

BAB I PENDAHULUAN

1.1 Latar Belakang

Kanker payudara menduduki posisi pertama sebagai penyebab kematian tertinggi pada wanita akibat kanker di Indonesia. Angka kematian akibat kanker payudara cukup tinggi karena banyak pasien datang dengan kondisi terlambat. Menemukan kanker payudara secara dini bukanlah suatu faktor kebetulan atau nasib, melainkan adalah tanggung jawab dari para wanita dan dokter[PUR-11]. Banyak penelitian membuktikan bahwa deteksi dini kanker payudara dapat menyelamatkan ribuan bahkan jutaan nyawa setiap tahunnya. Ada berbagai metode yang dapat digunakan untuk deteksi dini kanker payudara salah satunya adalah mamografi[MAG-12]. Mamografi merupakan metode yang paling efektif untuk *screening* kanker payudara. Mamografi berarti bahwa payudara di rontgen oleh sinar-X khusus[ELS-07].

Tim dokter menggunakan *computer-aided-diagnosis* (CAD) untuk mendiagnosa tingkat keganasan kanker payudara guna membantu melakukan keputusan operasi. CAD merupakan suatu prosedur dalam bidang kesehatan dimana alat ini nantinya akan membantu dokter dalam melakukan interpretasi suatu gambar medis. Dari interpretasi hasil *screening* mamografi, akan diperoleh hasil yang menyatakan apakah kanker yang diderita pasien termasuk jinak atau ganas[ELS-07]. Salah satu hasil *screening* mamografi penderita kanker payudara adalah *Mammographic mass data set* yang dimiliki oleh Prof. Dr. Rodrigo Schulz-Wendtland, seorang ahli *Radiology* dari Universitas Erlangen-Nuremberg dan Matthias Elter, seorang ahli *image processing* dari Erlangen. *Mammographic mass data set* merupakan data set yang mendeskripsikan tingkat keganasan penyakit kanker payudara dan merupakan interpretasi gambaran mamografi berdasarkan atribut BI-rads dan umur pasien[ELT-07]. Penelitian terkait data mamografi ini pernah dilakukan oleh Sahar dan Alaa menggunakan metode data mining Decision Tree (DT), Artificial Neural Network (ANN), dan Support Vector Machine (SVM) untuk memprediksi keparahan kanker payudara.

Dari hasil pengujian didapatkan akurasi tertinggi 81,25% dengan metode Support Vector Machine (SVM)[MOK-13].

Dalam interpretasi sebuah mamografi terkadang dokter mengalami kesulitan dalam menentukan keparahan (jinak dan ganas) sehingga dapat menyebabkan sekitar 70% biopsi yang tidak perlu dilakukan karena kanker itu bersifat jinak[MOK-13]. Biopsi seharusnya dilakukan apabila kanker bersifat ganas[ELS-07]. Oleh karena itu diperlukan aturan yang dapat digunakan untuk mengelompokkan pasien ke salah satu kelompok jinak yang tidak memiliki kanker payudara atau kelompok ganas yang memiliki bukti kuat memiliki kanker payudara sehingga dapat mengurangi biopsi yang tidak perlu.

Suatu aturan dapat dibangkitkan secara otomatis melalui suatu algoritma tertentu berdasarkan data yang disediakan pakar[ARA-10]. Aturan *fuzzy* dapat diekstraksi dengan menggunakan beberapa teknik seperti algoritma genetika, jaringan saraf tiruan, dan *clustering*. Salah satu algoritma *clustering* adalah *Fuzzy C-Means* (FCM) yang merupakan adaptasi dari algoritma *k-means* dengan fungsi keanggotaan halus. FCM memungkinkan suatu titik data menjadi bagian untuk semua pusat[ABA-03]. Oleh karena itu, FCM dapat digunakan untuk membangkitkan aturan *fuzzy* dari sekumpulan data. Aturan yang dihasilkan oleh FCM kemudian digunakan untuk penentuan tingkat risiko penyakit kanker payudara menggunakan sistem inferensi *fuzzy*. Sistem inferensi *fuzzy* yang digunakan adalah sugeno orde satu. Pemilihan metode sugeno orde satu dikarenakan konsekuen(*output*) berupa kumpulan konstanta[KUS-10:46]. Sedangkan untuk menghitung nilai derajat keanggotaan pada masing – masing aturan dapat menggunakan fungsi *gauss* dengan bantuan nilai pusat *cluster* dan standar deviasi[KUS-10:152].

Penelitian terkait pembangkitan aturan secara otomatis pernah dilakukan oleh Resti Ludviani. Pada penelitiannya Resti Ludviani menggunakan metode *fuzzy c-means* sebagai media pembelajaran dalam membangkitkan aturan. Aturan yang dihasilkan oleh FCM kemudian digunakan untuk diagnosa penyakit jantung koroner menggunakan sistem inferensi *fuzzy* sugeno orde-satu. Dari hasil penelitian didapatkan akurasi tertinggi 50% [LUD-11].

