

THE ARTIFICIAL NEURAL NETWORK (ANN) ALGORITHM IMPLEMENTATION FOR PREDICTING THE AMOUNT OF BOOK SALES AT ERLANGGA PUBLISHER PEMATANGSIANTAR

Hotmalina Silitonga¹, Indra Gunawan², Bahrudi Efendi Damanik³

^{1,2)} STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

³ AMIK Tunas Bangsa, Pematangsiantar, Indonesia

Corresponding author's: hotmalinasilitonga@gmail.com

ABSTRACT

Selling is one of the main goals of a company after producing its goods. The more goods sold, the more economic value the company is selling. Therefore, the purpose of this study is to determine how much the rate of increase or decrease in the number of book sales at the publisher of Erlangga Pematangsiantar is in the form of prediction. This study uses an Artificial Neural Network (ANN) with the Backpropagation method. Backpropagation is a method that is often used for prediction. The research data is secondary data (sales data) sourced from PT. Publisher Erlangga Pematangsiantar from 2013 to 2017. Data is divided into 2 parts, namely training data and testing data. There are 5 architectural models used in this study, 3-9-1, 3-11-1, 3-15-1, 3-30-1, and 3-31-1. Of the 5 (five) architectural models used, the best architecture is 3-11-1 with an accuracy rate of 80% and MSE 0.13001601. So this model is good for predicting the number of book sales at PT. Publisher Erlangga Pematangsiantar.

Keywords : Prediction, Backpropagation, ANN, Book Sales, Erlangga Publisher .

INTRODUCTION

Sales in the business world affect the economic assets of a company or agency. In the process of selling or providing goods and services to the buyer the level of desire of a commodity for a certain price, sales can be done through methods such as direct selling and through sales agents, especially sales. Sales activity is one of the main goals of a company after producing its goods. Often the sales manager takes on the dual responsibility of managing the sales team and selling goods to customers. This has resulted in a considerable time allocation and of course it can affect sales performance in general (Rapp, Petersen, Hughes, & Ogilvie, 2020). PT. Penerbit Erlangga Pematangsiantar is one of the companies engaged in the printing of books ranging from kindergarten, elementary, junior high, high school, to tertiary education levels. Books are an excellent means to participate in educating the nation. For this reason, the availability of books needs to be optimized for the progress and success of the world of education in particular. However, unstable book sales, especially at PT. Erlangga Pematangsiantar causes book availability is not optimal. Therefore it is necessary to predict book sales for the future, so that the management of PT. Erlangga has a reference to optimize more book stock, especially the best-selling books on the market. Moreover, sales predictions are an important measure of national economic development trends (Wang, Lin, & Wang, 2019).

The prediction algorithm proposed in this article is the backpropagation algorithm. The backpropagation algorithm is one of the ANN algorithms that is able to work systematically by training multiplayer networks using mathematical science based on developed network

architecture models (Febriyati & Gs, 2020; Ginantra, Hanafiah, Wanto, Winanjaya, & Okprana, 2021; Siregar & Wanto, 2017; Wanto, Windarto, Hartama, & Parlina, 2017; Windarto et al., 2020). Selain itu algoritma Backpropagation mampu melakukan prediksi berdasarkan data time-series (Bhawika et al., 2019; Gultom, Wanto, Gunawan, Lubis, & Kirana, 2021; Purba et al., 2019; Saputra, Hardinata, & Wanto, 2019; Setti & Wanto, 2018).

