

PREDIKSI DAN ANALISIS PENYAKIT KARDIOVASKULAR DENGAN ALGORITMA NEURAL NETWORK

Ardi Ramadhan Sukma¹; Riqadri Halfis²

Program Pascasarjana Magister Ilmu Komputer
STMIK Nusa Mandiri
<http://nusamandiri.ac.id/>
ardirsukma40@gmail.com¹, arizhboys@gmail.com²



Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi-NonKomersial 4.0 Internasional.

Abstract—*In the medical and health world it is very necessary to predict one of the cardiovascular diseases in patients. On thus occasion a dataset was used in the study of 220 records and several attributes. Prediction is done by using a neural network algorithm. Then perform calculations using Rpaidminer from the cardiovascular disease dataset. The Dataset will be predicted and analyzed with two trials on Rapidminer to determine the optimal level of predictions. The reseluts of first trial on Rapidminer were 68.81% preset value, accuracy with a 65.85% preset value, 79.41% recall, and 0.75 AUC and the results of the second trial were 77.27% accuracy, with 77.14% precision, 79.41% recall, and 0.764 AUC values.*

Intisari—*Dalam dunia medis dan kesehatan sangat perlu untuk melakukan prediksi salah satu nya penyakit kardiovaskular pada pasien. Pada kesempatan kali ini digunakan berupa dataset di dalam penelitian sebanyak 220 record dan beberapa atribut. Prediksi dilakukan dengan menggunakan algoritma neural network. Lalu melakukan perhitungan menggunakan Rapidminer dari dataset penyakit kardiovaskular. Dataset tersebut akan di prediksi dan analisis dengan dua uji coba di Rapidminer untuk menentukan tingakt optimal dari prediksi. Hasil dari uji coba pertama di Rapidminer adalah akurasi 68.81% dengan nilai presese 65.85%, recall 79.41%, dan nilai AUC 0.745 dan hasil uji coba kedua adalah akurasi 77.27%, dengan nilai presisi 77.14%, recall 79.41%, dan nilai AUC 0.764.*

Kata Kunci: *cardiovascular, machine learning, neural network, Rapidminer.*

PENDAHULUAN

Dalam dunia medis dan kesehatan sangat lah perlu untuk mengetahui prediksi suatu penyakit dan keputusan yang efektif dalam mengambil

keputusan dan anlisa terhadap pasien yang sakit.(Kementerian, n.d.)

Menurut definisi kardiovaskular dari WHO, penyakit kardiovaskuler adalah penyakit yang disebabkan gangguan fungsi jantung dan pembuluh darah. Ada banyak macam penyakit kardiovaskuler, tetapi yang paling umum dan paling terkenal adalah penyakit jantung koroner dan stroke. (Kementerian, n.d.)

Metode prediksi penyakit banyak diusulkan dengan menggunakan *Genetic* Algoritim(Ratnakar, Rajeswari, & Jacob, 2013), *naïve bayes* dan *decision tree* (Palaniappan & Awang, 2008), *Nueral Network* ,*Multilayer Perceptron*(Rodrigo & Tsokos, 2017).

Pada studi kali ini akan dilakukan pengujian menggunakan algoritma *neural network* untuk memprediksi penyakit kardiovaskular terhadap pasien yang mengalami dengan beberapa atribut.

Algoritma Neural Network adalah suatu usaha untuk meniru fungsi otak manusia. Otak manusia diyakini terdiri dari jutaan unit pengolahan kecil yang disebut neuron, yang berkerja secara paralel. Neuron saling terhubung satu sama lain melalui koneksi neuron. Setiap individu neuron mengambil input dari satu set neuron. Kemudian memproses input tersebut dan melewati output untuk satu set neuron. Keluran dikumpulkan oleh neuron lain untuk diproses lebih lanjut. (Shukla, 2010)

Tujuan penelitian kali ini untuk memprediksi akurat penyakit kardiovaskular dengan label “0” atau “1” atau memiliki penyakit kardiovaskular “ya” atau “tidak”.

