



Hyperparameter Tuning Model Artificial Neural Network untuk Klasifikasi Penyakit Breast Cancer

M. Nejatullah Sidqi^{1✉}, Ahmad Marsehan², Rio³, Rani Oktafelani⁴

Teknologi Informasi, Fakultas Sains dan Teknologi, Universitas PGRI Silampari, Lubukliggau^(1,2)

Sistem Informasi, Fakultas Sains dan Teknologi, Universitas PGRI Silampari, Lubuklinggau⁽³⁾

Teknologi Informasi, Fakultas Sains dan Teknologi, Universitas PGRI Silampari, Lubukliggau⁽⁴⁾

DOI: 10.31004/jutin.v7i3.31469

✉ Corresponding author:
[neja.sidqi97@gmail.com]

Article Info

Abstrak

Kata kunci:
Kanker Payudara;
Pemrosesan;
ReLU;
Sigmoid;
ANN

Kanker merupakan salah satu penyakit mematikan yang seringkali dialami secara tidak terduga. Kanker tidak hanya menyerang orang dewasa dan orang lanjut usia, bahkan bayi baru lahir pun bisa terkena kanker. Kanker payudara penyakit kanker nomor satu di Indonesia dan merupakan salah satu penyebab utama kematian akibat kanker. Menurut Globocan 2020, jumlah kasus baru kanker payudara di Indonesia meningkat dari 396.914 kasus kanker baru menjadi 68.858 kasus (16,6%). Terdapat uji coba diagnosis dini kanker payudara yang dilakukan pada dataset *Breast Cancer* yang berjumlah 569 pasien dengan kategori target yaitu *Malignant* (kanker ganas) dan *Benign* (kanker jinak). Diagnosis ini bertujuan melakukan klasifikasi dengan data mining untuk mengekstrak informasi serta pengetahuan baru. Metode yang diterapkan menggunakan algoritma *Artificial Neural Network* (ANN), dengan menggunakan parameter, 1 layer dan 250 sampai 500 epoch. Sedangkan fungsi aktivasi yang digunakan adalah aktivasi ReLU dan *Sigmoid*. Kemudian hasil klasifikasi terbaik dihasilkan oleh epoch sebanyak 250 dengan akurasi sebesar 98,24%.

Keywords:
Breast cancer;
Preprocessing;
ReLU;
Sigmoid;
ANN

Abstract

Cancer is a deadly disease that is often experienced unexpectedly. Cancer does not only attack adults and the elderly, even newborn babies can get cancer. Breast cancer is the number one cancer disease in Indonesia and is one of the main causes of death due to cancer. According to Globocan 2020, the number of new cases of breast cancer in Indonesia increased from 396,914 new cancer cases to 68,858 cases (16.6%). There was a trial for early diagnosis of breast cancer carried out on the Breast Cancer dataset, totaling 569 patients with target categories namely Malignant (malignant cancer) and Benign (benign cancer). This diagnosis aims to carry out classification using data mining to extract new information and knowledge. The method applied uses the Artificial Neural Network (ANN) algorithm, using parameters, 1 layer and 250 to 500 epochs. Meanwhile, the

activation functions used are ReLU and Sigmoid activation. Then the best classification results were produced by 250 epochs with an accuracy of 98.24%.

1. INTRODUCTION

Kanker merupakan salah satu penyakit mematikan yang seringkali dialami secara tidak terduga. Kanker tidak hanya menyerang orang dewasa dan orang lanjut usia, bahkan bayi baru lahir pun bisa terkena kanker. Ada banyak penyakit kanker dengan gejala yang hampir sama namun jenisnya berbeda, dan tingkat keparahan (bahaya) kanker bervariasi dari kanker biasa hingga kanker ganas yang memiliki perubahan signifikan pada tubuh (Kraugusteeliana dkk., 2023).

Kanker terdapat suatu sel tubuh yang mengalami proses pertumbuhan tidak normal bahkan dapat tumbuh (menyebarkan) ke bagian lain, pertumbuhan tersebut awalnya normal dan akhirnya bersifat ganas dan dapat menyebabkan kematian. Proses pertumbuhan ini sangat cepat sehingga dapat menimbulkan perubahan yang tidak dapat dikendalikan lagi dan disebut dengan mutasi sel tubuh. Menghilangkan kanker memang tidak mudah, ketika kanker sudah cukup umur maka naluri kanker tersebut akan hilang (mati), bahkan kanker terus menyebar dan terus menyerang sel-sel normal sehingga menyebabkan sel-sel normal tersebut mati akibat penyakit tersebut (Achmad, 2022).