Studies on the use of ANN algorithms to solve complex problems have developed widely and have been widely carried out. (Zhang & Mu, 2021) proposed a recharging decision model with the multivariate regression analysis method of backpropagation neural networks. With a regular pattern between sales and individual variables, coupled with the empirical safety stock formula, an accurate filling amount can be obtained. In a case analysis, this paper takes the pharmacy sales situation as an example and tests the accuracy and stability of the model. The results show that the model has good predictive accuracy that can be entered into the smart pharmacy system and used in refilling smart pharmacies to prevent over-stock or under-stock, thereby improving the financial situation, reducing the workforce burden of typical retail pharmacies, and helping residents buy drugs. (Yu & Zhao, 2019) proposed a backpropagation neural network model based on genetic algorithms (GA) to predict the properties of biodiesel fuel according to the composition of FAME (Fatty Acid Methyl Esters). The BPNN-GA hybrid model has five inputs (methyl palmitate, methyl stearate, methyl oleate, methyl linoleate and methyl linolenic acid) which correspond to the composition and output of the FAME with an estimate of fuel properties, with GA assisting the training to find out the local minimum deviation and weighting configuration updated. It was found that the BPNN-GA hybrid model made it possible to map the non-linear relationship between the FAME composition and the main properties of the biodiesel fuel with a fairly good agreement. The predicted value of fuel properties corresponds to the values measured by the R-square up to 96%, along with lower values (less than 10%) above the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE) compared to the values. other empirical. Next (Aritonang & Sihombing, 2019) in his paper discusses the prediction of product sales (rice) in the Rice Milling Unit so that it can be seen the amount of raw material needed so as to avoid a time lag. Forecasting / predicting optimal carried out by; comparing 2 (two) forecasting methods, Linear Regression, and Neural Networks with the Backpropagation algorithm. The results show that the MSE value in the linear regression method is 0.214, while when using Artificial Neural Networks, the MSE value is 0.00099713. Based on the MSE value, the smallest MSE is predicted by the Backpropagation algorithm.

Based on the background and related studies that have been described, this article is proposed to predict book sales at PT. Erlangga in the years to come. The goal is that the results of this research can be used as information and input for related parties concerned, especially in advising management to be more able to maintain the inventory of books so they don't run out, especially books that are best-selling on the market, don't run out.

MATERIALS AND METHODS

Method of collecting data

This research uses quantitative methods. In general, data collection methods to solve problems in this study use 3 (three) methods, namely:

1. Interview

At this stage, interviews are conducted with the sales manager to obtain data on book sales at PT. Erlangga

2. Observation

Make direct observations to the sales department to obtain the necessary data.

3. Study of literature

Looking for theoretical references that are relevant to the specified case or problem. These references can be found from books, journals, research report articles, and websites on the internet. The output of this literature study is the collection of references relevant to the problem topic.

Table 1. Data on Number of Book Sales (Sample Data)

No	Book Name	Author's Name	Publication Year	Number of Sales / Pieces				
				2013	2014	2015	2016	2017
1	Bahasa Indonesia	Nurhadi	2009	99	69	54	83	94
2	Matematika	Wono Setya Budhi	2007	75	73	44	55	78
3	IPA Terpadu	Tim Abdi Guru	2009	58	17	22	87	25
4	IPS Terpadu	Tim Abdi Guru	2009	20	68	15	25	42
5	Pendidikan Kewarganegaraan	Tim Abdi Guru	2005	47	34	25	13	20
6	Seni Budaya	Tim Abdi Guru	2004	26	18	12	20	15
7	Inggris on Sky	Mukarto	2008	70	30	82	47	80
8	Penjasorkes	Roji	2009	15	13	46	18	27
9	Maestro Olimpiade SMP	Ibnul Mubarak	2006	35	42	13	27	17
10	Seri Permit UN SMP	Tim Abdi Guru	2006	45	20	15	35	55

Source: PT. Erlangga Pematangsiantar

Research Stages

The research stages proposed and presented in this article are the general stages of the book sales prediction process at PT. Erlangga. These stages can be seen in figure 1.