BAHAN DAN METODE

Dalam Penelitian ini penulis menggunakan pengumpulan dengan metode :

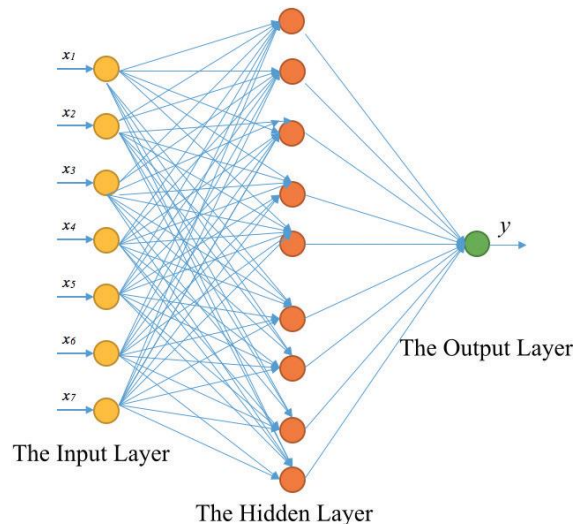
1. Pengumpulan Data
Mengumpulkan data berupa dataset yang ingin digunakan dalam suatu metode nantinya.

2. Pengolahan Data
Mengolah dataset berupa pengurangan beberapa atribut.
3. Metode yang digunakan
Metode yang digunakan adalah Algoritma *neural network* yang merupakan bagian dari *deep learning*. *Deep Learning* merupakan bagian dari *Machine Learning*.
4. Pengujian Data dengan Metode yang digunakan
Menguji *dataset* penyakit kardiovaskular dengan algoritma *neural network* menggunakan aplikasi *Rapidminer*.
5. Pengujian dan Hasil Validasi
Setelah melakukan pengujian maka di dapat hasil validasi prediksi dari penyakit kardiovaskular.

Machine Learning adalah salah satu cabang dari ilmu Kecerdasan Buatan (*Artificial Intelligence*) yang berfokus pada pengembangan sebuah sistem berdasarkan data. *Machine Learning* mempunyai empat pokok hal yang dipelajari yaitu pembelajaran terarah (*supervised learning*), pembelajaran tak terarah (*unsupervised learning*), pembelajaran semi terarah (*semi-supervised learning*), *reinforcement learning*. (Rodrigo & Tsokos, 2017)

Permasalahan yang di ambil dari penelitian adalah bagaimana mengimplementasi algoritma *neural network* untuk memprediksi penyakit akurasi, precision, recall, AUC pada dataset yang di dapat di <https://www.kaggle.com/sulianova/cardiovascular-disease-dataset> yaitu dataset berupa data penyakit kardiovaskular yang akan di prediksi dengan atribuk age, gender, height, weight, an_hi, an_lo, cholesterol, gluc, smok, alco, active, dan cadriovaskular. Label pada dataset tersebut terdapat di cardiovascular yaitu "0" atau "1" bisa juga di sebut seorang pasien mengalami penyakit kardiovaskular atau tidak. Dari label tersebut akan menguji titik akurasi penyakit kardiovaskular seorang pasien dengan *Rapidminer*.

Algoritma Neural Network adalah suatu usaha untuk meniru fungsi otak manusia. Otak manusia diyakini terdiri dari jutaan unit pengolahan kecil yang disebut neuron, yang berkerja secara paralel. Neuron saling terhubung satu sama lain melalui koneksi neuron. Setiap individu neuron mengambil input dari satu set neuron. Kemudian memproses input tersebut dan melewati output untuk satu set neuron. Keluran dikumpulkan oleh neuron lain untuk diproses lebih lanjut. (Shukla, 2010)



Sumber: (Prathama, Aminullah, Saputra, Teknik, & Mada, 2017)

Gambar1. Arsitektur Nerual Network

Arsitektur di atas biasa nya disebut sebagai *Multi Layer Perceptron (MLP)* atau *Fully-Connected Layer*. Arsitektur pertama mempunyai 7 neuron pada Input Layer dan 1 Output Layer. Di antara Input dan Output, terdapat Hidden Layer dengan jumlah 9 neuron.

