

# **BAB I**

## **PENDAHULUAN**

### **A. Latar Belakang**

Kanker payudara adalah tumor ganas yang berasal dari sel kelenjar, saluran kelenjar dan jaringan penunjang payudara, tidak termasuk kulit payudara (Departemen Kesehatan, 2009). *World Health Organization* (2016) menyebutkan kanker payudara adalah kanker paling umum terjadi pada wanita baik di negara maju dan berkembang. Diperkirakan bahwa di seluruh dunia lebih dari 508.000 wanita meninggal pada tahun 2011 karena kanker payudara. Meskipun kanker payudara dianggap penyakit dari negara maju, hampir 50% dari kasus kanker payudara dan 58% kematian terjadi di negara-negara kurang berkembang. Seperti wanita, pria memiliki jaringan payudara, meskipun dalam jumlah yang lebih kecil. Ini berarti bahwa pria juga dapat terkena kanker payudara, meskipun tidak banyak. Risiko seorang pria didiagnosa menderita kanker payudara sebelum usia 75 tahun adalah satu dari 1258 orang. Sedangkan, risiko seorang wanita didiagnosa menderita kanker payudara sebelum usia 85 tahun adalah satu dari delapan orang (*Breast Cancer Network Australia*, 2014).

Penyebab pasti kanker payudara tidak diketahui. Meskipun demikian, riset mengidentifikasi sejumlah faktor yang dapat meningkatkan risiko pada individu tertentu, yang meliputi: keluarga yang memiliki riwayat penyakit serupa, usia yang makin bertambah, tidak memiliki anak, kehamilan pertama pada usia di atas 30 tahun, periode menstruasi yang lebih lama (menstruasi pertama lebih awal

atau menopause lebih lambat), faktor hormonal (baik estrogen maupun androgen) (Moh. Yamin Darsyah, 2013: 15).

Kanker payudara adalah kanker yang umumnya terjadi pada wanita, baik di negara maju maupun berkembang. Insiden kanker payudara meningkat di negara berkembang karena meningkatnya harapan hidup, meningkatnya urbanisasi dan adopsi gaya hidup Barat. Meskipun beberapa pengurangan risiko dapat dicapai dengan pencegahan, strategi ini tidak dapat menghilangkan sebagian besar kanker payudara yang berkembang di negara-negara berpenghasilan rendah dan menengah di mana kanker payudara didiagnosis pada tahap yang sangat terlambat (*World Health Organization*, 2016). Oleh karena itu, pemeriksaan deteksi dini perlu dilakukan sehingga kanker payudara dapat disembuhkan.

Pemeriksaan awal dapat dilakukan dengan memeriksa payudara sendiri (SADARI) dan pemeriksaan klinis yang dilakukan oleh tenaga medis terlatih. Bila dibutuhkan, akan dilakukan tes untuk mendiagnosa kondisi payudara. Menurut *American Cancer Society* (2016), tes pencitraan menggunakan x-ray, medan magnet, gelombang suara, atau zat radioaktif dapat membuat gambar dari bagian dalam tubuh. Tes pencitraan dapat dilakukan untuk sejumlah alasan, termasuk untuk membantu mengetahui area mencurigakan yang memungkinkan menjadi kanker, mengetahui seberapa jauh kanker mungkin telah menyebar, dan untuk membantu menentukan apakah pengobatan bekerja. Tes pencitraan ini yaitu *Mammogram*, *Ultrasound*, dan *Magneting Resonance Imaging* (MRI). Selain tes pencitraan tersebut *National Breast Cancer Foundation* (2015) menyebutkan terdapat tes lain yang dapat mendiagnosa kondisi payudara yaitu biopsi payudara

yang terbagi menjadi tiga jenis, *Fine-needle aspiration* (FNA), *Core-needle biopsy*, dan *Surgical biopsy*. Dari sejumlah tes tersebut, skrining payudara dengan mamografi adalah satu-satunya metode skrining yang telah terbukti efektif (*World Health Organization*, 2016).