Selain itu kanker juga dapat menyerang perempuan yang bisa terdapat pada payudara wanita atau sering disebut kanker payudara (Anggriani dkk., 2023). Kanker payudara merupakan salah satu penyebab kematian utama pada wanita, nomor dua setelah kanker paru-paru. Kanker payudara salah satu penyakit non-kulit paling berbahaya pada wanita yang disebabkan oleh berbagai faktor selain kulit payudara, mulai dari sel dan saluran hingga jaringan pendukung payudara (Muntiari & Hanif, 2022). Kemudian ada juga yang dapat memicu tumbuhnya kanker payudara antara lain penuaan, jenis kelamin, ras, riwayat keluarga, genetika atau perilaku pribadi seperti merokok, minum alkohol (Oktavianto & Handri, 2020). Selain itu, sel kanker payudara dapat menyerang lemak, pembuluh darah, dan saraf payudara, sehingga mudah menyebar ke bagian tubuh lain jika tidak segera ditangani. Prosedurnya dimulai ketika sel-sel jaringan payudara tumbuh terlalu cepat. Selanjutnya kanker yang terjadi pada sel payudara dan dapat berasal dari komponen kelenjar seperti epitel duktal dan lobular, serta dari komponen non kelenjar seperti lemak, pembuluh darah, dan persarafan (A'yunan dkk., 2023).

Kanker payudara penyakit kanker nomor satu di Indonesia dan merupakan salah satu penyebab utama kematian akibat kanker. Menurut Globocan 2020, jumlah kasus baru kanker payudara di Indonesia meningkat dari 396.914 kasus kanker baru menjadi 68.858 kasus (16,6%). Di saat yang sama, jumlah kematian lebih dari 22 ribu (Kemenkes, 2022). Menurut WHO (Organisasi Kesehatan Dunia), kanker payudara menyumbang 42,5% dari seluruh kasus kanker di dunia pada tahun 2018, dan sekitar 9,3 wanita meninggal karenanya setiap tahun (Mubarog dkk., 2021). Diagnosis dini kanker payudara dapat dilakukan dengan menggunakan metode data mining. Tujuan dari proses ini adalah mengekstrak informasi baru dari kumpulan data dan menggunakan teknik seperti statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk menemukan informasi dan wawasan berguna dari database. Dalam hal ini, klasifikasi kanker payudara dengan menggunakan data mining berperan penting dalam menentukan tingkat keparahan dan menciptakan rencana pengobatan yang tepat (Meilani & Nurdiawan, 2023).

Data mining adalah proses yang menggabungkan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin (istilah yang digunakan untuk mesin yang dapat beroperasi secara mandiri tanpa menerima instruksi dari pengguna) untuk mengekstrak informasi yang relevan dan berguna dari data dalam jumlah besar. Data mining umumnya digunakan untuk menganalisis dan mengeksplorasi data berukuran besar untuk memperoleh informasi yang berguna. Dengan bantuan data mining, banyak informasi yang dapat diperoleh dari pengolahan, misalnya pada saat menganalisis pembelian konsumen atau menentukan proses produksi serta mendeteksi penyakit (Srirahayu & Pribadie, 2023). Proses klasifikasi data mining dapat menggunakan jenis algoritma Artificial Neural Network (ANN), dengan sistem kerja yang terdiri dari neuron-neuron buatan yang saling berhubungan untuk memproses informasi dan menggabungkannya untuk proses komputasi (Hananti & Sari, 2021).

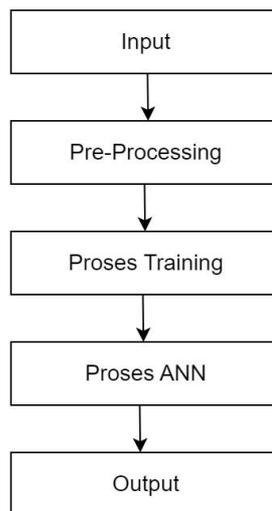
Terdapat penelitian sebelumnya dengan melakukan analisis performa metode klasifikasi Naive Bayes dengan menerapkan unbalanced dataset breast cancer menghasilkan tingkat akurasi sebesar 88% dengan menerapkan pembagian dataset menggunakan cross validation sebanyak 5 kali (Ericha Apriliyani & Salim, 2022). Sedangkan terdapat penelitian lain dengan melakukan fitur seleksi pada dataset Breast Cancer dengan menggunakan algoritma binary PSO yang digunakan untuk melakukan fitur seleksi dengan mereduksi dari 30 fitur

