

PREDIKSI KELULUSAN MAHASISWA PROGRAM STUDI MATEMATIKA UNIVERSITAS UDAYANA MENGGUNAKAN BACKPROPAGATION NEURAL NETWORK

Muhammad Rizki Nur Ihsan^{1§}, G. K. Gandhiadi², Luh Putu Ida Harini³

¹Program Studi Matematika – FMIPA – Universitas Udayana [e-mail: rizkinurihsan@student.unud.ac.id]

²Program Studi Matematika – FMIPA – Universitas Udayana [e-mail: gandhiadi@unud.ac.id]

³Program Studi Matematika – FMIPA – Universitas Udayana [e-mail: ballidah@unud.ac.id]

[§]Corresponding Author

ABSTRACT

Based on the rules of the BAN-PT Study Program Accreditation Instrument 4.0 (IAPS 4.0), one of the study program accreditation assessment indicators is the percentage of students who graduate on time. The percentage of Udayana University Mathematics Study Program students who graduated on time during the graduation period from 2002 to 2019 was 52.5%. It can be seen that many students fail to complete their studies on time, which has an impact on the study program accreditation assessment. Based on these problems, this study aims to help academics increase the percentage of Udayana University Mathematics Study Program graduates by predicting the graduation of mathematics study program students using the backpropagation neural network method. This study uses data on alumni of students of the 2002-2017 mathematics study program. With the BNN 5-3-1 architecture, the predicted results of graduation for students of the Udayana University mathematics study program are 73%.

Keywords: Accreditation, graduation, percentage, prediction, BNN.

1. PENDAHULUAN

Menurut buku Statistik Pendidikan Tinggi Indonesia tahun 2020 yang dipublikasi oleh Dirjen Dikti, dikatakan bahwa program sarjana di Indonesia pada tahun 2020 menerima mahasiswa baru sebanyak 1.617.869 mahasiswa dan meluluskan mahasiswanya sebanyak 1.042.844 mahasiswa. Hal ini dapat kita cermati bahwa mahasiswa yang lulus tidak sesuai dengan mahasiswa yang masuk. Sehingga, dapat disimpulkan bahwa tidak semua mahasiswa lulus tepat waktu. Berdasarkan standar pendidikan kompetensi lulusan FMIPA apabila mahasiswa program sarjana menyelesaikan masa studi kurang dari 5 tahun atau 10 semester maka mahasiswa tersebut dinyatakan lulus tepat waktu.

Mahasiswa Program Studi Matematika Fakultas MIPA Universitas Udayana yang lulus pada periode wisuda dari tahun 2002 hingga 2019 sebanyak 527 mahasiswa dan presentase mahasiswa yang lulus tepat waktu sebesar 52,5% (Srinadi & Nilakusmawati, 2020). Presentase mahasiswa yang lulus tepat waktu

ini menjadi salah satu indikator penilaian akreditasi program studi. Hal ini diatur BAN-PT pada paparan Instrumen Akreditasi Program Studi 4.0 (IAPS 4.0) kriteria Luaran dan Capaian Tridharma. Berdasarkan peraturan tersebut, untuk meningkatkan nilai mutu akreditasi suatu program studi, mengharuskan program studi di semua perguruan tinggi melakukan pemantauan dan evaluasi kegiatan akademik untuk mengambil langkah yang konkret agar mahasiswanya lulus tepat waktu. Dengan demikian, diperlukan suatu metode yang dapat memprediksi seorang mahasiswa dapat lulus dengan tepat waktu atau tidak.

