

---

## PENERAPAN MULTILAYER PERCEPTRON UNTUK IDENTIFIKASI KANKER PAYUDARA

Oleh

Nimas Ratna Sari<sup>1\*</sup>, Yulaikha Mar'atullatifah<sup>2</sup>

<sup>1</sup>Institut Teknologi Sains dan Kesehatan Sugeng Hartono

<sup>2</sup>Perairan, Universitas Satya Wiyata Mandala

Email: <sup>1\*</sup>[nimasratna@sugenghartono.ac.id](mailto:nimasratna@sugenghartono.ac.id), <sup>2</sup>[yulaikhalatifah@sugenghartono.ac.id](mailto:yulaikhalatifah@sugenghartono.ac.id)

---

### Article History:

Received: 18-03-2023

Revised: 28-03-2023

Accepted: 21-04-2023

### Keywords:

Artificial Intelligence,  
Multilayer Perceptron,  
kanker payudara,  
klasifikasi

**Abstract:** *Artificial Intelligence telah banyak diimplementasikan diberbagai bidang. Pada bidang Kesehatan, AI dapat digunakan untuk membantu tenaga Kesehatan dalam melakukan analisis dan memberikan support system. Pada permasalahan kanker payudara merupakan salah satu masalah kesehatan yang terus meningkat setiap tahunnya. Proses identifikasi yang merupakan langkah awal penanganan dapat memakan waktu dan biaya yang cukup terutama untuk daerah yang belum mempunyai fasilitas dan sumber daya yang cukup. Oleh karena itu, penelitian ini bertujuan untuk mendapatkan model multilayer perceptron untuk proses klasifikasi kanker payudara.*

---

## PENDAHULUAN

Kanker payudara menjadi salah satu masalah Kesehatan yang besar di setiap negara dikarenakan jumlah kasus yang terus meningkat setiap tahunnya [1], [2]. Kanker payudara adalah jenis kanker yang berkembang pada sel payudara dan merupakan kanker yang paling umum ditemukan pada wanita, meskipun masih ada peluang untuk tumbuh pada pria. Berbagai faktor risiko dapat berkontribusi pada tumbuhnya sel tersebut seperti: faktor genetik, kelebihan berat badan, kurangnya aktivitas fisik, konsumsi alkohol, dan paparan radiasi [3], [4]. Faktor risiko ini dapat meningkatkan kemungkinan berkembangnya keganasan ini. Namun, untuk mendiagnosis kanker payudara, umumnya dokter akan melakukan beberapa pemeriksaan seperti mammogram, foto rontgen payudara, USG, dan biopsi. Hasil tes ini akan menjadi pertimbangan apakah sel tersebut termasuk pada jenis sel tumor ganas atau jinak.

Interpretasi hasil tes merupakan bagian penting dari proses deteksi kanker payudara, karena membantu menentukan keberadaan dan karakteristik jaringan payudara yang abnormal. Hasil tes biasanya ditafsirkan oleh spesialis, seperti ahli radiologi atau ahli patologi, dengan pelatihan khusus dalam menganalisis hasil tes dan sampel jaringan. Akan tetapi seiring dengan perkembangan teknologi, kecerdasan buatan banyak diimplementasikan untuk sistem pendukung keputusan.

Artificial Neural Network (ANN) atau jaringan saraf tiruan merupakan sebuah cabang ilmu Kecerdasan Buatan yang banyak diaplikasikan untuk berbagai bidang ilmu sebagai sistem pendukung keputusan [5]. Salah satu bidang ilmu yang cukup populer untuk implementasi ANN adalah kesehatan. Hal ini didasari oleh jumlah kebutuhan tenaga kesehatan yang terus meningkat serta kebutuhan untuk dapat menjangkau berbagai tempat

akan tetapi ketersediaannya sangat terbatas. Meskipun belum dapat sepenuhnya menggantikan tenaga kesehatan untuk menangani pasien, akan tetapi ANN diharapkan dapat membantu pada screening dan identifikasi kebutuhan untuk penanganan pasien. Identifikasi awal merupakan langkah penting untuk meminimalisir resiko yang lebih besar pada pasien [6].

## LANDASAN TEORI

Kecerdasan buatan (Artificial Intelligence) (AI) mempunyai arti yang tertuju pada entitas buatan. Ilmu komputer mempunyai banyak cabang akan tetapi beda halnya dengan AI cabang ini berusaha untuk membuat komputer dapat beroperasi layaknya seperti manusia yang bisa belajar dan berpikir [7]. Sistem prediksi, rekomendasi, dan juga otomasi adalah sistem yang populer dalam penerapan AI. Sistem AI ini banyak dikembangkan dan diterapkan diberbagai bidang untuk membantu kerja manusia termasuk pada bidang kesehatan.

