International Journal of Computational Intelligence Systems
Vol. **12(2)**, 2019, *pp.* 1575–1584
DOI: https://doi.org/10.2991/ijcis.d.191101.003; ISSN: 1875-6891; eISSN: 1875-6883
https://www.atlantis-press.com/journals/ijcis/

Special Issue

Building an Artificial Neural Network with Backpropagation Algorithm to Determine Teacher Engagement Based on the Indonesian Teacher Engagement Index and Presenting the Data in a Web-Based GIS

Sasmoko Buddhtha[1,2*], Christina Natasha[3], Edy Irwansyah[3], Widodo Budiharto[3]

[1]*Primary Teacher Education Department, Faculty of Humanities, Bina Nusantara University, Jakarta, Indonesia 11480*
[2]*Research Interest Group in Educational Technology, Bina Nusantara University, Jakarta, Indonesia 11480*
[3]*Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480*

ARTICLE INFO

Article History
Received 27 May 2019
Accepted 14 Aug 2019

Keywords
Artificial neural networks
Backpropagation
Stochastic learning
Steepest gradient descent
Indonesian Teacher Engagement Index
Executive information system

ABSTRACT

Teacher engagement is a newly-emerged concept in the field of Indonesian teacher education. To support this concept, we designed an artificial neural network (ANN) using backpropagation, stochastic learning, and steepest gradient descent algorithms to determine teacher engagement based on the Indonesian Teacher Engagement Index (ITEI). The resulting ANN may be used in a data-gathering website for teachers to use for self-evaluation and self-intervention. The optimal architecture for the ANN has 44 input nodes, 26 first hidden layer nodes, 20 second hidden layer nodes, and 7 output nodes, with a learning rate of 0.05 and trained over 5000 iterations. The sample data used for training was gathered by ITEI researchers and the Executive Board of Indonesian Teachers Association (*Pengurus Besar Persatuan Guru Republik Indonesia*, PB-PGRI) and includes data of teachers from all around Indonesia. The maximum accuracy of this ANN was 97.98%. The sample data were then used to create an executive information system presented in the form of a map created using ArcGIS Pro software.

1. INTRODUCTION

The government of Indonesia supervises teacher competence and performance based on pedagogic competence, personality competence, social competence, and professional competence through a teacher certification program [1]. But teachers should also be engaged in their profession, not only competent and performing. To evaluate the level of teacher engagement that the teacher has, a tool called the *Indonesian Teacher Engagement Index* (ITEI) was developed.

ITEI is based on the concept of engagement in the context of education to complement the performance instruments that have been made by the Indonesian government. A teacher in Indonesia in carrying out his profession will produce optimal performance when the teacher has a positive psychological condition so that he can form a positive educational culture. Teachers are also expected to have basic competencies according to government standards and apply values in accordance with the basic philosophy of the Indonesian state and be able to apply nationalism leadership engagement in carrying out their profession. [2]. The ITEI has 6 dimensions. These dimensions are expressed in 22 indicators, with every indicator having 2 statements each, totaling 44 statements. For evaluation, the teacher is given a questionnaire with these 44 statements, to which he or she should express agreement on according to the Likert scale [3]. From these statements, the ITEI score can be determined. ITEI has 7 levels of scoring, from 1 (Disengagement), 2 (Frustrated), 3 (Burn Out), 4 (Dependent Engagement), 5 (Self-Interest Engagement), 6 (Critical Engagement), and finally 7 (Full Engagement). The data gathered by ITEI researchers and the Executive Board of Indonesian Teachers Association (*Pengurus Besar Persatuan Guru Republik Indonesia*, PB-PGRI) includes data of teachers from all around Indonesia and can potentially be categorized as big data in the future. Big data itself is characterized by high volume, velocity, and variety [4].

In addition, an executive information system is needed to present teacher engagement diagnosis in a district level, as intervention should not be top-to-bottom (from central government to district) nor bottom-to-top (from district to central government); instead, it should be in context for each district. Hopefully, with this executive system, context-based intervention based on the diagnosis is possible.

*Corresponding author. Email: sasmoko.36@gmail.com

2. RELATED WORKS

The ITEI was developed since 2015 through a grant by the Indonesian Ministry of Research and Higher Education (*Kementerian Riset dan Pendidikan Tinggi*, Kemenristekdikti) partnered with the Directorate General of Teachers and Educational Manpower (*Direktorat Jenderal Guru dan Tenaga Kependidikan*), the Indonesian Ministry of Education and Culture (*Kementerian Pendidikan dan Kebudayaan Republik Indonesia*), and the Executive Board of Indonesian Teachers Association (*Pengurus Besar Persatuan Guru Republik Indonesia*, PB-PGRI). A few applications were developed to facilitate data gathering and engagement evaluation of Indonesian teachers, such as a mobile application [5] and a website, http://www.itei.me. The teacher engagement data gathered by ITEI researchers and PB-PGRI on teachers from various regions around Indonesia has reached 10,642 validated data and around 2,000 unvalidated data. If the ITEI system is officially implemented, it may reach the status of big data due to its volume, containing millions of data of teachers from all regions of Indonesia. Big data has an important role in education, especially in teacher evaluation. It brings objectivity to the subjective and intuition-based work that is teacher evaluation [6].

While the aforementioned website uses an artificial neural network (ANN) to predict the ITEI score of a teacher, the development of the ANN used the deep learning libraries *Keras* and *Tensorflow*, resulting in an ANN with an accuracy of 97.65%. In this research, we aim to develop a more accurate ANN using the backpropagation, stochastic learning, and gradient descent algorithms [7] using less powerful libraries [9–11].

3. METHODS

3.1. System Architecture and Dataset

In building the ANN, the learning method used was supervised stochastic learning with backpropagation and steepest gradient descent algorithm, with 80% training set and 20% test set from 10,642 sample data which represent most of the regions in Indonesia.

