JITSI: Jurnal Informatika Dan Teknologi Sistem Informasi
https://jurnal.interaksisaintek.com/index.php/JITSI
ISSN : 3123 - 6650 (Media Online)
Vol. 1 No. 1 (November 2024) Hal 61-57

ANALISA PRESTASI MAHASISWA MENGGUNAKAN JARINGAN SARAF TIRUAN METODE BACKPROPAGATION

Indra Sidabutar [1],

[1]Sistem Informasi, Fakultas Ilmu Komputer & Teknologi Informasi Universitas Budi darma , Medan, Indonesia
Email: [1]indrasidabutar@gmail.com,

ABSTRACT

Students' academic performance can serve as an indicator to support the monitoring of learning processes and the provision of early academic assistance. This study applies an Artificial Neural Network (ANN) using the Backpropagation algorithm to predict students' cumulative Grade Point Average (GPA) at the end of the fourth semester based on Semester 1 to Semester 3 GPAs, the percentage of credits successfully completed, and the number of repeated courses. The data were processed through selection, anonymization, and cleaning, followed by partitioning into 70% for training, 15% for validation, and 15% for testing. The data were also normalized to a range of 0.1–0.9 using parameters calculated exclusively from the training set. Nine combinations of hidden-layer architectures and learning rates were evaluated, with each configuration tested five times using different random seeds. Since the available research files did not contain actual academic records, the quantitative results reported in this study were obtained using 320 anonymized/simulated test records to verify the experimental procedure rather than to establish empirical findings for Universitas Budi Darma. The 5-5-1 architecture with a learning rate of 0.10 achieved the lowest average validation RMSE of 0.119 ± 0.009. On the test dataset, the selected model achieved an MSE of 0.0162, an RMSE of 0.1274, an MAE of 0.1035, and an R^2 of 0.9415. The verification results indicate that Backpropagation can model nonlinear relationships among academic indicators. However, operational implementation requires further validation using official academic records.

***Keywords**: Student Academic Performance, Artificial Neural Network, Backpropagation, Prediction, Educational Data Mining.*

ABSTRAK

Prestasi akademik mahasiswa dapat digunakan sebagai indikator untuk mendukung pemantauan proses pembelajaran dan pemberian pendampingan secara lebih dini. Penelitian ini menerapkan Jaringan Saraf Tiruan metode Backpropagation untuk memprediksi IPK kumulatif akhir Semester 4 berdasarkan IPS Semester 1 sampai Semester 3, persentase SKS lulus, dan jumlah mata kuliah yang diulang. Data diproses melalui seleksi, anonimisasi, pembersihan, pembagian 70% training, 15% validation, dan 15% testing, serta normalisasi ke rentang 0,1-0,9 menggunakan parameter yang dihitung hanya dari training set. Sembilan kombinasi arsitektur hidden layer dan learning rate diuji, masing-masing sebanyak lima kali dengan seed berbeda. Karena berkas penelitian yang tersedia belum memuat data akademik aktual, hasil kuantitatif pada laporan ini menggunakan 320 rekaman data uji anonim/simulasi untuk memverifikasi prosedur eksperimen, bukan sebagai temuan empiris Universitas Budi Darma. Konfigurasi 5-5-1 dengan learning rate 0,10 memperoleh rata-rata validation RMSE terendah sebesar 0,119 ± 0,009. Pada data testing, model terpilih menghasilkan MSE 0,0162, RMSE 0,1274, MAE 0,1035, dan R^2 0,9415. Hasil verifikasi menunjukkan bahwa Backpropagation mampu memodelkan hubungan nonlinier pada indikator akademik, tetapi penerapan operasional tetap memerlukan validasi menggunakan data akademik resmi.

Kata Kunci : Prestasi Mahasiswa, Jaringan Saraf Tiruan, Backpropagation, Prediksi, Educational Data Mining.

