

Penerapan Multi-Layer Perceptron untuk Mengklasifikasi Penduduk Kurang Mampu

Senang Hati Gulo¹, Andre Hasudungan Lubis^{2*}

^{1,2}Teknik Informatika, Fakultas Teknik, Universitas Medan Area

¹senanghatigulo2000@gmail.com, ^{2*}andrelubis2201@gmail.com

^{*)} andrelubis2201@gmail.com

Abstrak—Dewasa ini, klasifikasi penduduk kurang mampu di Kecamatan Afulu masih menggunakan cara yang manual, sehingga hal ini mengakibatkan membutuhkan waktu yang lama dalam pendistribusian bantuan sosial kepada masyarakat. Untuk mengatasi masalah tersebut, penelitian ini berupaya mengembangkan suatu praktik untuk klasifikasi data penduduk, dengan tujuan utama untuk mempercepat pemrosesan data penduduk di Kecamatan Afulu. Penelitian ini akan berfokus pada sembilan desa di kecamatan tersebut, yang mencakup total populasi 11.722 orang, dengan jumlah sampel 386 orang. Penelitian ini menggunakan Multilayer Perceptron (MLP), sebuah algoritma klasik yang merupakan algoritma yang paling banyak digunakan di berbagai penelitian. Hasil dari penelitian ini menunjukkan bahwa dari total jumlah sampel, terdapat 152 jiwa diklasifikasikan sebagai penduduk mampu, 86 jiwa merupakan penduduk cukup mampu, dan sejumlah 148 jiwa adalah penduduk kurang mampu. Penelitian ini menggunakan confusion matrix sebagai alat untuk mengevaluasi hasil klasifikasi. Dimana, hasil evaluasi menunjukkan bahwa Arsitektur 3-5-1, yang terdiri dari 3 lapisan input, 5 lapisan tersembunyi, dan 1 lapisan output, adalah arsitektur MLP yang paling unggul. Arsitektur ini menunjukkan nilai akurasi sebesar 96,9%, nilai recall 92%, nilai presisi 98,5%, dan nilai F-score 94,9%. Penjelasan rinci mengenai parameter yang digunakan, formula yang digunakan, dan beberapa komputasi yang dilakukan dijelaskan lebih lanjut.

Kata Kunci: Multilayer Perceptron, Jaringan Syaraf Tiruan, Klasifikasi, Data penduduk, Kecamatan Afulu

Abstract—The classification of the less capable population in Afulu Sub-district is currently reliant on a manual system, resulting in prolonged processing times. To address this issue, this research endeavors to develop a practical application for the classification of population data, with the primary objective of expediting the processing of population data in Afulu Sub-district. The study will focus on nine villages within the sub-district, encompassing a total population of 11,722 individuals, with a sample size of 386. The present study utilizes the Multilayer Perceptron, a classical algorithm that continues to be the most widely employed method in numerous researches. The findings of the present study indicate that out of the total sample size, 152 individuals were classified as capable, 86 individuals were classified as moderately capable, and a substantial number of 148 individuals were classified as less capable. The classification results were evaluated using a confusion matrix. The 3-5-1 architecture, comprising of 3 input layers, 5 hidden layers, and 1 output layer, was found to be the most superior. This architecture demonstrated an accuracy value of 96.9%, a recall value of 92%, a precision value of 98.5%, and an F-score value of 94.9%. A detailed elucidation of the parameters employed, the formulas utilized, and several computations performed are explained further.

Keywords: Multilayer Perceptron, Artificial Neural Network, Classification, Population data, Afulu Sub-district

1. PENDAHULUAN

Kecamatan Afulu merupakan wilayah dari Kabupaten Nias Utara yang berada di pulau Nias, Indonesia. Kecamatan Afulu memiliki sebanyak 9 desa yang tersebar di wilayah tersebut, meliputi Desa Afulu, Desa Lauru Fadoro, Desa Faekhuna'a, Desa Sifaoro'asi, Desa Harewakhe, Desa Ombolata Afulu, Desa Lauru Lahewa, Desa Lauru I, dan Desa Sisobahili. Menurut Badan Pusat Statistika (BPS), jumlah penduduknya cukup padat yakni sebesar 11,772 jiwa dengan luas daerah 149,78 Km² di tahun 2019. Sehingga menjadikan kecamatan ini merupakan daerah yang cukup padat penduduknya. Perekonomian masyarakat Afulu sebagian besar bergantung pada sektor pertanian, perikanan, dan kerajinan lokal. Pertanian menjadi mata pencaharian utama, dengan hasil bumi seperti padi, kelapa, dan rempah-rempah yang menjadi sumber kehidupan. Selain itu, keberlanjutan tradisi perikanan tradisional juga memberikan kontribusi penting terhadap perekonomian lokal [1].