The architecture of the ANN is as shown in Figure 1, and the algorithm flowchart is as shown in Figure 2:

The dataset used was comprised of 1 ID field, 12 demographical fields, 44 questionnaire answer fields, and 1 label field. The 44 answer fields became the input for the ANN, while the label field became the target output. The label field was one-hot encoded before being fed into the ANN, because the resulting output using the sigmoid activation function was an array of values between 0 and 1. To evaluate the accuracy of the ANN, the test set was fed into the ANN and its output compared with the actual labels, resulting in a confusion matrix.

The equations used in feedforwarding and backpropagation is shown below:

Linear combinations:

$$o_j = \sum_{i=1}^{a} w_{ij} x_i + b_1 \tag{1}$$

$$p_n = \sum_{m=1}^{b} u_{mn} h_{1m} + b_2 \tag{2}$$

$$q_l = \sum_{k=1}^{c} v_{kl} h_{2k} + b_3 \tag{3}$$

with

o_j, p_n, q_l as linear combination results for the first hidden layer nodes, second hidden layer nodes, and output layer nodes, respectively

w_{ij} as weights from the input layer to the first hidden layer

u_{mn} as weights from the first hidden layer to the second hidden layer

v_{kl} as weights from the second hidden layer to the output layer

x_i as input nodes

h_{1m} as first hidden layer nodes

h_{2k} as second hidden layer nodes

b_1, b_2, b_3 as biases of the first hidden layer, second hidden layer, and output layer, respectively

i, m, k as the index of nodes in the input layer, first hidden layer, and second hidden layer, respectively

j, m, l as the index of nodes in the first hidden layer, second hidden layer, and output layer, respectively

a, b, c as the number of nodes in the input layer, first hidden layer, and second hidden layer, respectively

Sigmoid activation function:

$$h_{1j} = \frac{1}{1 + e^{\left(-o_j\right)}} \tag{4}$$

$$h_{2n} = \frac{1}{1 + e^{\left(-p_n\right)}} \tag{5}$$

$$y_l = \frac{1}{1 + e^{\left(-q_l\right)}} \tag{6}$$

with

h_{1j} as the first hidden layer nodes' output

h_{2n} as the second hidden layer nodes' output

y_l as the output layer nodes' output

o_j, p_n, q_l as linear combination results for the first hidden layer nodes, second hidden layer nodes, and output layer nodes, respectively

Mean squared error:

$$E = \frac{1}{2}\left(t - y_l\right)^2 \tag{7}$$

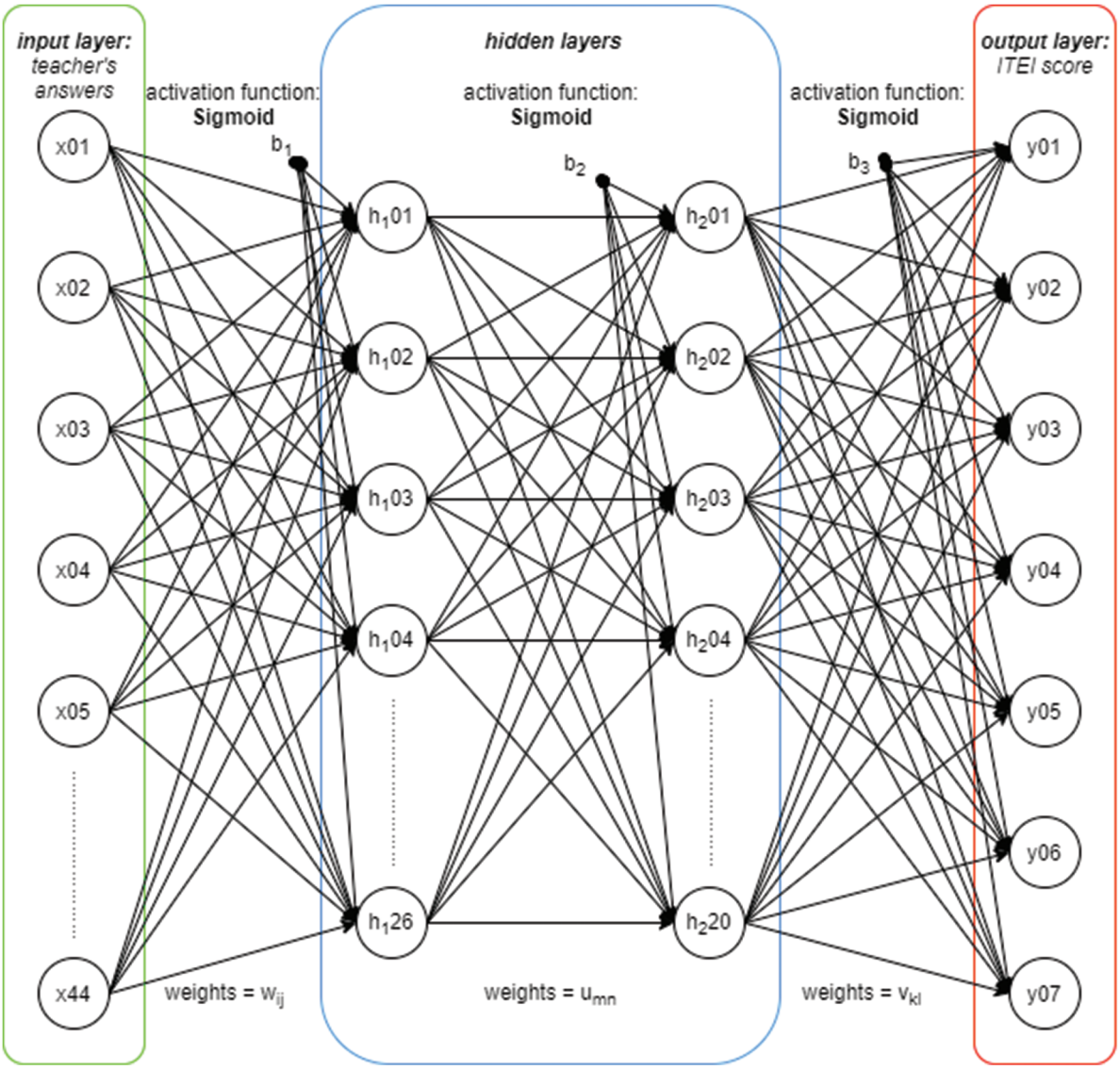


Figure 1 | Structure of the artificial neural network (ANN).

