking for inconsistent data, and correcting errors in data, such as printing errors and missing values.

a. Missing value

The missing value is one of the main problems in pre-processing data. Missing values can be found due to human error, machine error, or lack of updates to the data. The most popular method of overcoming the missing value is to change it to mean and mode [7].

b. Categorical Encoding

Two encodings are often used: a label encoder and one hot encoder [7]. Encoder labels are appropriate when the labels have different grades. The following is an example of using an encoder label in Table 1.

Table 1. Example of Label Encoder

Size	Label Encoder
Small	1
Medium	2
Big	3

3. Data Transformation

Data transformation is changing the measurement scale from the original form into another. Some methods used in data transformation are min-max normalisation, z-score normalisation, and decimal scaling. In this article, the z-score normalisation is used as a normalisation method based on the data's mean (mean value) and standard deviation [8]. The equation for z-score normalisation is:

$$Z = \frac{x - \mu}{\sigma} \quad (1)$$

4. Data Mining

Data mining is finding interesting patterns or information in selected data using specific techniques or methods.

5. Interpretation / Evaluation

Information patterns generated from the data mining process need to be displayed in a form that interested parties easily understand.

2.3 Resampling Data

Some data has imbalanced, which happens because some data classes are biased, or some classes have a lot of data and others have little data. Problems with imbalanced datasets can be overcome using resampling.

The resampling technique is divided into 3: oversampling, undersampling, and a hybrid of the two techniques. Oversampling is used to add the amount of data in the minority class until the amount of data in the minority class and majority class is balanced, for example, random oversampling. Undersampling eliminates some data in the majority class considered less relevant, such as random undersampling. A hybrid is a combination of the two samplings according to the needs and characteristics of the data [9].

2.3.1 Random Oversampling

Random oversampling is one way to overcome an imbalanced dataset by balancing the amount of data on the majority class and minority class. In contrast to random undersampling, which randomly eliminates data in the majority class, the random oversampling replication will be carried out on the data in the minority class until the minority class gets the same amount of data as the majority class.

2.3.2 Random Undersampling

Random undersampling is one way to overcome an imbalanced dataset by eliminating data. Random undersampling will perform random elimination on the data in the majority class until the majority class gets the same amount of data as the minority class.

2.4 Artificial Neural Network

One of the most commonly used methods for classification is Artificial Neural Network. Inspired by the human brain, ANN also consists of several neuron connections between neurons [10]. The design of the ANN algorithm on machine learning has two data differences, namely training (train/training) and testing/prediction (test/testing) [5].

Then the artificial neural network is determined by three things [11], namely:

1. The pattern of relationships between neurons (network architecture)
2. Training method
3. Activation function

2.5 Activation Function

The activation function is used to activate neurons [10].

a. Activation Function Rectified Linear Unit (ReLU)

ReLU is an activation function used to normalise the values generated by the ReLU. Every input given will always be mapped to 0, and every input will be retained its value, so there are no negative results. The equation for this activation function is:

$$f(x) = \max(0, x) \quad (2)$$

The graph of the ReLU activation function can be seen in Figure 1.

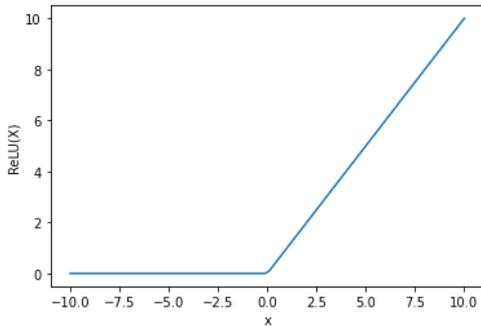


Figure 1. ReLU Activation Function.

b. Softmax Activation Function

Softmax is an activation function used for multiclass. The output of the activation function is an opportunity value in each class with a range of 0 to 1. If the opportunity values are added, the value will produce a value equal to 1. The following is the equation for the softmax.

