

Penggunaan Jaringan Syaraf Tiruan – *Backpropagation* dalam Memprediksi Jumlah Mahasiswa Baru

Ajrina Febri Suahati^{1*}, Ahmad Arif Nurrahman², Otong Rukmana³

^{1,2,3} Teknik Industri Universitas Islam Bandung
Jl Tamansari No. 1, Bandung Wetan, Kota Bandung, Jawa Barat

^{1*} afsuahati@gmail.com

² nurrahman18@yahoo.com

^{3*} otongrukmana@gmail.com

Predicting Number of New Student Using Artificial Neural Network - Backpropagation

Dikirimkan : 07, 2021. Diterima : 02, 2022. Dipublikasikan : 03, 2022

Abstract—Educational data is important information owned by a university. Large amounts of data can be used to identify specific patterns. This study aims to predict the number of new students by identifying data patterns using an artificial neural network (ANN). ANN is an artificial intelligence that has characteristics such as biological neural networks and works as an information processing system, one of the types is backpropagation model. Backpropagation trains the network to be able to recognize and identify patterns that are given during training and provide a response in the form of predicting similar patterns, therefore backpropagation can be used in the field of forecasting. Data that uses as input for the training is new student data from 2000 to 2020, and the expected output is the number of new students in the following year, which is 2021. This study tested 4 different ANN architecture models, with MSE 0.0001, learning rate 0.01, with a maximum of 1000 iterations. The training process uses a combination of the tansig-purelin and logsig-purelin as the activation functions, and traingdx, traingda traingdm as the acceleration functions. Result of the training process, it was found that the best network architecture is the 12-8-1 pattern, which means using 12 inputs, 8 hidden neurons, and 1 output. The pattern uses the logsig-purelin activation function, the traingda acceleration function, with the MSE result of 0.0001 after going through 94 iterations.

Keywords—Artificial Neural Network (ANN); Backpropagation; Forecasting

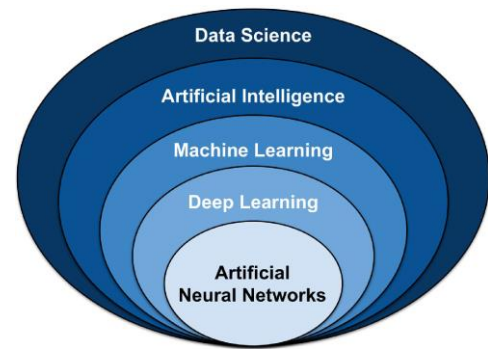
Abstrak—Data mahasiswa merupakan informasi penting yang dimiliki oleh suatu perguruan tinggi. Data dalam jumlah besar dapat digunakan untuk mengidentifikasi suatu pola tertentu. Penelitian ini bertujuan untuk memprediksi jumlah mahasiswa baru yang akan datang dengan mengidentifikasi pola data menggunakan Jaringan Syaraf Tiruan (JST). JST merupakan suatu kecerdasan buatan yang memiliki karakteristik seperti jaringan syaraf biologi dan berfungsi sebagai sistem pemroses informasi, salah satunya adalah model *backpropagation*. *Backpropagation* melatih jaringan agar mampu mengenali dan mengidentifikasi pola yang diberikan pada saat pelatihan dan memberikan respon berupa prediksi terhadap pola serupa, oleh karena itu *backpropagation* dapat digunakan dengan baik dalam bidang peramalan. Data masukan yang digunakan untuk pelatihan adalah data mahasiswa baru pada tahun 2000 sampai 2020, keluaran yang diinginkan adalah jumlah mahasiswa baru pada tahun berikutnya. Penelitian ini menguji 4 model arsitektur JST berbeda, dengan MSE 0.0001, *learning rate* 0.01, dengan maksimal 1000 iterasi. Proses pelatihan menggunakan kombinasi fungsi aktivasi *tansig-purelin* dan *logsig-purelin*, serta fungsi percepatan *traingdx*, *traingda*, dan *traingdm*. Dari pelatihan yang dilakukan, didapatkan bahwa arsitektur jaringan terbaik adalah pola 12-8-1 yang berarti menggunakan 12 masukan, 8 *hidden neuron*, dan 1 keluaran. Pola tersebut menggunakan fungsi aktivasi *logsig-purelin*, fungsi percepatan *traingda*, dengan hasil MSE 0.0001 setelah melalui 94 iterasi.

Kata kunci— Jaringan Syaraf Tiruan (JST); *Backpropagation*; Peramalan

I. PENDAHULUAN

Berkembangnya teknologi informasi menjadikan kumpulan data sebagai aset bagi suatu organisasi. Ketersediaan data tidak akan memberikan manfaat apabila data tidak diproses dengan metode analisis tertentu dan mengubahnya menjadi suatu informasi penting yang dapat mendukung suatu pengambilan keputusan [1]. Perguruan tinggi memiliki aset penting berupa data mahasiswa. Penting bagi suatu perguruan tinggi untuk dapat mengubah kumpulan data mahasiswa menjadi suatu informasi yang dapat membantu dosen, tenaga kependidikan, dan pembuat kebijakan sebagai analisis untuk mendukung pengambilan keputusan [2]. Perguruan tinggi saat ini memiliki persaingan yang ketat dan sistem yang sangat kompleks, oleh karena itu penting untuk dapat menganalisis keadaan dan mengembangkan strategi yang akan datang [3]. Memprediksi jumlah mahasiswa yang akan datang berdasarkan data historis dapat membantu perguruan tinggi untuk menentukan strategi dan mempersiapkan sumber daya yang digunakan baik sarana prasarana kegiatan belajar mengajar serta dosen dan tenaga kependidikan sebagai bagian dari kegiatan belajar mengajar. Tidak hanya itu, memprediksi jumlah mahasiswa juga dapat bermanfaat dalam menentukan strategi pemasaran ke depannya.

