Prediction of Open Unemployment Rate of Tuban Regency in 2022 with Backpropagation Method

Evi Septya Putri¹, Abdulloh Hamid², Dian Candra Rini Novitasari³

¹, ², ³ Sunan Ampel State Islamic University Surabaya, Indonesia
Correspondence: eviseptya557@gmail.com, diancrini@uinsby.ac.id

<table>
<thead>
<tr>
<th>Article Info</th>
<th>Abstract</th>
</tr>
</thead>
<tbody>
<tr>
<td>Received: 14-01-2022</td>
<td>Unemployment is also included in the severe and primary labor problem. No matter how poor or rich the country is, it will still experience unchecked unemployment problems. The Open Unemployment Rate is the number of unemployed from the total labor force in an area or region in the form of a percentage. In Tuban Regency, unemployment has experienced a significant increase from 2019 to 2020. Therefore, it is necessary to make predictions in 2021 and 2022 to find out the description of the Open Unemployment Rate in that year so that the government can take actions to deal with this. By using the Backpropagation method, this study obtained predictions that in 2021 it will be 4.58% of the total workforce in Tuban Regency 2021, and in 2022, the open unemployment rate will be 4.7% of the entire workforce in Tuban Regency 2022, with network architecture 2-70-1-1, the learning rate used is 0.1, and the MAPE value is 8.363%. The test was stopped at the 276th epoch. At the time of testing between the output and the target, there was a correlation (R) of 0.993, where the best result was 1. The results of forecasting the Open Unemployment Rate of Tuban Regency in 2020-2022 showed an increase from 2020 of 4.39% the number of Labor Force until 2022 4.7% the number Labor Force.</td>
</tr>
<tr>
<td>Revised: 19-04-2022</td>
<td></td>
</tr>
<tr>
<td>Accepted: 28-05-2022</td>
<td></td>
</tr>
<tr>
<td>Keywords:</td>
<td></td>
</tr>
<tr>
<td>Backpropagation Method; Predictions; Open Unemployment; Tuban Regency</td>
<td></td>
</tr>
</tbody>
</table>

INTRODUCTION

In this era of globalization, unemployment has become a global issue and phenomenon worldwide [1]–[4]. In addition, unemployment is also included in the severe and primary labor problem; no matter how poor or rich the country is, it will still experience unchecked unemployment problems [5][6]. Unemployment is a labor force person looking for a full-time (not part-time) job but does not earn it. Unemployment is divided into unemployment against his will and unemployment that is not against his wishes. Unemployment on his wish is unemployment obtained for his desire to find a new job while unemployment is not on his passion, namely unemployment brought from job cuts so that it is required to find a new job [1].

Unemployment Rate or TPT is the number of unemployed from the number of employment in one region or region in percentages [5]. The category of individuals who fall into the Open Education Level are individuals who are trying to find work but do not earn it and do not have a fixed income but have expenses, while children, retirees, students, part-time workers, and people who are not looking for work are not included in the open unemployment rate [1]. If there is no
handling of unemployment, the problem will continue to grow in each period, as happened in Tuban Regency of East Java Province, where the labor force and unemployment rate increase every year. In 2019 the workforce in Tuban Regency had as many as 645,156 people with an Open Unemployment Rate of 2.76%, while in 2020 the crew of 677,759 people with the OpenExrate Nomination Rate of 4.81% means that the number of Open Departure Rates in Tuban Regency will increase [7]. Even according to [8], unemployment dramatically affects the amount of poverty. Therefore, it stems from a problem that requires open unemployment level data from time to time as a basis for taking a policy to overcome unemployment by carrying out predictions that are considered suitable to be used in solving the problem of increasing the number of unemployed in Tuban Regency [9]. Prediction is a guessing or estimating activity entwined in the future that wants to arrive at later information that uses scientific procedures [10]. The prediction in this study is the prediction of the Open Unemployment Rate of Tuban Regency in 2021 and 2022.

Prediction of open unemployment rates can be made by various methods, such as Simple Linear Regression [5], Regression Tree [11], ANFIS [12], and Backpropagation [6]. While in this study uses the Backpropagation method because it has been proven to work well in various problems, such as regression, pattern recognition, and prediction [13]. As for the predictions that have been done using the Backpropagation method, namely [14] With, the best architectural network model is 18-5-1, obtained MSE of 0.000998685. On research [15], the best architectural model is 4-19-1 with an MSE of 0.00099982. Study [16] Acquired the best architectural model 4-50-1 with an MSE of 0.000997867. Predictions [17] obtained the best architectural model 4-14-1 with an MSE of 0.00274166, and in research [18], The best architectural model is 12-2-1 with an MSE of 0.04149487.