Penelitian kali menggunakan *Dataset* dengan record 220, dengan 12 atribut. Tabel 1 berikut merupakan dataset yang di gunakan.

Tabel 1. Dataset Kardiovaskular

Id	Age	Gender	Height	Weight	Ap_hi	Ap_lo	cholesterol	Gluc	Smoke	Alco	Active	cardio
1	20228	1	156	85	140	90	3	1	0	0	1	1
2	18857	1	165	64	130	70	3	1	0	0	0	1
3	17623	2	169	82	150	100	1	1	0	0	1	1
...
299	21179	1	164	67	120	80	1	1	0	0	1	1

Sumber: (Sukma & Halfis, 2019)

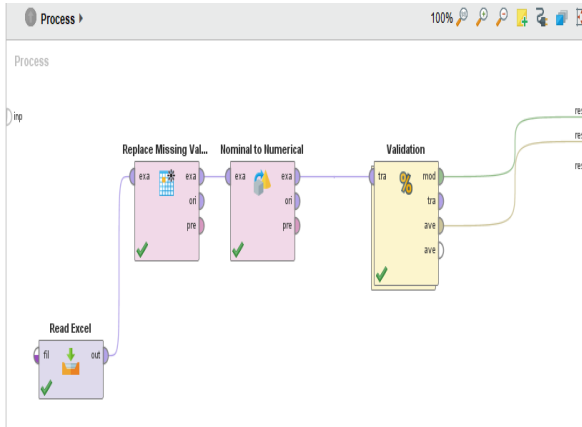
Dari *dataset* di atas akan digunakan output dari atribut cardio yaitu "0" dan "1" untuk memprediksi pasien sakit atau tidak nantinya.

Dalam penelitian akan menggunakan *Microsoft Excel 2010* sebagai dataset dan menggunakan aplikasi *Rapidminer* sebagai alat ukur prediksi.

HASIL DAN PEMBAHASAN

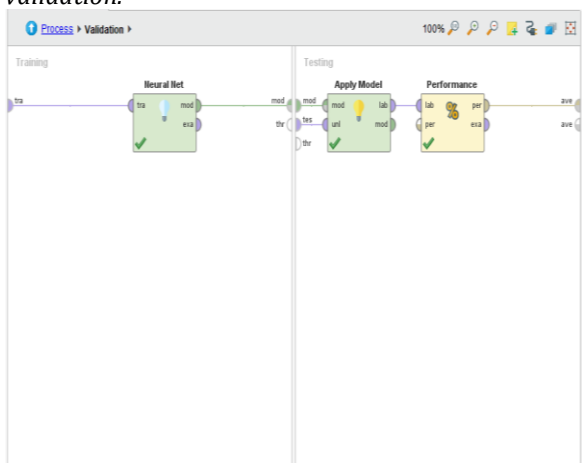
Implementasi dengan Rapidminer

Pada perhitungan ini akan dilakukan pengukuran prediksi dari *Dataset* menggunakan Algoritma *neural network* menggunakan alat bantu *Rapidminer*. Berikut modela Algoritma *neural network* pada *framework Rapidminer* :



Sumber: (Sukma & Halfis, 2019)
Gambar2. Main Process Rapidminer

Pada proses uji di Rapidminer pertama dataset di input ke *Read Excel* lalu di lakukan proses *proses replace missing value*, *nominal to numeric*, dan *cross validation*. Proses *replace missing value* digunakan ketika data mengandung duplikasi. Setelah melakukan proses tersebut maka di lanjutkan proses *nominal to numeric* untuk mengubah tipe atribut non-numerik ke tipe numerik dan juga memetakan semua nilai atribut ke nilai numerik dimana nilai atribut tersebut 0 dan 1. Lalu untuk pengujiannya di lakukan *cross validation*.