Data mahasiswa yang dimiliki oleh perguruan tinggi dapat diolah menggunakan berbagai metode, salah satunya menggunakan *machine learning*. *Machine learning* merupakan bagian dari kecerdasan buatan (*artificial intelligence*) yang mengeksplorasi konstruksi algoritma yang dapat mempelajari dan membuat prediksi dari suatu data [4]. *Machine learning* digunakan untuk mengajarkan mesin atau komputer untuk menangani dan memproses suatu data dengan efisien [5], menginterpretasikan data tidak dapat dilakukan hanya dengan melihat, data harus diolah untuk dapat menghasilkan informasi yang bermanfaat. Ketika algoritma dari *machine learning* diterapkan, menunjukkan bahwa model yang diimplementasikan telah menghasilkan *output* berupa informasi yang sesuai dengan data *input* yang diberikan. *Machine learning* telah berkembang sangat pesat pada saat ini sehingga menghasilkan bidang-bidang lain yang lebih spesifik. Salah satu cabang pembelajaran dari *machine learning* adalah metode Jaringan Syaraf Tiruan (JST) atau dikenal dengan istilah *Artificial Neural Network* (ANN) dapat dilihat pada Gambar 1.



Gambar 1. *Artificial neural network* dalam *artificial intelligence*

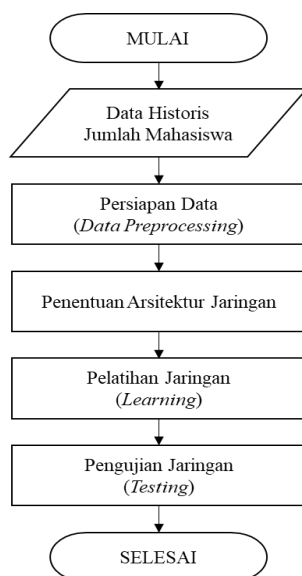
JST adalah algoritma *machine learning* yang terinspirasi dari jaringan syaraf biologi [6]. JST banyak digunakan pada bidang *data mining*, metode ini dapat digunakan pada *clustering*, *regression*, *classification*, *time series forecasting*, dan *visualization* [7]. JST memiliki kemampuan menganalisis kumpulan data rumit yang tidak mudah disederhanakan melalui metode statistik konvensional [4]. Struktur dasar JST terdiri dari tiga layer; *input*, *output*, dan *hidden layer*. *Input layer* merupakan lapisan neuron yang menerima data dari program pengguna, sedangkan *output layer* akan mengirimkan data keluaran ke dalam program pengguna. *Hidden layer* berada di antara *input* dan *output layer* dan hanya berhubungan antar neuron, tanpa berhubungan langsung dengan program [8]. Setiap neuron pada layer memiliki kemungkinan untuk mempengaruhi hasil keluaran karena pemrosesan data terjadi di setiap neuron pada jaringan syaraf tiruan [9]. Jumlah layer minimum pada jaringan adalah dua yaitu *input* dan *output layer*, sedangkan *hidden layer* bersifat opsional, dapat tidak digunakan jika memang lebih baik tidak menggunakannya [10]. Dalam JST yang sederhana, *hidden layer* berjumlah antara 0 sampai 3, sedangkan yang kompleks dapat mengandung puluhan sampai ratusan *hidden layer* [6]. Proses pelatihan pada JST dapat dilakukan secara *supervised* maupun *unsupervised*. Pelatihan *supervised* merupakan bentuk paling umum pada JST, pada metode ini pelatihan dilakukan dengan memberikan input berupa data yang sudah diberi label bersama dengan *output* yang ditentukan, sedangkan pada *unsupervised*, proses pelatihan dilakukan tanpa memberikan *output* yang diharapkan [12]. Salah satu model jaringan syaraf tiruan yang digunakan dalam peramalan adalah *Backpropagation*. *Backpropagation* memiliki kelebihan dalam kemampuan pembelajarannya. Metode ini melatih jaringan untuk mendapatkan keseimbangan antara kemampuan jaringan untuk mengenali pola yang digunakan selama pelatihan dengan kemampuan jaringan untuk memberikan

respon yang benar terhadap pola masukan yang serupa dengan pola yang digunakan saat pelatihan. *Backpropagation* adalah *algoritma iteratif* sederhana yang dapat bekerja dengan baik meskipun dengan data yang kompleks dan termasuk ke dalam model *supervised learning* karena mempelajari pola data berdasarkan masukan (*input*) dan keluarannya (*output*) [11].