Pengelolaan data kependudukan di Kecamatan Afulu merupakan komponen penting dalam perencanaan pembangunan di wilayah tersebut yang dikelola oleh pihak kantor camat. Namun, pengumpulan data kependudukannya masih dilakukan secara manual mengakibatkan lambatnya distribusi bantuan sosial kepada masyarakat. Hal ini dikarenakan sulitnya mengklasifikasikan penduduk mana yang kurang mampu, cukup mampu, dan mampu. Sehingga dibutuhkan pengklasifikasian data penduduk tersebut yang dapat membantu pihak kantor camat untuk mendapatkan calon penerima bantuan sosial secara efisien dan efektif.

Dalam mengklasifikasi, terdapat beberapa algoritma yang sudah banyak digunakan di berbagai penelitian untuk tujuan-tujuan tertentu, yaitu *decision tree*, *Support Vector Machines (SVMs)*, *K-Nearest-Neighbors (KNN)*

dan *Artificial Neural Networks* (ANNs) [2]. ANN merupakan algoritma yang meniru struktur otak manusia yang berfokus pada berbagai lapisan neuron yang saling berhubungan, yang disebut sebagai node dalam jaringan. Dimana node tersebut diatur dalam struktur hirarkis, dengan lapisan yang berbeda yang memfasilitasi pemrosesan dan transmisi informasi [3]. Setiap node digunakan untuk memproses data, yang kemudian ditransmisikan ke lapisan berikutnya dalam jaringan. Dengan demikian, hal ini dapat membuat keputusan yang serupa dengan yang dibuat oleh manusia dalam sistem komputasi [4]. Algoritma ANN sendiri memiliki banyak keunggulan dibandingkan algoritma klasifikasi lainnya. Salah satu keunggulannya adalah kemampuannya untuk memahami hubungan dan pola yang rumit di dalam data [5], dan keunggulan dalam mengelola dataset yang tidak seimbang [6].

Lebih lanjut, algoritma ANN telah banyak diterapkan di berbagai bidang, seperti digunakan untuk perencanaan energi [7], pemasaran [8], dan pengambilan keputusan organisasi [9]. Selain itu, ANN bahkan dapat digunakan dalam proses pendidikan untuk secara otomatis menganalisis tinjauan kritis dan menilai tipe psikologis siswa [10]. Sehingga, dapat disimpulkan bahwa algoritma ANN telah menunjukkan keefektifannya untuk dimanfaatkan di berbagai penerapan, dan dapat menyajikan prediksi yang akurat serta dapat meningkatkan proses pengambilan keputusan. ANN memiliki berbagai bentuk yang dirancang untuk menangani berbagai tujuan-tujuan khusus. Sebagai contoh terdapat bentuk ANN yang disebut dengan *Single-Layer Perceptron* (SLP), yang cocok digunakan untuk mengklasifikasi data yang berbentuk biner [11]. Contoh lainnya adalah *Convolutional Neural Networks* (CNN), yang merupakan bentuk ANN yang dirancang khusus untuk data berstruktur seperti citra [12]. Selain itu, juga ada bentuk ANN yang disebut dengan *Recurrent Neural Networks* (RNN) yang cocok digunakan untuk data berurutan seperti deret waktu (*time series*) atau bahasa alami (*natural language*) [13]. Pemilihan jenis jaringan saraf yang sesuai sangat tergantung pada karakteristik dan sifat data, serta tujuan analisis atau prediksi yang diinginkan. Dengan beragam bentuknya, ANN memberikan fleksibilitas dan adaptabilitas yang luas dalam menangani berbagai tugas pemrosesan dan analisis data [14]. Diantara bentuk ANN tersebut, terdapat bentuk yang paling umum digunakan, yakni *Multi-Layer Perceptron* (MLP).