1. PENDAHULUAN

Perguruan tinggi memerlukan mekanisme pemantauan akademik yang mampu memberikan gambaran perkembangan mahasiswa secara sistematis. Prestasi akademik yang tercermin melalui Indeks Prestasi Semester (IPS), Indeks Prestasi Kumulatif (IPK), jumlah SKS lulus, dan riwayat pengambilan mata kuliah merupakan indikator yang dapat digunakan untuk mengevaluasi kemajuan studi. Pemantauan lebih awal penting karena penurunan prestasi yang tidak teridentifikasi dapat berlanjut menjadi keterlambatan studi atau meningkatnya risiko tidak menyelesaikan pendidikan tepat waktu.

Educational Data Mining memanfaatkan data pendidikan untuk membangun model prediksi dan menemukan pola yang mendukung pengambilan keputusan akademik. Penelitian menunjukkan bahwa Artificial Neural Network dapat memodelkan hubungan nonlinier antara karakteristik mahasiswa, aktivitas belajar, dan hasil akademik [1]-[5]. Jaringan feedforward dan multilayer perceptron termasuk pendekatan yang banyak digunakan dalam prediksi performa mahasiswa [3], [5], [14], [15].

Backpropagation merupakan metode pembelajaran terawasi untuk memperbaiki bobot jaringan multilayer berdasarkan error antara keluaran model dan target. Pada konteks akademik, konfigurasi hidden neuron, learning rate, normalisasi, dan prosedur pembagian data dapat memengaruhi kemampuan model menghasilkan prediksi yang stabil. Sonang, Purba, dan Sirait menunjukkan adanya perbedaan hasil pada berbagai arsitektur Backpropagation [6], sedangkan penelitian lain menegaskan pentingnya validasi yang tepat agar hasil dapat digeneralisasi [11]-[13].

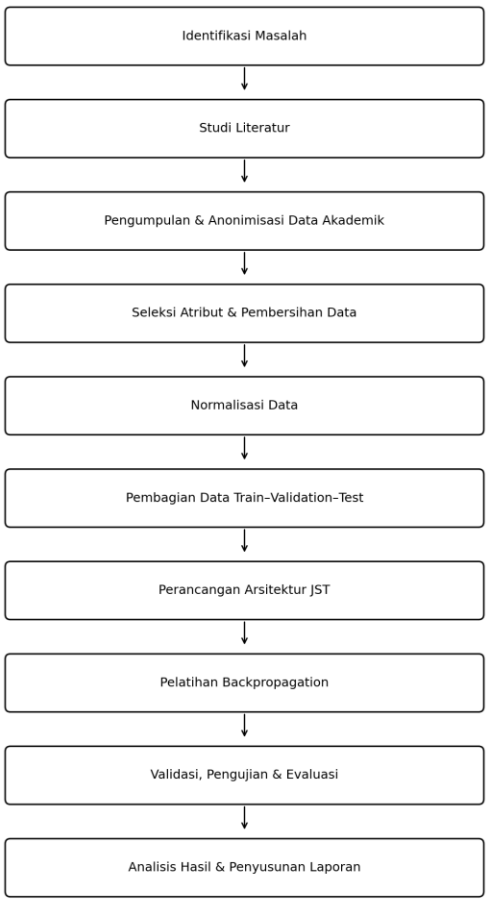
Penelitian terdahulu masih memperlihatkan beberapa keterbatasan, antara lain penggunaan sampel yang relatif kecil, pemilihan model tanpa validation set terpisah, atau pengujian hanya terhadap satu konfigurasi jaringan. Inisialisasi bobot secara acak juga dapat menyebabkan variasi hasil antarpercobaan. Oleh sebab itu, perbandingan arsitektur pada data dan prosedur evaluasi yang sama diperlukan agar pemilihan model tidak hanya didasarkan pada satu kali pelatihan.