diturunkan menjadi 5 fitur yang terpilih. Proses pemilihan fitur seleksi untuk menurangi komputasi atau proses perhitungan algoritma menjadi lebih ringan, kemudian untuk algoritma yang digunakan untuk dengan merapkan kombinasi antara model K-NN+BPSO menghasikan tingkat akurasi terbaik sebesar 95,32% pada k=9, selanjutnya pemilihan fitur yang digunakan di antaranya compactness se, symmetry se, texture worst, perimeter worst, dan symmetry worst (Hidayat dkk., 2023). Selain dari itu terdapat penelitian pada peningkatan algoritma C4.5 berbasis PSO pada penyakit kanker payudara menggunakan 116 jumlah dataset pasien yang didapat dari data public, mendapatkan nilai sebesar hasil akurasi terbaik sebesar 84.44% (Nurchahyo dkk., 2023)

Berdasarkan penelitian sebelumnya, maka penelitian ini akan melakukan klasifikasi penyakit kanker payudara menggunakan algoritma Artificial Neural Network untuk melakukan klasifikasi penyakit kanker payudara dengan indentifikasi jenis kanker apakah jinak atau ganas. Dataset yang digunakan kan berasal dari kaggle yang bernama breast cancer wisconsin dengan jumlah data pasien sebesar 568 record serta memiliki 32 fitur. Kemudian akan dilakukan pengoptimalan hyperparameter tuning pada epoch, batch size serta fungsi aktivasinya sehingga proses klasifikasi dapat mendapatkan akurasi terbaik.

2. METHODS

Metode penelitian yang akan dilakukan pada penelitian ini, terdiri dari proses input, proses dan output, dapat dilihat pada Gambar 1.



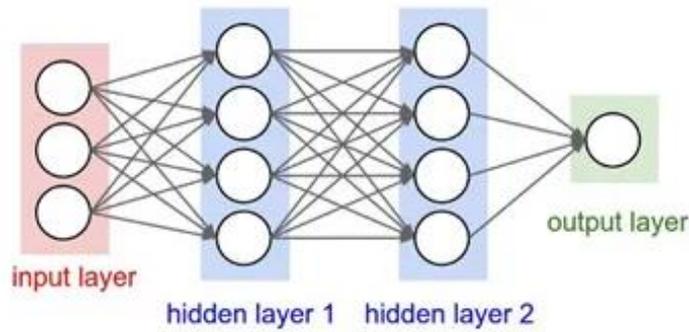
Gambar 1. Alur Penelitian

Proses input pada dataset Breast Cancer Wisconsin yang memiliki 32 fitur terdiri dari 568 record pasien yang dilakukan preprocessing untuk membersihkan data yang memiliki noise. Kemudian dilakukan proses training dengan melakukan split data sebesar 80 : 20 secara acak, selanjutnya akan dilakukan proses generate menggunakan proses model Artificial Neural Network (ANN), yang terdiri dari proses input, hidden layer dan output layer. Hasil akhir dari output layer berupa klasifikasi biner untuk menentukan penyakit kanker pada pasien.

2.1 Preprocessing

2.2 Artificial Neural Network

Algoritma ANN (Artificial Neural Network) merupakan model penalaran yang didasarkan pada otak manusia. ANN terdiri dari beberapa prosesor yang sangat sederhana dan saling berhubungan yang disebut neuron. Neuron yang dihubungkan oleh skala mengirimkan sinyal dari satu neuron ke neuron lainnya. ANN adalah model yang meniru perilaku jaringan saraf biologis. Selama proses pembelajaran, jaringan saraf tiruan dapat mengatur dirinya sendiri untuk memberikan respons yang konsisten terhadap banyak masukan. Jaringan saraf tiruan dirancang dan dilatih agar memiliki karakteristik mirip manusia (Sihombing & Hendarsin, 2020). Setiap neuron dapat memiliki banyak masukan dan satu keluaran. Jalur masukan ke suatu neuron mungkin berisi data mentah atau data yang diproses oleh neuron sebelumnya. Pada saat yang sama, produksi neutron dapat berupa produk akhir atau sebagai hasil akhir pada klasifikasi (Efsa dkk., 2021). Algoritma Artificial Neural Network memiliki konsep arsitektur jaringan syaraf tiruan seperti pada Gambar 2.



