



Literatur Review: Klasifikasi Penyakit Kanker Payudara Dengan Metode K-Nearest Neighbors

Lukman^{1*}, Sayuti Malik², Wildi Baydowi³, Yoggi Dwi Saputra⁴, Perani Rosyani⁵

^{1,2,3,4,5}Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

Email : ^{1*}lukmanrakenzu22@gmail.com, ²wildibaydowi@gmail.com, ³sayutimalik012@gmail.com, ⁴yoggisaputraaaa@gmail.com, ⁵dosen00837@unpam.ac.id

(* : coresponding author)

Abstrak - Kanker payudara adalah kanker nomor satu pada wanita. Sekitar 2.296.840 kasus baru kanker payudara terjadi pada wanita (World Cancer Fund, 2024). Kanker payudara adalah kondisi ganas yang menyerang jaringan payudara, berasal dari sel-sel kelenjar, saluran kelenjar, dan jaringan penunjang payudara (Lilieki Pratiwi dkk., 2024, hlmn 4). Dalam bidang kesehatan, klasifikasi kanker payudara menjadi sangat penting untuk menentukan jenis perawatan yang tepat bagi pasien. Deteksi kanker sejak dini sangat penting untuk mendeteksi agar penyakit kanker payudara dapat ditangani lebih awal dan meningkatkan angka kesembuhan, namun metode tradisional memiliki keterbatasan dalam hal akurasi. Kesalahan diagnosis dapat menyebabkan pengobatan yang tidak sesuai dan berdampak pada kualitas hidup pasien. Oleh karena itu, diperlukan pendekatan teknologi berbasis machine learning yang lebih akurat untuk mendukung diagnosis kanker payudara. Salah satu metode yang sering digunakan adalah *K-Nearest Neighbors* (KNN). *K-Nearest Neighbors* (KNN) merupakan salah satu metode yang sering digunakan dalam klasifikasi kanker payudara karena kesederhanaan dan efektivitasnya dalam mengelompokkan data berdasarkan kedekatan jarak antar titik untuk mengulas beberapa penelitian terkait penerapan metode KNN pada klasifikasi kanker payudara, dengan fokus pada optimalisasi kinerja algoritma KNN, seperti penyesuaian parameter bobot, integrasi dengan metode lain, dan analisis perbandingan performa dengan algoritma alternatif. Namun, beberapa catatan keterbatasan pada KNN, seperti sensitivitas terhadap data yang tidak seimbang dan efektivitas yang lebih rendah pada data besar. Dengan memperhatikan kelebihan dan kekurangannya, penelitian ini memberikan wawasan bagi pengembangan lebih lanjut dalam penggunaan metode KNN untuk klasifikasi kanker payudara.

Kata Kunci: Kanker Payudara, *K-Nearest Neighbors*, Klasifikasi, Akurasi, *Literatur Review*.

Abstract - Breast cancer is the number one cancer in women. About 2,296,840 new cases of breast cancer occur in women (World Cancer Fund, 2024). Breast cancer is a malignant condition that attacks breast tissue, originating from glandular cells, glandular ducts, and breast supporting tissues (Lilieki Pratiwi et al., 2024, pn 4). In the health sector, the classification of breast cancer is very important to determine the right type of treatment for patients. Early cancer detection is very important to detect so that breast cancer can be treated early and increase the cure rate, but traditional methods have limitations in terms of accuracy. Misdiagnosis can lead to inappropriate treatment and impact the patient's quality of life. Therefore, a more accurate machine learning-based technology approach is needed to support breast cancer diagnosis. One method that is often used is *K-Nearest Neighbors* (KNN). *K-Nearest Neighbors* (KNN) is one of the methods often used in breast cancer classification due to its simplicity and effectiveness in grouping data based on the closeness of the distance between points. To review some research related to the application of the KNN method in breast cancer classification, with a focus on optimizing the performance of the KNN algorithm, such as adjusting weight parameters, integration with other methods, and analyzing performance comparisons with alternative algorithms. However, some note the limitations of KNN, such as sensitivity to unbalanced data and lower effectiveness on large data. Taking into account the advantages and disadvantages, this study provides insights for further development in the use of the KNN method for breast cancer classification.

Keywords: Breast Cancer, *K-Nearest Neighbors*, Classification, Accuracy, *Literature Review*.