Untuk memecahkan masalah klinikal yang kompleks, seringkali para tenaga kesehatan dituntut untuk dapat menganalisis serta mengaplikasikan berbagai macam ilmu secara bersamaan. AI dalam pelayanan kesehatan didesain untuk membantu tenaga kesehatan untuk memanipulasi data sekaligus menerapkan ilmu kesehatan. AI juga dapat digunakan untuk membuat diagnosis, memprediksi risiko penyakit, dan membantu dokter dalam perencanaan tindakan pencegahan secara lebih efektif [8]. Selain itu, AI juga dapat membantu mengelola data pasien dan memberikan perawatan yang lebih personal[9].

### Artificial Neural Networks

Artificial Neural Networks (ANN) atau Jaringan Saraf Tiruan didasarkan pada struktur otak manusia yang bertujuan untuk meniru bagaimana respon otak terhadap pembelajaran. Seperti halnya susunan otak manusia, ANN terdiri dari kumpulan neuron yang saling terhubung dan dapat membawa sinyal aktivasi[10]. Dengan menggunakan arsitektur yang kompleks, ANN dapat memperkirakan fungsi yang sangat kompleks dan non-linear. ANN harus dilatih secara bertahap dengan memproses data pelatihan dan melakukan penyesuaian parameter model berdasarkan beberapa metode pelatihan sampai output model mencapai[7], [10]. Dengan menggunakan model ini, ANN memiliki efektivitas yang tinggi dalam berbagai aplikasi seperti estimasi, klasifikasi, dan pengenalan.

#### 1. Single Layer Perceptron (SLP)

Perceptron yang hanya memiliki satu lapisan variabel bobot yang dapat dilatih dan satu lapisan neuron keluaran. Input langsung terhubung ke keluaran melalui lapisan bobot. Koneksi sinaptik ini membawa bobot menghubungkan setiap input ke output [10]. Neuron keluaran dapat bervariasi dari satu neuron hingga banyak. SLP dapat digunakan untuk masalah klasifikasi pola ketika pola tersebut secara linear dapat dipisahkan, terlepas dari bentuk nonlinieritas yang digunakan. Pemisahan linier membutuhkan pola diklasifikasikan secara terpisah agar batas keputusan berupa hiperbidang. Untuk memecahkan masalah non-linear, lebih baik menggunakan Multilayer Perceptron (MLP).

#### 2. Multilayer Perceptron (MLP)

Multilayer neural networks atau multilayer perceptron (MLP) terdiri dari beberapa lapisan yang mengandung beberapa node. MLP dapat diatur dalam berbagai cara. Network MLP terdiri dari tiga struktur lapisan utama, yaitu lapisan *input*, *hidden*, dan *output* [11].

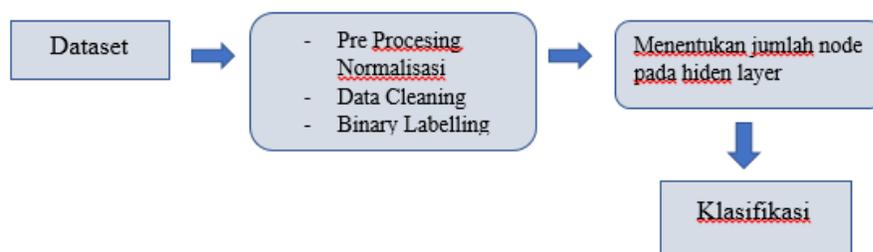
Lapisan masukan adalah lapisan di mana sinyal akan diinputkan. Lapisan ini terdiri dari  $n$  jumlah neuron yang mewakili  $n$  jumlah variabel dari dataset. Lapisan selanjutnya disebut *hidden layer* yang terdiri dari beberapa node tersembunyi. Setiap node tersembunyi tidak menerima masukan langsung dari sinyal masukan. Setiap node pada lapisan masukan akan meneruskan sinyal ke setiap node pada lapisan tersembunyi. Lapisan output terdiri dari  $m$  node output yang didedikasikan untuk jumlah kelas dari dataset. Arsitektur yang lebih kompleks ini juga terkait dengan bangunan non-linear menggunakan fungsi aktivasi sigmoidal atau aktivasi neuron buatan[7].