Gambar 2. Arsitektur Artificial Neural Network

Jaringan pada algoritma Artificial Neural Network (ANN) memiliki susunan atau lapisan yang terdiri dari 3 lapisan yaitu

- Lapisan input (Input Layer) sebagai lapisan masukan bertindak sebagai penghubung jaringan ke dunia luar (sumber data). Neuron-neuron ini tidak melakukan apa pun terhadap data, mereka hanya meneruskan data tersebut ke lapisan berikutnya.
- Lapisan tersembunyi (Hidden Layer) sebagai suatu jaringan yang dapat memiliki lebih dari satu lapisan tersembunyi, atau bahkan tidak memiliki lapisan sama sekali. Jika ada beberapa lapisan tersembunyi dalam jaringan, lapisan tersembunyi bagian bawah menerima masukan dari lapisan masukan.
- Lapisan keluaran (Output Layer) sebagai prinsip kerja lapisan ini sama dengan prinsip kerja lapisan tersembunyi dan juga digunakan fungsi sigmoid. Namun hasil dari lapisan ini sudah dianggap sebagai hasil proses (Mutia, 2024) (Ihsani dkk., 2020).

Hasil klasifikasi ANN menggunakan parameter yang akan di uji validitasnya menggunakan sebuah Confusion Matriz berupa Akurasi sebagai Akurasi didefinisikan sebagai tingkat kedekatan antara nilai prediksi dengan nilai actual. Berikut rumus mencari hasil akurasi (Riva dkk., 2022).

$$\text{Akurasi} = ((\text{TP} + \text{TN})) / ((\text{TP} + \text{TN} + \text{FP} + \text{FN}))$$

Keterangan Variabel:

TP : True Positive

TN : True Negative

FP : False Positive

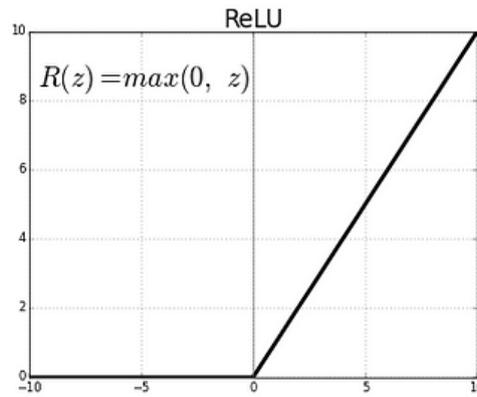
FN : False Negative

2.3 Fungsi Aktivasi

Fungsi aktivasi dalam jaringan saraf dan pembelajaran mendalam berperan penting dalam menyalakan simpul tersembunyi untuk menghasilkan keluaran yang lebih diinginkan. Tujuan utama fungsi aktivasi adalah untuk memperkenalkan sifat nonlinier ke dalam model. Pada penelitian ini menggunakan 2 jenis fungsi aktivasi yaitu aktivasi relu dan aktivasi sigmoid (Sudarsono, 2016).

1) Aktivasi ReLU.

ReLU adalah fungsi aktivasi yang telah banyak digunakan saat ini. Karena digunakan di hampir semua jaringan saraf konvolusional atau pembelajaran mendalam. Fungsi aktivasi ini digunakan pada hidden layer. Berikut gambar 3 dari fungsi aktvias ReLU

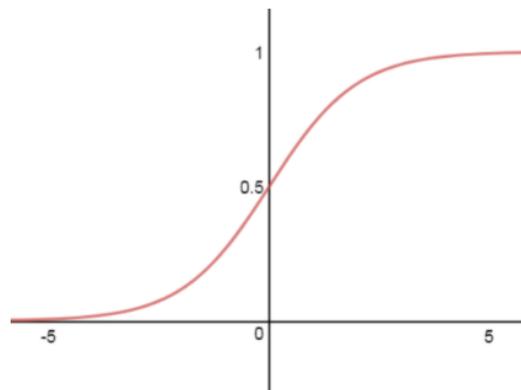


Gambar 3. Fungsi Aktivasi ReLU

ReLU setengah diperbaiki (dari bawah). $f(z)$ adalah nol ketika z kurang dari nol dan $f(z)$ sama dengan z ketika z di atas atau sama dengan nol. Namun masalahnya adalah semua nilai negatif langsung menjadi nol yang mengurangi kemampuan model untuk menyesuaikan atau melatih data dengan benar. Itu berarti setiap masukan negatif yang diberikan ke fungsi aktivasi ReLU langsung mengubah nilai menjadi nol dalam grafik, yang pada gilirannya memengaruhi grafik yang dihasilkan dengan tidak memetakan nilai negatif dengan tepat.