Beberapa penelitian telah dikembangkan untuk membantu mendiagnosa kanker payudara. Pada tahun 2012, Saritas Ismail melakukan penelitian untuk mendiagnosis kanker payudara menggunakan *Artificial Neural Network*. Penelitian lain dilakukan oleh Njoroge, Waititu, dan Wanjoya (2014), mereka menggunakan model *Artificial Neural Network* untuk memprediksi tingkat keganasan kanker payudara dan menggunakan regresi logistik untuk menguji variabel independen mana yang signifikan. Pada tahun yang sama (2014), Singh, Sushmitha, Harini, dan Surabhi melakukan penelitian menggunakan *Artificial Neural Network* dengan teknik otomatis sebagai pengambil keputusan dalam mendiagnosis kanker payudara.

Para peneliti secara berkelanjutan terus meningkatkan hasil diagnosis kanker payudara dengan berbagai macam metode. *Radial Basis Function* (RBF) merupakan salah satu variasi metode *Artificial Neural Network* (ANN) yang dapat digunakan untuk mendiagnosa kanker payudara. *Radial Basis Function* adalah sebuah *Artificial Neural Network* yang menggunakan fungsi radial basis sebagai fungsi aktivasi. Arsitektur RBF adalah jaringan tiga lapisan yaitu lapisan input, lapisan tersembunyi, dan lapisan output. (Seenivasagam & Arumugadevi).

Model *Radial Basis Function Neural Network* (RBFNN) memiliki beberapa keunggulan, RBFNN hanya memiliki satu lapisan tersembunyi.

Daripada menggunakan fungsi *sigmoid* dalam fungsi aktivasinya seperti pada metode *backpropagation*, lapisan tersembunyi pada RBF menggunakan fungsi *Gaussian* atau fungsi basis kernel lainnya. Ciri khas RBFNN ini menyebabkan proses kerja RBFNN lebih cepat dibandingkan algoritma NN yang lain (Halici, 2004:139).

Penelitian dalam mendiagnosis kanker payudara dengan model *Radial Basis Function* telah dilakukan oleh banyak peneliti. Janghel, Shukla, Tiwan, dan Kala (2010) menerapkan empat model *Neural Network* yaitu algoritma *Backpropagation*, *Radial Basis Function*, Pembelajaran vektor *Quantization*, dan Jaringan Pembelajaran Kompetitif dalam mendiagnosis kanker payudara. Ririn Uktarini, Isye Arieshanti, dan Ahmad Saikhu (2012) mengembangkan perangkat lunak menggunakan metode *Radial Basis Function* dengan *K-Means clustering* dan mengklasifikasikan data penyakit kanker payudara menjadi tumor jinak (*benign*) atau ganas (*malignant*). Data penyakit kanker payudara, diambil dari data *Wisconsin Breast Cancer*.

Pada tahun yang sama (2012) Ali Raad, Ali Kalakech, dan Mohammad Ayache melaporkan hasil penelitiannya mengenai klasifikasi kanker payudara menggunakan pendekatan *Neural Network* berupa *Multi Layer Perceptron* (MLP) dan *Radial Basis Function* (RBF). Penelitiannya mengevaluasi benjolan dan massa payudara berdasarkan hasil tes *Fine Needle Aspiration* (FNA). Kumpulan data terdiri dari sembilan fitur yang mewakili lapisan input ke *Neural Network*. Kemudian, *Neural Network* akan mengklasifikasikan fitur input menjadi dua jenis kanker (jinak dan ganas). Penelitian lain dilakukan oleh Umi dan Puteh (2012)

yang mengklasifikasi kanker payudara pada data *microarray* dengan menggunakan *Radial Basis Function*. RBF pada penelitian ini digunakan sebagai *classifier* yang membedakan antara sel-sel kanker dan yang bukan kanker. Pada tahun 2013, Azar dan El-Said menggunakan tiga metode dalam mendeteksi dan mengklasifikasi kanker payudara, yaitu *Multilayer Perceptron* (MLP), *Radial Basis Function* (RBF), dan *Probabilistic Neural Network* (PNN). Data kanker yang digunakan diambil dari data *Wisconsin Breast Cancer*.