$$S_k = \frac{e^{y_k}}{\sum_{k=1}^m e^{y_k}} \quad (3)$$

2.6 Backpropagation Algorithm

The backpropagation algorithm is one method that is very good in dealing with the problem of recognising complex patterns [12]. The backpropagation algorithm is part of the multi-layer perceptron, so a hidden layer must be used and noticed in the network structure. The backpropagation technique supervised learning that has a target as output by changing and adjusting the value of the weights connected to the neurons in the hidden layer. Backpropagation has two main ways of error-correction learning: forward and backward movement. In the forward movement process, the input vector is applied to the input layer in each network, which will affect all networks on a layer-by-layer basis.

Furthermore, the neurons are activated at this stage using the activation function. Then the output error or the difference between the values after being activated or the target is obtained. Then the error will be minimised by changing the weights and biases generated by moving backwards from the output previously generated. The weight of each layer in the forward movement is fixed, while in the backward movement, the weight in each layer changes based on the error correction rule [3]. Here is the equation in the advanced stage:

$$Y_k = f_2 [W_m + \sum_{j=1}^p W_{jk} \cdot f_1 [V_p + \sum_{i=1}^n V_{ij} X_i]] \quad (4)$$

The training process on the backpropagation includes the following steps:

1. The Advanced Stage

The input layer will receive the initial data $X_i, i = 1, 2, \dots, n$, which will be forwarded to

the hidden layer, and then each node in the hidden layer $Z_j, j = 1, 2, \dots, p$ will add up the weighted signals,

$$Z_j = V_p + \sum_{i=1}^n V_{ij} X_i \quad (5)$$

use the activation function to calculate the output in the hidden layer.

$$Z_j = f_1 \cdot z_j \quad (6)$$

Then the signal will be forwarded to the *output layer* and each *node* in the *output layer* $Y_k, k = 1, 2, \dots, m$ will add up the weighted signals,

$$y_k = [W_m + \sum_{j=1}^p W_{jk}] \cdot Z_j \quad (7)$$

use the activation function to calculate the output in the output layer.

$$Y_k = f_2 \cdot y_k \quad (8)$$

2. The Reverse Stage

Each *node* in the *output layer* $Y_k, k = 1, 2, \dots, m$ will receive the pattern target corresponding to the input pattern and each hidden layer $Z_j, j = 1, 2, \dots, p$ will calculate δ_j They were then used to calculate the corrected weight and bias between the input and hidden layers.

$$\frac{\partial E}{\partial W} = -(t_j - Y_k) * Y_k(1 - Y_k) * X_i \quad (9)$$

3. Modify the weight

of each *node* in the *output layer* $Y_k, k = 1, 2, \dots, m$ will update the weight and bias values $j = 1, 2, \dots, p$ and each *node* in the hidden layer updates the weight and bias values $i = 1, 2, \dots, n$ so that the new weight and bias values that are obtained.

$$Bobot_{baru} = Bobot_{lama} - \eta \frac{\partial E}{\partial W} \quad (10)$$

$$Bias_{baru} = Bias_{lama} - \eta \frac{\partial E}{\partial W} \quad (11)$$

2.7 Learning Rate, Epoch and Batch Size

Learning rate is one of the parameters used for the backward stage in backpropagation that affects the system to reduce the loss process training.

Epoch is the number of iterations during the training process that provides input from the network and updates the network weights [13].

Batching is one of the common approaches to speed up the computation of neural networks. This process involves calculating a gradient containing multiple pieces of training in one feed-forward/feedback [14].

2.8 Confusion Matrix

A confusion matrix is a matrix that is used to display the classification predictions and the actual classification. The confusion matrix with 2x2 dimensions consists of rows and columns where the row is the prediction class, and the column is the actual class. The class used in the fundamental matrix is divided into positive and negative classes

[15]. The confusion matrix can be seen in figure 2 below:

		Predicted Values	
		POSITIVE	NEGATIVE
Actual Values	POSITIVE	TP (True positive)	FN (False negative)
	NEGATIVE	FP (False positive)	TN (True negative)

Figure 2. Basic of Confusion Matrix.