Sumber: (Sukma & Halfis, 2019)
Gambar3. Proses Validation Neural Network

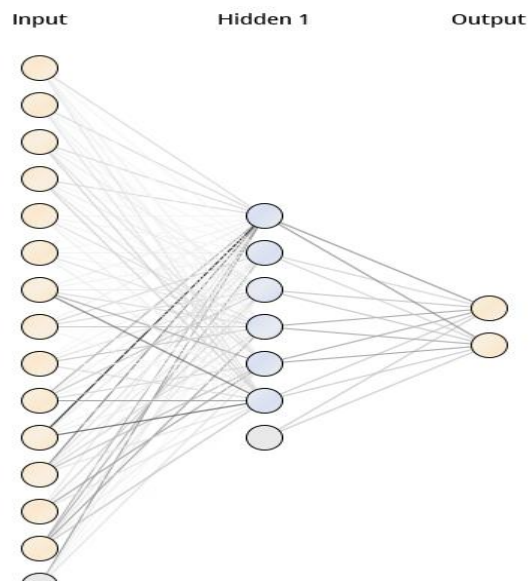
Pada Proses *cross validation* ada operation *neural network* dimana akan digunakan uji coba pertama dan uji coba kedua dengan *training cycle*,

learning rate, *momentum* yang berbeda dengan output berupa arsitektur *neural network*. *Apply Model* dan *Performance* digunakan untuk mendapatkan nilai *accuracy*, *Precision*, *recall*, dan *AUC*.

Model algoritma *neural network* dilakukan dengan proses training model dengan memberikan 2 uji coba yaitu:

- a. Uji Coba Pertama
 - 1. *Training cycles* : 200
 - 2. *Learing rate* : 0.5
 - 3. *Momentum* : 0.1
- b. Uji Coba ke dua
 - 1. *Training cycles* : 1500
 - 2. *Learing rate* : 0.5
 - 3. *Momentum* : 0.8

Dari penelitian uji coba pertama diatas maka dapat arsitektur *neural network* dengan menghasilkan hidden layer dengan dua output :



Sumber: (Sukma & Halfis, 2019)
Gambar4. Arsitektur Neural Network Kardiovaskular Uji Coba Pertama

Pada uji coba pertama di dapat arsitektur *neural network* berupa *Input* dan *Output Layer*. Di antara *Input* dan *Output* terdapat *Hidden Layer*. *Input Layer* memiliki 14 neuron, 1 bias dan *Output Layer* memiliki 2 neuron di antaranya ada 6 hidden layer dan 1 bias. Masing-masing memiliki nilai berbeda dengan uji coba pertama *training cycle* : 200, *learning rate* : 0.5, dan *momentum* : 0.1.

Dengan nilai setiap node-node algoritma *neural network* yang di dapatkan dari proses *Rapidminer* :

Tabel 2. Nilai Node Input Hidden Layer Uji Coba Pertama

	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6
Smoke = 0	-2.759	0.693	0.137	-1.129	0.958	1.880
Smoke = 1	2.763	-0.715	-0.156	1.086	-0.952	-1.953
Alco = 0	2.899	-0.757	-1.469	0.489	-2.280	3.089
Alco = 1	-2.855	0.702	1.407	-0.484	2.366	-3.123
Active = 0	-0.640	0.898	-0.835	-2.227	-0.497	-1.271
Active = 1	0.656	-0.820	0.761	2.264	0.497	1.318
Age	1.126	-0.994	-1.208	-2.752	-6.178	9.533
Gender	-1.150	-3.111	-2.690	-3.603	-0.241	0.751
Height	-0.437	0.993	2.359	4.055	-0.263	-2.672
Weight	5.823	-1.187	-4.275	1.057	-2.991	6.300
Ap_hi	16.375	1.782	0.096	-0.901	-1.153	10.751
Ap_lo	8.813	-2.228	-1.699	-2.864	-4.731	2.034
Cholesterol	-1.792	0.114	3.328	0.632	6.720	-2.059
Gluc	-5.518	5.823	1.904	6.887	1.166	4.282
Bias	-7.424	-1.476	-0.771	-0.509	-2.155	-0.216

Sumber: (Sukma & Halfis, 2019)

Dari nilai node hidden layer maka didapat output yang dihasilkan oleh algoritma *neural network* sebagai berikut :

