

PENERAPAN *ARTIFICIAL NEURAL NETWORK* UNTUK MENDUGA PROGRAM STUDI DI UNIVERSITAS UDAYANA BERDASARKAN NILAI RAPOR

I Putu Bayu Sulaksana Putra^{1§}, I Putu Eka N. Kencana², Luh Putu Ida Harini³

¹Program studi Matematika, Fakultas MIPA – Universitas Udayana [Email: bayu.aksana.p@gmail.com]

²Program studi Matematika, Fakultas MIPA – Universitas Udayana [Email: i.putu.enk@unud.ac.id]

³Program studi Matematika, Fakultas MIPA – Universitas Udayana [Email: ballidah@unud.ac.id]

§ *Corresponding Author*

ABSTRACT

This research aims to develop an artificial neural network-based estimation system to predict the suitable study program at Udayana University for high school students in Denpasar City based on their report cards. The research is divided into four stages: system overview, user interface design, implementation of the artificial neural network in the system, and system testing. System testing results on report card data for science and social science classes demonstrate that the developed model has good accuracy with an error rate below 7%.

Keywords: *Artificial neural network, model.*

1. PENDAHULUAN

SNMPTN (Seleksi Nasional Masuk Perguruan Tinggi Negeri) merupakan salah satu jalur penerimaan mahasiswa baru di perguruan tinggi negeri di Indonesia (Permenristekdikti, 2016). SNMPTN dimulai dengan pengumpulan nilai rapor pendaftar dan dilanjutkan dengan seleksi berdasarkan program studi yang dipilih. Memilih program studi yang tepat sangat penting karena dapat mempengaruhi peluang diterima di perguruan tinggi. Oleh karena itu, diperlukan analisis nilai rapor yang teliti untuk menentukan program studi yang memiliki kemungkinan kelulusan terbaik. Namun, analisis nilai rapor dapat menjadi sangat rumit dan beragam, sehingga sulit dilakukan oleh orang awam.

Untuk mengatasi kendala tersebut, penulis akan menggunakan metode *artificial neural network* dalam penelitian ini. *Artificial neural network* umumnya digunakan untuk memprediksi dan mempelajari pola dari kumpulan data, hasil penelitian Dreiseitl et al (2002) pada kasus pengklasifikasian dengan menggunakan satu *hidden layer* mampu menghasilkan tingkat prediksi yang baik pada hampir semua jenis *dataset*. *Hidden layer* adalah *layer* atau lapisan yang berada pada *input layer* dan *output layer* yang berfungsi membantu *network* mengenali pola-pola kompleks

(Heaton, 2011). Meskipun demikian, perlu diperhatikan bahwa tingkat kesalahan *artificial neural network* sangat dipengaruhi oleh metode pembelajaran yang digunakan dalam pelatihan model. Metode pembelajaran adalah proses dimana model diperbaiki dan disesuaikan dengan data yang diberikan untuk mencapai kemampuan pengenalan pola atau pemecahan masalah yang lebih baik (Goodfellow et al., 2016).

Beberapa penelitian sebelumnya mengenai metode *artificial neural network* telah terbukti memiliki performa yang baik dalam kasus klasifikasi, di antaranya adalah Kahar et al (2015) yang menganalisis penerapan metode *artificial neural network* model *perceptron* dan *backpropagation* dalam penentuan bidang kompetensi skripsi mahasiswa pada program studi Teknik Informatika STMIK Nuridin Hamzah Jambi. Selain itu, Hizham et al (2018) menggunakan metode BNN untuk mengklasifikasi ketepatan waktu kelulusan mahasiswa di Program Studi Sistem Informasi Universitas Jember. Riani et al (2021) membuat model seleksi peminatan sesuai dengan kebijakan terbaru pada Program Studi Sistem Informasi Universitas Telkom 2020 dengan bantuan *machine learning*. Ahmad et al (2021) memprediksi performa siswa dengan

menggunakan metode *artificial neural network*. Lau et al (2019) memodelkan performa akademik di salah satu universitas di Cina dengan menggunakan metode *artificial neural network*. Saravanan et al (2014) juga menunjukkan bahwa model *artificial neural network* yang dilatih dengan metode *backpropagation* mampu menghasilkan tingkat keakuratan yang lebih baik jika dibandingkan dengan metode *maximum likelihood*.