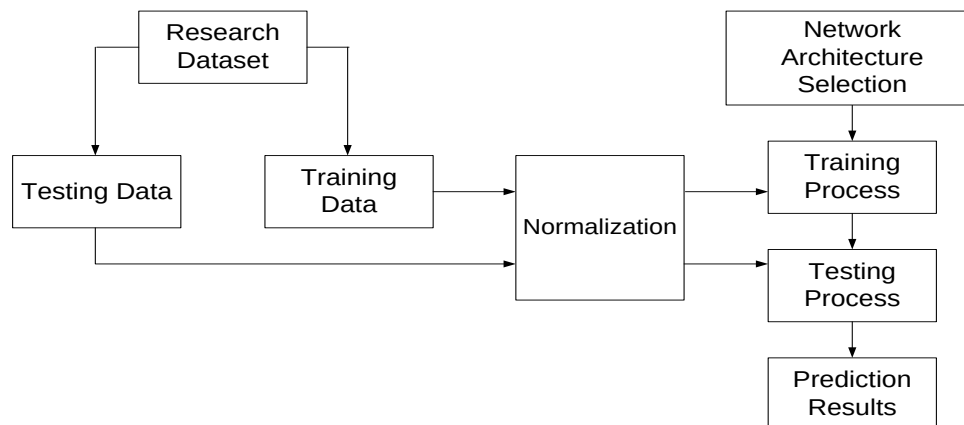


Figure 1. Research Stages

The research stages proposed in this article begin with the collection of the research dataset. The research dataset used is book sales data at PT. Erlangga Pematangsiantar, 2013-2017. Then the data preprocessing is carried out and divides the data into several parts, namely the data used for training and the data used for testing, after which the data is normalized first so that it can be processed and calculated using the Matlab 2011b application.

The next step is to determine the network architecture model that will be used for the training process and the testing process. Then the training and testing process will be carried out using a predetermined architectural model. Furthermore, from several architectural models used, the best one will be selected based on a higher level of accuracy and a smaller MSE value. After that, predictions will be made using the best architectural model that has been selected.

RESULTS AND DISCUSSIONS

Data Normalization

The research dataset presented in table 1 will be normalized using equation (1) (Afriliansyah et al., 2019; Lubis, Saputra, Wanto, Andani, & Poningsih, 2019; Parulian et al., 2019; Wanto & Hardinata, 2020; Wanto et al., 2019):

$$x' = \frac{0.8(x - a)}{b - a} + 0.1 \quad (1)$$

Explanation :

x' = Normalization results

x = Data to be normalized

a = The smallest data from the dataset

b = The largest data set from the dataset

Data that has been normalized using equation (1) can be seen in table 2.

Table 2. Normalization Results

Data	2013	2014	2015	2016	2017
1	0,90000	0,62414	0,48621	0,79268	0,90000
2	0,67931	0,66092	0,39425	0,51951	0,74390
3	0,52299	0,14598	0,19195	0,83171	0,22683
4	0,17356	0,61494	0,12759	0,22683	0,39268
5	0,42184	0,30230	0,21954	0,10976	0,17805
6	0,22874	0,15517	0,10000	0,17805	0,12927
7	0,63333	0,26552	0,74368	0,44146	0,76341
8	0,12759	0,10920	0,41264	0,15854	0,24634
9	0,31149	0,37586	0,10920	0,24634	0,14878
10	0,40345	0,17356	0,12759	0,32439	0,51951

The results of normalization in table 2 will be divided into 2 parts, namely training data and testing data. The training input data uses data from 2013 to 2015 with a target of 2016. As for the input data testing uses data from 2014 to 2016 with a target of 2017.

Best Architectural Model

There are 5 architectural models used in this study, including: 3-9-1, 3-11-1, 3-15-1, 3-30-1, and 3-31-1. Based on these 5 models, model 3-11-1 is the best model chosen because of the higher level of accuracy (80%) compared to other models. The way to determine the best architectural model with the Backpropagation algorithm is to look at the highest level of accuracy of each model. The error parameter used was 0.3-0.001. The analysis process uses Matlab and Microsoft Excel tools. Based on the results of training and testing using the MATLAB application and calculations using Microsoft Excel, the best architectural model of the five models used is 3-11-1. The results of the training and testing process for the 3-11-1 model can be seen in table 3 and table 4.