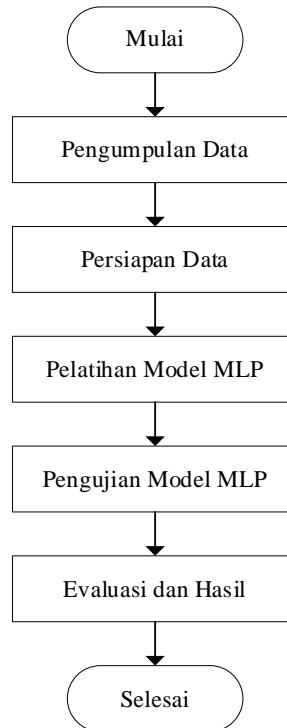
MLP adalah jenis arsitektur jaringan saraf tiruan dengan banyak lapisan neuron, dimana setidaknya satu lapisan tersembunyi di antara lapisan input dan output [15]. MLP merupakan jaringan saraf *feedforward*, yang artinya algoritma ini menghitung jumlah tertimbang dari input dan menerapkan fungsi aktivasi untuk menghasilkan output. Jaringan ini efektif dalam memperkirakan fungsi yang kompleks dan tugas data terstruktur [16]. MLP memiliki keunggulan pada kemampuannya untuk memodelkan hubungan yang kompleks dan non-linier dalam data. Selain itu, MLP juga memiliki kemampuan untuk mengekstraksi fitur secara otomatis dan mengatasi masalah yang tidak dapat dipecahkan oleh model linear. Sehingga, dengan kelebihan dan kemampuannya tersebut MLP menjadi salah satu algoritma yang cukup populer dalam berbagai aplikasi pembelajaran mesin [17]. Dalam hal mengklasifikasikan data, arsitektur pada MLP, yang terdiri dari neuron-neuron yang saling terhubung dapat mempermudah untuk menangkap hubungan yang rumit dalam dataset [18]. Mempertimbangkan masalah yang terjadi dan tersedianya algoritma yang cocok untuk digunakan sebagai solusi, maka penelitian ini akan menggunakan algoritma MLP untuk mengklasifikasi penduduk kurang mampu pada Kecamatan Afulu.

2. METODE PENELITIAN

Bagian ini akan menjelaskan metode dan teknik untuk mengumpulkan dan menganalisis data yang akan dilalui untuk mendapatkan hasil penelitian. Bagian ini terdiri dari tahapan penelitian, proses pengumpulan dan persiapan data, dan penjelasan tentang MLP. Bagian akhir akan memaparkan teknik yang digunakan untuk mengevaluasi hasil pengujian

2.1 Tahapan Penelitian

Proses pada penelitian ini terdiri dari serangkaian tahapan yang berurutan dan sistematis guna memperoleh hasil dan menjadi solusi bagi permasalahan. Penelitian ini dimulai dengan proses pengumpulan data, persiapan data, pelatihan dan pengujian model *perceptron*, sampai dengan tahapan evaluasi dan hasil. Tahapan-tahapan tersebut dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.2 Pengumpulan dan Persiapan Data

Data yang digunakan pada penelitian ini adalah data primer, yaitu data yang dikumpulkan secara langsung pada lokasi penelitian [19]. Sehingga, pengumpulan data dalam penelitian ini di ambil langsung di tempat penelitian yaitu kantor camat Afulu. Dari jumlah populasi sebanyak 11,772, penelitian ini mengambil sampel dengan menggunakan rumus 1 [20].

$$n = \frac{N}{1 + Ne^2} \quad (1)$$

Pada rumus (1), n merupakan jumlah sampel yang diambil. Nilai N adalah jumlah populasi data dan e adalah tingkat presisi atau toleransi yang bernilai 0.05.

$$n = \frac{11,772}{1 + 11,772(0.05)^2} = 386.291 \approx 386$$

Dari hasil perhitungan rumus Slovin data yang dikumpulkan adalah sebanyak 386. Adapun atribut yang didapat berupa nama, penghasilan, tingkat pendidikan, kepemilikan aset dan klasifikasi penduduk yang dijadikan sebagai atribut pada penelitian ini. Rangkuman dari data yang telah dikumpulkan pada penelitian ini dapat dilihat pada tabel 1.

Tabel 1. Data Penelitian

No	Nama	Penghasilan (Rp)	Tingkat Pendidikan	Kepemilikan Aset	Klasifikasi Penduduk
1	Folo`o Gulo	5.000.000	SMP	Milik Sendiri	Mampu
2	Folala Daeli	2.000.000	SMP	Milik Sendiri	Cukup Mampu
3	Izaaki Gea	2.450.000	SMA	Milik Sendiri	Cukup Mampu
⋮	⋮	⋮	⋮	⋮	⋮
386	Idhan Gulo	1.500.000	SD	Milik Sendiri	Kurang Mampu

Setelah data dikumpulkan, dalam penelitian ini data dibagi menjadi dua, yakni data latih dan data uji. Data latih digunakan untuk melatih model atau algoritma, sedangkan data uji digunakan untuk menguji seberapa baik kinerja model setelah dilatih [21]. Adapun pembagian data dapat dilihat pada tabel 2.

Tabel 2. Pembagian Proporsi Data

Pembagian	Jumlah Data Latih	Jumlah Data Uji
70:30	270	116
80:20	309	77
90:10	347	39

Kemudian, data tersebut dipersiapkan dengan cara mendeteksi data yang *missing*, melakukan perbaikan pada ketidakakuratan dalam data, dan transformasi data. Hal ini adalah dengan melakukan perubahan pada data yang bersifat kategorikal seperti tingkat pendidikan dan kepemilikan aset menjadi bentuk numerik yaitu dengan menggunakan *Label Encoding* [22]. Hasil transformasi datanya dapat dilihat pada tabel 3.