Oleh karena itu, penulis akan menggunakan metode ini untuk merancang sistem penduga yang dapat membantu siswa SMA di Kota Denpasar dalam memilih program studi yang sesuai di Universitas Udayana. Penelitian ini bertujuan untuk merancang sistem penduga dengan metode *artificial neural network* untuk menentukan program studi di Universitas Udayana yang sesuai untuk siswa SMA di Kota Denpasar berdasarkan nilai rapor. Melalui penelitian ini, diharapkan dapat memberikan kontribusi dalam meningkatkan efektivitas dan efisiensi dalam penentuan program studi yang tepat bagi siswa SMA.

2. METODE PENELITIAN

Penelitian ini dilakukan di Kota Denpasar dalam periode 2021-2022 dan menggunakan data nilai rapor siswa dari periode 2019-2020 sebagai data sekunder. Data tersebut didapat dari delapan SMA di Kota Denpasar. Variabel penelitian dalam penelitian ini adalah variabel *input* untuk model data siswa IPA (X_1) yang terdiri dari nilai rapor; Matematika, Bahasa Indonesia, Bahasa Inggris, peminatan Matematika, Kimia, Fisika, dan Biologi dalam 5 semester. Variabel *input* untuk model data siswa IPS (X_2) yang terdiri dari nilai rapor; Matematika, Bahasa Indonesia, Bahasa Inggris, Sejarah, Ekonomi, Geografi dan Sosiologi dalam 5 semester. Variabel *output* dari model data IPA (y_1) dan data IPS (y_2) adalah program studi yang diterima lulus jalur SNMPTN berdasarkan nilai rapor yang diinput dalam pendaftaran SNMPTN.

Metode analisis data yang digunakan dalam perancangan model *artificial neural network* adalah sebagai berikut:

- Penyamaan format, rasio, dan pembersihan data dari data yang tidak valid.
- Penentuan ambang error awal yang dapat diterima dalam penelitian ini adalah 10%. Tingkat ambang kesalahan umumnya dipengaruhi oleh kompleksitas jaringan dan

pentingnya keakuratan model terhadap hasil yang ingin dicapai. Menurut Loy, James (2019), model dapat dianggap layak digunakan jika tingkat keakuratannya mencapai minimal 75%, sementara tingkat keakuratan 95% dianggap sangat baik. Tujuan penelitian ini adalah menghasilkan prediksi program studi sebagai rekomendasi pendamping siswa dalam memilih program studi, bukan sebagai pilihan yang mutlak. Oleh karena itu, tingkat kesalahan sebesar 10% dipilih sebagai tingkat yang dapat diterima.

- Pembagian data menjadi data *training* dan data *testing*.
- Menentukan *neuron* pada *layer input* sebanyak tujuh mata pelajaran selama lima semester (35).
- Menentukan *neuron* pada *layer output* yang mewakili jenis program studi yang dihasilkan oleh model.
- Menentukan banyaknya *hidden layer*, dengan menggunakan satu *hidden layer*. Menurut Hassoun et al (1996), penggunaan satu *hidden layer* dianggap sebagai pendekatan yang paling umum dalam merancang arsitektur untuk metode *neural network classification*.
- Menentukan jumlah *neuron* pada *hidden layer*
- Penginisiasian bobot awal dilakukan secara acak (Nielsen, 2015), dengan bobot adalah anggota elemen bilangan real di antara rentang 0 sampai 1.
- Melakukan pembentukan model menggunakan data *training*.
- Menghitung besar *error* in-sample dengan mencari nilai *error* rata-rata antara prediksi data *training* dengan *output* sebenarnya dari data *training*.
- Meminimasi *loss function* untuk memperkecil besar *error*, kemudian melakukan penginisiasian bobot dengan bobot hasil minimasi *loss function*.
- Melakukan pengujian dengan data *testing*.

