Weather prediction system and recommendation of plant varieties as an effort to minimize harvest failure with android-based Backpropagation Artificial Neural Networks

M Ahsan1,*, W Setiyaningsih1, B Rinanto1, M Susilowati2 and I Sulistiyowati3

1Informatics Study Program, Faculty of Science and Technology, Universitas Kanjuruhan Malang, Jl. S. Supriadi, No. 48 Malang Jawa Timur, Indonesia
2Program Studi Sistem Informasi, Fakultas Sain dan Teknologi, Universitas Ma Chung Malang, Jl. Villa PuncakTidar Blok N No. 1, Dau, Malang, Jawa Timur, Indonesia
3Electrical Engineering Department, Universitas Muhammadiyah Sidoarjo

*ahsan@unikama.ac.id

Abstract. Based on Malang Regency agriculture data in 2017 presented by the Head of the Agriculture and Food Crops and Horticulture Office (DPTPH), there was a 2.5% decrease in yields. One hectare of paddy land can produce 7.71 tons during normal weather, but currently only produces 6.9 tons per hectare. Three farmers' groups in Malang Regency (2018), are of the opinion that the current weather greatly influences crop failure. Weather prediction used so far using conventional benchmarks for planting, is no longer relevant to current reality. These estimates are often incorrect, resulting in crop failures, which have an impact on the farmer's economy. This study aims to make the application of weather prediction systems and recommendations for suitable planting varieties based on data on temperature, humidity, time of sun exposure, wind speed, and rainfall using Backpropagation ANN method. Data used for Karangploso District weather prediction from 2009-2018. While the variables used in the application of Backpropagation ANN: temperature, humidity, time of sun exposure, wind speed, and rainfall. For output, the prediction of rainfall in the next 12 months and planting of varieties in accordance with the predicted rainfall, OPT and land type. From the calculation results of this application the MSE error value of 0.0299 is obtained.

1. Introduction

Almost all fields have been developed information system technology, such as transportation, education and social. But there are some fields that are still rarely or even have not been touched by this information system technology [1]. One of those fields is agriculture. Farmers still use the old method even though it is no longer relevant if it is used now. One of them is determining the plants to be planted later. Even though there has been a shift in the beginning of the season, farmers are still unable to anticipate the change of seasons [2]. As a result, farmers often experience losses and crop failures that are wrong in choosing the plants to be planted [3].

Uncertain weather in Indonesia, made Malang Regency farmers experience a decrease in production at harvest. According to the Head of the Agriculture and Food Crops and Horticulture Office (DPTPH) of Malang Regency, based on agricultural data in 2017 there was a 2.5% decline in
crop yields in almost all regions of Malang Regency. During normal weather, one hectare of paddy land can produce 7.71 tons, but currently only produces 6.9 tons per hectare.

The results of interviews with three farmer groups in Malang Regency in 2018, argued that the current weather greatly affected crop failure. Conventional weather prediction used as a benchmark for planting has long been no longer relevant to current realities. Often the estimates are not correct, resulting in crop failures, this certainly has an impact on the economy of Malang Regency farmers.

This study aims to make the application of weather prediction systems and planting varieties that are appropriate in supporting the planting process in Malang Regency based on data on temperature, humidity, sun exposure time, wind speed, and rainfall using the Backpropagation ANN method [4,5]. The method used is the calculation of backpropagation neural networks [6].

2. Object of research
Backpropagation is a learning algorithm for handling problems with complex patterns. Where the architecture has 3 layers, namely the input layer, hidden layer, and output layer [2].

2. Methods
The data used are sample data from the weather history of Karangploso sub-district from 2009-2018. Data sets used in the application of Backpropagation ANN include: temperature, humidity, sun exposure time, wind speed, and rainfall. The output system is in the form of rainfall prediction for the next 12 months and planting varieties that are in accordance with the predicted rainfall results, OPT and the type of land. According to BMKG rainfall data are categorized into four, namely: low (0 mm-100 mm), moderate (101 mm-300 mm), high (301 mm-500 mm), very high (above 500 mm). While the appropriate varieties of planting varieties according to the predicted rainfall in this research are rice, corn, cassava and sweet potatoes. This kind of variety is used as an appropriate alternative plant because it matches the majority of the market share of the Malang Regency farmer groups and according to the type of land.