Tabel 3. Nilai Node Output Uji Coba Pertama

	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Thereshold
Class = 0	-7.545	-3.261	-4.126	5.947	-6.199	-3.875	3.889
Class =1	7.544	3.260	4.169	-5.950	6.195	3.874	-3.889

Sumber: (Sukma & Halfis, 2019)

Dari analisa penguian dengan algoritma *neural network* maka di peroleh nilai *accuracy*, *presicion*, *recall*, dan *AUC*. Berikut nilai *accuracy*, *presicion*, *recall*, dan *AUC* :

Tabel 4. Nilai Algoritma *Neural Network* Uji Coba Pertama

	Neural Network
Accuracy	68.81%
Presicion	65.85%
Recall	79.41%
AUC	0.745

Sumber: (Sukma & Halfis, 2019)

accuracy: 68.18%

	true 0	true 1	class precision
pred. 0	18	7	72.00%
pred. 1	14	27	65.85%
class recall	56.25%	79.41%	

Sumber: (Sukma & Halfis, 2019)

Gambar5. Hasil Output Accuracy Uji Coba Pertama

precision: 65.85% (positive class: 1)

	true 0	true 1	class precision
pred. 0	18	7	72.00%
pred. 1	14	27	65.85%
class recall	56.25%	79.41%	

Sumber: (Sukma & Halfis, 2019)

Gambar6. Hasil Output Precision Uji Coba Pertama

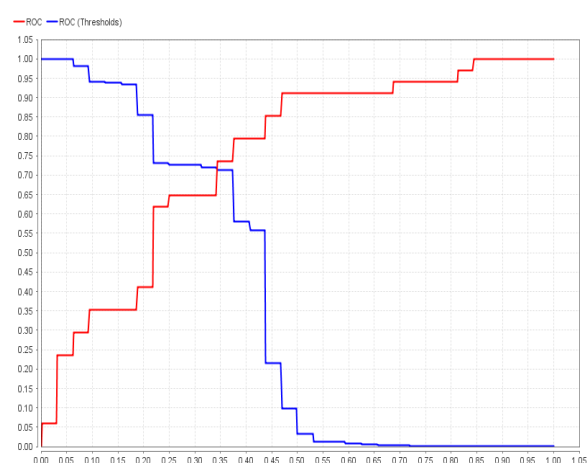
recall: 79.41% (positive class: 1)

	true 0	true 1	class precision
pred. 0	18	7	72.00%
pred. 1	14	27	65.85%
class recall	56.25%	79.41%	

Sumber: (Sukma & Halfis, 2019)

Gambar7. Hasil Output Recall Uji Coba Pertama

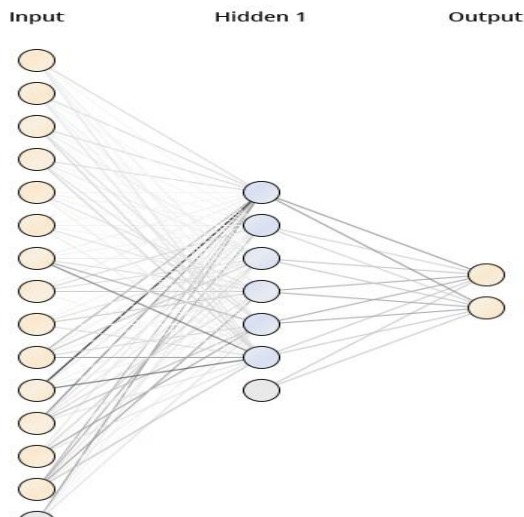
AUC: 0.747 (positive class: 1)



Sumber: (Sukma & Halfis, 2019)

Gambar8. Hasil Output AUC Uji Coba Pertama

Berikut Uji Coba Ke dua yang di lakukan dengan training cycles, learning rate, dan momentum yang berbeda :



Sumber: (Sukma & Halfis, 2019)
Gambar9. Arsitektur Neural Network Kardiovaskular Uji Coba Kedua

Pada uji coba kedua di dapat arsitektur *neural network* berupa *Input* dan *Output Layer*. Di antara *Input* dan *Output* terdapat *Hidden Layer*. *Input Layer* memiliki 14 neuron, 1 bias dan *Output Layer* memiliki 2 neuron di antaranya ada 6 hidden layer dan 1 bias. Masing-masing memiliki nilai berbeda dengan uji coba kedua *training cycle* : 1500, *learning rate* : 0.5, dan *momentum* : 0.8.