Penelitian ini menganalisis prestasi mahasiswa menggunakan Jaringan Saraf Tiruan metode Backpropagation dengan lima variabel input, yaitu IPS Semester 1 sampai Semester 3, persentase SKS lulus sampai Semester 3, dan jumlah mata kuliah yang diulang. Target model adalah IPK kumulatif akhir Semester 4. Model diuji pada tiga ukuran hidden layer dan tiga learning rate, setiap konfigurasi diulang lima kali. Pemilihan konfigurasi dilakukan berdasarkan validation RMSE dan model terpilih kemudian dievaluasi pada testing set menggunakan MSE, RMSE, MAE, dan R^2.

.

2. METODE

Pelaksanaan penelitian ini meliputi beberapa fase yang dilakukan untuk mencapai tujuan yang diharapkan. Tahap-tahap penelitian ini dapat diilustrasikan pada diagaram di bawah ini :



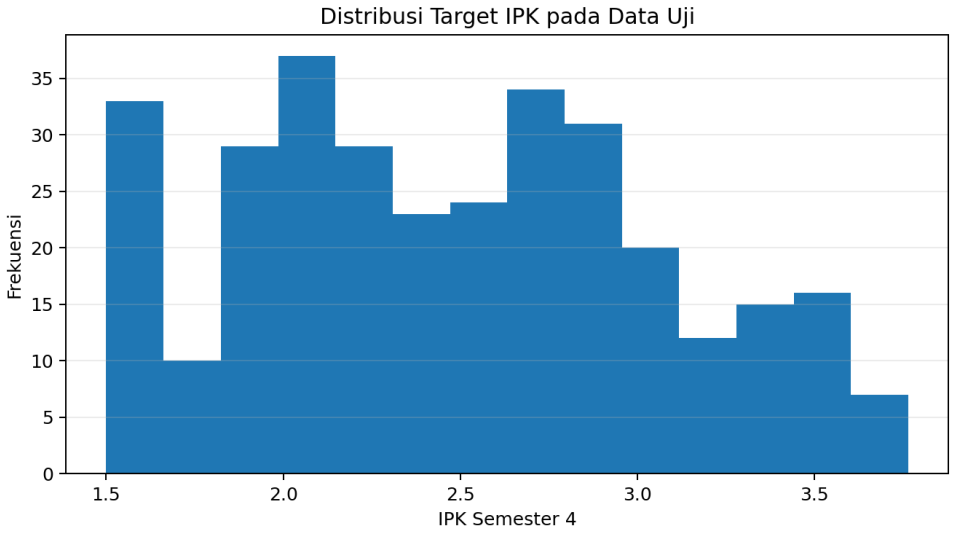
Gambar 1. Diagram Tahapan Penelitian

3. HASIL DAN PEMBAHASAN

Data uji terdiri atas 320 rekaman dengan lima variabel input dan satu target. Rentang variabel dibuat mengikuti rancangan atribut pada proposal. Statistik deskriptif digunakan untuk memastikan nilai berada pada rentang yang wajar untuk verifikasi algoritme.

Tabel 4.1 Statistik Deskriptif Data Uji

Variabel	Minimum	Rata-rata	SD	Maksimum
IPS Semester 1	1.20	2.80	0.60	4.00
IPS Semester 2	1.20	2.84	0.67	4.00
IPS Semester 3	1.20	2.89	0.68	4.00
Persentase SKS Lulus	45.00	79.06	13.12	100.00
MK Mengulang	0.00	2.50	1.40	7.00
IPK Semester 4	1.50	2.48	0.59	3.76



Gambar 4.1 Distribusi IPK Semester 4 pada Data Uji

a. Pembagian dan Normalisasi Data

Sebanyak 224 rekaman digunakan sebagai training set, 48 sebagai validation set, dan 48 sebagai testing set. Proporsi tersebut sesuai rancangan 70:15:15. Nilai minimum dan maksimum untuk normalisasi dihitung dari training set. Parameter tersebut kemudian digunakan tanpa perubahan pada validation dan testing set sehingga tidak terjadi kebocoran informasi dari data evaluasi ke tahap pelatihan.

