

KOMPARASI FUNGSI AKTIVASI RELU DAN TANH PADA MULTILAYER PERCEPTRON

Ichsan Firmansyah¹, dan B. Herawan Hayadi²

^{1,2}Ilmu Komputer Universitas Potensi Utama, Medan, Sumatera Utara

Email: ichsanfirmansyah1989@gmail.com¹, b.herawan.hayadi@email.com²

Abstrak

Neural network merupakan metode yang populer digunakan dalam penelitian di bidang machine, dan fungsi aktivasi, khususnya ReLu dan Tanh, memiliki fungsi yang sangat penting pada neural network, untuk meminimalisir nilai error antara lapisan output dan kelas target. Dengan variasi jumlah hidden layer, serta jumlah neuron pada masing-masing hidden layer yang berbeda, penelitian ini menganalisa 8 model untuk mengklasifikasikan dataset Titanic's Survivor. Diperoleh hasil bahwa fungsi ReLu memiliki kinerja yang lebih baik daripada Tanh, dilihat dari nilai rata-rata akurasi dan presisinya yang lebih tinggi dibandingkan fungsi aktivasi Tanh. Penambahan jumlah hidden layer tidak berpengaruh pada meningkatnya kinerja hasil klasifikasi, terlihat dari menurunnya rata-rata akurasi dan presisi dari model yang menggunakan 3 hidden layer dan model yang menggunakan 4 hidden layer. Nilai akurasi tertinggi diperoleh pada model yang menggunakan fungsi aktivasi ReLu dengan 4 hidden layer dan 50 neuron pada masing-masing hidden layer, sedangkan nilai presisi tertinggi diperoleh pada model yang menggunakan fungsi aktivasi ReLu dengan 4 hidden layer dan 100 neuron pada masing-masing hidden layer.

Kata Kunci: Hidden layer, multilayer perceptron, neuron, relu, tanh

Abstract

Neural network is a popular method used in machine research, and activation functions, especially ReLu and Tanh, have a very important function in neural networks, to minimize the error value between the output layer and the target class. With variations in the number of hidden layers, as well as the number of neurons in each different hidden layer, this study analyzes 8 models to classify the Titanic's Survivor dataset. The result is that the ReLu function has a better performance than the Tanh function, seen from the average value of accuracy and precision which is higher than the Tanh activation function. The addition of the number of hidden layers has no effect on increasing the performance of the classification results, it can be seen from the decrease in the average accuracy and precision of the models that use 3 hidden layers and models that use 4 hidden layers. The highest accuracy value was obtained in the model using the ReLu activation function with 4 hidden layers and 50 neurons in each hidden layer, while the highest precision value was obtained in the model using the ReLu activation function with 4 hidden layers and 100 neurons in each hidden layer.

KeyWords : Hidden layer, multilayer perceptron, neuron, relu, tanh

I. PENDAHULUAN

Neural network merupakan metode yang cukup populer digunakan di dalam penelitian di bidang *machine learning*, karena memiliki kinerja yang efektif, terutama dalam melakukan transformasi terhadap data sehingga dapat diproses dengan lebih mudah [1]. Perceptron dengan model backpropagation merupakan salah satu algoritma *neural network* yang mampu memprediksi data berdasarkan proses pelatihan, baik itu dari data yang sudah diketahui sebelumnya ataupun yang belum diketahui [2]. *Multilayer perceptron* (MLP) merupakan pengembangan dari algoritma perceptron, dimana yang membedakan algoritma ini dari algoritma sebelumnya adalah dalam arsitekturnya dapat memiliki lebih dari satu *hidden layer* [3]. Dalam algoritma MLP, *backpropagation* digunakan untuk menurunkan nilai *error* antara lapisan *output* dan kelas target [4], dan algoritma ini juga memanfaatkan fungsi aktivasi pada hidden layer untuk mengurangi nilai *error* dari *output* yang dihasilkan masing-masing neuron [5]. Fungsi aktivasi ReLu [6] dan Tanh [7] merupakan dua fungsi aktivasi yang cukup populer dan banyak digunakan pada arsitektur MLP, karena dapat menghasilkan waktu pelatihan data yang lebih cepat. Penelitian ini bertujuan untuk menganalisa hasil klasifikasi algoritma MLP dengan menggunakan kedua fungsi aktivasi tersebut dan variasi jumlah *hidden layer* yang berbeda. Dengan menggunakan aplikasi *data mining Orange 3* dan dataset *Titanic's Survivor* yang secara *default* disediakan oleh aplikasi tersebut, hasil klasifikasi dianalisa menggunakan evaluasi *confusion matrix* untuk memperoleh nilai akurasi dan presisi masing-masing model. Hasil inilah kemudian yang digunakan untuk menentukan model mana yang memiliki kinerja terbaik dan merupakan model yang paling ideal untuk diterapkan dalam masalah klasifikasi pada penelitian ini.