From the information obtained, the author wants to predict the Open Unemployment Rate of Tuban Regency with the Backpropagation method because it can set the architectural model to get the best model with a minor error. So the model obtained will be more accurately used to make predictions.

METHOD

This study consists of several steps, namely collecting data obtained from https://tubankab.bps.go.id. The data that has been accepted will then be normalized and form a time series pattern, then divided into 70% for training data, namely in 2001-2013, or data that will be used for the search for the best model and 30% data testing in 2014-2019, namely to represent the model that has been formed from training data before being represented again. After the model is used to predict, the results of the predictions used are data in 2020-2021. After obtaining the following results, denormalization and data testing are carried out. Other method explanations are as follows:

Data Collection

This study uses Tuban regency open unemployment rate data from 2001-202. Dataadiper obtained from the Office of The Central Statistics CenterKabupaten Tuban, East Java Province.
Table 1. Tuban Regency Open Unemployment Rate Data for 2001-2021

<table>
<thead>
<tr>
<th>Nbr.</th>
<th>Year</th>
<th>TPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2001</td>
<td>3.41</td>
</tr>
<tr>
<td>2</td>
<td>2002</td>
<td>4.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>2020</td>
<td>4.81</td>
</tr>
<tr>
<td>10</td>
<td>2021</td>
<td>4.68</td>
</tr>
</tbody>
</table>

**Data Normalization**

The search data is in the table. 1 is normalized using equations (2.1), after which it creates an information value of a range between 0-1. But the value of data must not be equal to or less than zero and must not be similar to or more than one because it is a condition of normalization of information.[19].

\[
X' = \frac{0.8(x-a)}{b-a} + 0.1 \tag{2.1}
\]

Information:
- \(X'\) = data that has been normalized
- \(X\) = data to be normalized
- \(a\) = lowest data
- \(b\) = highest data

**Time Series Shape**

Data is divided into two variables: time series and one target data. After that, the information is built in a pattern of sorts in table 3 as follows: [6]

Table 2 Time Series Pattern Formation Schemes

<table>
<thead>
<tr>
<th>Nbr.</th>
<th>(X_1)</th>
<th>(X_2)</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Data 2001</td>
<td>Data 2002</td>
<td>Data 2003</td>
</tr>
<tr>
<td>2</td>
<td>Data 2002</td>
<td>Data 2003</td>
<td>Data 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Data 2018</td>
<td>Data 2019</td>
<td>Data 2020</td>
</tr>
<tr>
<td>4</td>
<td>Data 2019</td>
<td>Data 2020</td>
<td>Data 2021</td>
</tr>
</tbody>
</table>

**Data Sharing Training and Testing**

Normalized data is then divided into 2 data. Training in 2001-2013 or as much as 70% and testing in 2014-2019 or 30% [20].
Table 3. Data Training

<table>
<thead>
<tr>
<th>Year</th>
<th>(X_1)</th>
<th>(X_2)</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>2002</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>2003</td>
<td>0.38</td>
<td>0.38</td>
<td>0.38</td>
</tr>
<tr>
<td>2004</td>
<td>0.90</td>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>2005</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>2006</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>2007</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
</tr>
<tr>
<td>2008</td>
<td>0.56</td>
<td>0.56</td>
<td>0.56</td>
</tr>
<tr>
<td>2009</td>
<td>0.32</td>
<td>0.32</td>
<td>0.32</td>
</tr>
<tr>
<td>2010</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>2011</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>2012</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>2013</td>
<td>0.34</td>
<td>0.34</td>
<td>0.34</td>
</tr>
</tbody>
</table>

Data training is data that will be used to form the Backpropagation model.

Table 4. Data Testing

<table>
<thead>
<tr>
<th>Year</th>
<th>(X_1)</th>
<th>(X_2)</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>0.23</td>
<td>0.14</td>
<td>0.34</td>
</tr>
<tr>
<td>2015</td>
<td>0.14</td>
<td>0.34</td>
<td>0.20</td>
</tr>
<tr>
<td>2016</td>
<td>0.34</td>
<td>0.20</td>
<td>0.11</td>
</tr>
<tr>
<td>2017</td>
<td>0.20</td>
<td>0.11</td>
<td>0.10</td>
</tr>
<tr>
<td>2018</td>
<td>0.11</td>
<td>0.10</td>
<td>0.41</td>
</tr>
<tr>
<td>2019</td>
<td>0.10</td>
<td>0.41</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Data testing is data used to simulate models that have been formed from training data.

**Backpropagation Algorithm**

Backpropagation is a straightforward iterative algorithm that works well, even when complex data is used. Here is the algorithm from Backpropagation.[21] :

**Backpropagation**

1. Initialize weights and biases.
2. When the conditions stop unfulfilled, stages 3-10 are carried out.
3. For each training pair, stages 4-9 are carried out.