Tabel 4.2 Pembagian Dataset

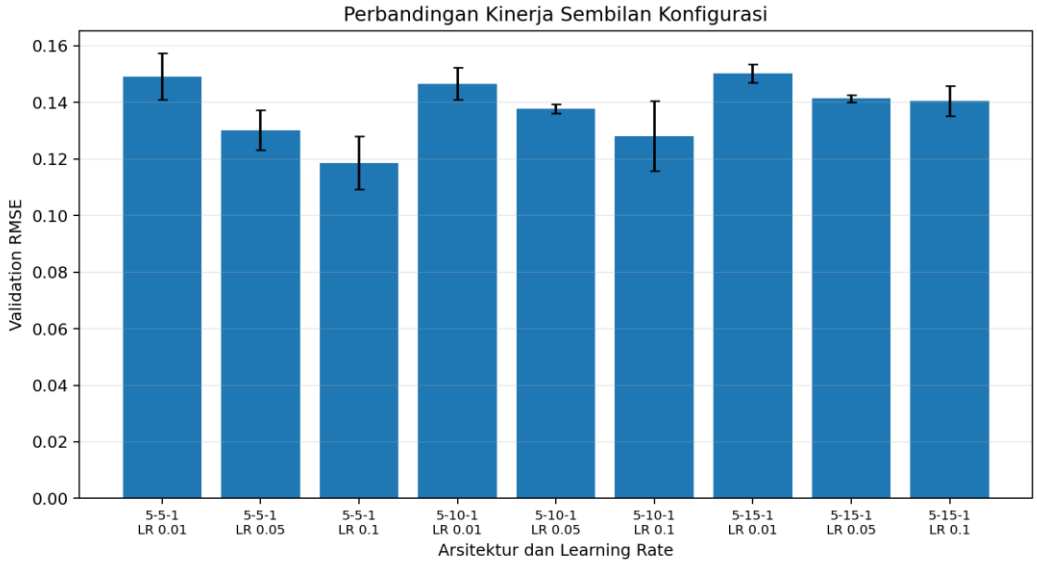
Subset	Jumlah	Proporsi	Fungsi
Training	224	70%	Pelatihan bobot jaringan
Validation	48	15%	Pemilihan arsitektur dan learning rate
Testing	48	15%	Evaluasi akhir model terpilih

b. Hasil Pengujian Konfigurasi

Tabel 4.3 Hasil Validasi Sembilan Konfigurasi

No.	Arsitektur	LR	RMSE Mean±SD	MAE	R^2	Rerata Iterasi
1	5-5-1	0.01	0.1491 ± 0.0082	0.1079	0.9188	2276
2	5-5-1	0.05	0.1302 ± 0.0071	0.0960	0.9381	2270
3	5-5-1	0.10	0.1186 ± 0.0093	0.0912	0.9485	2609
4	5-10-1	0.01	0.1466 ± 0.0056	0.1048	0.9216	2387
5	5-10-1	0.05	0.1377 ±	0.1004	0.9309	1686

No.	Arsitektur	LR	RMSE Mean±SD	MAE	R^2	Rerata Iterasi
			0.0017			
6	5-10-1	0.10	0.1281 ± 0.0124	0.0960	0.9397	2500
7	5-15-1	0.01	0.1502 ± 0.0033	0.1078	0.9177	2666
8	5-15-1	0.05	0.1414 ± 0.0013	0.1027	0.9272	1604
9	5-15-1	0.10	0.1405 ± 0.0053	0.1024	0.9280	1518