Maka tujuan penelitian ini adalah memprediksi waktu kelulusan mahasiswa Program Studi Matematika Universitas Udayana menggunakan metode *Backpropagation Neural Network* (BNN). *Backpropagation* atau perambatan balik adalah salah satu ANN yang sering digunakan untuk menyelesaikan masalah peramalan atau prediksi. Metode backpropagation ini dikenal sebagai metode pelatihan ANN dengan

supervisi (Sutikno et al., 2016). Metode *backpropagation* menggunakan data pelatihan untuk menyesuaikan bobot dan ambang batas neuron sehingga dapat meminimalkan kesalahan pada prediksi jaringan (Gershenson, 2003). Penelitian ini menggunakan lima buah variabel *input* berupa nilai Indeks Prestasi Semester 1-5, dan satu buah variabel *output* berupa lama masa studi. Pada penelitian ini arsitektur BNN yang dipilih untuk diujikan digunakan *Rules of Thumb* yang akan menentukan jumlah *neuron* pada *hidden layer* (Heaton, 2008). Beberapa arsitektur BNN yang didapat menggunakan *Rules of Thumb* selanjutnya akan dilakukan pelatihan untuk mendapatkan satu arsitektur BNN yang digunakan untuk menguji akurasi model prediksi.

2. METODE PENELITIAN

2.1. Jenis dan Sumber Data

Data penelitian ini merupakan data sekunder yang bersumber dari data arsip yang dimiliki oleh Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Udayana (FMIPA UNUD).

2.2. Variabel Penelitian

Penelitian menggunakan lima variabel *input* dan satu variabel *output* yang ditunjukkan pada tabel berikut.

Tabel 1. Variabel Penelitian

Variabel	Keterangan	Satuan
Variabel <i>Input</i>		
X_1	IP Semester I	Angka (0-4)
X_2	IP Semester II	Angka (0-4)
X_3	IP Semester III	Angka (0-4)
X_4	IP Semester IV	Angka (0-4)
X_5	IP Semester V	Angka (0-4)
Variabel <i>Output</i>		
Y	Lama Masa Studi	Semester

2.3. Teknik Analisis Data

Berdasarkan jenis data yang diperoleh, sehingga tahapan teknik analisis data sebagai berikut:

1. Tahap data *preparation*: penghapusan data yang tidak lengkap.
2. Tahap pengelompokan data: mengelompokkan 630 data menjadi data *train* dan data *test*. Penelitian ini menggunakan kombinasi 80% data *train* dan 20% data *test*.
3. Tahap normalisasi data: Menormalisasi data menjadi rentang 0-1 untuk mengaktifkan fungsi aktivasi yang akan digunakan.
4. Tahap penentuan arsitektur BNN: menentukan arsitektur BNN yang memiliki 5 *neuron* masukan, 1 *neuron* luaran dan variasi banyaknya *neuron* pada *hidden layer* yang ditentukan oleh *Rules of Thumb*.
5. Tahap penentuan fungsi aktivasi, optimizer dan *stopping condition*: menentukan fungsi aktivasi yang akan digunakan, karena penelitian ini berupa prediksi maka, penggunaan fungsi aktivasi sigmoid biner sangat cocok karena memiliki *range* nilai 0 sampai 1. optimizer atau *learning rate* dengan nilai random sebesar 0.1, dan setelah beberapa percobaan pelatihan dengan berbagai jumlah epoch, didapat epoch sebanyak 100 epoch yang menghasilkan *output* yang optimal.
6. Tahap pelatihan BNN: menerapkan algoritma *BNN* pada proses data latih beberapa arsitektur yang telah ditentukan pada tahap ke-3 dan memilih arsitektur BNN dengan nilai *Loss* terkecil.
7. Tahap pengujian dan perhitungan akurasi model BNN: menggunakan arsitektur yang telah terpilih pada tahap ke-5 untuk menguji model BNN dengan memasukkan data uji pada tahap ke-2. Kemudian dihitung nilai akurasi model tersebut dengan metode *confusion matrix*.

3. HASIL DAN PEMBAHASAN

3.1. Data Preparation

Data *preparation* yang dilakukan berupa identifikasi data, menghapus data yang hilang. Data yang diperoleh merupakan data alumni mahasiswa Program Studi Matematika Universitas Udayana dari angkatan tahun 2001 sampai 2017.