The classification model will obtain accuracy, precision, recall and f1-score from the confusion matrix. Accuracy is the ability of the classifier to predict the class correctly [16]. Then, precision is the level of accuracy between the information requested by the researcher and the output provided by the system. Furthermore, recall is the success rate of the system in retrieving information. The f1-score compares the average precision value with the recall that can be used as a reference in seeing the accuracy of the prediction results and the actual value when the amount of FN and FP data differs.

$$Accuracy = \frac{TP+TN}{Total} \quad (12)$$

$$Precision = \frac{TP}{TP+FP} \quad (13)$$

$$Recall = \frac{TP}{TP+FN} \quad (14)$$

The multiclass confusion matrix classification is different from the fundamental confusion matrix, where the multiclass confusion matrix classification has positive and negative classes. The picture multiclass confusion matrix is as follows:

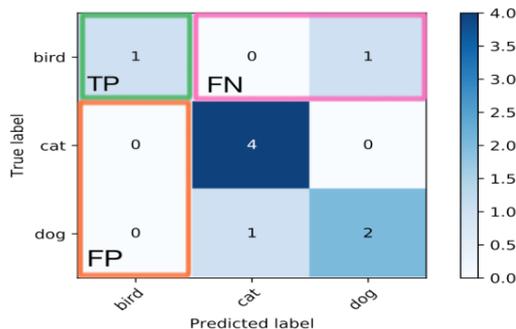


Figure 3. Multiclass Confusion Matrix

3. RESULTS AND DISCUSSION

The first step that needs to be done is pre-processing the data. At this stage, categorical encoding will be carried out, a labelling process to convert categorical data into numeric data using a label encoder. Actual will be labelled one, and false will be labelled 0. Next, the data will be transformed

using z-score normalisation on the variables X_1 , and X_2 . transformed data can be seen in table 2.

Table 2. Data Transformation

X_1	X_2	X_3	X_4
-0,761505	0,97584015	0	1
-0,5170831	0,96510765	0	1
-0,761505	-1,02758042	1	1
...
-0,84297897	0,97584015	0	0
-1,10835132	-1,03371328	0	0
-0,86625725	0,97584015	0	0

After transforming the data, the next step is to check the number of members in each class on the Y variable. As seen in figure 4, the data used is still *unbalanced*, with a striking difference in the number of ratings per class. So it is necessary to *resample the data* to get good classification results. It can be seen in figure 4, which is a plot of rating assessment data that has not applied *random oversampling*.

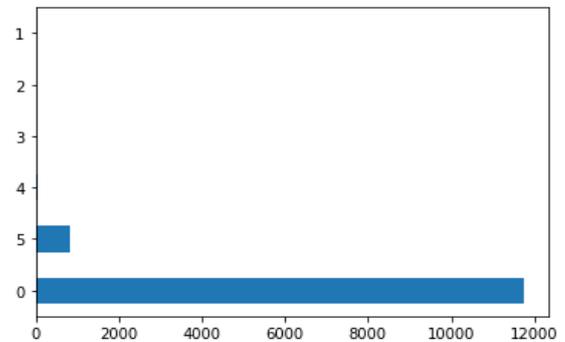


Figure 4. The plot of Rating Data.

To overcome the problem of unbalanced *random oversampling* data, where this method duplicates the amount of data that is lacking. Can be seen in table 3 is the result of *resampling* data.

Table 3. Resampling Data

Rating	Imbalanced Data	Balanced Data
0	11768	11768
1	2	11768
2	5	11768
3	7	11768
4	47	11768
5	844	11768

Before random oversampling, the data amounted to 12673 with products that have not received a rating assessment or a rating of 0, which has the most significant amount of data of 11768 data and a product with a rating of 1 that has the smallest amount of data of 2 data. Technique random oversampling, the data will be 70608, with each class consisting of 11768 data.