Table 3. Training Data Model 3-11-1

Table 4. Testing Data Model 3-11-1

Pattern	Target	Output	Error	SSE	Pattern	Target	Output	Error	SSE	Results
1	0,75287	0,74970	0,00317	0,00001007	1	0,90000	1,04070	-0,14070	0,01979649	1
2	0,49540	0,49940	-0,00400	0,00001598	2	0,74390	0,28380	0,46010	0,21169425	0
3	0,78966	0,76980	0,01986	0,00039423	3	0,22683	-0,03080	0,25763	0,06637284	1
4	0,21954	0,21900	0,00054	0,00000029	4	0,39268	0,76190	-0,36922	0,13632125	1
5	0,10920	0,17000	-0,06080	0,00369720	5	0,17805	0,16710	0,01095	0,00011988	1
6	0,17356	0,15910	0,01446	0,00020918	6	0,12927	0,36570	-0,23643	0,05589995	1
7	0,42184	0,41830	0,00354	0,00001253	7	0,76341	0,29060	0,47281	0,22355368	0
8	0,15517	0,15810	-0,00293	0,00000857	8	0,24634	0,24730	-0,00096	0,00000092	1
9	0,23793	0,17260	0,06533	0,00426814	9	0,14878	0,91390	-0,76512	0,58540787	1
10	0,31149	0,34870	-0,03721	0,00138427	10	0,51951	0,48800	0,03151	0,00099302	1
Sum SSE				0,01000046	Sum SSE				1,30016014	80%
MSE				0,00100005	MSE				0,13001601	

Tables 3 and 4 can be seen the results of the accuracy and MSE levels of the best architectural models, namely 3-11-1. Table 4 is created and calculated using Microsoft Excel. The description can be seen as follows:

Target = Obtained from target training data and target test data (based on table 2)
Output = Obtained from the results of calculations with Matlab
Error = Obtained from Target-Output
SSE = Obtained from Error ²
Sum SSE = The total SSE generated from the pattern 1 - 10
MSE = Obtained from Number of SSE / 10 (10 is the number of patterns)
Results = If the error value in the test data is <= 0.3 then the result is correct (1). If not then false (0).
Accuracy = Obtained from the number of correct results ((pattern / 10) * 100), yields 80% accuracy. Value 1 in table 4 is another word for "True", while 0 = "False"

The results of architectural model training 3-11-1 using Matlab can be seen in Figure 2

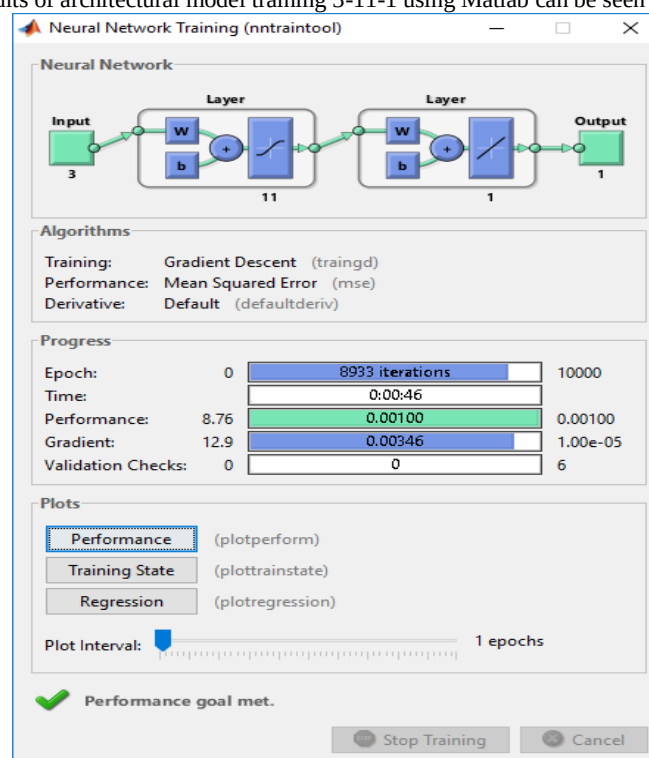


Figure 2. Architectural model training 3-11-1

In Figure 2 it can be explained that in the training input data model 3-11-1 uses 3 input layers, 11 neurons hidden layer and 1 neuron output layer. The resulting epoch is 8933 iterations within 46 seconds.

Comparison of Architectural Models Used

The comparison of the results of the training and testing process with the architectural model used can be seen in table 5.