with

E as error (or loss)

t as the target output

y_l as the output layer nodes' output

Steepest gradient descent:

$$w_{new} = w_{old} + \Delta w \tag{8}$$

$$\Delta w = -\alpha g_n(w) \tag{9}$$

$$b_{new} = b_{old} + \alpha (y - t)(1 - y)(y) \tag{10}$$

with

w_{new} as the new weight value

w_{old} as the current weight value

Δw. as the adjustment made to the current weight value

α as learning te

$g_n(w)$ as a function of weight

b_{new} as the new bias value

b_{old} as the current bias value

y as a vector of the output layer nodes' output

t as a vector of the target output

Chain rule for the second hidden layer → output weights (v):

$$g_n(v) = \frac{\partial E}{\partial v}$$

$$\frac{\partial E}{\partial v} = \frac{\partial E}{\partial y}\frac{\partial y}{\partial q}\frac{\partial q}{\partial v}$$

$$\frac{\partial E}{\partial v} = (y - t)(1 - y)(y)(h_2) \tag{11}$$

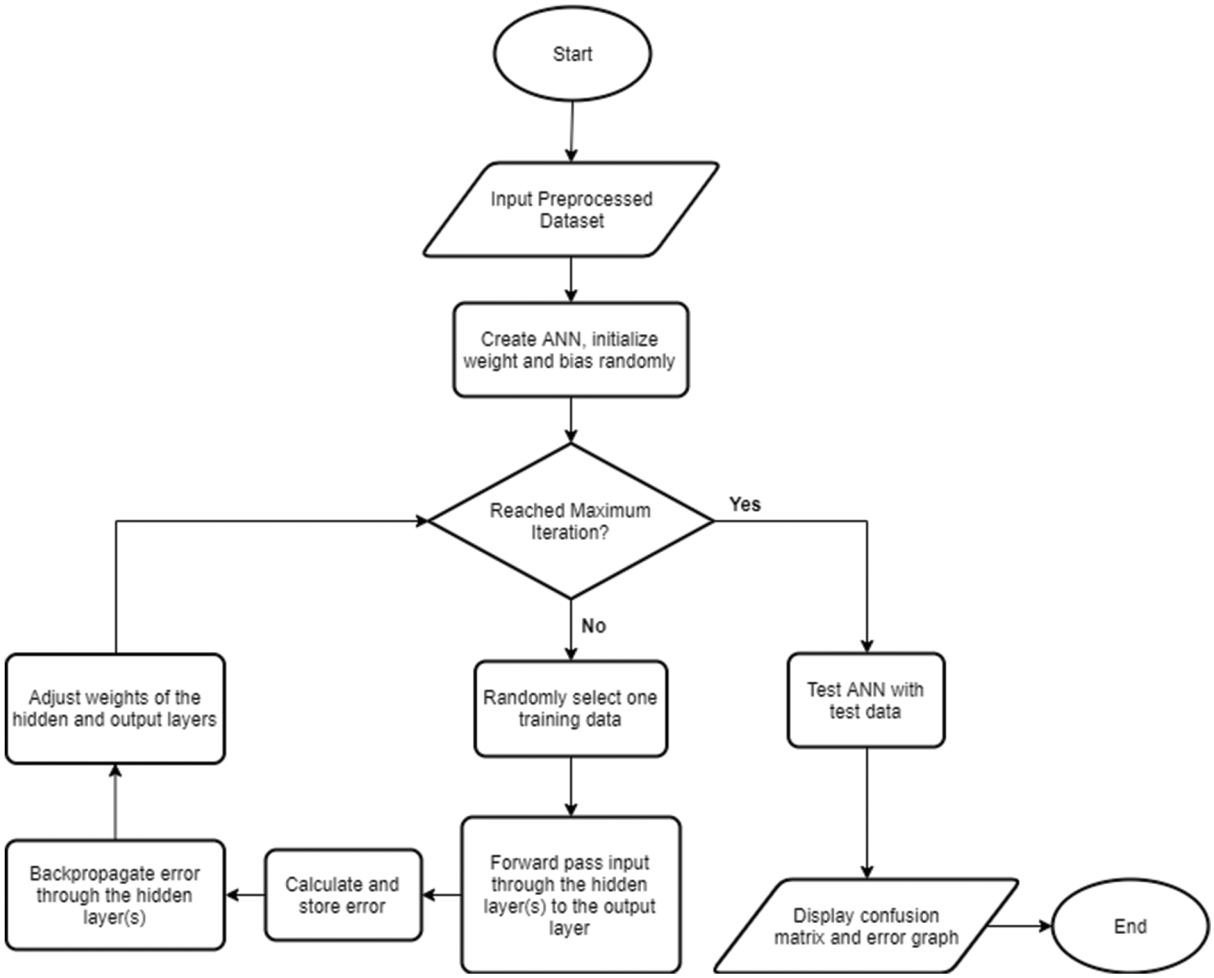


Figure 2 Flowchart of stochastic learning with backpropagation algorithm.

with

$g_n(v)$ as a function of the second hidden layer → output layer weights

$\frac{\partial E}{\partial v}$ as the partial derivative of Error against the second hidden layer → output layer weights

$\frac{\partial E}{\partial y}$ as the partial derivative of Error against the output layer