After resampling the data, the data will be divided into two parts: training data and testing. It can be seen in table 4 that the accuracy obtained from the percentage distribution of training data and testing is different.

Table 4. Distribution of Training Data and Testing Data

Distribution Data	Training	Testing	Accuracy
60% dan 40%	42364	28244	0,81
70% dan 30%	49425	21183	0,83
75% dan 25%	52986	17652	0,84
80% dan 20%	56486	14122	0,83
90% dan 10%	63547	7061	0,85

In the table above, the accuracy is 85% with the data sharing of 90% training and 10% testing, then 63547 data will be used for training and 7061 for testing. After the training data distribution and testing, the next step will be constructing artificial neural networks. Model neural network algorithm backpropagation is built using one input layer with four nodes, three hidden layers with 128 nodes for the hidden layer, 64 nodes for the hidden layer and 32 nodes for the hidden layer the third output layer with six nodes, namely the rating on the product. Figure 5 is an image of the backpropagation built in this article.

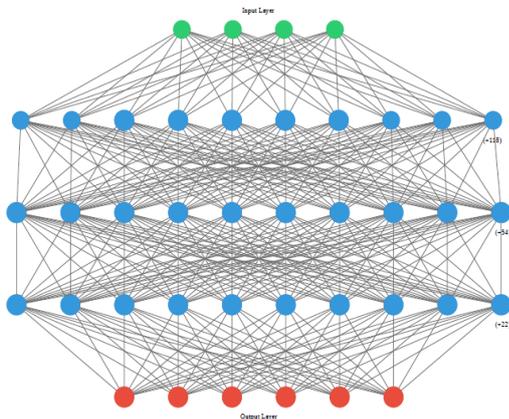


Figure 5. Backpropagation Architecture

In addition to the number of input, hidden, and output layers, an activation function is also needed to build an artificial neural network algorithm backpropagation. The ReLU activation function is used for the input layer, and the hidden layer activation function softmax is used for the output layer.

The next stage is determining the best parameter values by testing hyper-tuning. Parameters tested are learning rate, epoch and batch size. Learning rate was tested with values of 0.001, 0.01, and 0.1. The number of epochs is tested with values of 50, 100, and 1000. Furthermore, the batch sizes tested were 16, 32, and 64. After hyper tuning, the loss is 0.4129, and the learning rate value is obtained, the epoch value and the batch size, namely learning rate with a value of 0.001, epoch with a value of 50 and batch size with a value of 32. It can be seen in table 5-table 7 the difference in accuracy obtained with the number of learning rates, the number of epochs and the number of batch sizes differently.

Table 5. Comparison of Accuracy with Learning Rate 0,001

Epoch	Batch Size	Accuracy
50	16	0,83
50	32	0,84
50	64	0,80
100	16	0,82
100	32	0,80
100	64	0,83
1000	16	0,83
1000	32	0,80
1000	64	0,81

Table 6. Comparison of Accuracy with Learning Rate 0,01

Epoch	Batch Size	Accuracy
50	16	0,77
50	32	0,82
50	64	0,83
100	16	0,74
100	32	0,83
100	64	0,80
1000	16	0,81
1000	32	0,79
1000	64	0,83

Table 7. Comparison of Accuracy with Learning Rate 0,1

Epoch	Batch Size	Accuracy
50	16	0,17
50	32	0,17
50	64	0,17
100	16	0,25
100	32	0,17
100	64	0,17
1000	16	0,17
1000	32	0,25
1000	64	0,17

The next step is to test the model. At this stage, the loss smallest epoch worth 0.4129 with accuracy is 0.8448 and the validation loss worth 0.4381 with a validation accuracy of 0.8359. The minor loss obtained, the better the resulting classification, while the more significant the accuracy obtained, the better the resulting classification. The following is a graph of the model accuracy on data that does not apply random oversampling, presented in figure 6.