Tabel 3. Label Encoding pada Data

Kriteria	Klasifikasi	Transformasi
Kepemilikan Aset	Sewa	1
	Milik Sendiri	2
Tingkat Pendidikan	Tidak sekolah	1
	SD	2
	SMP	3
	SMA	4
	Pendidikan Tinggi	5

Data yang telah berbentuk numerikal berikutnya akan dilakukan normalisasi. Hal ini bertujuan untuk memastikan bahwa data dalam format yang konsisten dan sesuai untuk dianalisis, sehingga dapat meningkatkan keakuratan dari hasil klasifikasi [23]. Penelitian ini menggunakan *z-score* sebagai alat untuk menormalisasikan data. Adapun rumus yang digunakan adalah rumus 2 [24].

$$z = \frac{x - \mu}{\sigma} \quad (2)$$

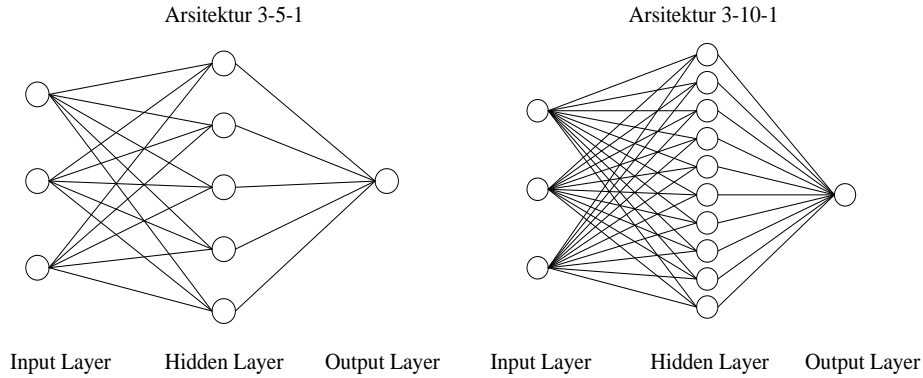
Dari rumus (2), z adalah hasil normalisasi, x adalah nilai awal pada data, μ adalah nilai rata-rata dari populasi data, dan σ adalah nilai dari standar deviasi. Sebagai contoh, hasil dari normalisasi pada data uji 10% pada proporsi data 90:10 sebanyak 10 data terakhir dapat dilihat pada tabel 4.

Tabel 4. Hasil Normalisasi Data Uji 10%

No	Penghasilan (Rp)	Tingkat Pendidikan	Kepemilikan Aset
376	0.1303	0.0384	-3.3718
377	-1.1625	-0.7162	0.2958
378	-1.1625	-0.7162	0.2958
379	1.4232	-0.7162	-3.3718
380	1.4232	6.4528	-3.3718
381	0.7767	2.6796	-3.3718
382	1.4232	4.1889	0.2958
383	-0.5161	0.0384	0.2958
384	-0.5161	0.0384	0.2958
385	-0.5161	0.7931	0.2958
386	-0.5161	0.7931	0.2958

2.3 Model MLP

Terdapat dua model MLP yang dikembangkan pada penelitian ini. Dimana, jumlah neuron pada *hidden layer* berkisar antara 5 dan 10, *input layer* sebanyak 3, dan 1 *output layer* pada arsitekturnya. Tampilan dari arsitektur tersebut dapat dilihat pada gambar 2.



Gambar 2. Model Arsitektur 3-5-1 dan 3-10-1

Pembelajaran pada MLP dilakukan dengan tahapan bobot balik (*backpropagation*), yakni dibagi atas tahap *Feedforward*, tahap *Backward*, dan tahap perubahan bobot [25]. Tahap *Feedforward* dimulai dengan proses penerimaan sinyal oleh setiap *input layer* dan meneruskannya ke *hidden layer*. Kemudian, lakukan penjumlahan pada bobot dari sinyal *input* (z_{inj}) yang dinyatakan dengan rumus (3).

$$z_{inj} = v_0j + \sum_{i=1}^n x_i v_{ij} \quad (3)$$

Selanjutnya, penelitian ini menggunakan fungsi aktivasi *sigmoid* untuk menghitung sinyal *output* yang telah dikirim oleh *hidden layer* (z_j) yaitu dengan menggunakan rumus (4) dibawah ini. Lalu, akan dilakukan perhitungan semua bobot sinyal pada jaringan *output layer* (y_{net}) dengan menggunakan rumus (5).

$$z_j = \frac{1}{1 + e^{-z_{inj}}} \quad (4)$$

$$y_{net} = w_0k + \sum_{j=1}^n z_j w_{jk} \quad (5)$$

Tahap *Backward* dimulai dengan menentukan nilai *error* (δ_k) atau kesalahan di setiap *output layer* dengan rumus (6), yang dilanjutkan dengan melakukan perhitungan perubahan bobot (Δw_{jk}) dengan menggunakan rumus (7). Adapun representasi dari rumus (6) dan (7) adalah sebagai berikut.