$\frac{\partial y}{\partial q}$ as the partial derivative of the output layer against the output layer's linear combination

$\frac{\partial q}{\partial v}$ as the partial derivative of the output layer's linear combination against the second hidden layer → output layer weights

Chain rule for the first hidden layer → second hidden layer weights (u):

$$g_n(u) = \frac{\partial E}{\partial u}$$

$$\frac{\partial E}{\partial u} = \frac{\partial E}{\partial y}\frac{\partial y}{\partial q}\frac{\partial q}{\partial h_2}\frac{\partial h_2}{\partial p}\frac{\partial p}{\partial u}$$

$$\frac{\partial E}{\partial u} = (y - t)(1 - y)(y)(v)(1 - h_2)(h_2)(h_1) \qquad (12)$$

with

$g_n(u)$ as a function of the first hidden layer → second hidden layer weights

$\frac{\partial E}{\partial u}$ as the partial derivative of Error against the first hidden layer → second hidden layer weights

$\frac{\partial q}{\partial h_2}$ as the partial derivative of the output layer's linear combination against the second hidden layer

$\frac{\partial h_2}{\partial p}$ as the partial derivative of the second hidden layer against the first hidden layer's linear combination

$\frac{\partial p}{\partial u}$ as the partial derivative of the second hidden layer's linear combination against the first hidden layer → second hidden layer weights

Chain rule for the input layer → first hidden layer weights (w):

$$g_n(w) = \frac{\partial E}{\partial w}$$

$$\frac{\partial E}{\partial w} = \frac{\partial E}{\partial y}\frac{\partial y}{\partial q}\frac{\partial q}{\partial h_2}\frac{\partial h_2}{\partial p}\frac{\partial p}{\partial h_1}\frac{\partial h_1}{\partial o}\frac{\partial o}{\partial w}$$

$$\frac{\partial E}{\partial w} = (y - t)(1 - y)(y)(v)(1 - h_2)(h_2)(u)(1 - h_1)(h_1)(x) \tag{13}$$

$g_n(w)$ as a function of the input layer → first hidden layer weights

$\frac{\partial E}{\partial w}$ as the partial derivative of Error against the input layer → first hidden layer weights

$\frac{\partial p}{\partial h_1}$ as the partial derivative of the second hidden layer's linear combination against the first hidden layer

$\frac{\partial h_1}{\partial o}$ as the partial derivative of the first hidden layer against the input layer's linear combination

$\frac{\partial o}{\partial w}$ as the partial derivative of the first hidden layer's linear combination against the input layer → first hidden layer weights

3.2. Data Presentation

The dataset containing the sample data was processed manually in Microsoft Excel using *PivotTable* to convert its individual primary key into regional primary key. The processed table was then imported into ArcGIS Pro and joined as attributes with a *Shapefile* table containing spatial data of Indonesian districts. The average of ITEI score for each district was then calculated and visualized in a GIS map layer.

4. RESULTS AND DISCUSSIONS

4.1. ANN with Backpropagation and Steepest Gradient Descent Algorithm

4.1.1. Initial architecture (44 input nodes, 26 hidden nodes, and 7 output nodes)

First, the ANN algorithm was tested without bias update and with the weights initialized randomly between values 0 to 1 using the *NumPy* function *random.rand()*. The initial architecture used was 44 input nodes, 26 hidden nodes, and 7 output nodes. This configuration yielded the results described in the darker bar of Figure 3 (LR = Learning Rate). As we can see, the resulting accuracy was unstable. We then added the bias update code, which improved the accuracy by a significant amount, as seen in the lighter bar of Figure 3.

Following the results above, we tried to use the *NumPy* function *random.randn()* to initialize the weights to a random Gaussian distribution of mean 0 and variance 1 [8] in order to improve the accuracy. Using this function yielded much better results with only a fraction of iterations than in the previous experiments, as can be seen in Figure 4.

4.1.2. Testing various hidden layer architectures

By this point, the accuracy is already decent, so we used the *randn()* function from this experiment forward. We then tried adding a second hidden layer to the architecture for complexity. The results are shown in Figure 5.

4.1.3. Lowering error vs. iteration graph oscillation

Even though the accuracy was decent, the error vs. iteration graph showed extreme oscillation. To minimize the oscillation, we tried lowering the learning rate. The following experiments used the architecture with 2 hidden layers that reached the highest accuracy, that is, with 20 nodes in the second hidden layer. The results were as seen in Figure 6.

The average accuracy shown in Figure 6 was obtained from doing 3 experiments for each configuration variations. The Error vs. Iteration graph of the experiments described in Figure 6 had much less oscillation than the experiments with the learning rate of 0.5, as can be seen in Figure 7.

From the data in Figure 6, we can see that the best-performing ANN was the one with a learning rate of 0.05 and trained over 5000

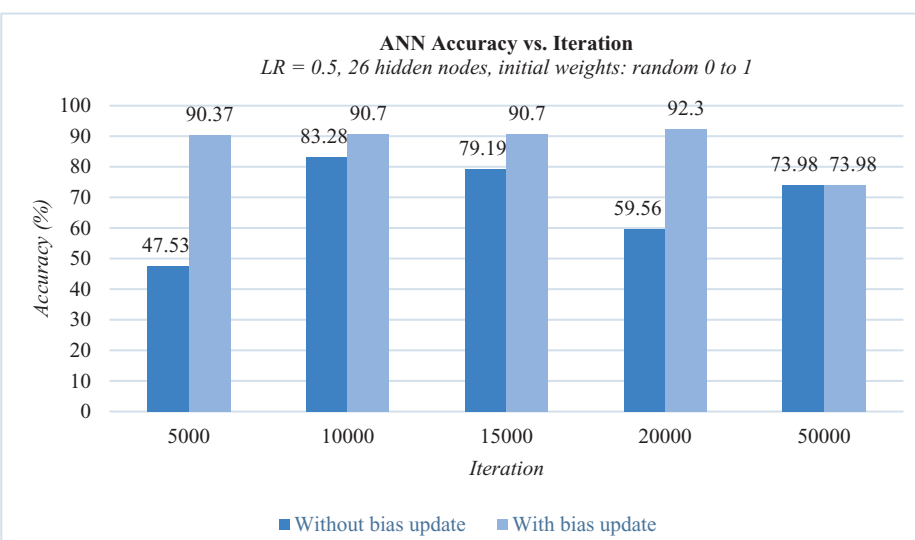


Figure 3 The accuracy of backpropagation gradient descent.