Tabel 2. Deskripsi Data

	IPS 1	IPS 2	IPS 3	IPS 4	IPS 5	Masa Studi
Count	646	641	637	638	637	646
Mean	3.163529	3.114540	3.078477	2.947273	3.126091	10.021672
Max	4	4	4	4	4	42
Min	2.06	1.83	1.5	1.33	0	7
Standar deviasi	0.344239	0.410015	0.431873	0.521131	0.557032	3.009467

Pada Tabel 2 baris Count dapat dilihat jumlah data pada data input tidak konsisten hal ini berarti terdapat beberapa *missing value*. Sehingga perlu dilakukan identifikasi data dan menghapus seluruh baris pada data yang hilang (Angga Ginanjar Maburr, 2012). Setelah dilakukan identifikasi data dan menghapus seluruh baris pada data yang hilang didapatkan deskripsi data yang baru.

Tabel 3. Hasil Preparasi Data

	IPS 1	IPS 2	IPS 3	IPS 4	IPS 5	Masa Studi
Count	631	631	631	631	631	631
Mean	3.166640	3.119002	3.081854	2.947353	3.128288	9.773376
Max	4	4	4	4	4	14
Min	2.06	1.83	1.8	1.33	0	7
Std	0.342526	0.407264	0.428563	0.518766	0.556138	1.553589

3.2. Normalisasi Data

Menormalisasi data menjadi rentang 0-1 untuk mengaktifkan fungsi aktivasi sigmoid biner. Sedangkan pada variabel target, nilai diubah menjadi 1 untuk masa studi kurang dari sama dengan 9 semester dan 0 untuk masa studi lebih dari 9 semester.

3.3. Pembagian Data

Sebanyak 631 data yang dihasilkan dari proses preparasi data selanjutnya dibagi menjadi data *train* dan data *test*. Penelitian ini menggunakan kombinasi 80% data *train* dan 20% data *test* dan menghasilkan sebanyak 504

data *train* dan 127 data *test*. Pembagian data dilakukan menggunakan *train_test_split* pada *python* sehingga pembagian data dilakukan secara *random*.

3.4. Penentuan Arsitektur BNN

Arsitektur BNN yang terdiri dari layar masukan, layar tersembunyi, dan layar luaran memiliki jumlah neuron yang berbeda. Jumlah neuron pada layar masukan sebanyak variabel input, jumlah neuron pada layar luaran sebanyak variabel target, sedangkan untuk banyaknya neuron pada layar tersembunyi, jumlah neuron ditentukan didasarkan pada tiga *rules of thumb* sebagai berikut (Heaton, 2008):

1. Rule yang pertama dikatakan “jumlah neuron pada hidden layer harus berada di antara ukuran input layer dan ukuran output layer”. Jumlah neuron layar masukan pada penelitian ini adalah 5 dan jumlah neuron layar luaran adalah 1. Maka banyaknya neuron pada layar tersembunyi berada pada interval [1,5];
2. Rule yang kedua dikatakan “jumlah neuron pada hidden layer harus sebanyak $\frac{2}{3}$ ukuran input layer, dan ditambahkan ukuran output layer”. Maka, banyak neuron pada layar tersembunyi harus sebanyak $\left(\frac{2}{3} \times 5\right) + 1 = 3.33 + 1 = 4.33$ neuron;
3. Rule yang ketiga dikatakan “jumlah neuron pada hidden layer harus kurang dari dua kali ukuran input layer”. Maka, jumlah neuron pada layar tersembunyi harus kurang dari 10 neuron. Sehingga banyaknya neuron pada layar tersembunyi berada pada interval [1,9].

Berdasarkan *rules of thumb*, $[1,5] \subset [1,9]$, dan $4.33 \in [1,9]$, maka arsitektur BNN yang digunakan adalah 5 neuron masukan, 1 neuron luaran, dan beragam jumlah neuron pada layar tersembunyi sebanyak 1 sampai 9 neuron.

3.5. Pemilihan Arsitektur BNN Terbaik

Sembilan arsitektur BNN yang didapat pada 3.4 selanjutnya dilatih menggunakan algoritma *Backpropagation Neural Network* dengan banyaknya iterasi sebanyak 100 iterasi dan learning rate sebesar 0,1. Proses pelatihan tersebut dilakukan menggunakan *software jupyter notebook*.