Penelitian ini menggunakan Jaringan syaraf Tiruan (JST) yang melalui proses pelatihan *supervised* dengan metode *Backpropagation* dan bertujuan untuk memprediksi jumlah mahasiswa baru pada tahun berikutnya berdasarkan data historis. Hal tersebut diharapkan dapat membantu perguruan tinggi untuk mempersiapkan metode pemasaran dan fasilitas perkuliahan sehingga dapat meningkatkan layanan terhadap mahasiswa. Penelitian ini terdiri dari metodologi penelitian yang mencakup penentuan arsitektur jaringan syaraf tiruan, hasil dari penelitian mengenai arsitektur jaringan terbaik dan *performansi algoritma*, hasil dan pembahasan penelitian, serta kesimpulan, dan saran penelitian.

II. METODOLOGI PENELITIAN

Metodologi penelitian menjelaskan tahapan penelitian yang dilakukan menggunakan metode Jaringan Syaraf Tiruan (JST) - *Backpropagation*. Prosedur memodelkan jaringan syaraf tiruan secara umum dapat dibagi menjadi empat langkah, yaitu: persiapan data, pemilihan arsitektur jaringan, pelatihan (*learning*), dan pengujian (*testing*) [12] seperti dilihat pada Gambar 2.



Gambar 2. Langkah-langkah jaringan syaraf tiruan

Penelitian ini dimulai dengan mempersiapkan data yang diperlukan, yaitu data historis jumlah

mahasiswa baru. Untuk memprediksi jumlah mahasiswa baru yang akan datang sesuai dengan urutan waktunya, maka dibutuhkan data *input* berupa data dalam suatu kurun waktu tertentu secara berurutan, yang nantinya akan menjadi *output* adalah data pada waktu berikutnya. Data historis jumlah mahasiswa diharapkan dapat menggambarkan suatu pola yang dapat dikenali dan dapat diprediksi pola ke depannya oleh jaringan syaraf tiruan. Data yang digunakan pada penelitian ini adalah data mahasiswa baru sejak tahun 2000 sampai dengan tahun 2020.

Pada metode jaringan syaraf tiruan (JST), data yang akan digunakan perlu dinormalisasi terlebih dahulu sebelum proses pengolahan data. Proses ini disebut *data preprocessing* yang bertujuan untuk menyesuaikan data yang digunakan ke dalam standar metode yang digunakan. Pada jaringan syaraf tiruan, *data preprocessing* dilakukan dengan metode normalisasi data menggunakan fungsi aktivasi *sigmoid*, data diubah ke dalam interval 0.1 - 0.9.

Persamaan yang digunakan dalam normalisasi data adalah persamaan (1) berikut ini:

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

Keterangan:

- a = Data bernilai minimum
- b = Data bernilai maksimum
- x = Data yang menjadi input
- x' = Hasil normalisasi data

Tahapan selanjutnya adalah menentukan perancangan arsitektur jaringan yang akan digunakan pada proses pelatihan. Perancangan arsitektur jaringan dimulai dengan menentukan *layer* yang akan digunakan; yaitu: *input*, *output*, dan *hidden layer*. *Input layer* berisi data yang sudah diberi label dan *output layer* berisi data keluaran yang diharapkan, *input* dan *output* ini nantinya menjadi bahan pembelajaran bagi jaringan syaraf tiruan pada proses pelatihan untuk dikenali bentuk polanya. *Hidden layer* dapat terdiri lebih dari satu, lebih banyak jumlah *hidden layer* maka akan menghasilkan jaringan yang besar dengan kompleksitas yang tinggi dalam proses pelatihan jaringan [13].

Pada proses pelatihan, penelitian ini menggunakan 2 macam pola *input* yaitu 6 dan 12. Hal ini menunjukkan bahwa pelatihan dilakukan berdasarkan data 6 tahun ke belakang untuk memprediksi jumlah mahasiswa baru tahun berikutnya, sebagai contoh: data tahun 2000 sampai 2005 dilatih untuk mengetahui jumlah mahasiswa pada tahun 2006, begitu pula seterusnya. Sedangkan *output* yang digunakan adalah 1, karena

hanya memprediksi satu tahun ke depan. *Hidden layer* yang digunakan adalah 8 dan 16 untuk mengetahui jumlah *hidden layer* yang optimal. Selain menentukan jumlah *input*, *hidden* dan *output layer*, parameter lain yang perlu ditentukan adalah fungsi aktivasi, fungsi percepatan, parameter pembelajaran (*learning rate*), target *error* berupa *Mean Square Error* (MSE) dan batas iterasi maksimal yang diperbolehkan (*epoch*). Fungsi aktivasi digunakan untuk menentukan keluaran dari neuron, dalam jaringan syaraf tiruan (JST) – *Backpropagation*, fungsi *aktivasi* yang umumnya digunakan adalah fungsi *aktivasi sigmoid*. Beberapa fungsi aktivasi yang dapat digunakan dalam proses pelatihan *backpropagation* menggunakan *software matlab* adalah *tansig* (fungsi *sigmoid bipolar*), *logsig* (fungsi *sigmoid biner*), dan *purelin* (fungsi identitas). Selain fungsi *aktivasi* terdapat pula fungsi percepatan yang bertujuan untuk mempercepat proses pelatihan. MSE adalah target kesalahan atau *error* yang diinginkan, semakin kecil atau mendekati 0 maka semakin baik. *Learning rate* adalah laju pembelajaran suatu jaringan, semakin besar nilainya makan semakin cepat proses pelatihan, namun apabila *learning rate* terlalu besar maka proses pelatihan menjadi tidak stabil. Iterasi atau *epoch* digunakan untuk menentukan banyaknya iterasi yang diperbolehkan, semakin banyak jumlah iterasi makan semakin lama proses pelatihan berlangsung. Proses iterasi akan berhenti jika nilai MSE telah mencapai target atau jumlah iterasi (*epoch*) telah mencapai batas maksimum yang telah ditentukan [8].