Table 5. Comparison of Architectural Models Used

No	Explanation	Training			Testing	
		Epoch	Times	MSE Training	MSE Testing	Accuracy
1	3-9-1	6818	00:34	0,00100076	0,42469013	60
2	3-11-1	8933	00:46	0,00100005	0,13001601	80
3	3-15-1	8251	00:40	0,00100004	0,34310119	50
4	3-30-1	1893	00:11	0,00099869	0,97122552	50
5	3-31-1	1347	00:08	0,00099942	1,49093257	30

In table 5, we can see the comparison of each of the architectural models used. Of the five trained and tested architectural models, the 3-11-1 architectural model is the best architectural model with an epoch of 8933 iterations and an accuracy rate of 80% (the highest compared to the other 4 architectural models) and MSE Testing 0.13001601 (the lowest compared to 4 other architectural models).

Prediction Results

Then the prediction will be carried out using the 3-11-1 model using the formula to return the value in equation (2):

$$x_n = \frac{(x - 0,1) * (b - a)}{0,8} + a \quad (2)$$

Explanation:

x_n = Prediction Results

x = Predicted Target

a = The smallest data from the dataset

b = The largest data set from the dataset

For the results of predictions for 2020 can be seen in table 6.

Table 6. Comparison of Preliminary Data with Prediction Result data

No	Book Name	Number of Sales / Pieces									
		2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
1	Bahasa Indonesia	99	69	54	83	94	94	92	90	86	79
2	Matematika	75	73	44	55	78	78	80	86	82	81
3	IPA Terpadu	58	17	22	87	25	28	37	46	55	69
4	IPS Terpadu	20	68	15	25	42	43	51	56	66	74
5	Pendidikan Kewarganegaraan	47	34	25	13	20	27	29	38	57	71
6	Seni Budaya	26	18	12	20	15	21	32	41	55	70
7	Inggris on Sky	70	30	82	47	80	81	85	80	85	83
8	Penjasorkes	15	13	46	18	27	24	27	42	56	68
9	Maestro Olimpiade SMP	35	42	13	27	17	23	34	44	55	69
10	Seri Permit UN SMP	45	20	15	35	55	51	56	62	69	74
Amount		490	384	328	410	453	471	523	585	665	739

CONCLUSSIONS

Based on the results and discussion described in this article, it can be concluded that the Backpropagation method can be used to predict book sales at PT. Erlangga Pematangsiantar with the best model 3-11-1. Based on the comparison of the initial data on the number of book sales (2013-2017) at PT. Erlangga with predictive data (2018-2022), sales are relatively stable but there is a slight increase.

REFERENCES

- Afriliansyah, T., Parulian, P., Ulva, A. F., Simanjuntak, M. Y., Wanto, A., Sihombing, D., ... Ginantra, N. (2019). Implementation of Bayesian Regulation Algorithm for Estimation of Production Index Level Micro and Small Industry. *Journal of Physics: Conference Series*, 1255(1), 1–6.
- Aritonang, M., & Sihombing, D. J. C. (2019). An Application of Backpropagation Neural Network for Sales Forecasting Rice Milling Unit. *IEEE Xplore*, 7–10. <https://doi.org/10.1109/ICoSNIKOM48755.2019.9111612>
- Bhawika, G. W., Purwantoro, P., GS, A. D., Sudrajat, D., Rahman, A., Makmur, M., ... Wanto, A. (2019). Implementation of ANN for Predicting the Percentage of Illiteracy in Indonesia by Age Group. *Journal of Physics: Conference Series*, 1255(1), 1–6.
- Febriyati, N. A., & Gs, A. D. (2020). *Analysis of Backpropagation Algorithm Using the Traingda Function for Export Prediction in East Java*. 4(36), 550–558.
- Ginantra, N. L. W. S. R., Hanafiah, M. A., Wanto, A., Winanjaya, R., & Okprana, H. (2021). Utilization of the Batch Training Method for Predicting Natural Disasters and Their Impacts. *IOP Conf. Series: Materials Science and Engineering*, 1071(012022), 1–7. <https://doi.org/10.1088/1757-899X/1071/1/012022>
- Gultom, W. T. C., Wanto, A., Gunawan, I., Lubis, M. R., & Kirana, I. O. (2021). Application of The Levenberg Marquardt Method In Predict The Amount of Criminality in Pematangsiantar City. *Journal of Computer Networks, Architecture, and High-Performance Computing*, 3(1), 21–29. <https://doi.org/10.47709/cnahpc.v3i1.926>
- Lubis, M. R., Saputra, W., Wanto, A., Andani, S. R., & Poningsih, P. (2019). Analysis of Artificial Neural Networks Method Backpropagation to Improve the Understanding Student in Algorithm and Programming. *Journal of Physics: Conference Series*, 1255(1), 1–6. <https://doi.org/10.1088/1742-6596/1255/1/012032>
- Parulian, P., Tinambunan, M. H., Ginting, S., Gibran, M. K., Wanto, A., Muharram, L. O., ... Bhawika, G. W. (2019). Analysis of Sequential Order Incremental Methods in Predicting the Number of Victims Affected by Disasters. *Journal of Physics: Conference Series*, 1255(1), 1–6.
- Purba, I. S., Wanto, A., Riansah, R. M., Ahmad, Y., Siregar, S. P., Winanjaya, R., ... Silitonga, H. (2019). Accuracy Level of Backpropagation Algorithm to Predict Livestock Population of Simalungun Regency in Indonesia. *Journal of Physics: Conference Series*, 1255(1), 1–6.
- Rapp, A. A., Petersen, J. A., Hughes, D. E., & Ogilvie, J. L. (2020). When time is sales: the impact of sales manager time allocation decisions on sales team performance. *Journal of Personal Selling and Sales Management*, 40(2), 132–148. <https://doi.org/10.1080/08853134.2020.1717961>
- Saputra, W., Hardinata, J. T., & Wanto, A. (2019). Implementation of Resilient Methods to Predict Open Unemployment in Indonesia According to Higher Education Completed. *JITE (Journal of Informatics and Telecommunication Engineering)*, 3(1), 163–174.