Berdasarkan hasil yang diperoleh di setiap arsitektur BNN akan dipilih satu arsitektur

dengan rata-rata nilai fungsi *loss* terkecil dengan nilai paling konstan dari seluruh arsitektur BNN (Sapitri et al., 2020).

Tabel 4. Rataan Nilai *Cross-Entropy Loss*

Arsitektur BNN	<i>Cross-Entropy Loss</i>		
	Percobaan 1	Percobaan 2	Percobaan 3
5-1-1	0.270307803 3	0.27027019 6	0.27033457 4
5-2-1	0.269482722 4	0.27144118 2	0.26862764 6
5-3-1	0.266603884 7	0.26618291 9	0.26922957 3
5-4-1	0.268853361 2	0.27045683 6	0.26979884 4
5-5-1	0.269646890 8	0.26958721 8	0.26880165 4
5-6-1	0.271100181 5	0.27108005 5	0.26912338 4
5-7-1	0.269761603 0	0.26813119 5	0.27003080 2
5-8-1	0.268638369 2	0.27034109 9	0.27019314 9
5-9-1	0.270979742 1	0.26887818 3	0.26952108 6

Berdasarkan tabel 4, diperoleh informasi bahwa arsitektur 5-3-1 memiliki rata-rata nilai *Cross-Entropy Loss* terkecil dan paling stabil pada setiap percobaan. Oleh karena itu, arsitektur 5-3-1 terpilih sebagai arsitektur terbaik dan akan digunakan dalam proses prediksi data uji.

Setelah didapatkan arsitektur terbaik pada tabel 4, kemudian bobot dan bias akhir pada setiap *layer* menghasilkan suatu persamaan sebagai berikut:

$$Z_j = f \left(\sum_{i=1}^5 w_{ij} x_i + b_{0j} \right)$$

Keterangan:

Z_j : Nilai *output* pada *hidden layer*

W_{ij} : Nilai bobot akhir *hidden layer*

X_i : Input

B_{0j} : Nilai bias akhir *hidden layer*

i = jumlah unit *input layer*

j = jumlah unit *hidden layer*

$$y_k = f \left(\sum_{j=1}^3 w_{jk} Z_j + b_{0k} \right)$$

Keterangan:

Y_k : Nilai *output*

W_{jk} : Nilai bobot akhir *output layer*

Z_j : Input *hidden layer*

B_{0k} : Nilai bias akhir *output layer*

j = jumlah unit *hidden layer*

k = jumlah unit *output layer*

3.6. Menghitung Nilai *CEL* Data *Testing*

Data *test* yang telah ditentukan sebelumnya kemudian digunakan untuk menghitung *Cross-Entropy Loss* dan proses BNN dengan arsitektur BNN 5-3-1 yang terpilih. Nilai *CEL* tersebut disajikan pada tabel sebagai berikut:

Tabel 5. Nilai *CEL*

k	Nilai Aktual (t_k)	Nilai Prediksi (y_k)	$t_k \log(y_k)$
1	0	0.60966198	0
2	1	0.78408838	-0.24323353
.	.	.	.
126	1	0.8389876500	-0.17555929
127	1	0.6371366400	-0.45077115
$\frac{-\sum_{k=1}^{127} t_k \log(y_k)}{127}$			0.20776363
Ket: t_k = nilai aktual y_k = nilai prediksi k = jumlah data			

3.7. Menghitung Akurasi Model Prediksi

Dari Arsitektur BNN 5-3-1 yang terpilih selanjutnya digunakan untuk menguji akurasi model pelatihan. Pengujian dilakukan dengan cara memasukkan hasil model prediksi pada data testing kedalam metode Confusion Matrix untuk mengukur performa prediksi yang diteliti dengan cara membandingkan nilai prediksi dan nilai aktual.