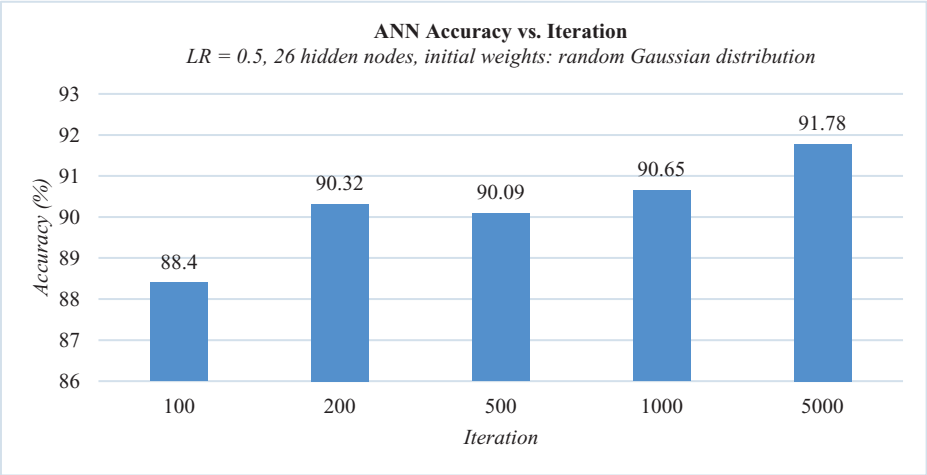


Figure 4 | The accuracy of backpropagation gradient descent with initial weights of random Gaussian distribution.

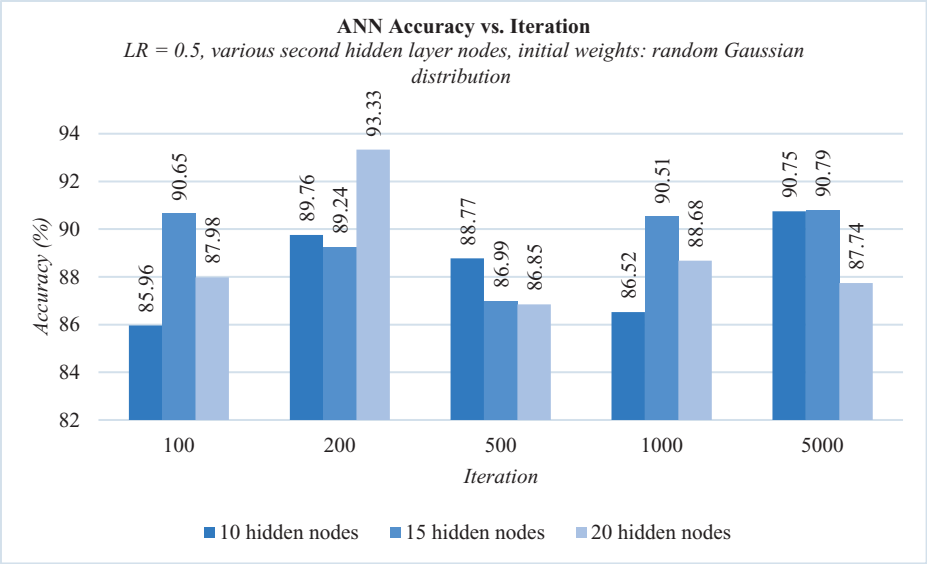


Figure 5 | The accuracy of backpropagation gradient descent with various nodes in the second hidden layer.

iterations. The maximum accuracy for this configuration was 97.98%, with the error vs. iteration graph as seen in Figure 8. The confusion matrix is as seen in Table 1, with colored cells showing true positive predictions.

As can be seen above, the ANN classifies data with the ITEI score of 1 and 2 as 4. This is most likely caused by the very little amount of data labeled with ITEI score of less than 4 in the training data.

4.2. GIS Map Presentation

A GIS map presenting the average ITEI score of each district in Indonesia was also created to allow executives to evaluate the sample data easily. The dataset containing individual data had to be processed manually in Microsoft Excel using the *PivotTable* feature to convert the individual primary key to regional primary key before joining it with spatial data in the ArcGIS Pro software, so the data in the map are not real-time data. An example of individual data is shown in Picture 1, while the processed data of Picture 1 is shown in Picture 2. The joined data can be seen in the screenshot shown in Picture 3.

The map was then published to the ArcGIS Online system and can be accessed as an embed map at http://www.itei.me/peta; the screenshot can be seen in Picture 4.