$$\delta_k = (t_k - y_k)y_k(1 - y_k) \quad (6)$$

$$\Delta w_{jk} = \alpha \delta_k z_j \quad (7)$$

Selanjutnya, dilakukan hal yang sama pada *hidden layer*, yakni menghitung *error* (δ_{net}) dengan rumus (8). Lalu, akan dilanjutkan dengan menentukan nilai informasi *error* dengan rumus (9). Lebih lanjut, diikuti dengan menghitung suku perubahan bobot (Δv_{ij}) dengan rumus (10). Rumus (8), (9) dan (10) dapat diwakilkan sebagai berikut.

$$\delta_{net} = v_0j + \sum_{k=1}^n \delta_k + w_{jk} \quad (8)$$

$$\delta_j = \delta_{net} z_j(1 - z_j) \quad (9)$$

$$\Delta v_{ij} = \alpha \delta_k x_i \quad (10)$$

Tahap perubahan bobot dilakukan dengan menghitung semua perubahan bobot yang menuju ke *output layer* sebagai langkah pertama. Perhitungannya dilakukan dengan menggunakan rumus (11). Kemudian, dilanjutkan dengan menghitung perubahan bobot garis yang menuju ke *hidden layer* dengan rumus (12).

$$w_{jk}(\text{baru}) = w_{jk}(\text{lama}) + \Delta w_{jk} \quad (11)$$

$$v_{ij}(\text{baru}) = v_{ij}(\text{lama}) + \Delta v_{ij} \quad (12)$$

2.4 Evaluasi dan Hasil

Tujuan dari penelitian ini adalah untuk mengklasifikasikan data penduduk kurang mampu dengan tiga tingkatan yakni kurang mampu, cukup mampu, dan mampu. Sehingga, penelitian ini akan menelusuri hasil klasifikasi yang dilakukan oleh MLP dengan penentuan bobot yang dihasilkan. Setelah melakukan proses pelatihan pada kedua model MLP, maka akan diperoleh bobot yang baru. Lebih lanjut, bobot baru tersebut akan dibandingkan dengan data yang telah ditentukan sebelumnya. Teknik yang digunakan dalam hal ini adalah *confusion matrix*, dengan menentukan nilai *accuracy*, *recall*, *precision* dan *F-score* [26].

3. HASIL DAN PEMBAHASAN

Hasil pada penelitian ini terdiri atas 3 bagian berdasarkan proporsi data yang telah dibagi sebelumnya, yaitu 70:30, 80:20, dan 90:10. Setiap proporsi data akan diterapkan pembelajaran model MLP 3-5-1 dan 3-10-1. Kemudian, evaluasi dilakukan untuk setiap model dengan menggunakan *confusion matrix*.

3.1 Hasil Klasifikasi Proporsi Data 70:30

Pada proporsi data 70:30, jumlah data untuk latihan adalah sebanyak 270 dan untuk uji adalah 116 dari total data sebanyak 386. Learning rate yang digunakan adalah 0.1 dengan nilai epoch 100 dan nilai alpha (α) adalah 0.1. Hasil dari klasifikasinya dapat dilihat pada tabel 5.

Tabel 5. Hasil Klasifikasi pada Proporsi Data 70:30

Arsitektur	Accuracy	Recall	Precision	F-score
3-5-1	96.9%	92%	98.5%	94.9%
3-10-1	93.8%	92%	94%	93.1%

Dari Tabel 5 dapat dilihat bahwa nilai dari akurasi pada arsitektur 3-5-1 (96.9%) lebih tinggi dibanding dengan arsitektur 3-10-1 yaitu 93.8%. Hal ini juga berlaku pada nilai *precision* dan *F-score*, dimana arsitektur 3-5-1 lebih unggul daripada arsitektur 3-10-1. Lebih lanjut, kedua arsitektur memiliki nilai *recall* yang sama yaitu 92%.

3.2 Hasil Klasifikasi Proporsi Data 80:20

Selanjutnya, penelitian ini akan melakukan klasifikasi pada proporsi data 80:20. Dimana, terdapat jumlah data untuk latihan adalah sebanyak 309 dan untuk uji adalah sebanyak 77. Dari sisi parameter, hal yang sama juga diterapkan pada proporsi data ini, yakni learning rate dengan nilai 0.1, epoch sebanyak 100 dan nilai alpha (α) adalah 0.1. Hasil dari klasifikasinya dapat dilihat pada tabel 6.