The android version used in the development of this rainfall prediction system is version 5.0 Lollipop. The weather prediction system from the farmer side only enters location data to get the prediction data.

The stages in Backpropagation are divided into 3 phases, namely the feedforward phase, the backforward phase, and the weight change phase [7].

- Stage 0: Initialize the weights;
- Stage 1: If conditions are not met, work on steps 2-8;
- Stage 2: For each pair of training data, work on steps 3-8.

Phase 1: Feedforward
- Stage 3: Each unit receives the input signal and forwards to the hidden layer above;
- Stage 4: Calculate all outputs in $Z_j$ hidden layer units ($j = 1, 2, \ldots, p$)
\[ z_{\text{net}}_j = v_{jo} + \sum_{i=1}^{n} x_i v_{kj} \]  
\[ z_j = f(z_{\text{net}}_j) = \frac{1}{1+e^{-z_{\text{net}}_j}} \]  

- Stage 5: Calculate all network outputs on the output unit \( Y_k \) (\( k = 1,2, \ldots, m \))

\[ z_{\text{net}}_k = w_{ko} + \sum_{j=1}^{p} z_j w_{kj} \]
\[ y_k = f(y_{\text{net}}_k) = \frac{1}{1+e^{-y_{\text{net}}_k}} \]  

| Phase 2: Backward |
|-------------------|
|                   |
| - Step 6: Calculate the factor \( \delta \) on the unit output based on the unit output error \( Y_k \) (\( k = 1,2, \ldots, m \))

\[ \delta_k = (t_k - y_k)f'(y_{\text{net}}k) = (t_k - y_k)y_k(1-y_k) \]  
Where, \( t_k \) is the target output, \( \delta \) is the unit of output that will be used in the layer below. Calculate the weight change in \( W_{kj} \), with the learning rate \( \alpha \).

\[ \delta W_{ji} = \alpha \delta_k z_j, \quad k = 1,2, \ldots, m; \quad j = 0,1, \ldots, p \]  

- Stage 7: Calculate the factor hidden units based on the error in each hidden unit \( Z_j \) (\( j = 1,2, \ldots, p \))

\[ \delta_{\text{net}}_j = \sum_{k=1}^{m} \delta_k w_{kj} \]

Calculate the factor \( \delta \) hidden unit

\[ \delta_j = \delta_{\text{net}} j f'(z_{\text{net}}j) = \delta_{\text{net}} j z_j (1-z_j) \]  
Calculate the change in weights \( V_{ji} \)

\[ \delta V_{ji} = \alpha \delta_k z_j, \quad k = 1,2, \ldots, m; \quad j = 0,1, \ldots, n \]  

| Phase 3: Modify new weights |
|-----------------------------|
| - Stage 8: Calculate all the change weights that are in the output unit

\[ W_{kj}(\text{new}) = W_{kj}(\text{old}) + \delta W_{ji}; \quad (k = 1,2, \ldots, p; \quad j = 0,1, \ldots, n) \]  
Calculate all the change weights that are in the hidden unit

\[ V_{kj}(\text{new}) = V_{kj}(\text{old}) + \delta V_{ji}; \quad (j = 1,2, \ldots, p; \quad w = 0,1, \ldots, n) \]  

Mean Squared Error (MSE) is a method of calculating forecast errors where for each error in a calculation is squared. The results of the calculation of errors with MSE if the error is small, it will get smaller, but if it is large, it will get bigger too [3].

\[ \text{MSE} = \frac{\sum (A_t - F_t)^2}{n} \]  

Information:

At : actual investment value in period \( t \)
Ft : forecasting period investment value \( t \)
n : number of forecasting periods involved

4. Results and discussion

For the implementation of this application, BMKG acts as a manager of weather data, BPTP as a manager of location data, pests, plants and land types, and farmers as end users who can see the
predicted results of crop recommendations. For the prediction calculation process in the application, there are BMKG (Meteorology, Climatology, and Geophysics) web users. Where the input dataset is there too. For the data set used in the calculation part can be seen in Table 1.