Pengukuran jumlah ketepatan algoritma yang dipakai menggunakan matrix 2x2 yang disebut confusion matrix (Hizham et al., 2018). Disajikan pada tabel 6 sebagai berikut (Widodo et al., 2013):

Tabel 6. Confusion Matrix

		Kelas Aktual	
		True	False
Kelas Prediksi	True	TP (True Positive)	FP (False Positive)
	False	FN (False Negative)	TN (True Negative)

Keterangan:

1. True Positive (TP): Jika data prediksi memiliki nilai positif yang sesuai dengan nilai sebenarnya (positif)
2. False Positive (FP): Jika data prediksi tidak sesuai dengan nilai sebenarnya
3. False Negative (FN): Jika data prediksi bernilai negatif dan data sebenarnya bernilai positif
4. True Negative (TN): jika data prediksi dan data sebenarnya bernilai negative.

Hasil dari pengujian model prediksi menggunakan python terhadap data *test* menggunakan metode confusion matrix dapat disajikan sebagai berikut:

Tabel 7. Hasil Akurasi

		Kelas Aktual	
		True	False
Kelas Prediksi	True	33	24
	False	10	60

Selanjutnya akan dilakukan uji performansi berdasarkan confusion matrix dari model prediksi, dengan rumus sebagai berikut:

1. Akurasi = $\frac{(TP + TN)}{(TP + FP + FN + TN)}$
2. Presisi = $\frac{TP}{(TP + FP)}$
3. Recall = $\frac{TP}{(TP + FN)}$
4. F-Measure = $\frac{(2 \times recall \times presisi)}{(recall + presisi)}$

Perhitungan akurasi, presisi, recall dan F-Measure sebagai nilai uji performansi tersebut adalah sebagai berikut:

$$Akurasi = \frac{(TP + TN)}{(TP + FP + FN + TN)} = \frac{33 + 60}{33 + 24 + 10 + 60} = \frac{93}{127} = 0.732283465 = 73.2283465\% \approx 73\%$$

$$Presisi = \frac{TP}{(TP + FP)} = \frac{33}{33 + 24} = \frac{33}{57} = 0.57894736842 = 57.894736842\% \approx 58\%$$

$$Recall = \frac{TP}{(TP + FN)} = \frac{33}{33 + 10} = \frac{33}{43} = 0.76744186046 = 76.744186046\% \approx 77\%$$

$$F-Measure = \frac{(2 \times recall \times presisi)}{(recall + presisi)} = \frac{2 \times 0.76744186046 \times 0.57894736842}{0.76744186046 + 0.57894736842} = \frac{0.8886168911}{1.346389229} = 0.66 = 66\%$$

Berdasarkan perhitungan uji performansi diatas dapat disimpulkan bahwa penggunaan metode *Backpropagation Neural Network* untuk memprediksi waktu kelulusan mahasiswa Prodi Matematika Universitas Udayana memiliki 73% nilai akurasi, 58% nilai presisi, 77% nilai recall, dan 66% nilai F-Measure. Hal tersebut dapat diartikan bahwa model prediksi tersebut dapat digunakan untuk memprediksi kelulusan mahasiswa Program Studi Matematika FMIPA Unud.

3.7. Prediksi Kelulusan Mahasiswa Matematika Unud

Arsitektur BNN 5-3-1 yang telah terpilih selanjutnya digunakan untuk memprediksi waktu kelulusan mahasiswa Prodi Matematika Unud. Prediksi dilakukan dengan cara memasukkan data mahasiswa Program Studi Matematika Unud angkatan 2018. Data yang terkumpul sebanyak 5 data.

Hasil prediksi yang dilakukan dengan metode *Backpropagation* yang menggunakan arsitektur terpilih 5-3-1 adalah sebagai berikut:

Tabel 8. Hasil Prediksi

Data ke-	Nilai Aktual	Nilai Prediksi
1	1	0.667067234 \approx 1
2	1	0.059419126 \approx 0
3	1	0.054674362 \approx 0
4	1	0.884415635 \approx 1
5	1	0.181567513 \approx 0

Berdasarkan pada Tabel 7, dari kelima data yang digunakan dalam proses prediksi, hanya ada 2 data yang sesuai dengan nilai aktual, sehingga akurasi prediksi didapatkan hanya sebesar 40%.