### Klasifikasi Biner

Klasifikasi biner adalah metode klasifikasi sederhana namun populer. Klasifikasi biner bertujuan untuk mengelompokkan satu set data training menjadi satu set kelas di mana kelas memiliki dua nilai. Sebagai contoh dalam diagnosis medis, sebuah sistem dapat digunakan untuk memprediksi apakah seorang pasien memiliki kanker. Kemudian, sistem akan memeriksa set fitur dari data berlabel untuk belajar atau memetakan data dan sistem akan menjawab 'ya' atau 'tidak'. Meskipun ini adalah tugas yang sederhana, klasifikasi binari dapat menangkap banyak fitur utama dari masalah umum.

### METODOLOGI PENELITIAN

Data yang digunakan merupakan data sekunder mengenai *feature extraction* inti sel yang tersedia pada repositori datase UCI ("UCI Machine Learning Repository: Breast Cancer Wisconsin [12] dan juga Kaggle [13]. Dalam dataset ini, terdapat 32 atribut yang mendeskripsikan karakteristik dari inti sel pada masing-masing kasus. Dataset ini juga memberikan label kasus tersebut dengan M yang berarti Malignant atau kanker, atau B yang berarti Bening atau dapat juga diartikan sebagai tumor. Dataset tersebut kemudian dilakukan pre-processing sebelum digunakan untuk membentuk model MLP untuk deteksi kanker payudara.



Gambar 1 Alur kerja penelitian

### Data Pre-Processing

Sebelum data digunakan untuk proses training dan tes dalam ANN, data tersebut dilakukan preprocessing terlebih dahulu. Hal ini dilakukan untuk mendapatkan hasil dengan akurasi yang lebih baik. Data preprocessing menghasilkan data yang lebih bersih, terorganisasi, dan siap untuk digunakan dalam algoritma AI. Beberapa tahap data preprocessing yang dilakukan pada penelitian ini adalah:

#### Data cleaning

Data mentah yang tersedia mungkin saja mengandung nilai - nilai kosong yang

dapat menjari bias pada proses training dan evaluasi model. Oleh karena itu sebelum melakukan pemrosesan, data yang ada dilakukan analisa akan adanya data kosong, tidak lengkap atau tidak sesuai format terlebih dahulu. Dalam penelitian ini, sistem akan menghapus baris apabila terdapat data yang kosong. Kemudian, dilakukan pemisahan antara label dan attributes yang akan digunakan untuk proses training.

### Normalisasi data

Pada tahap normalisasi data, dilakukan perubahan rentang nilai untuk menghindari perbedaan skala yang sangat besar antara nilai satu dengan yang lain. Hal ini dilakukan untuk mendapatkan data yang lebih akurat dan efisien untuk diproses oleh algoritma. Teknik normalisasi data yang digunakan pada penelitian ini adalah teknik min - max scalling.

Dalam normalisasi min - max scalling ini, nilai attributes data diubah dalam rentang 0 sampai 1. Nilai maximum dari atribut tersebut akan diubah menjadi 1. Sedangkan nilai minimum nya akan diubah menjadi 0. Perubahan nilai ini menggunakan persamaan berikut:

$$x' = \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

Dimana  $x'$  adalah nilai hasil normalisasi,  $x_i$  adalah nilai sebelum normalisasi,  $x_{min}$  adalah nilai  $x$  terendah dari atribut tersebut, dan  $x_{max}$  adalah nilai tertinggi pada atribut tersebut.

**Table 1 Sampel data sebelum normalisasi**

id	radius_me an	texture_me an	perimeter_me an	area_me an	smoothness_me an
842302	17.99	10.38	122.8	1001	0.1184
842517	20.57	17.77	132.9	1326	0.08474
8430090 3	19.69	21.25	130	1203	0.1096
8434830 1	11.42	20.38	77.58	386.1	0.1425
8435840 2	20.29	14.34	135.1	1297	0.1003

**Table 2 Sampel data setelah dilakukan normalisasi min - max scalling**

id	radius_me an	texture_me an	perimeter_me an	area_me an	smoothness_me an
842302	0.521037	0.022658	0.545989	0.363733	0.593753
842517	0.643144	0.272574	0.615783	0.501591	0.28988
8430090 3	0.601496	0.39026	0.595743	0.449417	0.514309
8434830 1	0.21009	0.360839	0.233501	0.102906	0.811321
8435840 2	0.629893	0.156578	0.630986	0.48929	0.430351

### Binary labelling

Untuk melakukan klasifikasi menggunakan MLP, label yang ada pada dataset harus diubah menjadi bentuk angka. Hal ini dikarenakan metode training ANN yang berupa punishment untuk prediksi yang salah menggunakan Mean Square Error. Pada penelitian ini, data dengan label M atau malignant diisyaratkan dengan angka 1. Sedangkan data dengan label B atau benign diisyaratkan dengan angka 0.