1. PENDAHULUAN

Kanker payudara merupakan salah satu jenis kanker yang paling banyak diderita wanita di seluruh dunia dan menjadi penyebab utama kematian akibat kanker. Menurut data dari World Health Organization (WHO), jumlah kasus kanker payudara terus meningkat setiap tahunnya, sehingga diperlukan metode deteksi dan klasifikasi yang efektif dan akurat. Salah satu pendekatan yang populer digunakan dalam klasifikasi kanker payudara adalah metode *K-Nearest Neighbors* (KNN). Algoritma ini bekerja dengan prinsip sederhana, yaitu mengklasifikasikan data berdasarkan



kedekatannya dengan sejumlah data tetangga terdekat, yang ditentukan oleh parameter **K** (Fahrurrozi et al., 2023).

Metode KNN memiliki beberapa kelebihan, antara lain kemudahan implementasi serta efisiensi pada data berukuran kecil hingga menengah (Aprilia et al., 2024). Namun, KNN juga memiliki keterbatasan, seperti sensitivitas terhadap data yang tidak seimbang dan kebutuhan untuk pengaturan parameter yang tepat agar menghasilkan hasil yang optimal. Oleh karena itu, beberapa penelitian telah berfokus untuk meningkatkan performa KNN melalui modifikasi parameter, seperti penambahan bobot (weighted KNN), atau menggabungkannya dengan algoritma lain seperti Decision Tree dan Naive Bayes (Cahyanti et al., 2020).

Penelitian ini bertujuan untuk memberikan tinjauan komprehensif terhadap berbagai studi terkait klasifikasi kanker payudara menggunakan metode KNN. Tinjauan ini mencakup analisis perbandingan hasil akurasi antara metode KNN dan metode alternatif yang sering digunakan dalam klasifikasi kanker, seperti Decision Tree, Naive Bayes, dan algoritma berbasis optimasi lainnya. Dengan adanya tinjauan literatur ini, diharapkan dapat memberikan wawasan mengenai efektivitas dan keterbatasan metode KNN dalam mendeteksi kanker payudara serta potensi pengembangannya lebih lanjut dalam bidang klasifikasi kanker.

2. METODOLOGI PENELITIAN

2.1 Systematic Literature Review (SLR)

Literatur review adalah sebuah metode yang sistematis, eksplisit dan reproduibel untuk melakukan identifikasi, evaluasi dan sintesis terhadap karya-karya hasil penelitian dan hasil pemikiran yang sudah dihasilkan oleh para peneliti dan praktisi (Zulfikar., 2020). Dalam sebuah penelitian yang akan dijalankan, tentunya seorang peneliti harus memiliki wawasan yang luas terkait objek yang akan diteliti. Jika tidak, maka dapat dipastikan dalam presentasi yang besar bahwa penelitian tersebut akan gagal. Sumber-sumber yang diteliti pun tidak boleh sembarangan. (Ahzril et al., 2023)

2.2 K-Nearest Neighbors (KNN)

K-Nearest Neighbors merupakan salah satu metode untuk mengambil keputusan dengan cara mengklasifikasikan berdasarkan data terdekat dalam data nilai (Adistya et al., 2023). Algoritma KNN merupakan algoritma yang banyak digunakan dalam melakukan klasifikasi. KNN merupakan algoritma yang sederhana untuk diimplementasikan tetapi menghasilkan akurasi yang baik (Wahyono, 2020). Algoritma K-Nearest Neighbor (K-NN) mengklasifikasikan didasarkan pada pembelajaran dari data yang telah terklasifikasi sebelumnya. Dalam supervised learning, hasil pertanyaan instance yang baru diklasifikasikan berdasarkan mayoritas kedekatan jarak dari kategori yang ada dalam K-NN.

2.3 Research Question

Research question adalah sebuah pernyataan yang merumuskan pertanyaan utama atau fokus dari penelitian (Fernianti et al., 2020). Sebuah pernyataan keingintahuan tentang suatu topik yang diperoleh secara iterative dan sistematis melalui pengaruh latar belakang keilmuan peneliti, sehingga menginspirasi proses pencarian yang terarah dan merangsang diskusi tentang solusi potensial (Irawan., 2020). untuk membantu dan mengarahkan proses penelitian, serta menentukan batasan dan tujuan dari penelitian. Penelitian ini telah merumuskan sejumlah pertanyaan kunci yang akan dieksplorasi lebih lanjut. :

ID	Pertanyaan Peneliti
001	Bagaimana performa metode <i>K-Nearest Neighbors</i> (KNN) dalam klasifikasi kanker payudara dibandingkan dengan algoritma machine learning lainnya?
002	Faktor-faktor apa saja yang mempengaruhi akurasi model <i>K-Nearest Neighbors</i> (KNN) dalam klasifikasi kanker payudara?