**Feedforward Stage**

4. Calculate the value on the remote unit \(z_j\). Then the \(z_j\) value is entered into the activation function with the binary sigmoid.

\[
z_{inj} = v_{0j} + \sum_i x_i v_{ij}
\]

\[
z_j = \frac{1}{1 + \exp(-z_{inj})}
\]

(2.1)  
(2.2)
5. Charge the unit output value \((y_k)\). Then the value of \(y_k\) is entered into the binary sigmoid activation function.

\[
y_{\text{in}}_k = w_{0j} + \sum_l z_l w_{lj}
\]
\[
y_k = \frac{1}{1 + \exp(-y_{\text{in}}_k)}
\]

(2.3) (2.4)

**Backpropagation Error Stage**

6. Calculate the value of \(\delta\) in the unit output to determine the error rate.

\[
\delta_k = (t_k - y_k)f'(y_{\text{in}}_k)
\]

(2.5)

7. Find the correct value of weights to update weight and bias values \(w_{jk}\).

\[
\Delta w_{jk} = \alpha \delta_k z_j
\]
\[
\Delta w_{0k} = \alpha \delta_k
\]

(2.6) (2.7)

8. Calculates the value of the \(\delta_j\) in the remote unit to calculate the error rate.

\[
\delta_{\text{in}}_j = \sum_{k=1}^m \delta_k w_{jk}
\]
\[
\delta_j = \delta_{\text{in}}_j f'(z_{\text{in}}_j)
\]

(2.8) (2.9)

9. Find weight truth values to update weight and bias values \(v_{ij}\).

\[
\Delta v_{ij} = \alpha \delta_j x_i
\]
\[
\Delta v_{0j} = \alpha \delta_j
\]

(2.10) (2.11)

**Weight and Bias Update Stage**

10. Calculate new weights and \(w_{jk}\) biases:

\[
w_{jk}(\text{new}) = w_{jk}(\text{old}) + \Delta w_{jk}
\]

(2.12)

11. Looking for new weight values and biases \(v_{ij}\):

\[
v_{ij}(\text{new}) = v_{ij}(\text{old}) + \Delta v_{ij}
\]

(2.13)

**Data Testing**

Data testing is done to measure the correctness of a model that has been built to predict. In the training process, obtained weights that iteratively will minimize the error value. The error will be calculated from the average value of the square of the error (MSE). MSE is also used as the basis of calculations for the work of the activation function. MSE is calculated using formulas (2.14) [14]:

\[
MSE = \frac{\sum_{i=1}^n e_i^2}{n}
\]

(2.14)

Information:

\(e_i^2 = \) difference between target and tilapia prediction output

\(n = \) amount of learning data

**RESULT AND DISCUSSION**

**Analysis**

To obtain results by expectations, it must go through a training and testing process where the parameters to be used have been determined [22]. These parameters include [15]:

---

Copyright © 2022, Numerical: Jurnal Matematika dan Pendidikan Matematika
Print ISSN: 2580-3573, Online ISSN: 2580-2437

29
a. Determine errors  
b. Fault tolerance  
c. Specify activation function  
d. Defining epoch (iteration)  
e. Determine the hidden layer and output layer  
f. Specify network training functions  
g. Determine the learning rate  

Backpropagation analysis in the training stage is repeated to get the best weight with a minor error. If the best weight and a little error have been obtained, an analysis will be used at the testing stage.  

**Result**  

In the Backpropagation method, the architecture model influences the target or prediction to be targeted, and not all problems can be solved with the same architectural model. The architecture model consists of inputs - hidden layers - outputs that the user of the system can determine by trial and error until obtaining the best architectural model. This study uses architectural models: Input Layer: 2 neurons, Hidden Layer: 70-1 neuron, Output Layer: 1 neuron, with maximum epoch/iteration: 2000. Goal/Target (MSE): 0.000001. From the results of the training process/training data in Table 3 obtained the results of training/training as in Figure 1, the following:

**Figure 1. Neuron Network Training Process**
Here are the results of the process of analysis of training data conducted based on data in Table 3:

![Figure 2. Plot Performance](image)

In Figure 2, the learning process is shown in each iteration. In this learning process, epoch/iteration is stopped in the 276th epoch/iteration because the desired epoch/iteration limit has been reached, namely the MSE value of 0.00099235, where the MSE value is an MSE value that appears when the training/termination process has been completed by the epoch/iteration specified.

![Figure 3. Plot Regression](image)
Figure 3. shows the relationship between the target and the output/network results in the training/training data. From the process of the training/training data for the match between the output/network results with the target obtained a correlation coefficient (R) of 0.99273, where the best outcome is worth 1, with a correlation coefficient of 0.99273 showing that the network can predict well by existing data. After conducting training and testing (testing), the output/output prediction value was obtained in the output training and testing section.