5. CONCLUSION

From the results, we can conclude that the optimal architecture of the ANN is 44 input nodes, 26 and 20 nodes in the first and

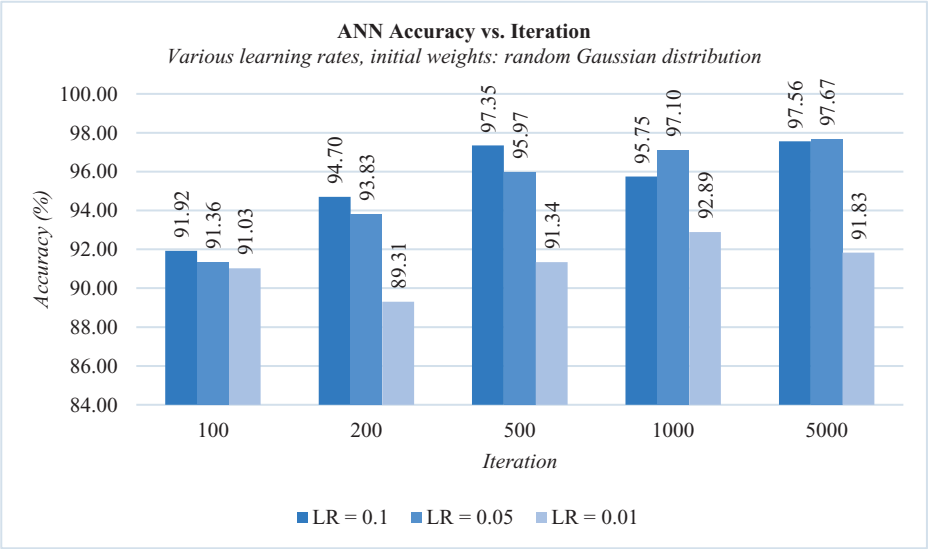


Figure 6 | The average accuracy of backpropagation gradient descent with various learning rates.

Figure 7 | Error vs. iteration for 2 hidden layers (26 and 20 nodes each), LR 0.5 (solid), and LR 0.05 (dashed), trained over 200 iterations.

second hidden layer respectively, and 7 output nodes. The optimal learning rate is 0.05, especially when trained over 5000 iterations. This configuration yielded an average accuracy of 97.67% and the maximum accuracy of 97.98%. The GIS map created to visualize the data was also able to present the data accurately, even though it must be updated manually every time a batch of new data comes in.

For future work, the ANN built in this research may be tested on more real data.

6. DECLARATIONS

List of abbreviations:

- ANN: Artificial Neural Network
- LR: Learning Rate
- ITEI: Indonesian Teacher Engagement Index