003	Faktor-faktor apa saja yang mempengaruhi akurasi model <i>K-Nearest Neighbors</i> (KNN) dalam klasifikasi kanker payudara?
004	Apa saja keunggulan dan kelemahan <i>K-Nearest Neighbors</i> (KNN) dibandingkan dengan algoritma lain dalam klasifikasi kanker payudara?

2.4 Search Process

Penulis melakukan pencarian metodis pada mesin pencarian google untuk mengumpulkan data yang berkaitan dengan penelitian. Penulis mengacu pada jurnal dan buku sebagai sumber informasi untuk dikaji. Selain mesin pencarian google, penulis memanfaatkan scholar.google.com menggunakan frase “Kanker Payudara”, “*K-Nearest Neighbors*”, “klasifikasi” dan “Literature Review”

2.5 Inclusion and Exclusion Criteria

Penulis menilai artikel yang ditemukan selama fase pencarian, dengan mempertimbangkan kriteria inklusi dan eksklusi yang telah ditentukan sebelumnya. (Septian et al., 2023): data yang digunakan dalam periode 2020 - 2024 data yang digunakan berasal dari scholar.google.com dan mesin pencari google.com

2.6 Data Collection

Data collection mencakup langkah-langkah untuk mengumpulkan literatur yang relevan, memilih jurnal yang memenuhi kriteria inklusi, dan mengekstraksi data yang relevan dari jurnal yang dipilih (Prasetya et al., 2022). Pengumpulan data melibatkan penggunaan data primer dan data sekunder.

3. ANALISA PEMBAHASAN

Bagian ini penulis memaparkan temuan, serta hasil pembahasan dari topik penelitian

No	Nama Peneliti dan Tahun	Metode yang digunakan	Tujuan Penelitannya	Hasil yang didapat
1	Dewi Cahyantia, Alifah Rahmayania, Syafira Ainy Husniara, 2020	Kualitatif Deskriptif	Mengembangkan model klasifikasi kanker payudara dan menganalisis kinerja KNN dibandingkan metode lain	KNN menunjukkan hasil akurasi tinggi untuk klasifikasi kanker payudara dengan keunggulan pada kecepatan komputasi.
2	Ahvan Muharam, Enda Suhadi, Tiki Ramdhani, Imam Azmi, Fitriyani (2022)	Kuantitatif	Mengoptimalkan akurasi klasifikasi kanker payudara dengan menggunakan metode KNN yang dilengkapi optimasi bobot (forward)	KNN dengan optimasi bobot (forward) menghasilkan akurasi 85,30%. Model ini memberikan akurasi yang lebih baik dibanding KNN tanpa optimasi, yang memiliki akurasi 50,86%.
3	Micha Annata	Kuantitatif	Mengklasifikasikan	Nilai akurasi



	Shinami, Saiful Bahri (2023)		penyakit kanker payudara berdasarkan data untuk membedakan kategori kanker jinak dan ganas	klasifikasi sebesar 77%, presisi 76%, dan recall 71%. Confusion Matrix menunjukkan 62 sampel data dalam kelas positif dengan prediksi yang benar.
4	Tresi Aprilia (2024)	Kuantitatif	Menganalisis efektivitas Naive Bayes dan KNN dalam klasifikasi kanker payudara	<i>Naive Bayes</i> menunjukkan akurasi 96%, lebih tinggi daripada KNN yang hanya mencapai 73%
5	Fahrurrozi & Wasilah (2023)	Kuantitatif	Membandingkan dua algoritma data mining (KNN dan Decision Tree C4.5) untuk klasifikasi kanker payudara	<i>Decision Tree C4.5</i> mencapai akurasi 96,49%, lebih tinggi daripada KNN yang mencapai 94,73%

4. KESIMPULAN

Berdasarkan hasil tinjauan literatur, metode *K-Nearest Neighbors* (KNN) terbukti efektif dan banyak digunakan dalam mengklasifikasikan penyakit kanker payudara, yang merupakan salah satu penyebab utama kematian di kalangan wanita. KNN bekerja dengan mengklasifikasikan data baru berdasarkan kedekatan dengan data yang sudah terklasifikasi sebelumnya, sehingga dapat dengan akurat membedakan antara jenis kanker payudara yang jinak dan ganas.