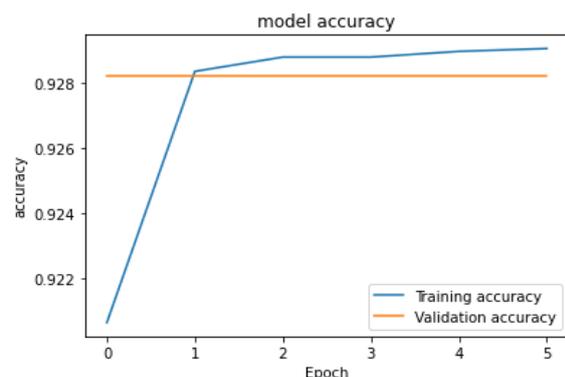


Figure 6. Graph of Model Accuracy Without ROS.

It can be seen in picture 6 the graph model accuracy on data that does not apply random oversampling experiences overfitting. That is when the training can classify the data well, but the testing cannot classify the data correctly because the amount of data is not balanced in each class. Then the following is a graph of a loss that does not apply random oversampling, presented in figure 7.

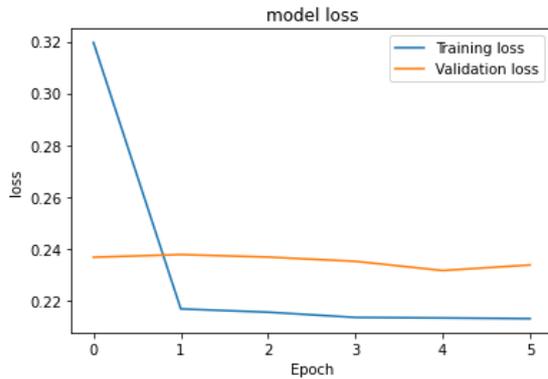


Figure 7. Graph of a Loss without ROS.

From that figure, model graph loss is on data that does not apply random oversampling experiences overfitting, when training can reduce errors in classifying data correctly. However, data testing cannot reduce errors in classifying data correctly because the amount of data is not balanced in each class. Then the following is a graph of the accuracy that applies random oversampling, which is presented in figure 8.

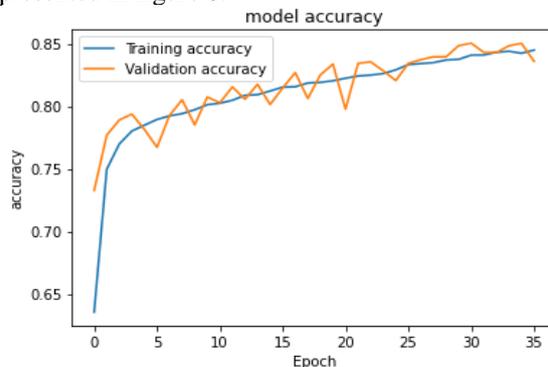


Figure 8. Graph of the Accuracy ROS.

Figure 8 shows that the accuracy of the graph model on the data uses random oversampling and does not experience overfitting. The accuracy increases from 64% to 84%. The following is a graph of losses that apply random oversampling, which is presented in Figure 9

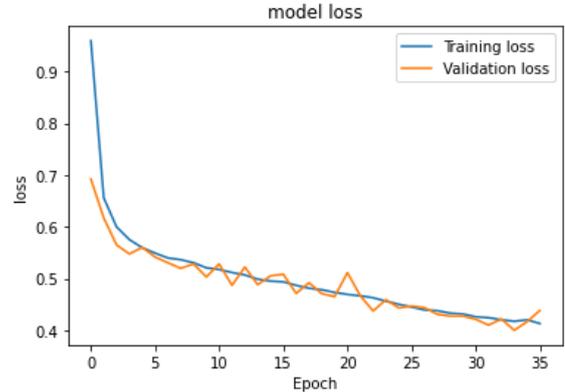


Figure 9. Graph of the Loss ROS.