Dari seluruh parameter yang digunakan pada Jaringan Syaraf Tiruan – *Backpropagation*, penelitian ini menggunakan beberapa parameter dalam pembentukan dan pelatihan jaringan, di antaranya sebagai berikut:

- Jumlah *input layer*: 6 dan 12
- Jumlah *hidden layer*: 8 dan 16
- Jumlah *output layer*: 1
- Fungsi aktivasi:
 - *Input - hidden layer*: *tansig* dan *logsig*
 - *Hidden layer - output*: *purelin*
- Fungsi percepatan: *trangdx*, *traingda*, *traingdm*
- Parameter pembelajaran atau *Learning Rate* (LR): 0.01
- Target *error* atau MSE (*Mean Square Error*): 0.0001
- Batas iterasi yang ditentukan (*epoch*): 1000

Terdapat 24 arsitektur jaringan yang dihasilkan dari kombinasi beberapa parameter yang telah ditentukan dapat dilihat pada Tabel I. Seluruh arsitektur jaringan tersebut akan melalui proses pelatihan untuk menemukan arsitektur jaringan terbaik yang akan digunakan.

TABEL I
KOMBINASI PARAMETER PADA PELATIHAN JARINGAN

Arsitektu r Jaringan No.	Input, Hidden, Output	Fungsi Aktivasi	Fungsi Percepatan
1	6 – 8 – 1	<i>tansig - purelin</i>	<i>traingdx</i>
2	6 – 8 – 1	<i>tansig - purelin</i>	<i>traingdm</i>
3	6 – 8 – 1	<i>tansig - purelin</i>	<i>traingda</i>
4	6 – 16 – 1	<i>tansig - purelin</i>	<i>traingdx</i>
5	6 – 16 – 1	<i>tansig - purelin</i>	<i>traingdm</i>
6	6 – 16 – 1	<i>tansig - purelin</i>	<i>traingda</i>
7	6 – 8 – 1	<i>logsig - purelin</i>	<i>traingdx</i>
8	6 – 8 – 1	<i>logsig - purelin</i>	<i>traingdm</i>
9	6 – 8 – 1	<i>logsig - purelin</i>	<i>traingda</i>
10	6 – 16 – 1	<i>logsig - purelin</i>	<i>traingdx</i>
11	6 – 16 – 1	<i>logsig - purelin</i>	<i>traingdm</i>
12	6 – 16 – 1	<i>logsig - purelin</i>	<i>traingda</i>
13	12 – 8 – 1	<i>tansig - purelin</i>	<i>traingdx</i>
14	12 – 8 – 1	<i>tansig - purelin</i>	<i>traingdm</i>
15	12 – 8 – 1	<i>tansig - purelin</i>	<i>traingda</i>
16	12 – 16 – 1	<i>tansig - purelin</i>	<i>traingdx</i>
17	12 – 16 – 1	<i>tansig - purelin</i>	<i>traingdm</i>
18	12 – 16 – 1	<i>tansig - purelin</i>	<i>traingda</i>
19	12 – 8 – 1	<i>logsig - purelin</i>	<i>traingdx</i>
20	12 – 8 – 1	<i>logsig - purelin</i>	<i>traingdm</i>
21	12 – 8 – 1	<i>logsig - purelin</i>	<i>traingda</i>
22	12 – 16 – 1	<i>logsig - purelin</i>	<i>traingdx</i>
23	12 – 16 – 1	<i>logsig - purelin</i>	<i>traingdm</i>
24	12 – 16 – 1	<i>logsig - purelin</i>	<i>traingda</i>

Pada proses pelatihan, jaringan syaraf tiruan akan mempelajari pola data historis sesuai dengan parameter yang telah ditentukan. Arsitektur jaringan terbaik akan dipilih berdasarkan nilai MSE yang memenuhi target serta tidak melebihi jumlah iterasi (*epoch*) yang telah ditentukan. Arsitektur jaringan terbaik akan digunakan pada tahapan selanjutnya yaitu tahap pengujian. Pengujian dilakukan untuk mengetahui apakah jaringan syaraf tiruan sudah dapat mengenali pola dengan baik. Jaringan akan diuji menggunakan data yang diambil secara acak pada proses pelatihan. Hasil pengujian akan menunjukkan nilai *error* dari perbandingan antara hasil prediksi jaringan syaraf tiruan dengan

data target aktual yang didapat dari data historis. Tahapan pelatihan dan pengujian jaringan dilakukan menggunakan *software matlab R2020*.