II. METODE

Proses yang dilakukan di dalam penelitian, seperti proses pengumpulan data, rancangan penelitian, analisis data serta perancangan bentuk model adalah sebagai berikut.

A. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini adalah *Titanic's Survivor*, yang merupakan dataset *default* yang disediakan pada aplikasi *Orange 3*. Dataset ini terdiri dari 2201 data penumpang kapal Titanic, yang terbagi berdasarkan kategori status penumpang pada kapal Titanic, rentang usia serta jenis kelamin. Kelas target pada dataset ini adalah apakah penumpang kapal Titanic tersebut berhasil selamatkan atau tidak.

B. Analisis Data

Dataset *Titanic's Survivor* terdiri dari 2201 data dengan 3 fitur, yaitu *status*, *age* dan *sex*; serta satu kelas target yaitu *survived*. Fitur *status* berisi informasi status penumpang kapal Titanic, yang terdiri dari data *first* (penumpang *first class*), *second* (penumpang *second class*), *third* (penumpang *third class*), dan *crew* (awak kapal). Fitur *age* berisi informasi rentang usia penumpang kapal Titanic, yang terdiri dari data *adult* (penumpang dewasa) dan *child* (penumpang anak-anak). Fitur *sex* berisi informasi jenis kelamin penumpang kapal Titanic, yang terdiri dari data *male* (penumpang laku-laki) dan *female* (penumpang perempuan). Kelas *target survived* berisi informasi apakah penumpang kapal Titanic tersebut selamat atau tidak, yang terdiri dari data *yes* (penumpang selamat) dan *no* (penumpang tidak selamat). Tabel I menunjukkan 20 sampel data yang diambil dari dataset *Titanic's Survivor* yang digunakan.

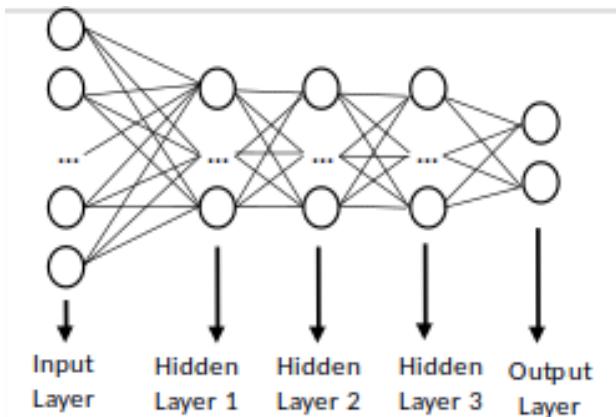
Tabel I: Sampel Dataset

Status	Age	Sex	Survived
first	adult	male	yes
first	adult	female	yes
first	child	male	yes
first	child	female	yes
first	adult	male	no
second	adult	male	yes
second	child	male	yes
second	child	female	yes
second	adult	male	no
second	adult	female	no
third	child	male	yes
third	child	female	yes
third	adult	male	yes
third	adult	male	no
third	adult	female	yes
third	adult	female	no
crew	adult	male	yes
crew	adult	female	yes
crew	adult	male	no
crew	adult	female	no