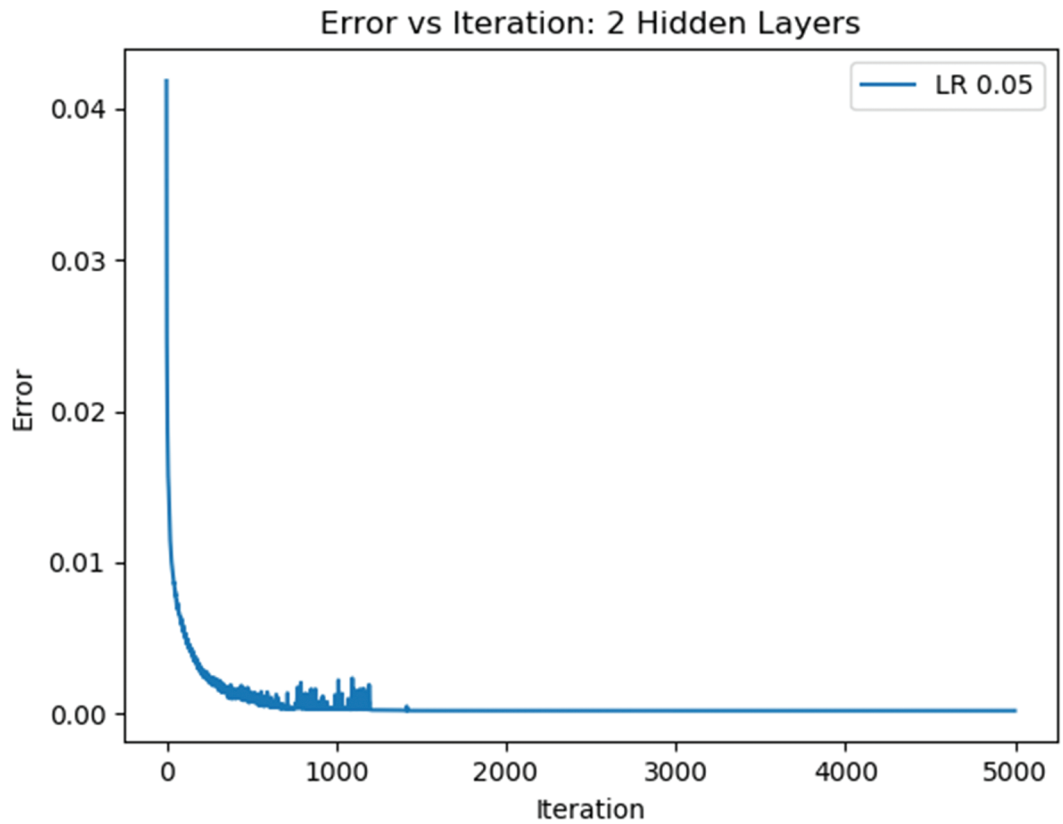


Figure 8 Error vs. iteration for 2 hidden layers (26 and 20 nodes each), LR 0.05, trained over 5000 iterations.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	id	name	provinc	district	status	certifica	gender	age	teachin	edu_ba	school_	school_	teacher
8374	8376	R8373	Nanggroe	Aceh Barat	Guru Teta	Telah Serti	Laki-laki	30-40 tahu	5-10 tahur	S1	Negeri	SMA	Guru Map
8375	8377	R8374	Nanggroe	Aceh Barat	Guru Teta	Telah Serti	Perempua	30-40 tahu	5-10 tahur	S1	Negeri	SMP	Guru Map
8376	8378	R8375	Nanggroe	Aceh Barat	Guru Teta	Belum Ser	Laki-laki	30-40 tahu	5-10 tahur	S1	Negeri	SMA	Guru Map
8377	8379	R8376	Nanggroe	Aceh Barat	Guru Teta	Belum Ser	Laki-laki	< 30 tahur	< 5 tahun	S1	Swasta	SMP	Guru Map
8461	8463	R8460	Nanggroe	Aceh Barat	Guru Teta	Telah Serti	Laki-laki	30-40 tahu	11-20 tahu	S2	Negeri	SMA	Guru Map
8462	8464	R8461	Nanggroe	Aceh Barat	Guru Teta	Belum Ser	Laki-laki	30-40 tahu	5-10 tahur	S1	Negeri	SMP	Guru Map
8463	8465	R8462	Nanggroe	Aceh Barat	Honorer	Belum Ser	Perempua	< 30 tahur	5-10 tahur	S1	Negeri	SMA	Guru Map
8464	8466	R8463	Nanggroe	Aceh Barat	Honorer	Belum Ser	Perempua	< 30 tahur	5-10 tahur	S1	Negeri	SMA	Guru Map
8482	8484	R8481	Nanggroe	Aceh Barat	Guru Teta	Telah Serti	Laki-laki	30-40 tahu	11-20 tahu	S1	Negeri	SMA	Guru Map
8483	8485	R8482	Nanggroe	Aceh Barat	Guru Teta	Belum Ser	Laki-laki	> 50 tahur	11-20 tahu	SMA	Negeri	SD	Guru Kelas
8484	8486	R8483	Nanggroe	Aceh Barat	Guru Teta	Telah Serti	Laki-laki	41-50 tahu	> 20 tahur	S1	Negeri	SD	Guru Kelas
8485	8487	R8484	Nanggroe	Aceh Barat	Guru Teta	Telah Serti	Laki-laki	> 50 tahur	11-20 tahu	SMA	Negeri	SD	Guru Kelas
8527	8529	R8526	Nanggroe	Aceh Barat	Guru Teta	Belum Ser	Laki-laki	30-40 tahu	5-10 tahur	S1	Swasta	SMA	Guru Map
8528	8530	R8527	Nanggroe	Aceh Barat	Guru Teta	Telah Serti	Laki-laki	30-40 tahu	5-10 tahur	S1	Negeri	SMA	Guru Map
8529	8531	R8528	Nanggroe	Aceh Barat	Guru Teta	Telah Serti	Perempua	30-40 tahu	5-10 tahur	S1	Negeri	SMP	Guru Map
8530	8532	R8529	Nanggroe	Aceh Barat	Guru Teta	Belum Ser	Laki-laki	30-40 tahu	5-10 tahur	S1	Negeri	SMA	Guru Map
8569	8571	R8568	Nanggroe	Aceh Barat	Guru Teta	Belum Ser	Perempua	41-50 tahu	5-10 tahur	S1	Negeri	SD	Guru Kelas
8570	8572	R8569	Nanggroe	Aceh Barat	Guru Teta	Belum Ser	Perempua	41-50 tahu	11-20 tahu	S1	Negeri	SD	Guru Kelas
8571	8573	R8570	Nanggroe	Aceh Barat	Guru Teta	Telah Serti	Perempua	30-40 tahu	11-20 tahu	S1	Negeri	SMA	Guru Map
8572	8574	R8571	Nanggroe	Aceh Barat	Guru Teta	Telah Serti	Laki-laki	41-50 tahu	11-20 tahu	S1	Negeri	SMK	Guru Map

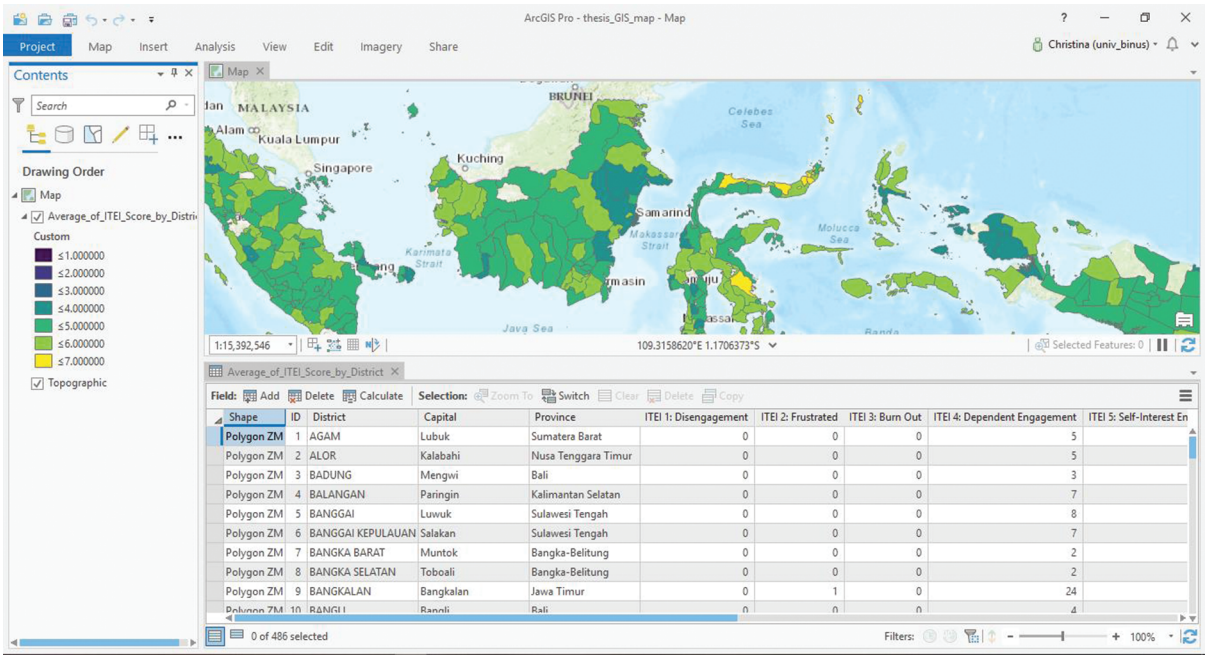
Picture 1 An example of the original dataset (individual data).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	ID	district	province	itei=1	itei=2	itei=3	itei=4	itei=5	itei=6	itei=7	age<30	age=30-40	age=41-50	age>50
2	293	Aceh Barat	Nanggroe Aceh Darussalam (NAD)	0	0	0	14	2	4	0	3	11	4	2