**Table 7. Training Results Data (Years 2001-2012)**

<table>
<thead>
<tr>
<th>No</th>
<th>Year</th>
<th>Target</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2001</td>
<td>0.38</td>
<td>0.37</td>
</tr>
<tr>
<td>2</td>
<td>2002</td>
<td>0.90</td>
<td>0.84</td>
</tr>
<tr>
<td>3</td>
<td>2003</td>
<td>0.13</td>
<td>1.28</td>
</tr>
<tr>
<td>4</td>
<td>2004</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>5</td>
<td>2005</td>
<td>0.63</td>
<td>0.64</td>
</tr>
<tr>
<td>6</td>
<td>2006</td>
<td>0.56</td>
<td>0.56</td>
</tr>
<tr>
<td>7</td>
<td>2007</td>
<td>0.32</td>
<td>0.32</td>
</tr>
<tr>
<td>8</td>
<td>2008</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>9</td>
<td>2009</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>10</td>
<td>2010</td>
<td>0.31</td>
<td>0.33</td>
</tr>
<tr>
<td>11</td>
<td>2011</td>
<td>0.34</td>
<td>0.31</td>
</tr>
<tr>
<td>12</td>
<td>2012</td>
<td>0.23</td>
<td>0.32</td>
</tr>
<tr>
<td>13</td>
<td>2013</td>
<td>0.14</td>
<td>0.12</td>
</tr>
</tbody>
</table>

A measure of the accuracy of forecasting/prediction using the average percentage of absolute errors (MAPE) with the formula [6]:

$$MAPE = \frac{\sum |e_i| \times 100}{\sum x_{astil}} (\%)$$  \hspace{1cm} (3.1)

Information:

- $e_i$ = difference between target and prediction results
- $x_{astil}$ = original data
- $n$ = amount of data

The results of the calculation obtained a mape value of 0.047895 %, and it can be said that the average success of prediction/forecasting of 99.9521 %
Figure 4 illustrates the comparison between the original target and the output of the JST network in training data with a learning rate ($\alpha$) = 0.1.

Table 8. Testing Results Data (Using The Year 2014-2019)

<table>
<thead>
<tr>
<th>No</th>
<th>Year</th>
<th>Target</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2014</td>
<td>0.34</td>
<td>0.13</td>
</tr>
<tr>
<td>2</td>
<td>2015</td>
<td>0.20</td>
<td>0.70</td>
</tr>
<tr>
<td>3</td>
<td>2016</td>
<td>0.11</td>
<td>0.06</td>
</tr>
<tr>
<td>4</td>
<td>2017</td>
<td>0.10</td>
<td>0.02</td>
</tr>
<tr>
<td>5</td>
<td>2018</td>
<td>0.41</td>
<td>0.35</td>
</tr>
<tr>
<td>6</td>
<td>2019</td>
<td>0.39</td>
<td>0.38</td>
</tr>
</tbody>
</table>

The accuracy rate for predictions using MAPE with formula (3.1). After accumulating obtained MAPE value of 8.36282% or predictive success in this study of 91.63718%. In the testing stage can be seen the results in table 8, then using a network of architectures that have been created represented to predict the future three years of Tuban Regency, namely in 2020-2022, which can be seen in the following table:

Table 9. Results of Forecasting Open Unemployment Rate (TPT) Tuban Regency

<table>
<thead>
<tr>
<th>Nbr</th>
<th>Year</th>
<th>Prediction (Normalization) (%)</th>
<th>Prediction (Denormalization) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2020</td>
<td>0.35</td>
<td>4.40</td>
</tr>
<tr>
<td>2</td>
<td>2021</td>
<td>0.38</td>
<td>4.59</td>
</tr>
<tr>
<td>3</td>
<td>2022</td>
<td>0.40</td>
<td>4.70</td>
</tr>
</tbody>
</table>
The prediction results state that in 2020 the unemployment rate will be 4.39%, which means that the open unemployment rate has 4.39% of the number of Tuban Regency Workforce in 2020. Similarly, in 2021 and 2022, 2021 has an Open Unemployment Rate of 4.58% of the number of Workers in Tuban Regency 2021, and 2022 has an Open Unemployment Rate of 4.7% of the Labor Force in Tuban Regency 2022.

CONCLUSIONS

The results showed that the average success of TPT predictions was 91.6371%, with MAPE at 8.36282%. The results were obtained when the learning rate value is 0.1 with the best network architecture or pattern 2-70-1-1, which means two input neurons, 70 hidden layer neurons, one hidden layer neuron, and one output neuron. Epoch/iteration: 276. The results of forecasting the Open Unemployment Rate of Tuban Regency in 2020-2022 showed an increase from 2020 of 4.39% the number of Labor Force until 2022 4.7% the number Labor Force.

REFERENCES