III. HASIL PENELITIAN

Penelitian dilakukan sesuai dengan tahapan yang dijelaskan pada metodologi bahwa metode Jaringan Syaraf Tiruan (JST) – *Backpropagation* dimulai dengan persiapan data. Data yang digunakan adalah data jumlah mahasiswa baru sejak tahun 2000 sampai dengan 2020. Jaringan syaraf tiruan – *Backpropagation* menggunakan fungsi *aktivasi sigmoid* sehingga data yang diolah diharuskan melalui proses normalisasi terlebih dahulu sebagai tahapan data *preprocessing*. Tabel II merupakan hasil normalisasi data jumlah mahasiswa baru. Normalisasi bertujuan untuk mengubah data aktual yang dimiliki ke dalam interval 0.1 – 0.9 menggunakan persamaan (1). Data dengan nilai maksimum memiliki nilai normalisasi 0.9 yaitu pada tahun 2000 dengan data aktual sebanyak 279 mahasiswa, sedangkan data dengan nilai minimum memiliki nilai 0.1 yaitu pada tahun 2006 dengan data aktual sebanyak 26 mahasiswa.

TABEL II
HASIL NORMALISASI DATA JUMLAH MAHASISWA

No.	Tahun	Jumlah Mahasiswa	Normalisasi
1	2000	279	0.900
2	2001	233	0.755
3	2002	202	0.657
4	2003	97	0.325
5	2004	80	0.271
6	2005	55	0.192
7	2006	26	0.100
8	2007	35	0.128
9	2008	28	0.106
10	2009	31	0.116
11	2010	31	0.116
12	2011	69	0.236
13	2012	137	0.451
14	2013	144	0.473
15	2014	171	0.558
16	2015	147	0.483
17	2016	156	0.511
18	2017	154	0.505
19	2018	159	0.521
20	2019	121	0.400
21	2020	63	0.217

Dari hasil normalisasi tersebut, data kemudian dibagi menjadi 2 pola *input*, yaitu pola data *input 6* dan 12. Pola data *input 6* menjadikan data 6 tahun sebelumnya sebagai proses pembelajaran untuk memprediksi jumlah mahasiswa baru pada tahun berikutnya sebagai *output*. Pola data *input 6* dapat dilihat pada Tabel III.

TABEL III
POLA DATA *INPUT 6*

1	2	3	4	5	6	Output
0.90	0.75	0.66	0.32	0.27	0.19	0.10
0.75	0.66	0.32	0.27	0.19	0.10	0.13
0.66	0.32	0.27	0.19	0.10	0.13	0.11
0.32	0.27	0.19	0.10	0.13	0.11	0.12
0.27	0.19	0.10	0.13	0.11	0.12	0.12
0.19	0.10	0.13	0.11	0.12	0.12	0.24
0.10	0.13	0.11	0.12	0.12	0.24	0.45
0.13	0.11	0.12	0.12	0.24	0.45	0.47
0.11	0.12	0.12	0.24	0.45	0.47	0.56
0.12	0.12	0.24	0.45	0.47	0.56	0.48
0.12	0.24	0.45	0.47	0.56	0.48	0.51
0.24	0.45	0.47	0.56	0.48	0.51	0.50
0.45	0.47	0.56	0.48	0.51	0.50	0.52
0.47	0.56	0.48	0.51	0.50	0.52	0.40
0.56	0.48	0.51	0.50	0.52	0.40	0.22
0.90	0.75	0.66	0.32	0.27	0.19	0.10
0.75	0.66	0.32	0.27	0.19	0.10	0.13

Sedangkan pola data *input 12* berarti menggunakan data 12 tahun sebelumnya sebagai proses pembelajaran untuk memprediksi jumlah mahasiswa baru pada tahun berikutnya sebagai *output*. Sedangkan pola data *input 12* dapat dilihat pada Tabel IV.

Setelah data dan parameter yang dibutuhkan telah tersedia maka tahapan selanjutnya adalah proses pelatihan jaringan. Pada proses pelatihan terdapat 24 kombinasi arsitektur jaringan berdasarkan kombinasi parameter yang telah ditentukan pada metodologi penelitian. Proses pelatihan menggunakan parameter *learning rate*, MSE, dan *epoch* yang sama untuk seluruh kombinasi. Proses pelatihan pada setiap arsitektur jaringan akan berhenti setelah mencapai MSE yang diinginkan yaitu 0.0001 atau telah mencapai batas maksimum iterasi (*epoch*) yang diperbolehkan yaitu sebanyak 1000 iterasi. Proses pelatihan seluruh arsitektur jaringan menggunakan *learning rate* sebesar 0.01. Hasil pelatihan jaringan dapat dilihat pada Tabel V.

Arsitektur jaringan terbaik dipilih berdasarkan nilai MSE yang tercapai yaitu 0.0001 dengan iterasi maksimum sebesar 1000 iterasi. Pada Tabel V dapat dilihat bahwa beberapa arsitektur jaringan mencapai nilai MSE sesuai target yang ditentukan yaitu arsitektur jaringan No. 1, 3, 4, 6, 7, 9, 10, 13, 15, 16, 18, 19, 21, 22, dan 24. Dari beberapa arsitektur jaringan tersebut, yang memiliki jumlah iterasi terkecil adalah arsitektur jaringan No. 21, yaitu sebesar 94 iterasi. Arsitektur jaringan No. 21 merupakan kombinasi dari (dapat dilihat pada Tabel 1) pola data *input 12*, *hidden neuron 8*, dan *output 1* dengan menggunakan fungsi *aktivasi logsig-purelin* dan fungsi percepatan *traingda*. Jaringan