Tabel 6. Hasil Klasifikasi pada Proporsi Data 80:20

Arsitektur	Accuracy	Recall	Precision	F-score
3-5-1	97%	94.4%	95%	97.1%
3-10-1	93.8%	95.2%	90.9%	94.2%

Berdasarkan hasil yang ditunjukkan pada tabel 6, terlihat adanya perbedaan nilai antara arsitektur 3-5-1 dan 3-10-1. Nilai akurasi pada arsitektur 3-5-1 adalah sebesar 97%, sedangkan pada arsitektur 3-10-1 hanya 93.8%. Hal yang sama juga terjadi pada perbandingan nilai *precision* dan *F-score*, dimana arsitektur 3-5-1 lebih unggul dibanding dengan 3-10-1. Namun, terdapat hal yang kontras jika ditinjau dari nilai *recall*. Arsitektur 3-10-1 memiliki nilai yang lebih besar (95.2%) dibanding dengan arsitektur yang hanya sebesar 94.4%.

3.3 Hasil Klasifikasi Proporsi Data 90:10

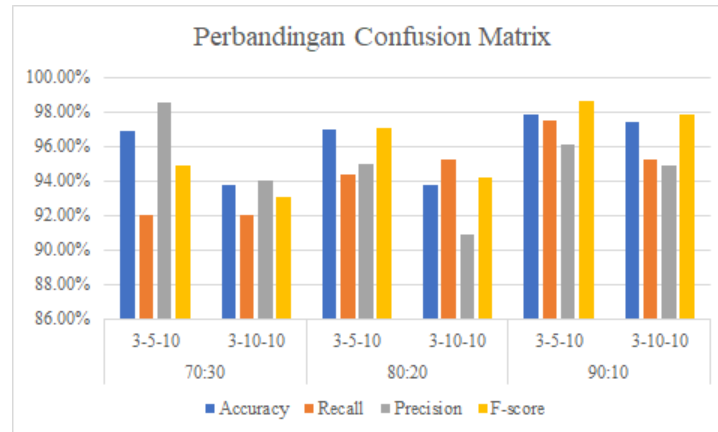
Pengujian klasifikasi yang terakhir pada penelitian ini adalah pada proporsi data 90:10. Pada proporsi data ini, data latih memiliki data yang cukup banyak yakni sebesar 347. Sedangkan, pada data uji adalah sebanyak 39. Parameter yang diterapkan masih menggunakan pengaturan yang sama, yakni learning rate dengan nilai 0.1, epoch sebanyak 100 dan nilai alpha (α) adalah 0.1. Hasil dari klasifikasinya dapat dilihat pada tabel 7.

Tabel 7. Hasil Klasifikasi pada Proporsi Data 90:10

Arsitektur	Accuracy	Recall	Precision	F-score
------------	----------	--------	-----------	---------

3-5-1	97.8%	97.5%	96.1%	98.6%
3-10-1	97.4%	95.2%	94.9%	97.8%

Tabel 7 menunjukkan bahwa arsitektur 3-5-1 menghasilkan performa yang baik. Hal ini dibuktikan dengan nilai *accuracy*, *recall*, *precision*, dan *F-score*-nya yang lebih unggul daripada arsitektur 3-10-1. Perbandingan nilai *confusion matrix* dari klasifikasi untuk setiap proporsi data dapat dilihat pada gambar 3.



Gambar 3. Perbandingan Confusion Matrix

3.4 Pembahasan

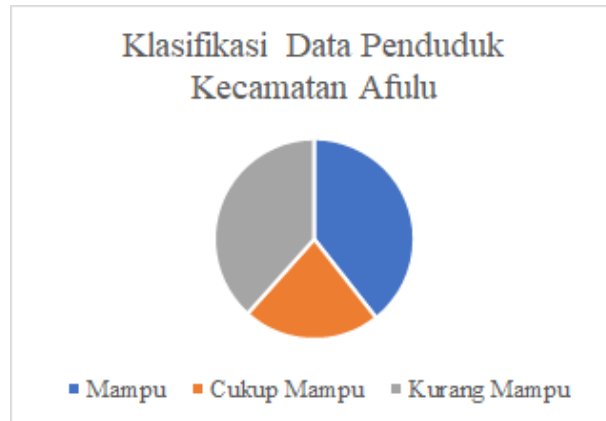
Hasil penelitian menunjukkan bahwa klasifikasi data penduduk kurang mampu diselesaikan dengan cukup baik. Hal ini dibuktikan dengan hasil evaluasi dari *confusion matrix* pada arsitektur MLP dengan nilai yang cukup signifikan. Hasil penelitian juga menunjukkan terdapat kenaikan dari nilai *accuracy*, *recall*, *precision*, dan *F-score* pada proporsi data. Semakin besar pembagian data latihnya, maka semakin tinggi pula nilai metrik-metrik kinerjanya.