3. HASIL DAN PEMBAHASAN

3.1 Perancangan Arsitektur

Tahapan perancangan arsitektur model ANN dibagi sebagai berikut:

- Cleaning* dan Pembagian Data

Tahap pertama dalam penelitian ini adalah mengumpulkan data nilai siswa IPA dan IPS dari 8 SMA di kota Denpasar yang memiliki

akreditasi yang seragam. Data yang dikumpulkan kemudian distandarisasi dengan membagi nilai-nilai dengan nilai maksimal yang mungkin dimiliki oleh variabel *input* tersebut. Variabel *input* yang digunakan adalah nilai rapor siswa dengan rentang 1-100, sedangkan variabel *output*nya adalah kelompok program studi. Proses standardisasi akan menghasilkan data dengan satuan yang sama pada rentang 0 sampai 1 (MinMaxScaler) (Géron, 2019). Mempertimbangkan komposisi data, pembagian data secara acak akan dilakukan menggunakan rasio 1:2 untuk data *testing* dan data *training*. Keterbatasan jumlah data menyebabkan rasio ini dipilih agar data *training* memiliki jumlah yang lebih besar. Pembagian data dengan rasio 7:3 adalah pendekatan yang umum digunakan, namun karena keterbatasan data, rasio 1:2 dipilih untuk penelitian ini.

b. Feedforward Propagation

Arsitektur model yang digunakan terdiri dari tiga lapisan yaitu *input layer*, satu *hidden layer*, dan *output layer* (Goodfellow et al., 2016). Banyaknya *neuron* pada *input layer* disesuaikan dengan jumlah variabel *input*, banyaknya *neuron* pada *hidden layer* sebesar 2/3 dari *input layer* (Heaton, 2011). Fungsi aktivasi pada *hidden layer* yang digunakan adalah fungsi *sigmoid*. Fungsi *sigmoid* adalah fungsi yang paling umum digunakan dalam pembangunan multilayer neural network dengan menggunakan metode pelatihan *backpropagation* (Du & Swamy, 2006) sehingga memungkinkan model untuk memetakan *input* menjadi *output* dalam rentang 0 hingga 1. Proses pembentukan model dapat ditulis dengan rumus sebagai berikut:

$$\hat{y} = f(f(XW^{(1)})W^{(2)}) \quad (1)$$

dimana \hat{y} adalah *output* dari model, f adalah fungsi aktivasi *sigmoid*, $W^{(1)}$ adalah bobot pada *input layer* dan *hidden layer*, $W^{(2)}$ adalah bobot pada *hidden layer* dan *output layer*, dan X adalah variabel *input*.

c. Backpropagation

Backpropagation dilakukan dengan tujuan untuk mencari kombinasi bobot terbaik yang dapat menghasilkan model dengan *loss function* yang seminimal mungkin. *Backpropagation* bekerja dengan menentukan bobot baru berdasarkan penurunan *loss function* (Loy, 2019). Proses dimulai dengan menghitung nilai *error* pada model awal menggunakan rumus *loss function*.

$$f(y, \hat{y}) = \sum \frac{1}{2} (y - \hat{y})^2 \quad (2)$$

Dengan menurunkan *loss function* yang dibangun berdasarkan fungsi *feedforward propagation* dapat didapatkan bobot-bobot baru yang dapat dicapai dengan menggunakan rumus berikut :

$$\delta^{(3)} = -(y - \hat{y})^2 f'(z^{(3)}) \quad (3)$$

$$\frac{\partial J}{\partial W^{(2)}} = \delta^{(3)} (a^{(2)})^T \quad (4)$$

$$\frac{\partial J}{\partial W^{(1)}} = \delta^{(3)} (W^{(2)})^T f'(z^{(2)}) X^T \quad (5)$$