4. SIMPULAN DAN SARAN

4.7. Simpulan

Berdasarkan hasil penelitian mengenai metode *Backpropagation Neural Network* dalam memprediksi waktu kelulusan mahasiswa Program Studi Matematika Universitas Udayana dapat disimpulkan:

1. Model BNN terbaik yang digunakan dalam memprediksi kelulusan mahasiswa Program Studi Matematika Universitas Udayana yaitu model BNN dengan arsitektur 5 neuron pada layar masukan, 3 neuron layar tersembunyi, dan 1 neuron layar luaran dengan menghasilkan nilai *cross-entropy loss* sebesar 0.20776363.
2. Metode confusion matrix digunakan untuk menghitung tingkat akurasi dalam penelitian ini. Nilai yang dihasilkan berupa nilai uji performansi dengan nilai akurasi sebesar 73%, presisi sebesar 58%, recall sebesar 77%, dan F-Measure sebesar 66%. Maka hasil tersebut berarti model prediksi dapat digunakan untuk memprediksi kelulusan mahasiswa Prodi Matematika FMIPA Unud.
3. Prediksi yang dilakukan pada data mahasiswa angkatan 2018 sebanyak lima mahasiswa mendapatkan hasil prediksi sebesar 40% hal ini berarti model yang digunakan untuk memprediksi kelulusan mahasiswa tepat waktu sebanyak lima data masih kurang baik.

4.8. Saran

Adapun saran yang dapat disampaikan berdasarkan temuan penelitian ini adalah sebagai berikut:

1. Atribut masukan yang digunakan dalam penelitian ini masih sangat terbatas, dan faktor yang memengaruhi mahasiswa lulus tepat waktu cukup banyak sehingga hal ini perlu dikaji ulang dengan atribut masukan yang lebih luas.
2. Disarankan untuk melakukan studi lebih lanjut mengenai prediksi kelulusan mahasiswa pada studi kasus yang lebih luas menggunakan metode BNN dengan jumlah variabel yang lebih banyak dan kompleks.

DAFTAR PUSTAKA

- Angga Ginanjar Mabrur, L. R. (2012). Penerapan Data Mining Untuk Memprediksi Kriteria Nasabah Kredit. *Jurnal Komputer Dan Informatika (KOMPUTA)*, 1(1), 53–57.
- BAN-PT. (2019). Instrumen Akreditasi Program Studi 4.0. *Risetdikti*, April, 1–69.
- Gershenson, C. (2003). *Artificial Neural Networks for Beginners*.

Heaton, J. (2008). *Introduction to Neural Networks with C#* (Vol. 99).

Hizham, F. A., Nurdiansyah, Y., & Firmansyah, D. M. (2018). Implementasi Metode Backpropagation Neural Network (BNN) dalam Sistem Klasifikasi Ketepatan Waktu Kelulusan Mahasiswa (Studi Kasus: Program Studi Sistem Informasi Universitas Jember). *Berkala Sainstek*, 6(2), 97.

PDDikti. (2020). *Higher Education Statistics 2020*. 81–85.

Sapitri, N. K. E., Kencana, I. P. E. N., & Harini, L. P. I. (2020). Penerapan Artificial Neural Network (ANN) untuk Memprediksi Perubahan Derajat Miopia pada Manusia. *Jurnal Matematika*, 10(1), 53.

Srinadi, I. G. A. M., & Nilakusmawati, D. P. E. (2020). Analisis Waktu Kelulusan Mahasiswa Fmipa Universitas Udayana Dan Faktor-Faktor Yang Memengaruhinya. *E-Jurnal Matematika*, 9(3), 205.

Sutikno, Indriyanti, Sukmawati, N. E., Priyo, S. S., Helmie, A. W., Indra, W., Nurdin, B., & K, T. W. (2016). *Backpropagation dan Aplikasinya*. 135–146.

Widodo, P. P., Handayanto, R. T., & Herlawati. (2013). Penerapan Data Mining dengan Matlab. In *Rekayasa Sains*.