Dalam hal klasifikasi data penduduk kurang mampu, penelitian ini menerapkan arsitektur 3-5-1 dengan proporsi data 70:30 untuk menentukan hasilnya. Hal ini dikarenakan proporsi data ini memberikan keseimbangan antara kecukupan data latih dan data uji, sehingga dapat meningkatkan performa prediktifnya [27]. Rangkuman dari hasil klasifikasinya dapat dilihat pada tabel 8.

Tabel 8. Klasifikasi Data Penduduk Kurang Mampu

No	Nama	Klasifikasi Penduduk	Prediksi
1	Folo`o Gulo	Mampu	Mampu
2	Folala Daeli	Cukup Mampu	Cukup Mampu
3	Izaaki Gea	Cukup Mampu	Kurang Mampu
⋮	⋮	⋮	⋮
386	Idhan Gulo	Kurang Mampu	Kurang Mampu

Tabel 8 menunjukkan adanya beberapa ketidaksamaan antara klasifikasi penduduk dengan prediksi yang dilakukan oleh MLP. Dimana, nilai akurasi nya adalah sebesar 96.9% seperti yang terlihat pada Tabel 5. Adapun hasil dari klasifikasi yang dilakukan adalah sebanyak 152 orang dikategorikan mampu, 86 orang dinyatakan cukup mampu, dan sebanyak 148 orang adalah kurang mampu. Visualisasi dari hasil klasifikasi data terlihat pada gambar 4.



Gambar 4. Klasifikasi Data Penduduk Kecamatan Afulu

4. KESIMPULAN

Penelitian ini membahas tentang klasifikasi data kependudukan di Kecamatan Afulu, Kabupaten Nias Utara, Indonesia, dalam rangka pendistribusian bantuan sosial yang efisien dan efektif. Penelitian ini menerapkan algoritma Multilayer Perceptron (MLP), yakni salah satu jenis jaringan syaraf tiruan atau artificial neural network untuk mengklasifikasikan penduduk yang kurang mampu ke dalam tiga kategori: tidak mampu, cukup mampu, dan mampu. Hasil klasifikasi menunjukkan bahwa 152 orang dikategorikan mampu, 86 orang dinyatakan cukup mampu, dan sebanyak 148 orang adalah kurang mampu. Penelitian ini menggunakan confusion matrix untuk mengevaluasi hasil klasifikasi. Arsitektur 3-5-1, yakni sebanyak 3 input layer, hidden layer berjumlah 5, dan 1 output layer adalah arsitektur yang paling unggul. Dimana, arsitektur ini memiliki nilai accuracy sebesar 96.9%, nilai recall sebesar 92%, nilai precision sebesar 98.5%, dan nilai F-score adalah 94.9%. Penelitian ini menyajikan bukti empiris bahwa Multilayer Perceptron (MLP) tidak hanya terbatas pada klasifikasi citra, tetapi juga dapat digunakan secara efektif untuk klasifikasi dalam bentuk kumpulan data.

REFERENCES

- [1] L. I. Lestari, A. A. Miftah, and B. Arisha, "Analisis Pengelolaan Anggaran Pendapatan dan Belanja Desa di Desa Sungai Ruan Iir Kecamatan Maro Sebo Ulu Kabupaten Batanghari Tahun 2020-2022," *J. Masharif Al-Syariah J. Ekon. dan Perbank. Syariah*, vol. 8, no. 1, 2023.
- [2] E. Y. Boateng, J. Otoo, and D. A. Abaye, "Basic tenets of classification algorithms K-nearest-neighbor, support vector machine, random forest and neural network: a review," *J. Data Anal. Inf. Process.*, vol. 8, no. 4, pp. 341–357, 2020.
- [3] A. A. Salih and A. M. Abdulazeez, "Evaluation of classification algorithms for intrusion detection system: A review," *J. Soft Comput. Data Min.*, vol. 2, no. 1, pp. 31–40, 2021.
- [4] A. Goel, A. K. Goel, and A. Kumar, "The role of artificial neural network and machine learning in utilizing spatial information," *Spat. Inf. Res.*, vol. 31, no. 3, pp. 275–285, 2023.
- [5] M. R. Choudhury, N. Muraleedharan, P. Acharjee, and A. T. George, "Network Traffic Classification Using Supervised Learning Algorithms," in *2023 International Conference on Computer, Electrical & Communication Engineering (ICCECE)*, 2023, pp. 1–6.
- [6] P. Rani, R. Lamba, R. K. Sachdeva, P. Bathla, and A. N. Aledaily, "Diabetes Prediction Using Machine Learning Classification Algorithms," in *2023 International Conference on Smart Computing and Application (ICSCA)*, 2023, pp. 1–5.
- [7] C. Pavlatos, E. Makris, G. Fotis, V. Vita, and V. Mladenov, "Utilization of Artificial Neural Networks for Precise Electrical Load Prediction," *Technologies*, vol. 11, no. 3, p. 70, 2023.
- [8] N. Syam and R. Kaul, "Neural Networks in Marketing and Sales," in *Machine Learning and Artificial Intelligence in Marketing and Sales*, Emerald Publishing Limited, 2021, pp. 25–64.
- [9] A. Psarras, T. Anagnostopoulos, I. Salmon, Y. Psaromiligkos, and L. Vryzidis, "A Change Management Approach with the Support of the Balanced Scorecard and the Utilization of Artificial Neural Networks," *Adm. Sci.*, vol. 12, no. 2, p. 63, 2022.
- [10] S. Ibragim, B. Akhat, M. Dinara, G. Anastasiya, K. Mariya, and M. Grigoriy, "Example of the use of artificial neural network in the educational process," in *Advances in Information and Communication: Proceedings of the 2020 Future of Information and Communication Conference (FICC), Volume 1*, 2020, pp. 420–430.