Selanjutnya adalah dengan menginisiasi ulang bobot dengan bobot baru yang dihasilkan dari $\frac{\partial J}{\partial W^{(1)}}$ dan $\frac{\partial J}{\partial W^{(2)}}$ sehingga didapatkan model baru, dan proses ini dilakukan secara berulang sampai batas iterasi tercapai.

d. Error in sample

Error in sample adalah besar *error* dari model dengan menggunakan data *training* sebagai dasar perhitungannya. *Error* dihitung dengan menggunakan *loss function* dari hasil prediksi dan data aktual dari data *training* (Hastie et al., 2009). Proses perhitungan *error in sample* dapat dihitung sebagai rata-rata dari $\frac{1}{2} (y - \hat{y})^2$, di mana y adalah *output* aktual dari data *training*, dan \hat{y} adalah *output* yang diprediksi oleh model.

e. Error out sample

Error out sample adalah besar *error* dari model dengan menggunakan data *testing* sebagai dasar perhitungannya. *Error* dihitung dengan menggunakan *loss function* dari hasil prediksi dan data aktual dari data *testing*. Proses perhitungan *error out sample* juga dapat dihitung sebagai rata-rata dari $\frac{1}{2} (y - \hat{y})^2$, di mana y adalah *output* aktual dari data *testing*, dan \hat{y} adalah *output* yang diprediksi oleh model.

f. Menyimpan Model

Model yang telah dibangun dan dilatih akan disimpan di penyimpanan lokal pengguna. Model tersebut akan digunakan kembali untuk menduga data baru tanpa harus melatih ulang model.

g. Pendugaan Data Baru

Setelah model dilatih dan disimpan, model dapat digunakan untuk menduga data baru di

luar dari data *training* dan data *testing*. Untuk melakukan pendugaan data baru, kita dapat menggunakan rumus yang terbentuk pada tahap feedforward propagation yaitu:

$$\hat{y} = f(f(XW^{(1)})W^{(2)}) \quad (6)$$

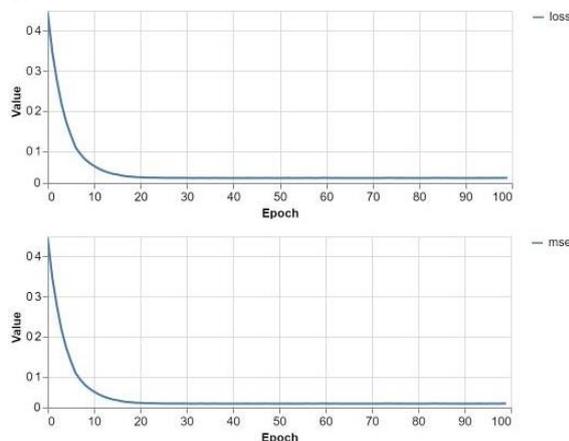
dimana \hat{y} adalah *output* dari model, f adalah fungsi aktivasi *sigmoid*, $W^{(1)}$ adalah bobot pada *input layer* yang telah melalui tahap pelatihan dan *hidden layer*, $W^{(2)}$ adalah bobot pada *hidden layer* dan *output layer* yang telah melalui tahap pelatihan, dan X adalah variabel *input*.