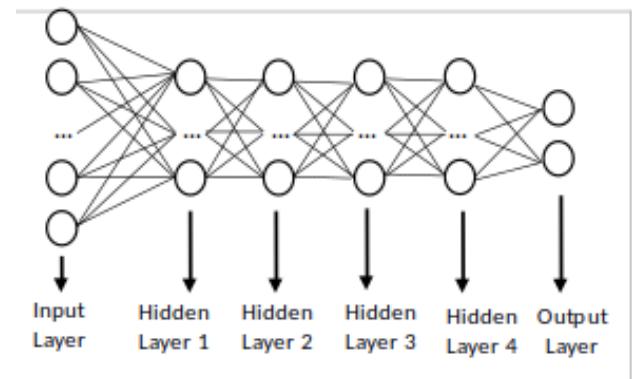
C. Multilayer Perceptron

Penelitian ini menggunakan algoritma *multilayer perceptron* karena algoritma ini terbukti memiliki tingkat akurasi yang tinggi dalam masalah klasifikasi [8]; kinerjanya lebih baik jika dibandingkan dengan algoritma klasifikasi lain seperti *naïve bayes*, *support vector machine*, *k-nearest neighbors*, dan *random forest* [9]; dapat dikalibrasi arsitekturnya untuk meningkatkan kinerjanya, terutama pada bagian jumlah *hidden layer* [10] dan jumlah *neuron* pada setiap *hidden layer* [11]; serta memiliki variasi fungsi aktivasi yang dapat diimplementasikan ke dalam masing-masing *hidden layer* [12].

Arsitektur MLP yang digunakan untuk mengklasifikasikan dataset *Titanic's Survivor* pada penelitian ini terdiri dari 4 variasi, yaitu 3 *hidden layer* yang masing-masing *hidden layer* nya memiliki 50 *neuron*, 3 *hidden layer* yang masing-masing *hidden layer* nya memiliki 100 *neuron*, 4 *hidden layer* yang masing-masing *hidden layer* nya memiliki 50 *neuron*, dan 4 *hidden layer* yang masing-masing *hidden layer* nya memiliki 100 *neuron*. Pada masing-masing variasi arsitektur MLP ini digunakan 2 variasi fungsi aktivasi, yaitu *ReLU* dan *Tanh*. Gambar 1 menunjukkan arsitektur MLP dengan 3 *hidden layer*, dan 4 *hidden layer* yang digunakan pada penelitian ini.



(a) Arsitektur MLP 3 Hidden Layer



(b) Arsitektur MLP 4 Hidden Layer

Gambar 1: Arsitektur MLP

D. Fungsi Aktivasi ReLu

Fungsi ReLu (*rectified liner unit*) merupakan fungsi *non-linier* dimana pengaktifan *neuron* tidak dilakukan secara bersamaan, dan hanya ketika *output* dari transformasi linier bernilai nol. Dalam penggunaannya, fungsi ReLu dituliskan menggunakan persamaan (1) [13]:

$$f(x) = \max(0, x) \quad (1)$$

Dimana :

x = Nilai data input,

$f(x)$ = Hasil output fungsi *ReLU* berupa nilai dalam bentuk 0 dan 1.

E. Fungsi Aktivasi Tanh

Fungsi aktivasi *Tahn* merupakan fungsi simetris yang berbentuk s (*sigmoid*), dimana *output* nya terletak pada rentang nilai -1 sampai +1. Dalam penggunaannya, fungsi *Tanh* dituliskan menggunakan persamaan (2) [14].

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2)$$

Dimana:

e^x = Nilai eksponensial positif dari nilai data input

e^{-x} = Nilai eksponensial negatif dari nilai data input

$f(x)$ = Hasil output fungsi *Tanh* berupa nilai dalam rentang -1 sampai 1.