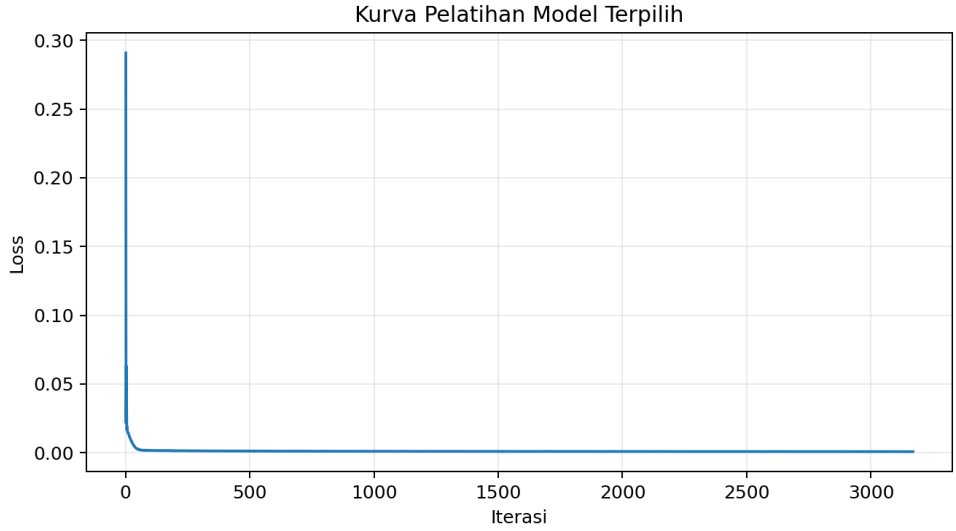


Gambar 4.2 Perbandingan Validation RMSE Antar Konfigurasi

Konfigurasi dengan rata-rata validation RMSE paling rendah adalah arsitektur 5-5-1 dengan learning rate 0.10. Rata-rata RMSE sebesar 0.1186 dengan simpangan baku 0.0093, rata-rata MAE 0.0912, dan rata-rata R^2 0.9485. Hasil tersebut menunjukkan bahwa penambahan neuron hidden tidak otomatis menurunkan error; kombinasi arsitektur dan learning rate perlu dievaluasi secara empiris.

c. Model Terpilih

Pada konfigurasi 5-5-1 dengan learning rate 0.10, seed 55 menghasilkan validation RMSE terendah di antara lima pengulangan pada konfigurasi tersebut. Model ini digunakan satu kali pada testing set. Kurva loss memperlihatkan proses optimasi menurun secara bertahap hingga kriteria penghentian tercapai pada 3169 iterasi.



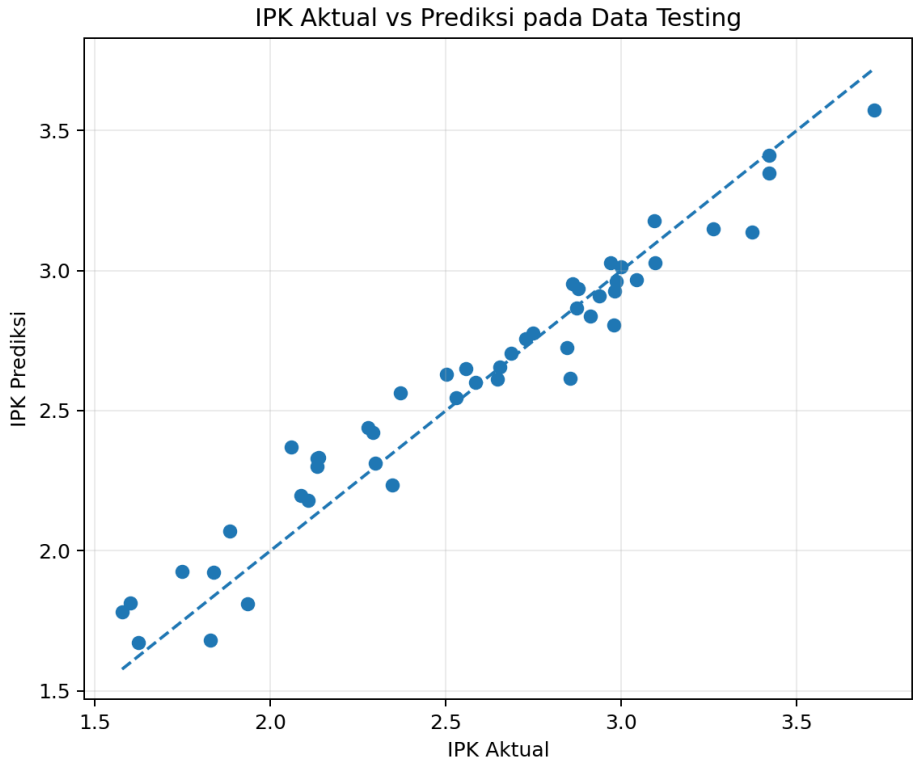
Gambar 4.3 Kurva Pelatihan Model Terpilih