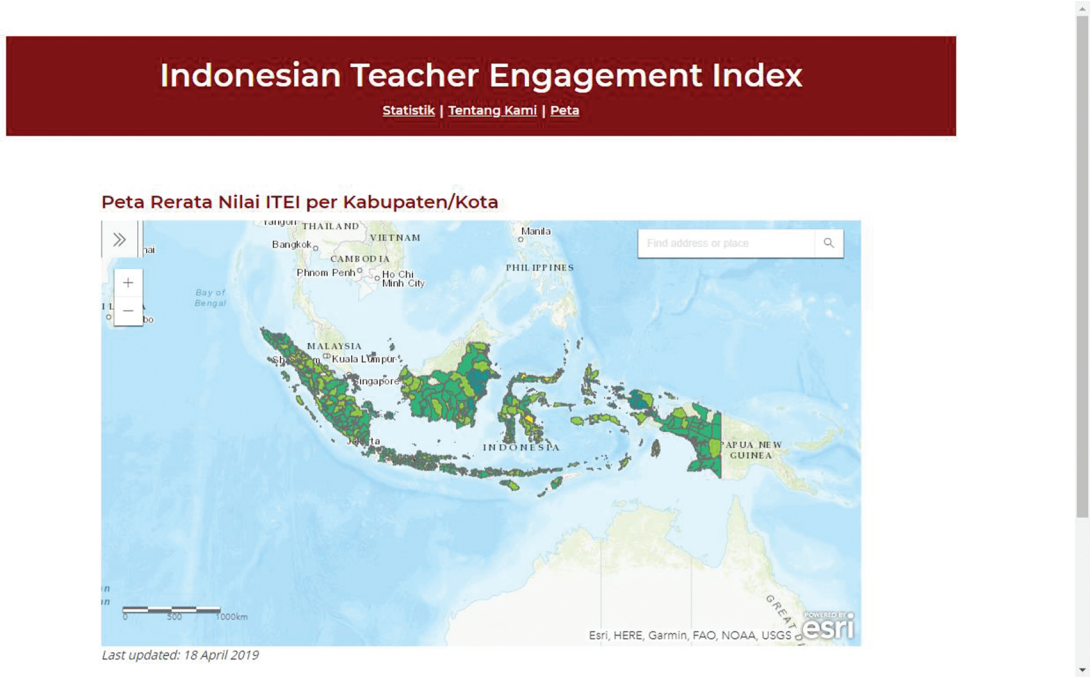
Picture 2 An example of the processed data (regional data).

Table 1 | Confusion matrix of the maximum-accuracy configuration.

		Predicted Class						
		1	*2*	*3*	*4*	*5*	*6*	*7*
Actual Class	*1*	**0**	0	0	3	0	0	0
	2	0	**0**	0	2	0	0	0
	3	0	0	**0**	0	0	0	0
	4	0	0	0	**988**	18	2	0
	5	0	0	0	13	**368**	1	0
	6	0	0	0	3	1	**641**	0
	7	0	0	0	0	0	0	**89**



Picture 3 | A screenshot of the data being processed in ArcGIS Pro.



Picture 4 | The embed GIS map as seen in http://www.itei.me/peta.

CONFLICT OF INTEREST

Authors have no conflict of interest

ACKNOWLEDGMENT

We say thanks to Bina Nusantara University, the Indonesian Ministry of Research and Higher Education for supporting this research. We also want to say thanks to the Indonesian Ministry of Education and Culture and the Indonesian Directorate of Teachers and Educational Manpower for the help in collecting the national data.

REFERENCES

[1] Presiden Republik Indonesia,Undang-Undang Republik Indonesia Nomor 14 Tahun 2005 tentang Guru dan Dosen, Indonesia, 2005.

[2] B. Sasmoko, F. Doringin, Y. Indrianti, A.M. Goni, P. Ruliana, Indonesian teacher engagement index (ITEI): an emerging concept of teacher engagement in Indonesia, IOP Conf. Ser. Mater. Sci. Eng. 306 (2018).

[3] B. Sasmoko, B.S. Abbas, Y. Indrianti, S.A. Widhoyoko, Indonesian teacher engagement index: a rasch model analysis, (2018).

[4] A. De Mauro, M. Greco, M. Grimaldi, A formal definition of big data based on its essential features, Libr. Rev. 65 (2016), 122–135.

[5] S.R. Manalu, B. Sasmoko, S.D. Permai, S.A. Widhoyoko, Y. Indrianti, Designing indonesian teacher engagement index (itei) applications based on android, J. Phys. Conf. Ser. 978 (2018).

[6] Q. Dai, Y. Chen, G. Hua, Relationship between big data and teaching evaluation, in 3rd International Conference on Education and Social Development, 2017.

[7] S.O. Haykin, Neural Networks and Learning Machines, third ed., Pearson Education, 2009.

[8] SciPy.org, numpy.random.randn, 2019. https://docs.scipy.org/doc/numpy/reference/generated/numpy.random.randn.html.

[9] K. Abdulsalam, O. Babatunde, Electrical energy demand forecasting model using artificial neural network: a case study of lagos state nigeria, Int. J. Data Netw. Sci. 3 (2019), 305–322.

[10] D. Aikhuele, A. Periola, D Ighravwe. Wind turbine systems operational state and reliability evaluation: an artificial neural network approach, Int. J. Data Netw. Sci. 3 (2019), 323–330.

[11] M. Purwadhi, The role of education management, learning teaching and institutional climate on quality of education: evidence from Indonesia, Manag. Sci. Lett. 9 (2019), 1507–1518.