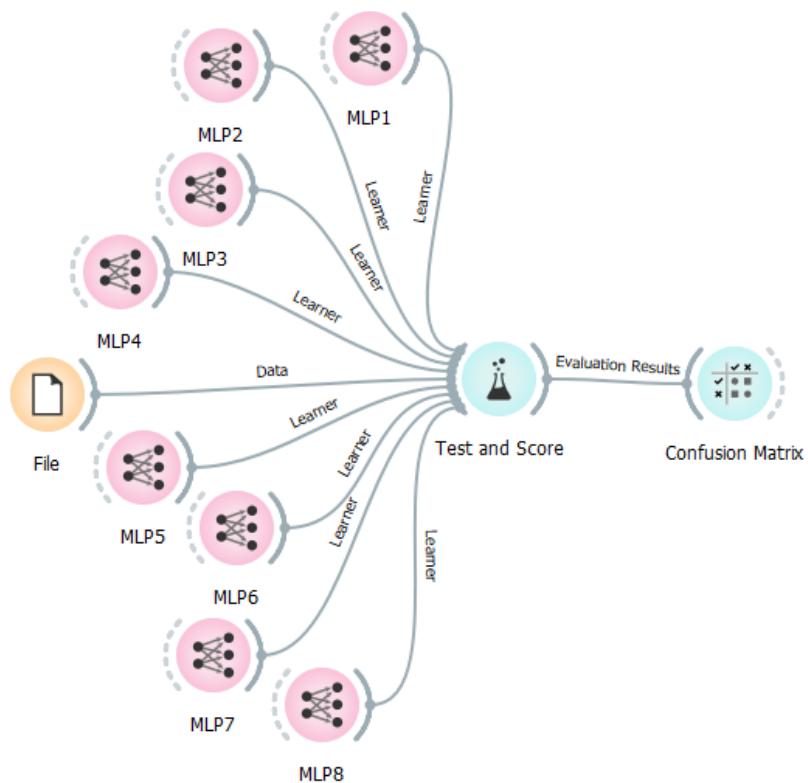
F. Rancangan Model

Model klasifikasi dirancang menggunakan algoritma MLP, dimana digunakan dua fungsi aktivasi yang berbeda, yaitu *ReLU* dan *Tanh*. Masing-masing model menggunakan kombinasi variasi parameter yang berbeda, seperti jumlah *hidden layer*, jumlah *neuron* per *hidden layer* dan fungsi optimasi, seperti terlihat pada Tabel II.

Tabel II: Parameter Model

Model	Hidden Layer	Neuron per Hidden Layer	Fungsi Aktivasi	Fungsi Optimasi
MLP1	3	50	ReLU	Adam
MLP2	3	100	ReLU	Adam
MLP3	4	50	ReLU	Adam
MLP4	4	100	ReLU	Adam
MLP5	3	50	Tanh	Adam
MLP6	3	100	Tanh	Adam
MLP7	4	50	Tanh	Adam
MLP8	4	100	Tanh	Adam

Desain model klasifikasi menggunakan aplikasi *data mining Orange 3*, dengan memanfaatkan *widget file*, *learner neural netowork*, *test & score* serta *confusion matrix*, seperti terlihat pada Gambar 2.