1) Aktivasi Sigmoid

Fungsi aktivasi yang paling umum adalah fungsi sigmoid. Fungsi sigmoid memetakan masukan yang masuk ke rentang antara 0 dan 1. Fungsi sigmoid di gunakan pada output layer, gambar dari fungsi aktivasi terdapat pada gambar 4.



Gambar 4. Fungsi Aktivasi Sigmoid

Sifat lain yang tidak diinginkan dari aktivasi sigmoid adalah bahwa keluaran fungsi tidak berpusat pada nol. Hal ini biasanya membuat pelatihan jaringan saraf menjadi lebih sulit dan tidak stabil (Suhartanto dkk., 2017).

3. RESULT AND DISCUSSION

3.1 Pegamblian Dataset

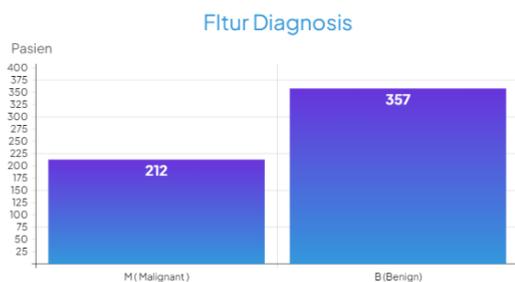
Dataset Breast Cancer Wisconsin memiliki data pasien sebesar 568, terdiri dari 31 atribut/fitur. Dari atribut target atau label terdapat pada tabel diagnosis, sedangkan dari atribut tersebut memiliki *record* yaitu M (Malignant) untuk menentukan diagnosa kanker ganas, sedangkan B (Benign) hasil diagnosa sebagai kanker jinak. Dataset di presentasikan dalam sampel data yang terdapat pada Tabel 9 selanjutnya pengambilan dataset Breast Cancer Wisconsin, yang memiliki jumlah data pasien sebanyak 569 serta dimiliki 31 kolom (Kaggle, 2021).

Tabel 9. Sampel Dataset

Nama Atribut	Pasien					
	1	2	3	4	5	6
diagnosis	M	M	M	M	M	M
radius_mean	12.45	18.25	13.71	13	15.85	13.73

Nama Atribut	Pasien					
	1	2	3	4	5	6
diagnosis	M	M	M	M	M	M
texture_mean	15.7	19.98	20.83	21.82	23.95	22.61
perimeter_mean	82.57	119.6	90.2	87.5	103.7	93.6
area_mean	477.1	1040	577.9	519.8	782.7	578.3
smoothness_mean	0.1278	0.09463	0.1189	0.1273	0.08401	0.1131
compactness_mean	0.17	0.109	0.1645	0.1932	0.1002	0.2293
concavity_mean	0.1578	0.1127	0.09366	0.1859	0.09938	0.2128
concave points_mean	0.08089	0.074	0.05985	0.09353	0.05364	0.08025
symmetry_mean	0.2087	0.1794	0.2196	0.235	0.1847	0.2069
fractal_dimension_mean	0.07613	0.05742	0.07451	0.07389	0.05338	0.07682
radius_se	0.3345	0.4467	0.5835	0.3063	0.4033	0.2121
texture_se	0.8902	0.7732	1.377	1.002	1.078	1.169
perimeter_se	2.217	3.18	3.856	2.406	2.903	2.061
area_se	27.19	53.91	50.96	24.32	36.58	19.21
smoothness_se	0.00751	0.004314	0.008805	0.005731	0.009769	0.006429
compactness_se	0.03345	0.01382	0.03029	0.03502	0.03126	0.05936
concavity_se	0.03672	0.02254	0.02488	0.03553	0.05051	0.05501
concave points_se	0.01137	0.01039	0.01448	0.01226	0.01992	0.01628
symmetry_se	0.02165	0.01369	0.01486	0.02143	0.02981	0.01961
fractal_dimension_se	0.005082	0.002179	0.005412	0.003749	0.003002	0.008093
radius_worst	15.47	22.88	17.06	15.49	16.84	15.03
texture_worst	23.75	27.66	28.14	30.73	27.66	32.01
perimeter_worst	103.4	153.2	110.6	106.2	112	108.8
area_worst	741.6	1606	897	739.3	876.5	697.7
smoothness_worst	0.1791	0.1442	0.1654	0.1703	0.1131	0.1651
compactness_worst	0.5249	0.2576	0.3682	0.5401	0.1924	0.7725
concavity_worst	0.5355	0.3784	0.2678	0.539	0.2322	0.6943
concave points_worst	0.1741	0.1932	0.1556	0.206	0.1119	0.2208
symmetry_worst	0.3985	0.3063	0.3196	0.4378	0.2809	0.3596
fractal_dimension_worst	0.1244	0.08368	0.1151	0.1072	0.06287	0.1431