3.2 Pengujian Sistem

Pengujian sistem adalah tahap untuk memeriksa apakah sistem yang dibangun sudah dapat berjalan sesuai dengan tujuan yang diinginkan dalam penelitian ini. Sistem akan diuji menggunakan nilai IPA dan IPS siswa-siswa di kota Denpasar pada tahun 2019 sampai dengan 2020. Berikut adalah rincian pengujian yang dilakukan.

a. Pengujian Data Nilai IPA

Pada tahap pengujian data nilai IPA, terdapat 51 data siswa dengan 35 jenis nilai mata pelajaran (tujuh mata pelajaran dalam lima semester) dan 17 program studi. Model dibangun dengan arsitektur fungsi aktivasi *sigmoid* dan satu *hidden layer*. Data dibagi dengan rasio 2:1, dengan proporsi dua data *training* dan satu data *testing*. Model dilatih dengan data *training* sebanyak 100 iterasi. Performa pelatihan model terlihat pada Gambar 1.



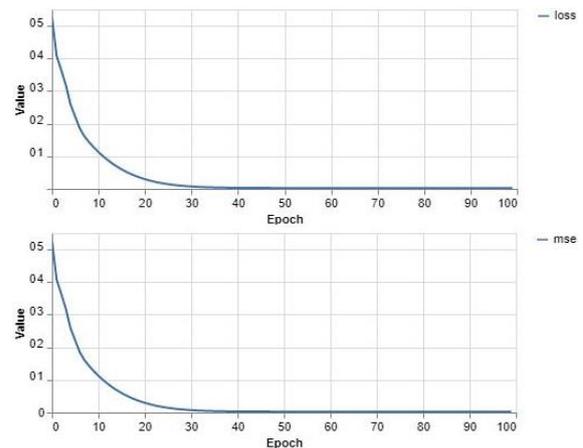
Gambar 1 Performa pelatihan model nilai IPA

Gambar 1 menunjukkan tingkat *error in sample*

pada tahap pelatihan sebesar 4.060%, sementara tingkat *error out sample* pada pengujian dengan data *testing* sebesar 4.049%. Selanjutnya, dilakukan pemeriksaan untuk melihat apakah model mengalami *overfitting* dengan menghitung selisih *error in sample* dan *out sample*, yang didapatkan sebesar 0.011% masih dibawah standar *overfitting* 5%. Hasil prediksi menunjukkan bahwa metode *artificial neural network* (ANN) dengan fungsi aktivasi *sigmoid* dapat digunakan untuk melakukan prediksi program studi untuk nilai IPA dengan performa yang baik.

b. Pengujian Data Nilai IPS

Pada tahap pengujian data nilai IPS, data nilai rapor yang digunakan adalah terdapat 36 data siswa dengan 35 jenis nilai mata pelajaran (tujuh mata pelajaran dalam lima semester) dan 12 program studi. Model dibangun dengan arsitektur fungsi *sigmoid* dan satu *hidden layer*. Data dibagi dengan rasio 2:1, dengan proporsi dua data *training* dan satu data *testing*. Model dilatih dengan data *training* sebanyak 100 iterasi. Performa pelatihan model terlihat pada Gambar 2.



Gambar 2 Performa pelatihan model nilai IPS

Gambar 2 menunjukkan tingkat *error in sample* pada tahap pelatihan sebesar 5.89%, yang memenuhi ambang tingkat *error* maksimal yang diperbolehkan dalam penelitian ini. Pengujian dengan data *testing* mendapatkan hasil *error out sample* sebesar 6.03%. Selanjutnya, dilakukan pemeriksaan untuk melihat apakah model mengalami *overfitting* dengan menghitung selisih *error in sample* dan *error out sample*,

yang didapatkan sebesar 0.54%, masih dibawah standar *overfitting* 5%. Hasil prediksi menunjukkan bahwa metode *artificial neural network* (ANN) dengan fungsi aktivasi *sigmoid* dapat digunakan untuk melakukan prediksi program studi untuk nilai IPS dengan performa yang baik.

4. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang telah dilakukan, sistem penduga kelulusan program studi di Universitas Udayana berdasarkan nilai rapor berbasis *artificial neural network* (ANN) berhasil dibangun. Hasil pengujian sistem pada data nilai rapor, untuk kategori kelas IPA dan IPS, mampu menghasilkan model prediksi kelulusan dengan tingkat *error* di bawah 7%. Dalam hal ini, disimpulkan bahwa metode *artificial neural network* (ANN) dengan fungsi aktivasi *sigmoid* dapat digunakan untuk melakukan prediksi program studi dengan performa yang baik.

Adapun saran yang dapat diberikan sebagai tindak lanjut dari penelitian ini adalah pertama, dapat dilakukan penelitian dengan menggunakan arsitektur model yang berbeda dari arsitektur yang telah diuji coba dalam penelitian ini. Kedua, penggunaan data terbaru diharapkan dapat menghasilkan model dengan tingkat reliabilitas yang lebih baik. Terakhir, sistem yang dikembangkan dalam penelitian ini memisahkan data nilai siswa IPA dan data nilai siswa IPS. Oleh karena itu, diharapkan ke depannya dapat dikembangkan model penelitian yang dapat merangkul kondisi di mana siswa IPA dapat memilih program studi IPS dan siswa IPS dapat memilih program studi IPA.

DAFTAR PUSTAKA

Ahmad, N., Hassan, N., Jaafar, H., & Enzai, N. I. M. 2021. "Students' Performance Prediction using Artificial Neural Network.". *IOP Conference Series: Materials Science and Engineering*, 1176(1), 1-9.

Dreiseitl, S., & Ohno, M. L. 2002. "Logistic regression and artificial neural network classification models: a methodology

review". *Journal of Biomedical Informatics*, 35 (5–6), 352–359.

- Du, K.-L., & Swamy, M. N. S. 2006. *Neural networks in a softcomputing framework*. London: Springer-Verlag.
- Géron, Aurelien. 2022. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. Sebastopol, California: O'Reilly Media.
- Goodfellow, I., Bengio, Y., & Courville, A. 2016. *Deep Learning*. Cambridge, MA: MIT Press.
- Hassoun, Mohamad. 1995. *Fundamentals of Artificial Neural Networks*. Cambridge, Massachusetts: The MIT Press.
- Hastie, T., Tibshirani, R., & Friedman, J. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer.
- Heaton, Jeff. (2011). *Programming Neural Networks with Encog3 in Java*. United States: Heaton Research, Inc.
- 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), 97-105.
- Kahar, N., Nurcahyo, G. W., & Defit, H. S. 2015. "Analisis Penerapan Artificial Neural Network dalam Penentuan Bidang Kompetensi Skripsi Mahasiswa (Studi Kasus di Program Studi Teknik Informatika Stmik Nurdin Hamzah Jambi)". *Senatkom*,1(2), 203–209.
- Lau, E. T., Sun, L., & Yang, Q. 2019. "Modelling, prediction and classification of student academic performance using artificial neural networks". *SN Applied Sciences*, 1(9), 1–10.
- Loy, James. 2019. *Neural Network Projects with Python*. Birmingham: Packt Publishing Ltd.
- Nielsen, Michael. 2015. *Neural Networks and Deep Learning*. Birmingham: Determination Press.

- Permenristekdikti. 2016. Peraturan Menteri Riset, Teknologi, dan Pendidikan Tinggi Nomor 126 Tahun 2016 Tentang Penerimaan Mahasiswa Baru Program Sarjana Pada Perguruan Tinggi Negeri.
- Riani, N. A., Andreswari, R., & Fauzi, R. 2021. "Implementasi Algoritma Artificial Neural Network Pada Proses Seleksi Pilih Peminatan (Studi Kasus : Prodi Sistem Informasi Universitas Telkom)". *Journal of Science and Social Research*, 4(3), 241.
- Saravanan, K., & Sasithra, S. (2014). "Review on Classification Based on Artificial Neural Networks." *The International Journal of Ambient Systems and Applications*, 2(4), 11–18.