| Wind velocity | Long exposure to the sun | Humidity | Temperature | Rainfall |
|---------------|--------------------------|----------|-------------|----------|
| 7.4           | 3.4                      | 85       | 23.5        | 258      |
| 9             | 3.3                      | 84       | 23.5        | 435      |
| ...           | ...                      | ...      | ...         | ...      |
| 4.4           | 7.4                      | 68       | 24.7        | 6        |
| 2.6           | 3.9                      | 85       | 24          | 406      |

After the data is entered then the training calculation process is carried out by repeating the feedforward and backforward processes 500 times. The final weight obtained will be carried out the testing process with test data prepared to see the value of MSE obtained [8]. From the training results obtained weights from the input layer to the hidden layer (V) and from the hidden layer to the output layer (W), which weights can be seen in Table 2.

| V1     | V2     | V3     | ...    | V238   | V239   | V240   |
|--------|--------|--------|--------|--------|--------|--------|
| 0.891  | -0.541 | -0.296 | ...    | -0.306 | -0.184 | -0.325 |
| 0.770  | -0.872 | -0.071 | ...    | -0.528 | 0.109  | -0.106 |
| ...    | ...    | ...    | ...    | ...    | ...    | ...    |
| 0.730  | -0.604 | 0.178  | ...    | 0.961  | -0.083 | 0.363  |
| 0.368  | -0.251 | 0.312  | ...    | 0.355  | -0.040 | -0.565 |

Next weights from the hidden layer to the output layer which can be seen in Table 3 below.

| W1    | W2     | W3     | ...    | W117   | W118   | W119   | W120   |
|-------|--------|--------|--------|--------|--------|--------|--------|
| 0.767 | -0.643 | -0.539 | ...    | -0.023 | 0.214  | 0.647  | -0.944 |
| 0.904 | -0.041 | -0.724 | ...    | -0.415 | 0.409  | -0.872 | -0.315 |
| -0.910| 0.356  | 0.026  | ...    | -0.499 | 0.308  | -0.763 | 0.097  |
| -0.108| 0.605  | -0.463 | ...    | 0.211  | -0.881 | -0.797 | 0.155  |

Next is the prediction process. The prediction process in this application consists of several stages. Stage 1 Take the input values from the normalized data and convert them into a matrix that will be inserted in each input layer neuron. For the weights used are the weights in Table 2. Which is the weight is the result of training that has been done previously in order to get the desired weight. From the results of the matrix made, it can be seen that the data used in the input layer is in Table 4. Where each neuron is represented by the letter X. The amount of data entered is a total of 240 data representing weather history data for 5 years.

| neuron | value | neuron | value | neuron | value | neuron | value |
|--------|-------|--------|-------|--------|-------|--------|-------|
| X1     | 0.38  | X5     | 0.38  | ...    | ...   | X237   | 0.25  |
| X2     | 0.10  | X6     | 0.14  | X234   | 0.55  | X238   | 0.35  |
| X3     | 0.81  | X7     | 0.74  | X235   | 0.70  | X239   | 0.86  |
| X4     | 0.32  | ...    | ...   | X236   | 0.44  | X240   | 0.36  |
Step 2 Calculation of the value in the hidden layer by multiplying each weight and then activating it. All 240 neurons in the input layer will be multiplied by their respective weights so that they fill the value of 120 neurons in the hidden layer.

Which then the process is carried out on all the neurons in the input layer that are so that each neuron in the hidden layer becomes like what is in Table 5.

| neuron | value | neuron | value | neuron | value | neuron | value |
|--------|-------|--------|-------|--------|-------|--------|-------|
| Z1     | 0.0013| Z7     | 0.9053| ...    | ...   | Z115   | 0.9878|
| Z2     | 0.9836| Z8     | 0.9876| Z110   | 0.9924| Z116   | 0.0003|
| Z3     | 0.9834| Z9     | 0.9266| Z111   | 0.9991| Z117   | 0.8830|
| Z4     | 0.1063| ...    | ...   | Z114   | 0.0113| Z120   | 0.9656|