### HASIL DAN PEMBAHASAN

Arsitektur Multi-Layer Perceptron (MLP) yang digunakan dalam penelitian ini menggunakan 3 layer, yaitu:

1. Input layer
2. Hidden layer
3. Output layer

Input layer memiliki 30 nodes untuk menampung 30 input variables, sedangkan output layer memiliki 2 nodes karena hasil yang diharapkan adalah 1 atau 0 yang menggambarkan prediksi sel tersebut termasuk Malignant (sel kanker) atau Benign (sel Tumor). Jumlah nodes pada hidden layer adalah variable yang harus dihitung dalam penelitian ini. Peneliti menentukan rentang antara 10 sampai dengan 30 nodes untuk hidden layer.

Untuk menemukan jumlah nodes yang digunakan, dataset yang telah dinormalisasi kemudian dibagi menjadi 2 bagian. Bagian pertama adalah dataset yang akan digunakan untuk proses training dengan proporsi 80% dari total data. Sedangkan bagian kedua adalah dataset yang digunakan untuk proses validasi dengan proporsi 20% dari total dataset.

Dari proses training dan validation diatas, didapatkan arsitektur multilayer perceptron seperti gambar berikut

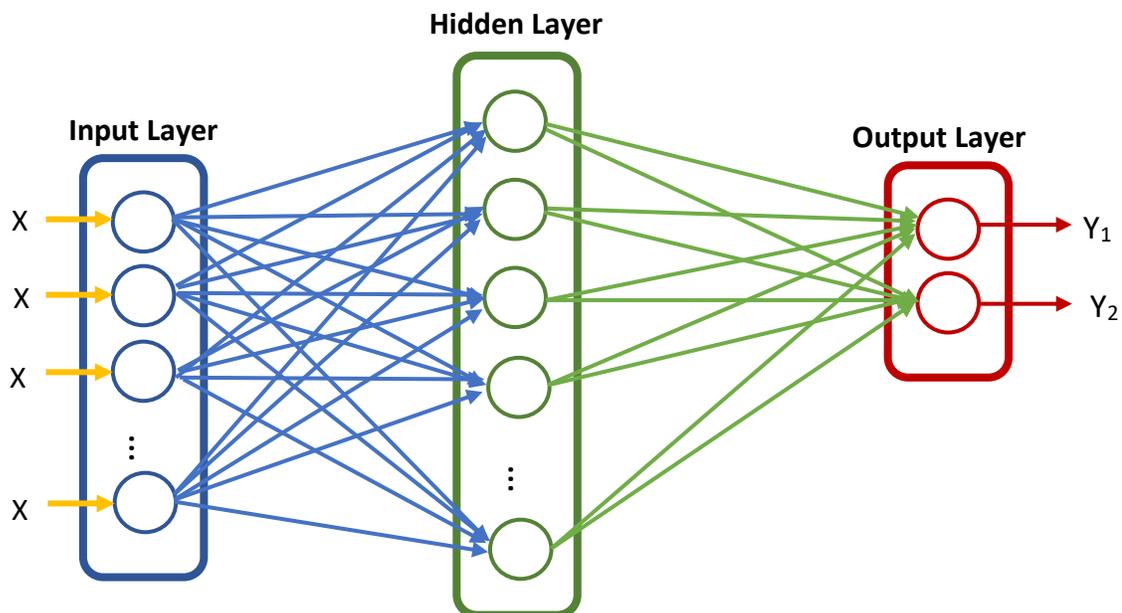


Figure 4. 1 Multilayer perceptron model as binary classifier

Jumlah Nodes	Error Rate
10	0.04386
11	0.04386
12	0.04386
13	0.04386
14	0.04386
15	0.04386
16	0.035088
17	0.04386
18	0.026316
19	0.04386
20	0.035088
21	0.04386
22	0.052632
23	0.035088
24	0.035088
25	0.04386
26	0.035088
27	0.04386
28	0.035088
29	0.035088
30	0.052632



Dalam rentang antara 10 sampai dengan 30 nodes, didapatkan error rate sebagai berikut:

Dari table dan grafik tersebut, dapat disimpulkan bahwa jumlah nodes 18 memiliki error rate yang paling rendah. Sehingga penulis menggunakan 18 nodes pada hidden layer untuk proses data training dan pembentukan model selanjutnya.