- Setti, S., & Wanto, A. (2018). Analysis of Backpropagation Algorithm in Predicting the Most Number of Internet Users in the World. *JOIN (Jurnal Online Informatika)*, 3(2), 110–115. <https://doi.org/10.15575/join>.
- Siregar, S. P., & Wanto, A. (2017). Analysis of Artificial Neural Network Accuracy Using Backpropagation Algorithm In Predicting Process (Forecasting). *International Journal Of Information System & Technology*, 1(1), 34–42.
- Wang, P. H., Lin, G. H., & Wang, Y. C. (2019). Application of Neural Networks to Explore Manufacturing Sales Prediction. *Applied Sciences*, 9(23), 1–14. <https://doi.org/10.3390/app9235107>
- Wanto, A., & Hardinata, J. T. (2020). Estimations of Indonesian poor people as poverty reduction efforts facing industrial revolution 4.0. *IOP Conference Series: Materials Science and Engineering*, 725(1), 1–8. <https://doi.org/10.1088/1757-899X/725/1/012114>
- Wanto, A., Hayadi, B. H., Subekti, P., Sudrajat, D., Wikansari, R., Bhawika, G. W., ... Surya, S. (2019). Forecasting the Export and Import Volume of Crude Oil , Oil Products and Gas Using ANN. *Journal of Physics: Conference Series*, 1255(1), 1–6. <https://doi.org/10.1088/1742-6596/1255/1/012016>
- Wanto, A., Windarto, A. P., Hartama, D., & Parlina, I. (2017). Use of Binary Sigmoid Function And Linear Identity In Artificial Neural Networks For Forecasting Population Density. *International Journal Of Information System & Technology*, 1(1), 43–54.
- Windarto, A. P., Nasution, D., Wanto, A., Tambunan, F., Hasibuan, M. S., Siregar, M. N. H., ... Nofriansyah, D. (2020). *Jaringan Saraf Tiruan: Algoritma Prediksi dan Implementasi*. Yayasan Kita Menulis.
- Yu, W., & Zhao, F. (2019). Prediction of critical properties of biodiesel fuels from FAMEs compositions using intelligent genetic algorithm-based back propagation neural network. *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, 0(0), 1–14. <https://doi.org/10.1080/15567036.2019.1641575>
- Zhang, H., & Mu, J.-H. (2021). A Back Propagation Neural Network-Based Method for Intelligent Decision-Making. *Complexity*, 2021, 1–11. <https://doi.org/10.1155/2021/6610797>