Penelitian menunjukkan bahwa KNN dapat mencapai tingkat akurasi yang tinggi, terutama ketika metode ini dioptimalkan lebih lanjut, misalnya dengan penyesuaian bobot dan pengolahan data yang tepat. Meskipun KNN memiliki keunggulan dalam hal kesederhanaan dan kecepatan, terdapat beberapa keterbatasan, seperti sensitivitas terhadap data yang tidak seimbang dan penurunan efektivitas pada dataset yang besar.

Oleh karena itu, diperlukan penelitian lebih lanjut untuk mengatasi tantangan-tantangan ini, termasuk mengeksplorasi integrasi KNN dengan algoritma lain serta pengaturan parameter yang lebih baik. Secara keseluruhan, KNN memiliki potensi besar dalam meningkatkan akurasi diagnosis kanker payudara, dan hasil penelitian ini memberikan wawasan berharga untuk pengembangan lebih lanjut penggunaan metode ini dalam bidang kesehatan.

REFERENCES

- Fahrurrozi, & Wasilah. (2023). *Deteksi Dini menggunakan Algoritma K-Nearest Neighbour (KNN) dan Decision Tree C-45*. Jurnal Teknika, 17(2), 427–434.
- Aprilia, T. (2024). *Klasifikasi Kanker Payudara Menggunakan Algoritma K-Nearest Neighbor dan Metode Naive Bayes*. SATESI (Jurnal Sains Teknologi dan Sistem Informasi), 4(2), 156–163. DOI: 10.54259/satesi.v4i2.3167.
- Cahyanti, D., Rahmayani, A., & Husniar, S. A. (2020). Analisis performa metode Knn pada Dataset pasien pengidap Kanker Payudara. *Indonesian Journal of Data and Science*, 1(2), 39–43.
- Muharam, A., Suhadi, E., Ramdhani, T., Azmi, I., & Fitriyani. (2022). *Optimasi Klasifikasi Kanker Payudara dengan KNN Berbasis Bobot*. Journal of Data Science, 10(1), 85–94.
- Shinami, M. A., & Bahri, S. (2023). *Klasifikasi Kanker Payudara untuk Menentukan Kategori Jinak dan Ganas*. Jurnal Teknik Informatika, 8(1), 77–83.



JRIIN : Jurnal Riset Informatika dan Inovasi

Volume 2, No. 8 Januari Tahun 2025

ISSN 3025-0919 (media online)

Hal 1451-1455

Teknomo, K. (2006). *What is K-Nearest Neighbor Algoritm ?*

Adisty, A. P., Lutfiyani, N., Tara, P., Rifaldi, R., Adriyan, R., & Rosyani, P. (2023). Klasterisasi Menggunakan Algoritma K-Means Clustering Untuk Memprediksi Kelulusan Mata Kuliah Mahasiswa. *OKTAL: Jurnal Ilmu Komputer dan Science*, 2(8), 2301–2306. ISSN 2828-2442.

Fernianti, A., Studi, P., Anak, P., Dini, U., Keguruan, F., Ilmu, D. A. N., & Surakarta, U. M. (2020). LITERATURE REVIEW : PENATAAN RUANG BELAJAR YANG MENARIK DAN PERKEMBANGAN KOGNITIF ANAK.

Septian, D. B., Rizal, F. N., Taufiqi, A., Salangka, S. I., & Rosyani, P. (2023). Literature review: Sistem pakar untuk mendiagnosis gangguan mental menggunakan metode certainty factor. *OKTAL: Jurnal Ilmu Komputer dan Science*, 2(6), 1655–1666.

Zulfikar S.U, Mayu R. (2020). *Panduan Penulisan Skripsi Literature Review*. Fakultas Kedokteran UIN Maulana Malik Ibrahim Malang.

Wahyono, W. (2020). Peningkatan Kecepatan Algoritma k-NN Untuk Sistem Pengklasifikasian Kendaraan Bermotor. *Techno. Com*, 19(2), 190-196.