Gambar 2: Rancangan Model

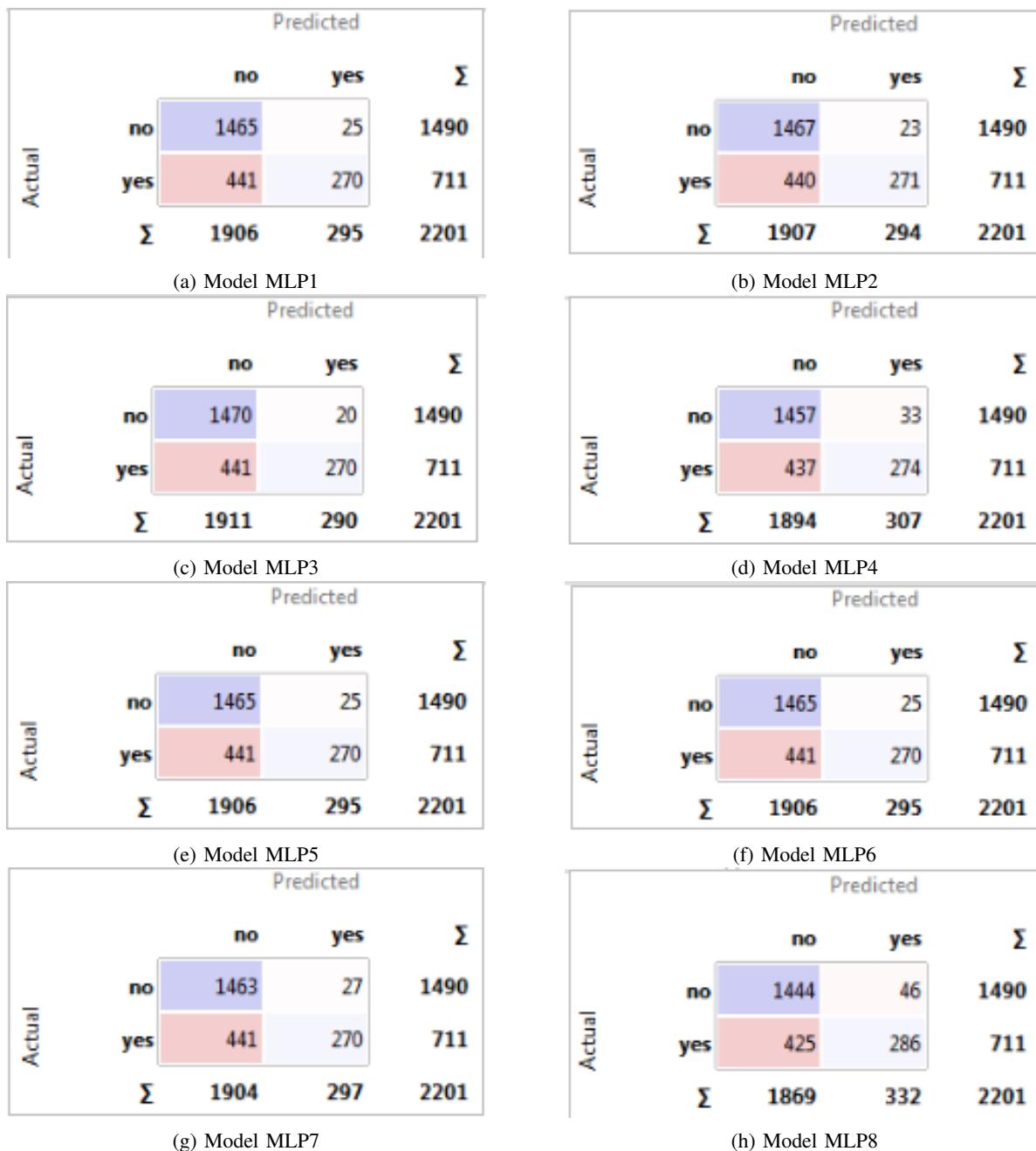
Widget *Orange 3* yang digunakan pada rancangan model dapat dijelaskan sebagai berikut:

- 1) Widget *File*: Berfungsi untuk membaca dataset *Titanic's Survivor* yang disediakan oleh aplikasi data *mining Orange 3*.
- 2) Widget *learner neural network*: Berfungsi sebagai *classifier* yang mengolah data sehingga menghasilkan klasifikasi data. Digunakan 8 widget *learner neural network* (MLP1, ..., MLP8), sesuai dengan konfigurasi parameter model pada Tabel II.
- 3) Widget *Test & Score*: Berfungsi untuk melatih dan menguji data menggunakan *classifier* yang sudah ditentukan.
- 4) Widget *Confusion Matrix*: Berfungsi untuk menampilkan nilai *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN) dari hasil klasifikasi masing-masing model.