Stage 3 is calculating the value of the output layer by multiplying the weight then activating it. The method used is the same as stage 2, but the final result here only produces 12 neurons in the output layer. Each of these neurons represents the moon in the prediction of the next 12 months. Prediction results can be seen in Table 6.

| Y1         |
|------------|
| 0.389717125|
| 0.857804079|
| 0.713880145|
| 0.230519965|

Mean Squared Error (MSE) calculation is carried out by the system by testing the results of the test input and the test target. The data used is a weather history data set 5 years before the data set is used for prediction. The data used can be seen in Table 7 below [9].

| Wind velocity | Long exposure to the sun | Humidity | Temperature |
|--------------|--------------------------|----------|-------------|
| 0.2882       | 0.2018                   | 0.8556   | 0.3486      |
| 0.2882       | 0.4055                   | 0.7889   | 0.3703      |
| ...          | ...                      | ...      | ...         |
| 0.1282       | 0.1291                   | 0.8111   | 0.3703      |
| 0.1471       | 0.1727                   | 0.7889   | 0.3595      |

From the prediction calculation using the test data in Table 7, the prediction results can be used to calculate the value of the MSE weights obtained from the training results. MSE calculations will be displayed in a new window as shown in Figure 2.
4.1. Calculation of Matlab prediction with applications.

In the matlab calculation process the calculation process is carried out with the same variables and conditions as the sample data performed by the application. It aims to see the comparison of the results of the running algorithm. Where architecture can be seen in Figure 3.

![Figure 3. Architecture neural network Matlab.](image)

In Figure 3, it can be seen that there are 8 neuron inputs, 4 hidden neurons and 1 output. For weights in matlab predictions can be seen in Figure 4.

![Figure 4. Weight in Matlab.](image)

5. Conclusion

Trials that have been carried out on the systems that have been made, it was concluded that the system created can produce weather predictions for the next year (12 months) with a relatively small MSE of
0.0299 and the predicted results can later help farmers choose the appropriate type of plant so that production can increase and crop failure can be minimized.

References
[1] Sulistyanto M P T, Pranata K B, Afandi A N, Sendari S and Sulistiyowati I 2019 Monitoring electrical energy in electronic energy audits through internet of things technology Journal of Physics: Conference Series vol 1402 (IOP Publishing) p 77067
[2] Muttaqin M 2018 PREDIKSI CURAH HUJAN MENGGUNAKAN JARINGAN SARAF TIRUAN BACKPROPAGATION
[3] Ahsan M, Setiyaningsih W, Susilowati M, Dijaya R and Tjahjanti P H 2019 Selecting multiple intelligences on children with weighted product, analytical hierarchy process, simple additive weighting and TOPSIS Journal of Physics: Conference Series vol 1402 (IOP Publishing) p 77033
[4] Kusuma M S D R, Priyadi A and Lystianingrum B P V 2019 Transient Stability Prediction Using algorithm Backpropagation Journal of Physics: Conference Series vol 1201 (IOP Publishing) p 12051
[5] Wanto A, Zarlis M and Hartama D 2017 Analysis of Artificial Neural Network Backpropagation Using Conjugate Gradient Fletcher Reeves In The Predicting Process JPhCS 930 12018
[6] Hadihardaja I K and Sutikno S 2007 Rainfall and stream flow discharge relation using artificial neural network ICSIIT 2007 1
[7] Salman N, Lawi A and Syarif S 2018 Artificial neural network backpropagation with particle swarm optimization for crude palm oil price prediction Journal of Physics Conference Series vol 1114 pp 1–8
[8] Prayudani S, Hizriadi A, Lase Y Y and Fatmi Y 2019 Analysis Accuracy Of Forecasting Measurement Technique On Random K-Nearest Neighbor (RKNN) Using MAPE And MSE Journal of Physics: Conference Series vol 1361 (IOP Publishing) p 12089
[9] Sriadhi S, Gultom S, Martiano M, Rahim R and Abdullah D 2018 K-means method with linear search algorithm to reduce Means Square Error (MSE) within data clustering MS&E 434 12032