Berdasarkan latar belakang yang telah diuraikan, maka akan diimplementasikan algoritma *fuzzy c-means clustering* untuk pembangkitan aturan *fuzzy* pada pengelompokan tingkat risiko penyakit kanker payudara dengan sistem inferensi *fuzzy* yang digunakan adalah metode sugeno orde-satu. Oleh karena itu skripsi ini diberi judul “IMPLEMENTASI ALGORITMA FUZZY C-MEANS CLUSTERING UNTUK PEMBANGKITAN ATURAN FUZZY PADA PENGELOMPOKAN TINGKAT RISIKO PENYAKIT KANKER PAYUDARA”.

1.2 Rumusan Masalah

Adapun permasalahan yang dapat dirumuskan dari latar belakang adalah sebagai berikut :

- 1) Bagaimana menerapkan algoritma *fuzzy c-means* (FCM) untuk membangkitkan aturan *fuzzy* pada pengelompokan tingkat risiko penyakit kanker payudara.
- 2) Berapa jumlah *cluster* ideal yang terbentuk berdasarkan perhitungan nilai batasan varian.
- 3) Berapa akurasi hasil sistem inferensi *fuzzy* sugeno berdasarkan pembangkitan aturan *fuzzy c-means* (FCM) untuk pengujian data tingkat risiko kanker payudara.

1.3 Batasan Masalah

Berdasarkan rumusan masalah, penulis perlu memberikan batasan masalah sebagai berikut :

- 1) Data yang digunakan pada skripsi ini adalah data mamografi (*Mammographic mass data set*) yang diambil dari UCI Machine Learning Repository <http://archive.ics.uci.edu/ml/datasets/Mammographic+Mass>.
- 2) Metode inferensi yang digunakan adalah model Sugeno orde satu dengan aturan yang dibangkitkan oleh algoritma *fuzzy c-means clustering*.
- 3) Data bermissing value tidak disertakan dalam penelitian ini.
- 4) Indeks kekaburan *fuzzy* (w) yang digunakan dalam pelatihan menggunakan *fuzzy c-means* ini adalah 2.

1.4 Tujuan

Tujuan dari penelitian ini, antara lain:

- 1) Mengimplementasikan algoritma *fuzzy c-means* (FCM) untuk membangkitkan aturan *fuzzy* pada pengelompokan tingkat risiko penyakit kanker payudara berdasarkan data mamografi.
- 2) Menghitung jumlah *cluster* ideal yang terbentuk berdasarkan perhitungan nilai batasan varian.
- 3) Menghitung akurasi hasil sistem inferensi *fuzzy* sugeno berdasarkan pembangkitan aturan *fuzzy c-means* (FCM) untuk pengujian data tingkat risiko kanker payudara.

1.5 Manfaat

Manfaat yang diharapkan dari penelitian ini adalah membantu pihak medis mengelompokkan tingkat risiko penyakit kanker payudara pasien (jinak dan ganas) agar dapat melakukan penanganan medis secara cepat dan tepat. Dengan penanganan yang tepat dan cepat diharapkan dapat mengurangi angka kematian pasien penderita kanker payudara. Manfaat lain yang diharapkan adalah untuk mengurangi tindakan biopsi yang tidak perlu.

1.6 Sistematika Penulisan

Pembuatan hasil penelitian yang didokumentasikan dalam bentuk skripsi ini berdasarkan sistematika penulisan sebagai berikut:

1) **BAB I PENDAHULUAN**

Bab pendahuluan ini memuat latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat, dan sistematika penulisan dari penelitian ini.

2) **BAB II TINJAUAN PUSTAKA**

Bab tinjauan pustaka ini berisi kajian pustaka dan pembahasan tentang teori dasar yang berkaitan dengan objek penelitian. Kajian pustaka adalah membahas penelitian yang telah ada dan yang diusulkan. Dasar teori membahas teori yang diperlukan untuk menyusun penelitian yang diusulkan.

3) **BAB III METODOLOGI**

Pada bab ini dijelaskan langkah-langkah yang akan dilakukan dalam pengerjaan penelitian.

4) **BAB IV PERANCANGAN**

Bab ini membahas tentang proses perancangan sistem yang dibangun, meliputi perancangan sistem, perhitungan manual, perancangan *database*, perancangan antarmuka untuk pengelompokan risiko penyakit kanker payudara pada data mamografi.

5) **BAB IV IMPLEMENTASI**

Bab ini berisi spesifikasi perangkat keras dan perangkat lunak, implementasi algoritma *fuzzy c-means clustering* (FCM) dan implementasi antarmuka untuk pembangkitan aturan *fuzzy* pada pengelompokan penyakit kanker payudara pada data mamografi.

6) **BAB V PENGUJIAN DAN ANALISIS**

Bab ini membahas tentang pengujian sistem dan memastikan bahwa program sudah ideal sesuai perancangan dan disertai analisis.

7) **BAB VI PENUTUP**

Bab terakhir sebagai penutup dimana didalamnya berisikan kesimpulan dan saran. Kesimpulan didasarkan atas pengujian dan analisis yang dilakukan dalam proses penelitian. Saran berisi rekomendasi apa saja yang perlu dikembangkan kedepannya.