III. HASIL

A. Hasil Klasifikasi

Dengan menggunakan *10-fold cross validation*, diperoleh hasil *confusion matrix* pada masing-masing model seperti terlihat pada Gambar 3.



Gambar 3: Koefisien filter Daubechies 2

B. Evaluasi Model

Hasil *confusion matrix* pada Gambar 3 di atas dievaluasi untuk menghitung nilai akurasi dan presisi masing-masing model menggunakan persamaan (3), (4) berikut [15]:

$$Akurasi = \frac{TP + TN}{TP + FN + FP + FN} \cdot 100\% \quad (3)$$

$$Presisi = \frac{TP}{TP + FP} \cdot 100\% \quad (4)$$

Dimana:

TP = Nilai *true positive*

FP = Nilai *false positive*

FN = Nilai *false negative*

TN = Nilai *true negative*

Dengan menggunakan persamaan (3), diperoleh nilai akurasi masing-masing model dengan rincian sebagai berikut:

$$\begin{aligned} Akurasi_{MLP1} &= \frac{1465 + 270}{1466 + 24 + 441 + 270} \times 100\% = \frac{1735}{2201} \times 100\% = 78,827\% \\ Akurasi_{MLP2} &= \frac{1467 + 271}{1467 + 23 + 441 + 271} \times 100\% = \frac{1738}{2201} \times 100\% = 78,964\% \\ Akurasi_{MLP3} &= \frac{1470 + 270}{1470 + 20 + 441 + 270} \times 100\% = \frac{1740}{2201} \times 100\% = 79,054\% \\ Akurasi_{MLP4} &= \frac{1457 + 274}{1457 + 33 + 437 + 274} \times 100\% = \frac{1731}{2201} \times 100\% = 78,646\% \\ Akurasi_{MLP5} &= \frac{1465 + 270}{1465 + 25 + 441 + 270} \times 100\% = \frac{1735}{2201} \times 100\% = 78,827\% \\ Akurasi_{MLP5} &= \frac{1465 + 270}{1465 + 25 + 441 + 270} \times 100\% = \frac{1735}{2201} \times 100\% = 78,827\% \\ Akurasi_{MLP7} &= \frac{1463 + 270}{1463 + 27 + 441 + 270} \times 100\% = \frac{1733}{2201} \times 100\% = 78,736\% \\ Akurasi_{MLP8} &= \frac{1444 + 286}{1444 + 46 + 425 + 286} \times 100\% = \frac{1730}{2201} \times 100\% = 78,6\% \end{aligned}$$

Dengan menggunakan persamaan (4), diperoleh nilai presisi masing-masing model dengan rincian sebagai berikut:

$$\begin{aligned} Presisi_{MPL1} &= \frac{1465}{1465 + 441} \times 100\% = \frac{1465}{1906} \times 100\% = 76,86253\% \\ Presisi_{MPL2} &= \frac{1467}{1467 + 440} \times 100\% = \frac{1467}{1907} \times 100\% = 76,92711\% \\ Presisi_{MPL3} &= \frac{1470}{1470 + 441} \times 100\% = \frac{1470}{1911} \times 100\% = 76,92307\% \\ Presisi_{MPL4} &= \frac{1457}{1457 + 437} \times 100\% = \frac{1457}{1894} \times 100\% = 76,92713\% \\ Presisi_{MPL5} &= \frac{1465}{1465 + 441} \times 100\% = \frac{1465}{1906} \times 100\% = 76,86253\% \\ Presisi_{MPL6} &= \frac{1465}{1465 + 441} \times 100\% = \frac{1465}{1906} \times 100\% = 76,86253\% \\ Presisi_{MPL7} &= \frac{1465}{1465 + 441} \times 100\% = \frac{1465}{1906} \times 100\% = 76,86253\% \\ Presisi_{MPL8} &= \frac{1444}{1444 + 425} \times 100\% = \frac{1444}{1896} \times 100\% = 76,16033\% \end{aligned}$$

IV. PEMBAHASAN

Hasil perhitungan nilai akurasi dan presisi masing-masing model kemudian ditabulasikan untuk mempermudah proses analisa. Tabel III menunjukkan hasil tabulasi evaluasi kinerja masing-masing model.