Dari hasil uji coba kedua maka di dapatkan node-node dari 14 neuron yang ada sebagai berikut :

Tabel 5. Nilai Node Input Hidden Layer Uji Coba Kedua

	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6
Smoke = 0	3.704	4.052	-2.852	-8.311	2.009	8.306
Smoke = 1	-3.700	-4.704	2.833	8.268	-2.004	-8.380
Alco = 0	-7.830	-3.558	-2.499	1.288	-11.217	-7.216
Alco = 1	7.874	3.503	2.436	-1.282	11.303	7.182
Active = 0	-0.526	5.030	-1.314	2.325	-3.533	-5.223
Active = 1	0.542	-4.953	1.241	-2.288	3.533	5.270
Age	-12.285	-3.528	0.210	14.758	-6.184	27.720
Gender	3.859	-6.174	0.823	-23.025	8.384	18.956
Height	-5.499	-2.699	4.469	9.010	0.589	-34.999
Weight	-9.425	-6.989	-18.075	-23.025	-17.029	-14.849
Ap_hi	0.205	11.948	62.911	-15.314	13.218	13.441
Ap_lo	-8.907	0.439	-1.733	41.201	16.183	7.438
Cholesterol	23.376	5.041	9.582	-22.617	3.698	-20.498
Gluc	4.381	18.612	-18.661	-15.600	21.012	31.701
Bias	-3.187	-9.873	-15.415	-32.010	21.596	-11.787

Sumber: (Sukma & Halfis, 2019)

Dari nilai node hidden layer maka didapat output yang dihasilkan oleh algoritma *neural network* sebagai berikut :

Tabel 6. Nilai Node Output Uji Coba Pertama

	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Thereshold
Class = 0	-15.624	-6.831	-15.133	-3.828	8.763	-4.108	2.529
Class = 1	15.624	6.831	15.133	3.828	-8.763	4.108	-2.529

Sumber: (Sukma & Halfis, 2019)

Dari analisa penguian dengan algoritma *neural network* maka di peroleh nilai *accuracy*, *presicion*, *recall*, dan *AUC*. Berikut nilai *accuracy*, *presicion*, *recall*, dan *AUC* :

Tabel 7. Nilai Algoritma Neural Network Uji Coba Kedua

	Neural Network
Accuracy	77.27%
Presicion	77.14%
Recall	79.41%
AUC	0.764

Sumber: (Sukma & Halfis, 2019)

accuracy: 77.27%

	true 0	true 1	class precision
pred. 0	24	7	77.42%
pred. 1	8	27	77.14%
class recall	75.00%	79.41%	

Sumber: (Sukma & Halfis, 2019)

Gambar10. Hasil Output Accuracy Uji Coba Kedua

precision: 77.14% (positive class: 1)

	true 0	true 1	class precision
pred. 0	24	7	77.42%
pred. 1	8	27	77.14%
class recall	75.00%	79.41%	

Sumber: (Sukma & Halfis, 2019)

Gambar11. Hasil Output Presicion Uji Coba Kedua

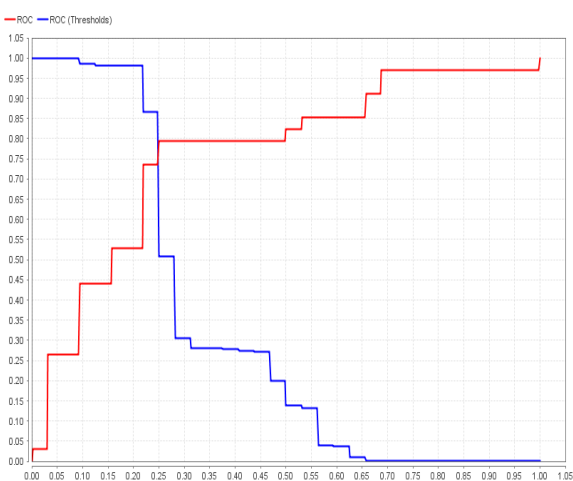
recall: 79.41% (positive class: 1)

	true 0	true 1	class precision
pred. 0	24	7	77.42%
pred. 1	8	27	77.14%
class recall	75.00%	79.41%	