Pada dataset Breast Cancer Wisconsin memiliki dua target label yang akan dilakukan proses klasifikasi menggunakan Artificial Neural Network (ANN) dengan jumlah kelas M (Malignant) berjumlah 212 pasien sedangkan kelas B (Benign) berjumlah 357 pasien. Kemudian data tersebut di presentasikan dalam bentuk Grafik pada gambar 5.



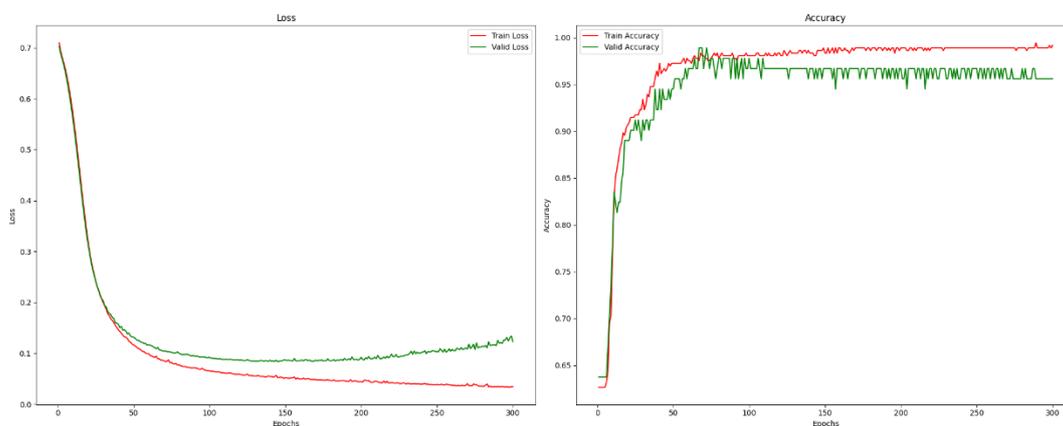
Gambar 5. Grafik Fitur Diagnosis

percobaan dari 1 hidden LAYER, selanjutnya untuk epoch yang digunakan setiap klasifikasi menggunakan 250 epoch, 300 epoch, 350 epoch, 400 epoch, 450 epoch dan 500 epoch,

Tabel 10. Hasil Akurasi

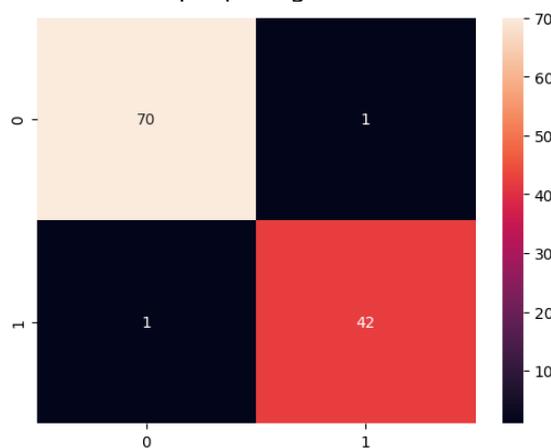
Epoch	Akurasi
250	96,80
300	98,24
350	95,61
400	95,60
450	93,85
500	93,80

Hasil penelitian yang telah dilakukan menggunakan algoritma artificial neural network (ann) dengan menggunakan epoch yang berbeda, hasil akurasi terendah pada sebesar 93,85% pada epoch 450, kemudian hasil akurasi tertinggi terdapat pada epoch 300, dengan hasil akurasi terbesar 98,24%. Sedangkan dari hasil akurasi tertinggi terdapat gambar plot loss dan akurasi dari hasil training dan testing yang telah dilakukan. Gambar plot loss dan plot akurasi dapat dilihat pada gambar 8.