- [11] A. A. Heidari, H. Faris, S. Mirjalili, I. Aljarah, and M. Mafarja, "Ant lion optimizer: theory, literature review, and application in multi-layer perceptron neural networks," *Nature-Inspired Optim. Theor. Lit. Rev. Appl.*, pp. 23–46, 2020.
- [12] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: analysis, applications, and prospects," *IEEE Trans. neural networks Learn. Syst.*, 2021.
- [13] H. Hewamalage, C. Bergmeir, and K. Bandara, "Recurrent neural networks for time series forecasting: Current status and future directions," *Int. J. Forecast.*, vol. 37, no. 1, pp. 388–427, 2021.
- [14] M. E. Shaik, M. M. Islam, and Q. S. Hossain, "A review on neural network techniques for the prediction of road traffic accident severity," *Asian Transp. Stud.*, vol. 7, p. 100040, 2021.
- [15] T. S. Fatayer and M. N. Azara, "IoT secure communication using ANN classification algorithms," in *2019 International Conference on Promising Electronic Technologies (ICPET)*, 2019, pp. 142–146.
- [16] H. Siqueira and I. Luna, "Performance comparison of feedforward neural networks applied to streamflow series forecasting," *Math. Eng. Sci. & Aerosp.*, vol. 10, no. 1, 2019.
- [17] A. H. Abd-elaziem and T. H. M. Soliman, "A Multi-Layer Perceptron (MLP) Neural Networks for Stellar Classification: A Review of Methods and Results," *Int. J. Adv. Appl. Comput. Intell.*, vol. 3, no. 10, pp. 54216, 2021.
- [18] A. Secilmis, N. Aksu, F. A. Dael, I. Shayea, and A. A. El-Saleh, "Machine Learning-Based Fire Detection: A Comprehensive Review and Evaluation of Classification Models," *JOIV Int. J. Informatics Vis.*, vol. 7, no. 3–2, pp. 1982–1988, 2023.
- [19] Z. Lubis, *Statistika terapan untuk ilmu-ilmu sosial dan ekonomi*. Medan: Penerbit Andi, 2021.
- [20] A. Susanti, R. A. A. Soemitro, H. Suprayitno, and V. Ratnasari, "Searching the appropriate minimum sample size calculation method for commuter train passenger travel behavior survey," *J. Infrastruct. & Facil. Asset Manag.*, vol. 1, no. 1, 2019.
- [21] K. Sugali, C. Sprunger, and V. Inukollu, "AI Testing: Ensuring a Good Data Split Between Data Sets (Training and Test) using K-means Clustering and Decision Tree Analysis," *Int. J. Soft Comput.*, vol. 12, pp. 1–11, 2021, doi: 10.5121/ijsc.2021.12101.
- [22] T. Al-Shehari and R. A. Alsowail, "An insider data leakage detection using one-hot encoding, synthetic minority oversampling and machine learning techniques," *Entropy*, vol. 23, no. 10, p. 1258, 2021.
- [23] D. Singh and B. Singh, "Investigating the impact of data normalization on classification performance," *Appl. Soft Comput.*, vol. 97, p. 105524, 2020.
- [24] V. Çetin and O. YILDIZ, "A comprehensive review on data preprocessing techniques in data analysis," *Pamukkale Üniversitesi Mühendislik Bilim. Derg.*, vol. 28, no. 2, pp. 299–312, 2022.
- [25] A. Rana, A. S. Rawat, A. Bijalwan, and H. Bahuguna, "Application of multi layer (perceptron) artificial neural network in the diagnosis system: a systematic review," in *2018 International conference on research in intelligent and computing in engineering (RICE)*, 2018, pp. 1–6.
- [26] J. Xu, Y. Zhang, and D. Miao, "Three-way confusion matrix for classification: A measure driven view," *Inf. Sci. (Ny)*, vol. 507, pp. 772–794, 2020.
- [27] I. Muraina, "Ideal dataset splitting ratios in machine learning algorithms: general concerns for data scientists and data analysts," 2022.