Sumber: (Sukma & Halfis, 2019)

Gambar12. Hasil Output Recall Uji Coba Kedua

AUC: 0.764 (positive class: 1)



Sumber: (Sukma & Halfis, 2019)

Gambar13. Hasil Output AUC Uji Coba Kedua

Hasil penelitian untuk penerapan prediksi penyakit kardiovaskular menggunakan algoritma *neural network* dengan 2 uji coba dimana uji coba pertama di dapat akurasi 68.81% dengan nilai presisi 65.85%, recall 79.41%, dan nilai AUC 0.745 dan ujicoba kedua di dapat akurasi 77.27% dengan nilai presisi 77.14%, recall 79.41%, dan nilai AUC 0.764.

KESIMPULAN

Algoritma *neural network* di gunakan dalam penelitian ini untuk melakukan prediksi terhadap penyakit kardiovaskular berdasarkan beberapa atribut dengan menggunakan aplikasi *Rapidminer* untuk menguji *dataset* dimana dengan output hasil prediksi penyakit kardiovaskular terdapat perbandingan dengan dua uji coba. Hasil dari pengujian *Rapidminer* didapat pada uji coba pertama berupa akurasi 68.81% dengan nilai presisi 65.85%, recall 79.41%, dan nilai AUC 0.745 dan pada uji coba kedua di dapat akurasi 77.27% dengan nilai presisi 77.14%, recall 79.41%, dan nilai AUC 0.764. Dengan demikian hasil pengujian dengan algoritma *neural network* masih belum optimal dan di diharapkan untuk menggunakan

algoritma lainnya. Hasil penelitian berupa prediksi dan analisis penyakit kardiovaskular pada pasien dengan dua ujicoba di *Rapidminer* menggunakan algoritma *neural network*. Hasil lebih optimal di uji coba kedua dengan *training cycles*, *learning rate*, dan *momentum* yang berbeda.

REFERENSI

- Kementerian, B. L. (n.d.). *Situasi Kesehatan Jantung*. Retrieved from <https://www.depkes.go.id/download.php?file=download/pusdatin/infodatin/infodatin-jantung.pdf>
- Palaniappan, S., & Awang, R. (2008). Intelligent heart disease prediction system using data mining techniques. *AICCSA 08 - 6th IEEE/ACS International Conference on Computer Systems and Applications*, (December), 108–115. <https://doi.org/10.1109/AICCSA.2008.4493524>
- Prathama, A. Y., Aminullah, A., Saputra, A., Teknik, D., & Mada, U. G. (2017). *UNTUK PENENTUAN PROSENTASE BOBOT PEKERJAAN PADA RUMAH SAKIT PRATAMA*. 7. <https://doi.org/10.22146/teknosains.30139>
- Ratnakar, S., Rajeswari, K., & Jacob, R. (2013). Prediction of Heart Disease Using Genetic Algorithm. *International Journal of Advanced Computational Engineering and Networking*, 1(2), 51–55.
- Rodrigo, H., & Tsokos, C. P. (2017). Artificial Neural Network Model for Predicting Lung Cancer Survival. *Journal of Data Analysis and Information Processing*, 05(01), 33–47. <https://doi.org/10.4236/jdaip.2017.51003>
- Shukla, A. (2010). Real Life Applications of Soft Computing. In *Real Life Applications of Soft Computing*. <https://doi.org/10.1201/ebk1439822876>
- Sukma, A. R., & Halfis, R. (2019). *PREDIKSI DAN ANALISIS PENYAKIT KARDIOVASKULAR DENGAN ALGORITMA NEURAL NETWORK*. 16(2), 1–6.