In the picture above, it can be seen that the graph of the loss on the data that applies random oversampling does not experience overfitting because the training and data testing can reduce errors in classifying data correctly where the loss obtained in training starts in the range of 0.9 and continues to move down to the range of 0.41 in epoch the 36/50 loss obtained in the validation loss starts in the range of 0.6. It continues to move until epoch the 36/50 validation loss up to 0.43.

The last stage is the evaluation of the model using a confusion matrix, and an accuracy value of 84% is obtained, which means the level of proximity of the predicted class to the actual class is 84%. The following is an image of the confusion matrix in the model.

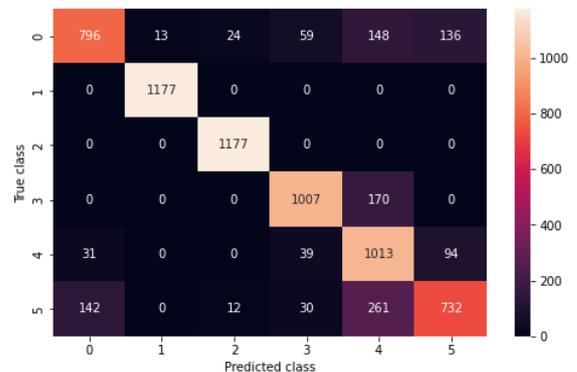


Figure 10. Confusion Matrix.

In addition to getting the accuracy value, it also obtained precision value, recall value and f1-score, which are as follows:

Table 8. Precision, Recall, and F1-score

Rating	Precision	Recall	F1-score
0	0,82	0,68	0,74
1	0,99	1	0,99
2	0,97	1	0,98
3	0,89	0,86	0,87
4	0,64	0,86	0,73
5	0,76	0,62	0,68

4. CONCLUSION

Based on the results and discussion of the method of artificial neural network algorithm Backpropagation for classification of rating ratings on Blackmores at tokopedia, it is concluded that :

1. The application of resampling using random oversampling affects the results of the study because data that did not apply random oversampling experienced overfitting. In contrast, data that applied random oversampling did not experience overfitting model graph accuracy and graph of loss obtained on research. Then several aspects used in this study, such as determining the amount of training data and testing data, learning rate, number of epochs and batch sizes, can affect the study's results.
2. By dividing 90% of data for training and 10% for testing, value can be obtained accurately compared to the percentage of other data sharing.
3. By applying hyper tuning in determining the learning rate, the number of epochs and batch size, learning rate, epoch and batch size, namely with a learning rate of 0.001, the number of epochs is 50, and the number of batch sizes is 16.
4. Furthermore, it can be concluded that the artificial neural network algorithm is backpropagation able to classify rating ratings on Blackmores on tokopedia well by getting an accuracy of 84%. So that by using the available data later, the model can classify the rating assessment on the product correctly.

