

BAB I

PENDAHULUAN

1.1 Later Belakang

Kanker payudara adalah salah satu penyakit kanker yang paling umum di dunia, terutama di kalangan wanita. Menurut data Organisasi Kesehatan Dunia (WHO), setiap tahunnya terjadi jutaan kasus baru kanker payudara di seluruh dunia, dan kanker payudara merupakan penyebab kematian tertinggi kedua akibat kanker di seluruh dunia.[1] Ini menunjukkan bahwa kanker payudara adalah masalah kesehatan global yang serius yang membutuhkan pendekatan diagnosis dan pengobatan yang efektif.

Meskipun telah ada banyak penelitian dalam bidang deteksi dan diagnosis kanker payudara, masih ada tantangan dalam mengidentifikasi kanker payudara secara tepat dan akurat. Sistem klasifikasi tradisional sering kali menghadapi masalah dalam mengenali pola yang rumit dan non-linear yang mungkin terdapat dalam data medis, seperti dataset kanker payudara.[2][3] Oleh karena itu, diperlukan pendekatan yang lebih canggih dan efektif untuk meningkatkan keakuratan diagnosis dan klasifikasi kanker payudara. [4][5]

Penelitian ini penting karena memiliki potensi untuk meningkatkan diagnosis dan klasifikasi kanker payudara secara signifikan. Dengan membandingkan kinerja ANN dan KNN dalam mengklasifikasikan dataset Breast Cancer, kita dapat mengevaluasi kelebihan dan kelemahan masing-masing pendekatan dan

menentukan pendekatan mana yang lebih efektif dalam mengatasi tantangan klasifikasi kanker payudara.[6]

ANN dan KNN dipilih karena keduanya merupakan pendekatan yang umum digunakan dalam klasifikasi data medis. ANN mampu menangani pola-pola yang kompleks dan non-linear dalam data, sedangkan KNN sederhana namun dapat memberikan hasil yang baik dalam kasus-kasus dengan pola yang jelas.[3]

Dengan merancang penelitian untuk membandingkan klasifikasi kanker payudara menggunakan pendekatan ANN dan KNN, diharapkan penelitian ini dapat memberikan wawasan yang berharga tentang kinerja empat algoritma tersebut dalam konteks klinis.[7][8] Hal ini dapat mengarah pada pengembangan model klasifikasi yang lebih akurat dan dapat diandalkan untuk mendukung diagnosis dan pengobatan kanker payudara secara efektif.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang diangkat, maka dibuat rumusan-rumusan masalah yang akan diselesaikan dalam penelitian ini, antara lain :

- a. Bagaimana perbandingan kinerja algoritma Artificial Neural Network (ANN), dan k-Nearest Neighbors (KNN) dalam mengklasifikasikan kanker payudara berdasarkan dataset, serta algoritma mana yang memberikan akurasi tertinggi?
- b. Apakah terdapat perbedaan signifikan dalam kemampuan algoritma Artificial Neural Network (ANN), dan k-Nearest Neighbors (KNN) dalam

mengklasifikasikan kanker payudara berdasarkan atribut-atribut yang ada dalam dataset Breast Cancer?

- c. Bagaimana pengaruh penggunaan algoritma Artificial Neural Network (ANN), dan k-Nearest Neighbors (KNN) terhadap akurasi klasifikasi kanker payudara, serta faktor-faktor apa yang mempengaruhi kinerja masing-masing algoritma dalam konteks dataset Breast Cancer?

1.3 Batasan Masalah

Untuk membatasi pembahasan agar terfokus pada aspek yang diangkat sebagai pembahasan utama dalam penelitian ini, maka perlu dibuat adanya batasan-batasan variabel, antara lain :

- a. Penelitian ini terbatas pada penggunaan dataset Breast Cancer yang tersedia di <https://www.kaggle.com/datasets/yasserh/breast-cancer-dataset>.
- b. Dataset yang digunakan harus diverifikasi keakuratannya dan sesuai dengan standar yang diterima secara ilmiah.
- c. Penelitian ini akan membatasi diri pada atribut-atribut yang tersedia dalam dataset Breast Cancer.
- d. Hanya atribut-atribut yang relevan dan dapat digunakan untuk klasifikasi kanker payudara yang akan dipertimbangkan.

- e. Penelitian ini akan membatasi penggunaan algoritma Artificial Neural Network (ANN) dan k-Nearest Neighbors (KNN) untuk klasifikasi kanker payudara.
- f. Evaluasi kinerja algoritma akan didasarkan pada metrik-metrik standar seperti akurasi, presisi, recall, dan F1-score.
- g. Penelitian ini akan memusatkan perhatian pada perbandingan kinerja kedua algoritma dalam konteks klasifikasi kanker payudara menggunakan dataset Breast Cancer.

1.4 Tujuan Penelitian

Berdasarkan latar belakang yang diangkat, maka dibuat tujuan yang akan diselesaikan dalam penelitian ini, antara lain :

- a. Menentukan akurasi masing-masing algoritma dalam mengklasifikasikan kanker payudara berdasarkan dataset Breast Cancer.
- b. Membandingkan performa kinerja keempat algoritma dalam mengklasifikasikan kanker payudara dengan memperhatikan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score.
- c. Mengidentifikasi kelebihan dan kelemahan masing-masing algoritma dalam konteks klasifikasi kanker payudara, termasuk kecenderungan untuk overfitting atau underfitting.

1.5 Manfaat Penelitian

Berdasarkan latar belakang yang diangkat, maka dibuat manfaat yang akan diselesaikan dalam penelitian ini, antara lain :

- a. Penelitian ini dapat membantu meningkatkan diagnosis dini kanker payudara dengan mengidentifikasi algoritma machine learning yang paling efektif dalam mengklasifikasikan data medis.
- b. Hasil dari penelitian ini dapat digunakan sebagai dasar untuk pengembangan model prediktif yang lebih akurat dalam klasifikasi kanker payudara.
- c. Penelitian ini dapat memberikan kontribusi yang berharga terhadap penelitian dan literatur medis dalam bidang klasifikasi kanker payudara menggunakan machine learning.

1.6 Sistematika Penulisan

- a. Bab I Pendahuluan :

Pada Bab I Pendahuluan, akan diuraikan tentang latar belakang, rumusan masalah, ruang lingkup, tujuan, manfaat penelitian, dan sistematika penulisan.

- b. Bab II Landasan Teori :

Pada Bab II atau pada Tinjauan Pustaka, akan diuraikan tentang teori, atau pendekatan teori, proposisi dan konsep yang relevan untuk digunakan dalam menyelesaikan masalah yang telah dirumuskan, untuk mencapai tujuan yang telah ditetapkan.

- c. Bab III Metode Penelitian :

Pada Bab III atau pada Metode Penelitian, akan diuraikan tentang rancangan dan penelitian, lokasi dan objek penelitian, sumber data, serta responden penelitian.