Gambar 8. Grafik Loss dan Akurasi

Pada gambar .. loss dan akurasi terdapat pada sisi kiri, merupakan grafik loss dari hasil model artificial neural network (ann), dengan menghasilkan grafik yang cukup stabil, grafik warna hijau merupakan data testing sedangkan pada grafik berwarna merah merupakan grafik data training. Kemudian terdapat grafik akurasi dengan kondisi cukup stabil dan baik, yang bersifat konsiten. Selanjutnya setelah dari hasil grafik loss dan akurasi terdapat matrik hasil data testing yang telah dilakukan pengujian dengan model algoritma artificial neural network (ann). Confusion matrik hasil pengujian tersebut terdapat pada gambar 9.



Gambar 9. Confusion Matrix Pengujian

Pengujian yang telah dilakukan dengan menggunakan data teting sebanyak 114 data testing yang dilakukan prediksi dengan masing-masing data, kelas 0 sebanyak 71 pasien kemudian kelas 1 sebanyak 43 pasien. Hasil dari confusion matrix yang didapatkan, menghasilkan data yang berhasil di prediksi dengan kelas 0 sebesar 70, sedangkan yang salah prediksi 1 pasien. Selanjutnya data yang di prediksi benar dengan kelas 1 sebesar 42 pasien, kemudian data di prediksi salah hanya 1 orang.

4. CONCLUSION

Penelitian dilakukan dengan pada klasifikasi penyakit breast cancer menggunakan model artificial neural network (ANN), dengan tahapan persiapan dataset. Selanjutnya dilakukan tahapan preprocessing dengan, kemudian dilakukan pemodelan algoritma ANN menggunakan epoch sebesar 250, 300, 350, 400, 450 dan 500 epoch dengan menggunakan satu hidden layer. Kemudian dengan menggunakan split data sebesar 80 : 20 dengan 42 random state, menghasilkan akurasi terbaik sebesar **98,24%** dengan menggunakan fungsi aktivasi ReLU dan fungsi keluaran sigmoid.