Tabel III: Hasil Evaluasi Model

Model 1	Akurasi	Presisi
MLP1	78,827	76,86253
MLP2	78,964	76,92711
MLP3	79,054	76,92307
MLP4	78,646	76,92713
MLP5	78,827	76,86253
MLP6	78,827	76,86253
MLP7	78,736	76,83823
MLP8	78,6	76,1603

Dari Tabel III, terlihat bahwa nilai akurasi tertinggi dicapai pada model MLP3, sedangkan nilai terendah dicapai pada model MLP8. Untuk nilai presisi, nilai tertinggi dicapai pada model MLP4, sedangkan nilai terendah dicapai pada model MLP8. Dari model MLP1 sampai MLP4, diperoleh nilai rata-rata akurasi untuk fungsi ReLu sebesar 78,8727% dan nilai rata-rata presisi sebesar 76,90996%. Dari model MLP5 sampai MLP8, diperoleh nilai rata-rata akurasi untuk fungsi Tanh sebesar 78,7475% dan nilai rata-rata presisi sebesar 76,6809%.

Dari model MLP1 dan MLP2, diperoleh nilai rata-rata akurasi untuk fungsi ReLu dengan 3 hidden layer sebesar 78,8955% dan rata-rata presisi sebesar 76,89482%. Dari model MLP3 dan MLP4, diperoleh nilai rata-rata akurasi untuk fungsi ReLu dengan 4 hidden layer sebesar 78,85% dan rata-rata presisi sebesar 76,9251%.

Dari model MLP5 dan MLP6, diperoleh nilai rata-rata akurasi untuk fungsi Tanh dengan 3 hidden layer sebesar 78,827% dan rata-rata presisi sebesar 76,86253%. Dari model MLP7 dan MLP8, diperoleh nilai rata-rata akurasi untuk fungsi Tanh dengan 4 hidden layer sebesar 78,668% dan rata-rata presisi sebesar 76,49926%.

V. SIMPULAN

Fungsi aktivasi ReLU memiliki kinerja yang lebih baik dibandingkan fungsi aktivasi Tanh untuk model penelitian ini, terlihat dari rata-rata akurasi dan presisinya yang lebih tinggi. Untuk model klasifikasi dataset Titanic's Survivor pada penelitian ini, penambahan jumlah hidden layer tidak meningkatkan kinerja hasil klasifikasi, terlihat dari menurunnya rata-rata akurasi dan presisi dari model yang menggunakan 3 hidden layer dan 4 hidden layer, baik pada fungsi ReLu maupun fungsi Tanh. Akurasi tertinggi diperoleh pada model MLP3, sehingga model ini merupakan yang terbaik untuk digunakan jika akurasi adalah nilai ukur terpenting dari hasil klasifikasi. Sedangkan nilai presisi tertinggi diperoleh oleh model MLP4, sehingga model ini merupakan yang terbaik untuk digunakan jika presisi adalah nilai ukur terpenting dari hasil klasifikasi.