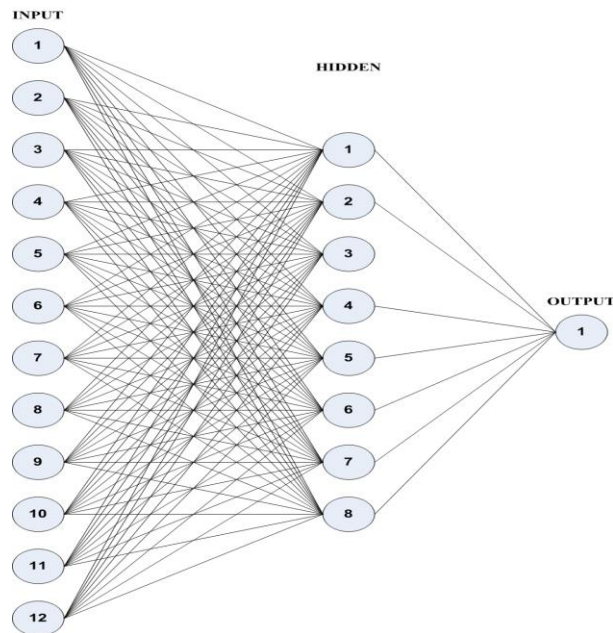
tersebut memenuhi target MSE 0.0001 dengan jumlah iterasi (*epoch*) 94 pada proses pelatihan, sehingga jaringan tersebut dinyatakan sebagai arsitektur jaringan terbaik. Arsitektur jaringan 12-8-1 dapat dilihat pada Gambar 3.

TABEL IV
POLA DATA INPUT 12

1	2	3	4	5	6	
0.90	0.75	0.66	0.32	0.27	0.19	
0.75	0.66	0.32	0.27	0.19	0.10	
0.66	0.32	0.27	0.19	0.10	0.13	
0.32	0.27	0.19	0.10	0.13	0.11	
0.27	0.19	0.10	0.13	0.11	0.12	
0.19	0.10	0.13	0.11	0.12	0.12	
0.10	0.13	0.11	0.12	0.12	0.24	
0.13	0.11	0.12	0.12	0.24	0.45	
0.11	0.12	0.12	0.24	0.45	0.47	
7	8	9	10	11	12	Output
0.10	0.13	0.11	0.12	0.12	0.24	0.45
0.13	0.11	0.12	0.12	0.24	0.45	0.47
0.11	0.12	0.12	0.24	0.45	0.47	0.56
0.12	0.12	0.24	0.45	0.47	0.56	0.48
0.12	0.24	0.45	0.47	0.56	0.48	0.51
0.24	0.45	0.47	0.56	0.48	0.51	0.50
0.45	0.47	0.56	0.48	0.51	0.50	0.52
0.47	0.56	0.48	0.51	0.50	0.52	0.40
0.56	0.48	0.51	0.50	0.52	0.40	0.22

TABEL V
HASIL PELATIHAN JARINGAN

Arsitektur Jaringan No.	MSE	Iterasi (<i>Epoch</i>)
1	0.0001	162
2	0.012	1000
3	0.0001	187
4	0.0001	162
5	0.0119	1000
6	0.0001	476
7	0.0001	385
8	0.0143	1000
9	0.0001	309
10	0.0001	274
11	0.0255	1000
12	0.000194	1000
13	0.0001	145
14	0.015	1000
15	0.0001	147
16	0.0001	132
17	0.00576	1000
18	0.0001	103
19	0.0001	195
20	0.0137	1000
21	0.0001	94
22	0.0001	158
23	0.015	1000
24	0.0001	177



Gambar 3. Arsitektur jaringan 12 – 8 - 1

Dari Gambar 3 dapat dilihat bahwa terdapat banyaknya neuron yang menghubungkan antar *layer*, yaitu *layer input-hidden* dan *hidden-output*. Pada setiap neuron terjadi suatu pemrosesan. Arsitektur jaringan tersebut nantinya akan dilanjutkan kepada proses pengujian. Proses pengujian akan menggunakan parameter sesuai dengan arsitektur jaringan terbaik pada proses pelatihan yaitu pola data *input 12*, *hidden neuron 8*, dan *output 1*, dengan menggunakan fungsi *aktivasi logsig-purelin*, dan fungsi percepatan *traingda*. Data *input* yang digunakan pada proses pengujian menggunakan sebagian data yang digunakan pada proses pelatihan untuk mengetahui apakah jaringan telah dapat mengenali pola dengan baik. Proses pengujian dapat dilihat pada Tabel VI.

TABEL VI
HASIL PENGUJIAN DAN PREDIKSI ARSITEKTUR JARINGAN TERBAIK

	Tahun	Hasil Prediksi	Data Aktual	Selisi	Error
Penguji an	2012	0.450	0.451	0.001	0.003
	2013	0.474	0.473	0.001	0.001
	2014	0.548	0.558	0.011	0.020
	2015	0.502	0.483	0.019	0.040
	2016	0.500	0.511	0.011	0.021
	2017	0.491	0.505	0.014	0.028
	2018	0.522	0.521	0.002	0.003
	2019	0.403	0.400	0.002	0.005
	2020	0.221	0.217	0.004	0.018
	Prediksi	2021	0.380	-	-