5. REFERENCES

- A'yunan, Y. A. D. K., Indahyanti, U., & Busono, S. (2023). Implementasi Data Mining dalam Klasifikasi Diagnosa Kanker Payudara menggunakan Algoritma Logistic Regression. *Jurnal TEKINKOM*, 6(2), 400–407. <https://doi.org/10.37600/tekinkom.v6i2.948>
- Achmad, A. D. (2022). Klasifikasi Breast Cancer Menggunakan Metode Logistic Regression. *JTRISTE*, 9(1), 143–148.
- Anggriani, B., Sitorus, R. J., Flora, R., & Octariyana, O. (2023). Perempuan Dan Penyakit Keganasan (Kanker Payudara Dan Kanker Serviks). *Electronic Journal Scientific of Environmental Health And Disease*, 3(2), 131–142. <https://doi.org/10.22437/esehad.v3i2.27654>
- Efsa, H. A. P. A.-S., Jondri, J., & Lhaksmana, K. M. (2021). Prediksi Retweet Menggunakan Fitur Berbasis Pengguna Dan Fitur Berbasis Konten Dengan Metode Klasifikasi Ann. *eProceedings of Engineering*, 8(5).
- Ericha Apriliyani, & Salim, Y. (2022). Analisis performa metode klasifikasi Naïve Bayes Classifier pada Unbalanced Dataset. *Indonesian Journal of Data and Science*, 3(2), 47–54. <https://doi.org/10.56705/ijodas.v3i2.45>
- Hananti, H., & Sari, K. (2021). Perbandingan Metode Support Vector Machine (SVM) dan Artificial Neural Network (ANN) pada Klasifikasi Gizi Balita. *Seminar Nasional Official Statistics, 2021(1)*, 1036–1043. <https://doi.org/10.34123/semnasoffstat.v2021i1.1014>
- Hidayat, R., Kartini, D., Mazdadi, M. I., Budiman, I., & Ramadhani, R. (2023). Implementasi Seleksi Fitur Binary Particle Swarm Optimization pada Algoritma K-NN untuk Klasifikasi Kanker Payudara. *Jurnal Sistem dan Teknologi Informasi (JustIN)*, 11(1), 135. <https://doi.org/10.26418/justin.v11i1.53608>
- Ihsani, D. A., Arifin, A., & Fatoni, M. H. (2020). Klasifikasi DNA Microarray Menggunakan Principal Component Analysis (PCA) dan Artificial Neural Network (ANN). *Jurnal Teknik ITS (SINTA: 4, IF: 1.1815)*, 9(1), A124–A129.
- Kaggle. (2021). *Breast Cancer Dataset*. Kaggle. 15-Juli--2024
- Kemkes. (2022). *Kanker Payudara Paling Banyak di Indonesia, Kemenkes Targetkan Pemerataan Layanan Kesehatan*. <https://sehatnegeriku.kemkes.go.id/baca/umum/20220202/1639254/kanker-payudara-paling-banyak-di-indonesia-kemenkes-targetkan-pemerataan-layanan-kesehatan/>
- Kraugusteeliana, K., Muis, S., Nugroho, F., Karim, A., & Siagian, Y. (2023). Data Mining Klasifikasi Breast Cancer Menerapkan Algoritma Gradient Boosted Trees. *JURNAL MEDIA INFORMATIKA BUDIDARMA Volume 7, Nomor 2, April 2023, Page 881-890*, 7(April), 881–890. <https://doi.org/10.30865/mib.v7i2.6095>
- Meilani, N., & Nurdiawan, O. (2023). Data Mining untuk Klasifikasi Penderita Kanker Payudara Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Wahana Informatika*, 2(1), 177–187.
- Mubarog, I., Setyanto, A., & Sismoro, H. (2021). Sistem klasifikasi pada penyakit breast cancer dengan menggunakan metode Naïve Bayes. *Creative Information Technology Journal*, 6(2), 109–118.
- Muntiari, N. R., & Hanif, K. H. (2022). Klasifikasi penyakit kanker payudara menggunakan perbandingan algoritma machine learning. *Jurnal Ilmu Komputer dan Teknologi*, 3(1), 1–6.
- Mutia, A. C. (2024). *Rencana Program Makan Siang Gratis Mulai 2025, Siapa Target Penerimaannya*. katadata. <https://databoks.katadata.co.id/datapublish/2024/02/27/rencana-program-makan-siang-gratis-mulai-2025-siapa-target-penerimaannya>
- Nurchahyo, R., Fanani, A. Z., Affandy, A., & Aziz, M. I. (2023). Peningkatan Algoritma C4. 5 Berbasis PSO Pada Penyakit Kanker Payudara. *Jurnal Media Informatika Budidarma*, 7(4), 1758–1765. <https://doi.org/10.30865/mib.v7i4.6841>

- Oktavianto, H., & Handri, R. P. (2020). Analisis Klasifikasi Kanker Payudara Menggunakan Algoritma Naive Bayes. *INFORMAL: Informatics Journal*, 4(3), 117. <https://doi.org/10.19184/isj.v4i3.14170>
- Riva, L. S., Febriana, F., Panghurina, M., Maulida, R. A., Najah, B., & Yasiin, A. D. (2022). Perbandingan Algoritma Cnn Dan Ann Dengan Projection Histogram Untuk Klasifikasi Citra Tulisan Tangan Berupa Angka. *Prosiding Seminar Nasional Mahasiswa Bidang Ilmu Komputer dan Aplikasinya*, 3(2), 899–913.
- Sihombing, P. R., & Hendarsin, O. P. (2020). Perbandingan Metode Artificial Neural Network (ANN) dan Support Vector Machine (SVM) untuk Klasifikasi Kinerja Perusahaan Daerah Air Minum (PDAM) di Indonesia. *Jurnal Ilmu Komputer*, 13(1), 9–20.
- Srirahayu, A., & Pribadie, L. S. (2023). Review Paper Data Mining Klasifikasi Data Mining. *Jurnal Ilmiah Informatika Global*, 14(1). <https://doi.org/10.36982/jiig.v14i1.2981>
- Sudarsono, A. (2016). Jaringan Syaraf Tiruan Untuk Memprediksi Laju Pertumbuhan Penduduk Menggunakan Metode Bacpropagation (Studi Kasus Di Kota Bengkulu). *Jurnal Media Infotama*, 12(1).
- Suhartanto, R. S., Dewi, C., & Muflikah, L. (2017). Implementasi Jaringan Syaraf Tiruan Backpropagation untuk Mendiagnosis Penyakit Kulit pada Anak. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 1(7), 555–562.