PUSTAKA

- [1] I. Ariyati, Ridwansyah, and Suhardjono, "Implementasi Particle Swarm Optimization untuk Optimalisasi Data Mining Dalam Evaluasi Kinerja Asisten Dosen," JIKO (Jurnal Inform. dan Komputer) STMIK AKAKOM, vol. 3, no. 2, pp. 70–75, 2018, doi: <http://dx.doi.org/10.26798/jiko.v3i2.127>.
- [2] A. Wanto et al., "Forecasting the Export and Import Volume of Crude Oil, Oil Products and Gas Using ANN," J. Phys. Conf. Ser., vol. 1255, no. 1, 2019, doi: 10.1088/1742-6596/1255/1/012016.
- [3] M. Handayani, M. Riandini, and Z. Situmorang, "Perbandingan Fungsi Optimasi Neural Network Dalam Klasifikasi Kelayakan Calon Suami," J. Inform., vol. 9, no. 1, pp. 78–84, 2022, doi: <https://doi.org/10.31294/inf.v9i1.12318>.
- [4] M. S. Simanjuntak, Wanayumiini, R. Rosnelly, and T. S. Gunawan, "The Activity Activation Function Of Multilayer Perceptron - Based Cardiac Abnormalities," J. Mantik, vol. 4, no. 1, pp. 555–561, 2020, [Online]. Available: <http://iocscience.org/ejournal/index.php/mantik/article/view/882/595>.
- [5] D. Pardede, B. H. Hayadi, and Iskandar, "Kajian literatur multi layer perceptron: seberapa baik performa algoritma ini," J. ICT Apl. Syst., vol. 1, no. 1, pp. 23–35, 2022, [Online]. Available: <https://e-jurnal.rokania.ac.id/index.php/jictas/article/view/127/84>.
- [6] J. Rynkiewicz, "Asymptotic statistics for multilayer perceptron with ReLU hidden units," Neurocomputing, vol. 342, pp. 16–23, 2019, doi: 10.1016/j.neucom.2018.11.097.
- [7] J. C. Chen and Y. M. Wang, "Comparing activation functions in modeling shoreline variation using multilayer perceptron neural network," Water (Switzerland), vol. 12, no. 5, 2020, doi: 10.3390/W12051281.
- [8] Hartono, M. Sadikin, D. M. Sari, N. Anzelina, S. Lestari, and W. Dari, "Implementation of Artificial Neural Networks with Multilayer Perceptron for Analysis of Acceptance of Permanent Lecturers," J. Mantik, vol. 4, no. 2, pp. 1389–1396, 2020, [Online]. Available: <https://iocscience.org/ejournal/index.php/mantik>.
- [9] M. Fachrie and A. P. Wibowo, "Pemanfaatan Jaringan Syaraf Tiruan Untuk Memprediksi Kinerja Satpam," JIKO (Jurnal Inform. dan Komputer), vol. 3, no. 1, p. 46, 2018, doi: 10.26798/jiko.2018.v3i1.80.
- [10] A. J. Mohammed, M. H. Arif, and A. A. Ali, "A multilayer perceptron artificial neural network approach for improving the accuracy of intrusion detection systems," IAES Int. J. Artif. Intell., vol. 9, no. 4, pp. 609–615, 2020, doi: 10.11591/ijai.v9.i4.pp609-615.
- [11] I. Gunawan, "Optimasi Model Artificial Neural Network untuk Klasifikasi Paket Jaringan," Simetris, vol. 14, no. 2, pp. 1–5, 2020, doi: 10.51901/simetris.v14i2.135.
- [12] B. Al-shargabi, F. Al-shami, R. S. Alkhawaldeh, and C. Information, "Enhancing Multi-Layer Perceptron for Breast Cancer Prediction," Int. J. Adv. Sci. Technol., vol. 130, pp. 11–20, 2019.
- [13] K. Sharma, S. College, and Rajhasthan, "Classification of IRIS Dataset using Weka," Int. J. Comput. Appl. Inf. Technol., vol. 12, no. 1, pp. 287–291, 2020, [Online]. Available: <http://www.ijcait.com/IJCAIT/vol121kalpanasharma.php>.
- [14] B. Karlik and A. V. Olgac, "Performance Analysis of Various Activation Functions in Artificial Neural Networks," Int. J. Artif. Intell. Expert Syst., vol. 1, no. 4, pp. 111–122, 2019, doi: 10.1088/1742-6596/1237/2/022030.
- [15] K. F. Margolang, M. M. Siregar, S. Riyadi, and Z. Situmorang, "Analisa Distance Metric Algoritma K-Nearest Neighbor Pada Klasifikasi Kredit Macet," J. Inf. Syst. Res., vol. 3, no. 2, pp. 118–124, 2022, doi: 10.47065/josh.v3i2.1262.