Tabel VI merupakan hasil proses pengujian terhadap data *input* baru yang diberikan. Karena jaringan terpilih menggunakan pola data *input 12*, yang berarti menggunakan data 12 tahun sebelumnya untuk memprediksi tahun yang akan datang, maka data tahun 2000–2011 dapat digunakan untuk memprediksi jumlah mahasiswa baru tahun 2012. Oleh karena itu jaringan diuji untuk memprediksi jumlah mahasiswa mulai tahun 2012 – 2020 menggunakan data sejak tahun 2020. Kemudian hasil prediksi akan dibandingkan dengan data aktual yang telah dimiliki. Hasil prediksi dan nilai *error* merupakan nilai yang dikeluarkan langsung oleh *software matlab*. Dari hasil pengujian pada Tabel VI tersebut dapat dilihat bahwa terdapat selisih antara hasil prediksi dengan data aktual. Namun hasil dari pengujian dan prediksi yang dikeluarkan langsung oleh *software matlab* adalah data yang telah melalui proses normalisasi sehingga data perlu dikembalikan terlebih dahulu melalui proses

denormalisasi. Proses *denormalisasi* mengubah data yang sebelumnya pada interval 0.1-0.9 ke dalam rentang yang sebenarnya sesuai dengan data aktual. Hasil *denormalisasi* dapat hasil prediksi dapat dilihat pada Tabel VII.

TABEL VII
DENORMALISASI HASIL PENGUJIAN DAN PREDIKSI

	Tahun	Hasil Prediksi	Data Aktual	Selisi	Error
Penguji an	2012	137	137	0	0.003
	2013	144	144	0	0.001
	2014	168	171	-3	0.020
	2015	153	147	6	0.040
	2016	153	156	-3	0.021
	2017	150	154	-4	0.028
	2018	159	159	0	0.003
	2019	122	121	1	0.005
	2020	64	63	1	0.018
	Prediksi	2021	115	-	-

Dari Tabel VII dapat dilihat bahwa terdapat selisih antara prediksi yang dihasilkan oleh jaringan syaraf tiruan dengan data aktual yang telah dimiliki sebelumnya. Besaran selisih antara hasil prediksi yang dihasilkan pada tahun 2013 sampai dengan 2020 bernilai sangat kecil sehingga nilai *error* berada di bawah nilai 0.05. Hal ini membuktikan bahwa jaringan syaraf tiruan telah dapat mengenali pola yang diberikan pada saat pelatihan dan pengujian.

Setelah hasil pengujian membuktikan bahwa jaringan dapat mengenali pola dengan baik, maka selanjutnya arsitektur jaringan terbaik juga digunakan untuk memprediksi jumlah mahasiswa baru pada tahun berikutnya yaitu pada tahun 2021. Proses prediksi juga dilakukan menggunakan *software matlab*. Hasil prediksi jaringan syaraf tiruan untuk jumlah mahasiswa baru pada tahun 2021 dapat dilihat pada Tabel VII yaitu sebesar 115 mahasiswa.

IV. PEMBAHASAN

Pada penelitian ini, Jaringan Syaraf Tiruan (JST) digunakan sebagai metode untuk mengenali pola data historis dan memprediksi pola data yang akan datang. Jaringan akan mempelajari *data set* (data historis) yang diberikan dalam bentuk *input-output* untuk dapat memprediksi data berikutnya. Sedangkan *backpropagation* diimplementasikan untuk mengevaluasi dan memodifikasi bobot yang digunakan dalam melatih jaringan untuk mengenali pola *input* dan *output* dengan baik.

Proses penelitian dilakukan mulai dari pengumpulan data, *data preprocessing*, pelatihan arsitektur jaringan sampai dengan pengujian. Berdasarkan penelitian yang telah dilakukan, ditemukan suatu arsitektur jaringan terbaik berdasarkan data *input* yang telah diberikan, yaitu data jumlah mahasiswa sejak tahun 2000 sampai dengan 2020. Arsitektur jaringan terbaik yang dihasilkan pada proses pelatihan menggunakan pola data *input* 12, *hidden neuron* 8, dan *output* 1, dengan menggunakan fungsi aktivasi *logsig-purelin*, dan fungsi percepatan *traingda*. Jaringan tersebut dapat mencapai target MSE yang telah ditentukan yaitu 0.0001 pada iterasi ke 94, yang mana lebih cepat dari arsitektur jaringan lain yang mencapai MSE sesuai target pada jumlah iterasi lebih dari 100, bahkan pada arsitektur jaringan lain tidak mencapai target MSE meskipun telah melalui batas maksimum iterasi yaitu 1000 iterasi.

Arsitektur jaringan terbaik tersebut selanjutnya diproses lebih lanjut ke tahap pengujian untuk mengetahui sejauh mana jaringan dapat mengenali *input* yang diberikan. *Input* pada proses pengujian merupakan data yang diambil sebagian pada proses pelatihan. Karena arsitektur jaringan terbaik dapat mengenali pola menggunakan data 12 tahun sebelumnya maka pengujian dimulai dengan memprediksi jumlah mahasiswa baru pada tahun 2012 berdasarkan dengan data jumlah mahasiswa baru pada tahun 2000-2011. Proses pengujian memprediksi jumlah mahasiswa baru pada tahun 2012-2020, serta tahun yang akan datang yaitu tahun 2021. Hasil pengujian dibandingkan dengan data aktual nilai *error*. Berdasarkan hasil pengujian jaringan ditemukan bahwa terdapat *error* yang sangat kecil antara hasil prediksi dengan data aktual., nilai *error* terbesar hanya bernilai 0.04 yakni hasil prediksi pada tahun 2015.