Setelah memiliki arsitektur MLP, arsitektur tersebut dapat digunakan untuk

melakukan training backpropagation kembali untuk mendapatkan model untuk menyelesaikan masalah klasifikasi dalam penentuan kanker payudara. Prediksi model yang sesuai dengan data tes, maka nilai dari fungsi indikator akan mengembalikan nilai 1. Sedangkan apabila nilai prediksi tersebut tidak sesuai, maka fungsi akan mengembalikan nilai 0.

Dari perhitungan presentase akurasi diatas, didapatkan akurasi dari model MLP dengan menggunakan backpropagation dan 18 hidden nodes adalah 93.59%.

## KESIMPULAN

Model artificial neural networks menggunakan Multilayer Perceptron dengan 3-layer dapat digunakan untuk memprediksi sel kanker payudara. Data karakteristik sel dapat digunakan sebagai data training untuk mendapatkan model yang sesuai. Training tersebut dilakukan setelah melakukan beberapa langkah preprocessing yang bertujuan untuk mendapatkan hasil yang lebih akurat. Model yang dihasilkan terbukti dapat menghasilkan prediksi dengan nilai akurasi 93,59%.

## PENGAKUAN/ACKNOWLEDGEMENTS

Saya mengucapkan terimakasih kepada berbagai pihak yang telah terlibat dan membantu terlaksananya penelitian ini. Terlebih kepada Institut Teknologi Sains dan Kesehatan Sugeng Hartono khususnya pada Program Studi S1 Informatika yang menajadi tempat saya melaksanakan Tri Dharma Perguruan Tinggi.

## DAFTAR PUSTAKA

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018, doi: 10.3322/caac.21492.
- [2] R. A. da Costa Vieira, G. Biller, G. Uemura, C. A. Ruiz, and M. P. Curado, "Breast cancer screening in developing countries," *Clinics (Sao Paulo)*, vol. 72, no. 4, pp. 244–253, Apr. 2017, doi: 10.6061/clinics/2017(04)09.
- [3] C. Hu *et al.*, "A Population-Based Study of Genes Previously Implicated in Breast Cancer," *New England Journal of Medicine*, vol. 384, no. 5, pp. 440–451, Feb. 2021, doi: 10.1056/NEJMoa2005936.
- [4] Z. Momenimovahed and H. Salehiniya, "Epidemiological characteristics of and risk factors for breast cancer in the world," *Breast Cancer (Dove Med Press)*, vol. 11, pp. 151–164, Apr. 2019, doi: 10.2147/BCTT.S176070.
- [5] R. Dastres and M. Soori, "Artificial Neural Network Systems".
- [6] S. Nardin *et al.*, "Breast Cancer Survivorship, Quality of Life, and Late Toxicities," *Frontiers in Oncology*, vol. 10, 2020, Accessed: Jan. 29, 2023. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fonc.2020.00864>
- [7] C. Zhang and Y. Lu, "Study on artificial intelligence: The state of the art and future prospects," *Journal of Industrial Information Integration*, vol. 23, p. 100224, Sep. 2021, doi: 10.1016/j.jii.2021.100224.
- [8] "History of artificial intelligence in medicine | Elsevier Enhanced Reader." <https://reader.elsevier.com/reader/sd/pii/S0016510720344667?token=ED3FBB7A>

7905683269BACF07AB9784409640781E83F1EEA383B0EFF8B9A945FC81C1A8073A412BC88CE01677B89CC360&originRegion=eu-west-1&originCreation=20230414071938 (accessed Apr. 14, 2023).

- [9] A. Haleem, M. Javaid, and I. H. Khan, "Current status and applications of Artificial Intelligence (AI) in medical field: An overview," *Current Medicine Research and Practice*, vol. 9, no. 6, pp. 231–237, Nov. 2019, doi: 10.1016/j.cmrp.2019.11.005.
- [10] S. Teixeira Zavadzki de Pauli, M. Kleina, and W. H. Bonat, "Comparing Artificial Neural Network Architectures for Brazilian Stock Market Prediction," *Ann. Data. Sci.*, vol. 7, no. 4, pp. 613–628, Dec. 2020, doi: 10.1007/s40745-020-00305-w.
- [11] K.-C. Ke and M.-S. Huang, "Quality Prediction for Injection Molding by Using a Multilayer Perceptron Neural Network," *Polymers*, vol. 12, no. 8, Art. no. 8, Aug. 2020, doi: 10.3390/polym12081812.
- [12] "UCI Machine Learning Repository: Breast Cancer Wisconsin (Diagnostic) Data Set." <https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29> (accessed Feb. 21, 2023).
- [13] Breast Cancer Wisconsin (Diagnostic) Data Set." <https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data> (accessed Feb. 21, 2023).