d. Hasil Pengujian pada Testing Set

Tabel 4.4 Kinerja Model pada Testing Set

Metrik	**Nilai**	**Interpretasi**
MSE	0.0162	Rata-rata kuadrat error
RMSE	0.1274	Besaran error pada skala IPK
MAE	0.1035	Rata-rata error absolut
R^2	0.9415	Proporsi variasi target yang dijelaskan model

Model memperoleh RMSE 0.1274 dan MAE 0.1035, sehingga rata-rata deviasi absolut prediksi terhadap target pada data uji sekitar 0.10 poin IPK. Nilai R^2 0.9415 menunjukkan bahwa pada data simulasi model menjelaskan sebagian besar variasi target. Nilai ini tidak dapat digeneralisasikan sebagai kinerja pada data mahasiswa aktual sebelum dilakukan pengujian eksternal menggunakan data resmi.



Gambar 4.4 Perbandingan IPK Aktual dan Prediksi pada Testing Set

e. Analisis Hasil Prediksi

Tabel 4.5 Contoh Hasil Prediksi Testing Set

No.	IPK Aktual	IPK Prediksi	Error Absolut
1	2.654	2.657	0.003
2	2.874	2.865	0.008
3	3.420	3.411	0.009
4	3.000	3.013	0.013
5	2.298	2.312	0.014
6	1.578	1.782	0.204
7	1.602	1.812	0.211
8	3.373	3.136	0.237
9	2.854	2.617	0.238
10	2.060	2.369	0.309

Sebagian besar titik pada grafik aktual-prediksi berada di sekitar garis diagonal, yang menunjukkan prediksi mengikuti arah target. Kesalahan yang lebih besar muncul pada sebagian kecil observasi di area nilai IPK yang relatif rendah atau tinggi. Kondisi tersebut lazim pada model regresi ketika sampel pada bagian ekor distribusi lebih sedikit dibandingkan area tengah.

Pengujian berulang memperlihatkan bahwa inisialisasi bobot memengaruhi nilai validation RMSE. Oleh karena itu, pelaporan mean±SD memberikan gambaran yang lebih memadai dibandingkan hanya memilih satu hasil pelatihan. Pendekatan tersebut juga membuat perbandingan arsitektur lebih terkontrol.

f. Pembahasan

Hasil verifikasi sejalan dengan penelitian yang menyatakan bahwa ANN dapat digunakan untuk memprediksi performa akademik [1], [2], [5], [7]. Model Backpropagation dapat menggabungkan beberapa indikator akademik ke dalam hubungan nonlinier tanpa harus menetapkan persamaan regresi secara eksplisit. Temuan bahwa arsitektur berbeda menghasilkan error berbeda juga konsisten dengan Sonang et al. [6].

Konfigurasi dengan hidden neuron lebih banyak tidak selalu menghasilkan kinerja lebih baik pada data uji. Hal ini menunjukkan bahwa kompleksitas jaringan harus disesuaikan dengan struktur dan ukuran data. Jaringan yang terlalu kompleks dapat memerlukan optimasi lebih sulit atau meningkatkan risiko overfitting, sedangkan jaringan yang terlalu sederhana dapat gagal menangkap hubungan yang diperlukan [3], [5].