Berdasarkan nilai *error* yang sangat kecil dapat dikatakan bahwa jaringan syaraf tiruan telah dapat mengenali pola data dengan baik, sehingga dilanjutkan dengan proses *forecasting* untuk jumlah mahasiswa baru pada tahun berikutnya. Hasil prediksi menggunakan jaringan syaraf tiruan – *backpropagation* yang dijalankan dengan *software matlab* menunjukkan bahwa jumlah mahasiswa baru pada tahun 2021 adalah sebanyak 115 mahasiswa.

Dengan mengetahui jumlah mahasiswa yang akan datang, pihak perguruan tinggi diharapkan dapat mempersiapkan aspek pendukung, baik akademis maupun non-akademis dengan lebih baik. Beberapa aspek tersebut di antaranya adalah penentuan jumlah kelas dan dosen yang dibutuhkan, jadwal perkuliahan, sarana dan prasarana, serta strategi lebih lanjut untuk meningkatkan jumlah mahasiswa baru.

V. KESIMPULAN

Saat ini, data adalah salah satu aspek penting bagi suatu perguruan tinggi, karena mencakup banyak informasi baik terkait mahasiswa, dosen, publikasi ilmiah, tenaga kependidikan, dan banyak cakupan lainnya. Pengolahan data yang tepat terhadap data yang dimiliki dapat membantu pihak pimpinan dan pengelola perguruan tinggi untuk mendapatkan informasi yang bermanfaat dan menentukan strategi ke depannya. Penelitian ini berfokus pada data mahasiswa. Hasil utama dari penelitian ini adalah *algoritma* jaringan syaraf tiruan yang dapat digunakan untuk mengenali pola jumlah mahasiswa baru. Arsitektur terbaik yang ditemukan adalah dengan menggunakan pola data *input* 12, *hidden neuron* 8, dan *output* 1, dengan menggunakan fungsi *aktivasi logsig-purelin*. Berdasarkan arsitektur jaringan tersebut, diprediksi jumlah mahasiswa di tahun yang akan datang yaitu 115 mahasiswa di tahun 2021.

Pengembangan penelitian yang akan datang diharapkan dapat menggali lebih dalam pemanfaatan data mahasiswa bagi suatu perguruan tinggi khususnya dalam bidang *educational data mining* (EDM).

UCAPAN TERIMA KASIH

Penelitian ini dapat dilaksanakan dengan baik dengan bantuan dan dukungan dari Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas Islam Bandung.

REFERENSI

- [1] Stefanova, K., & Kabakchieva, D. (2017, June). Educational data mining perspectives within university big data environment. In 2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC) (pp. 264-270). IEEE.
- [2] Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. *Computers & Education*, 113, 177-194.
- [3] Khasanah, A. U. (2017, June). A comparative study to predict student's performance using educational data mining techniques. In IOP Conference Series: Materials Science and Engineering (Vol. 215, No. 1, p. 012036). IOP Publishing.
- [4] Ongsulee, P. (2017, November). Artificial intelligence, machine learning and deep learning. In 2017 15th International Conference on ICT and Knowledge Engineering (ICT&KE) (pp. 1-6). IEEE.
- [5] Mahesh, B. (2020). *Machine Learning Algorithms-A Review. International Journal of Science and Research (IJSR)*. [Internet], 9, 381-386.
- [6] Choi, R. Y., Coyner, A. S., Kalpathy-Cramer, J., Chiang, M. F., & Campbell, J. P. (2020). Introduction to machine learning, neural networks, and deep learning. *Translational Vision Science & Technology*, 9(2), 14-14.
- [7] Salloum, S. A., Alshurideh, M.44, Elnagar, A., & Shaalan, K. (2020, April). Mining in Educational Data: Review and Future Directions. In AICV (pp. 92-102).

- [8] Siang, J. J. (2005). Jaringan Syaraf Tiruan & Pemrogramannya Menggunakan Matlab.
- [9] Salah, M., Altalla, K., Salah, A., & Abu-Naser, S. S. (2018). Predicting Medical Expenses Using Artificial Neural Network. *International Journal of Engineering and Information Systems (IJEAIS)*, 2(20), 11-17.
- [10] Qwaider, S. R., Abu-Naser, S. S., & Zaqout, I. S. (2020). Artificial Neural Network Prediction of the Academic Warning of Students in the Faculty of Engineering and Information Technology in Al-Azhar University-Gaza.
- [11] Setti, S., & Wanto, A. (2019). Analysis of Backpropagation Algorithm in Predicting the Most Number of Internet Users in the World. *Jurnal Online Informatika*, 3(2), 110-115.
- [12] Nurhani, L., Gunaryati, A., Andryana, S., & Fitri, I. (2018). Jaringan Syaraf Tiruan Dengan Metode Backpropagation Untuk Memprediksi Jumlah Mahasiswa Baru. *SEMNASSTEKNOMEDIA ONLINE*, 6(1), 2-12.
- [13] Sukhbaatar, O., Usagawa, T., & Choimaa, L. (2019). An artificial neural network based early prediction of failure-prone students in blended learning course. *International Journal of Emerging Technologies in Learning (IJET)*, 14(19), 77-92.