5. REFERENCES

- [1] P. R. Sihombing and A. M. Arsani, "Perbandingan Metode *Machine learning* Dalam Klasifikasi Kemiskinan Di Indonesia Tahun 2018," *Jurnal Teknik Informatika.*, vol. 2, no. 1, pp. 51-56, 2021, doi: 10.20884/1.jutif.2021.2.1.52.
- [2] J. J. Purnama, S. Rahayu, S. Nurdiani, T. Haryanti, and N. A. Mayangky, "Analisis Algoritma Klasifikasi *Neural Network* Untuk Diagnosis Penyakit Diabetes," *Indonesian Journal on Computer and Information Technology.*, vol. 5, no. 1, pp. 1-7, 2020.
- [3] N. Hadianto, H. B. Novitasari, and A. Rahmawati, "Klasifikasi Peminjaman Nasabah Bank Menggunakan Metode *Neural Network*," *Jurnal PILAR Nusa Mandiri.*, vol. 5, no. 2, pp. 163-170, 2019, doi: 10.33480/pilar.v15i2.658.
- [4] I. T. Julianto, D. Kurniadi, M. R. Nashrulloh, and A. Mulyani. Comparison of Classification Algorithm and Feature Selection in Bitcoin Sentiment Analysis. *Jurnal Teknik Informatika.* vol. 13, no. 3, pp. 739-744. 2022.
- [5] M. P. Deisenroth, A. A. Faisal, and C. S. Ong, *Mathematics for Machine Learning*. Cambridge University Press England.
- [6] Bramer and Max, *Principles of Data Mining*. Springer London.
- [7] N. Sharma, H. V. Bhandari, N. S. Yadav, and H. V. J. Shroff, "*Optimisation of IDS using Filter-Based Feature Selection and Machine learning Algorithms*" *International Journal of Innovative Technology and Exploring Engineering.*, vol. 10, no. 2, pp. 96-102, 2020.
- [8] D. A. Nasution, H. H. Khotimah, and N. Chamidah, N, "Perbandingan Normalisasi Data Untuk Klasifikasi Wine Menggunakan Algoritma K-NN." *Journal of Computer Engineering System and Science.*, vol. 4, no. 1, pp. 78-82, 2019, doi: 10.24114/cess.v4i1.11458.
- [9] A. M. Mahmood, "*Class Imbalance Learning in Data Mining – A Survey.*" *International Journal of Communication Technology for Social Networking Services.*, vol. 3, no. 2, pp. 17-38, 2015, DOI: 10.21742/ijtsns.2015.3.2.2.
- [10] A. W. Putri, A.W, "Implementasi *Artificial Neural Network* (ANN) *Backpropagation* Untuk Klasifikasi Jenis Penyakit Pada Daun Tanaman Tomat.," *Jurnal Ilmiah Matematika.*, vol. 9, no. 2, pp. 344-350, 2021, doi: 10.26740/mathunesa.v9n2.p344-350.
- [11] E. Y. Susantia and E. Maiyana, "Pemanfaatan ANN untuk Prediksi Penjualan *Online* Industri Rumahan selama Pandemi Covid-19," *Jurnal Sains dan Informatika.*, vol. 7, no. 1, pp. 1-7, 2021, doi: 10.22216/jsi.v7i1.234.
- [12] J. R. Prabowo, R. Santoso, and H. Yasin, "Implementasi Jaringan Syaraf Tiruan *Backpropagation* Dengan Algoritma *Conjugate Gradient* Untuk Klasifikasi Kondisi Rumah (Studi Kasus Di Kabupaten Cilacap Tahun 2018)," *Jurnal Gaussian.*, vol. 9, no. 1, pp. 41-49, 2020, doi: 10.14710/j.gauss.v9i1.27522.
- [13] V. Vijayalakshmi and K. Venkatachalapathy, "*Deep Neural Network for Multi-Class Prediction of Student Performance in Educational Data.*" *International Journal of Recent Technology and Engineering (IJRTE).*, vol. 8, no. 2, pp. 5073-5081, 2019, DOI: 10.35940/ijrte.B2155.078219.
- [14] N. D. Lewis, *Neural Network for Time Series Forecasting with R*. CreateSpace Independent Publishing Platform US.
- [15] A. A. Arifiyanti, R. M. Pradana, and I. F. Novian, "Klasifikasi produk retur dengan algoritma pohon keputusan C4.5," *Jurnal IPTEK.*, vol. 22, no. 1, pp. 79-86, 2018, doi: 10.31284/j.ipitek.2018.v22i1.243.
- [16] J., Lever, M., Krzywinski, and N., Altman, "Classification evaluation," *Nat Methods*, vol. 13, pp. 603–604, 2016, doi: 10.1038/nmeth.3945.