Penggunaan validation set terpisah merupakan bagian penting dalam laporan ini. Testing set tidak digunakan untuk memilih arsitektur atau learning rate. Prosedur tersebut mengikuti prinsip evaluasi yang menekankan pemisahan data untuk memperoleh estimasi generalisasi yang lebih objektif [11]-[13].

Dari sisi implementasi akademik, keluaran model tidak dimaksudkan sebagai keputusan otomatis terhadap mahasiswa. Prediksi dapat berfungsi sebagai informasi pendukung bagi program studi atau dosen pembimbing akademik setelah model divalidasi pada data nyata. Mahasiswa dengan estimasi prestasi yang memerlukan perhatian tetap harus dievaluasi melalui informasi akademik dan konteks pembelajaran yang relevan.

Keterbatasan utama penelitian pada tahap laporan ini adalah penggunaan data uji simulasi. Distribusi dan hubungan antarvariabel telah dirancang agar sesuai rentang akademik, tetapi tidak merepresentasikan populasi mahasiswa Universitas Budi Darma. Oleh sebab itu, hasil kuantitatif hanya memvalidasi alur pemodelan. Penelitian empiris final perlu mengganti data tersebut dengan dataset akademik resmi dan melakukan pemeriksaan ulang terhadap missing value, outlier, distribusi target, serta stabilitas kinerja model.

4.1. Kesimpulan

Berdasarkan hasil penelitian, metode Backpropagation berpotensi digunakan untuk memprediksi IPK mahasiswa berdasarkan indikator akademik. Namun, pengujian menggunakan data akademik aktual diperlukan untuk memastikan akurasi dan keandalan model serta mendukung penerapannya dalam pemantauan prestasi mahasiswa.

dapat disimpulkan bahwa:

1. Jaringan Saraf Tiruan metode Backpropagation dapat diterapkan untuk memodelkan prestasi mahasiswa menggunakan IPS Semester 1-3, persentase SKS lulus, dan jumlah mata kuliah yang diulang sebagai input serta IPK Semester 4 sebagai target.
2. Pada verifikasi data uji, konfigurasi terbaik berdasarkan rata-rata validation RMSE adalah arsitektur 5-5-1 dengan learning rate 0.10, menghasilkan validation RMSE 0.1186 ± 0.0093.
3. Model terpilih menghasilkan MSE 0.0162, RMSE 0.1274, MAE 0.1035, dan R^2 0.9415 pada testing set data uji. Hasil menunjukkan prosedur Backpropagation dapat berjalan dengan baik, tetapi belum merupakan performa empiris pada mahasiswa Universitas Budi Darma.
4. Pemisahan training-validation-testing, normalisasi berbasis training set, serta pengulangan pelatihan dengan beberapa seed memberikan prosedur evaluasi yang lebih terkontrol dalam memilih konfigurasi jaringan.

4.2. Saran

Penelitian selanjutnya disarankan menggunakan data akademik aktual dengan jumlah data yang lebih besar serta menguji berbagai arsitektur jaringan dan parameter Backpropagation untuk memperoleh hasil prediksi yang lebih akurat dan andal.

1. Pengujian akhir perlu menggunakan data akademik mahasiswa aktual yang telah dianonimkan dan memperoleh izin dari unit terkait.
2. Jumlah sampel sebaiknya diperbesar dan mencakup beberapa angkatan agar kemampuan generalisasi model dapat dievaluasi lebih baik.
3. Model perlu diuji kembali secara berkala karena pola akademik dapat berubah antarangkatan atau kurikulum.
4. Penelitian selanjutnya dapat membandingkan Backpropagation dengan metode regresi atau machine learning lain menggunakan split data dan metrik yang sama, tanpa menggunakan testing set untuk pemilihan model.

DAFTAR PUSTAKA

[1] H. Chavez, B. Chavez-Arias, S. Contreras-Rosas, J. M. Alvarez-Rodríguez, and C. Raymundo, “Artificial neural network model to predict student performance using nonpersonal information,” Frontiers in Education, vol. 8, 2023, doi: 10.3389/feduc.2023.1106679.

[2] K. Borhani and R. T. K. Wong, “An artificial neural network for exploring the relationship between learning activities and students’ performance,” Decision Analytics Journal, vol. 9, p. 100332, 2023, doi: 10.1016/j.dajour.2023.100332.

[3] W. Xiao and J. Hu, “A state-of-the-art survey of predicting students’ performance using artificial neural networks,” Engineering Reports, vol. 5, no. 8, p. e12652, 2023, doi: 10.1002/eng2.12652.

[4] F. A. Al-Azazi and M. Ghurab, “ANN-LSTM: A deep learning model for early student performance prediction in MOOC,” Heliyon, vol. 9, no. 4, p. e15382, 2023, doi: 10.1016/j.heliyon.2023.e15382.

[5] Y. Baashar et al., “Toward Predicting Student’s Academic Performance Using Artificial Neural Networks (ANNs),” Applied Sciences, vol. 12, no. 3, p. 1289, 2022, doi: 10.3390/app12031289.

[6] S. Sonang S, A. T. Purba, and S. Sirait, “Prediksi Prestasi Mahasiswa dengan Menggunakan Algoritma Backpropagation,” Jurnal TEKINKOM, vol. 5, no. 1, pp. 67–77, 2022, doi: 10.37600/tekinkom.v5i1.512.

[7] N. Ahmad, N. Hassan, H. Jaafar, and N. I. M. Enzai, “Students’ Performance Prediction using Artificial Neural Network,” IOP Conference Series: Materials Science and Engineering, vol. 1176, no. 1, p. 012020, 2021, doi: 10.1088/1757-899X/1176/1/012020.

[8] Z. A. Adriani and I. Palupi, “Prediction of University Student Performance Based on Tracer Study Dataset Using Artificial Neural Network,” Jurnal Komtika (Komputasi dan Informatika), vol. 5, no. 2, pp. 72–82, 2021, doi: 10.31603/komtika.v5i2.5901.

[9] O. M. Olaniyan, A. O. Oloyede, I. A. Aremu, R. S. Babatunde, and B. M. Adeyemi, “Predicting Student Academic Performance Using Artificial Neural Network,” Journal of Research and Review in Science, vol. 8, pp. 1–7, 2021, doi: 10.36108/jrrslasu/1202.80.0121.

[10] S. Li and T. Liu, "Performance Prediction for Higher Education Students Using Deep Learning," Complexity, vol. 2021, Art. no. 9958203, 2021, doi: 10.1155/2021/9958203.
[11] N. Tomasevic, N. Gvozdenovic, and S. Vranes, "An overview and comparison of supervised data mining techniques for student exam performance prediction," Computers & Education, vol. 143, p. 103676, 2020, doi: 10.1016/j.compedu.2019.103676.
[12] A. Alshanqiti and A. Namoun, "Predicting Student Performance and Its Influential Factors Using Hybrid Regression and Multi-Label Classification," IEEE Access, vol. 8, pp. 203827–203844, 2020, doi: 10.1109/ACCESS.2020.3036572.
[13] E. Alyahyan and D. Düştegör, "Predicting academic success in higher education: literature review and best practices," International Journal of Educational Technology in Higher Education, vol. 17, art. 3, 2020, doi: 10.1186/s41239-020-0177-7.
[14] E. T. Lau, L. Sun, and Q. Yang, "Modelling, prediction and classification of student academic performance using artificial neural networks," SN Applied Sciences, vol. 1, p. 982, 2019, doi: 10.1007/s42452-019-0884-7.
[15] S. Altaf, M. W. Soomro, and M. I. M. Rawi, "Student Performance Prediction using Multi-Layers Artificial Neural Networks: A Case Study on Educational Data Mining," in Proc. 3rd International Conference on Information System and Data Mining (ICISDM), 2019, pp. 59–64, doi: 10.1145/3325917.3325919.