Metode pengelompokan adalah suatu metode untuk memisahkan objek-objek ke dalam beberapa kelompok atau *cluster*, sehingga sifat dari objek dalam suatu kelompok akan mirip dan objek-objek berbeda kelompok akan mempunyai sifat berbeda. Jarak *Euclidean* umumnya digunakan untuk mengukur kemiripan dari objek-objek yang diamati. Semakin kecil nilai jarak *Euclidean*, semakin tinggi tingkat kemiripan, begitu pula sebaliknya, semakin besar nilai jarak *Euclidean* maka semakin rendah tingkat kemiripannya. Setelah ukuran kemiripan ditemukan, maka dapat dilakukan pengelompokan (Brodjol Sutijo, 2008). Metode pengelompokan pada RBFNN, diantaranya K-Means dan Fuzzy C-Means *clustering*. Metode K-Means dan Fuzzy C-Means *clustering* merupakan metode *clustering* yang paling sering digunakan.

Metode K-Means *cluster* atau pengelompokan dengan menggunakan metode K-Means adalah salah satu teknik *clustering* pada pemodelan RBFNN dan metode K-Means adalah salah satu bentuk metode pemetaan pada dirinya sendiri (*Self Organizing Map*) yang juga dikembangkan dalam pemodelan NN. Pada metode K-Means, data dipartisi ke dalam subgroup atau *cluster*, dimana pada

setiap *cluster* mempunyai sifat yang homogen serta antar *cluster* mempunyai ciri yang berbeda. Jika ada  $p$  buah unit pada RFBNN, maka akan ada  $\mu_i, i = 1, 2, 3, \dots, p$  nilai rata-rata atau pusat (Brodjol Sutijo dkk, 2006).

Penelitian mengenai Metode K-Means *clustering* dilakukan oleh Munnoli dan Bapat (2013) yang menggunakan K-Means dan algoritma fuzzy *clustering* untuk *Radial Basis Function Neural Network*. Pembelajaran RBFNN berdasarkan pada estimasi beberapa parameter seperti pusat *cluster*, jarak, bobot antar *neuron* dll. Pada tahun 2014, Fajarani menentukan nilai pusat dan varians dari variabel *input* di setiap *neuron* pada lapisan tersembunyi dengan metode K-Means *cluster* untuk membentuk model RBFNN dalam meramal kasus demam berdarah.

Berbeda dengan K-Means *clustering*, pada Fuzzy C-Means *cluster* atau pengelompokkan dengan menggunakan Fuzzy C-Means (FCM), keberadaan tiap-tiap datanya dalam suatu *cluster* ditentukan oleh derajat keanggotaan. Konsep dasar FCM, pertama kali adalah menentukan pusat *cluster*, yang akan menandai lokasi rata-rata untuk tiap-tiap *cluster*. (Sri Kusumadewi, 2002).

Penelitian mengenai Metode Fuzzy C-Means *clustering* dilakukan oleh Essam Al-Daoud (2010) melakukan penelitian untuk mendeteksi kanker payudara menggunakan metode Fuzzy C-Means *Radial Basis Function Network* yang dimodifikasi. Essam menggunakan tiga aturan yang selanjutnya dikomparasikan dengan metode *Adaptive Neuro Fuzzy Inference System* (ANFIS) yang diterapkan pada kumpulan data *Wisconsin Breast Cancer* yang merupakan hasil diagnosis massa payudara hanya berdasarkan tes *Fine Needle Aspiration* (FNA). Sedangkan di tahun 2013 Indira Muhic meneliti mengenai diagnosis kanker payudara dengan

analisis fuzzy menggunakan Fuzzy C-Means dan *Pattern Recognition* pada diagnosis massa payudara berdasarkan hasil tes *Fine Needle Aspiration*. Dalam penelitiannya, digunakan algoritma Fuzzy C-Means. Contohnya klinik dikelompokkan menjadi dua kelompok, satu dengan kasus jinak dan lainnya dengan kasus ganas. Selanjutnya, input data dibagi dalam data *training* dan data *testing* kemudian hasil dari masing-masing dievaluasi. Pada tahun 2015 Anand, Vinod, dan Rampure menggunakan algoritma Fuzzy C-Means dan *Self Organizing Map* (SOM) dalam membagi citra input dan kemudian mengkategorikan ke dalam citra payudara yang terkena tumor dan citra payudara normal.

Penggunaan metode K-Means dan Fuzzy C-Means *clustering* pada penelitian-penelitian sebelumnya juga menghasilkan nilai akurasi yang tinggi, diantaranya, penelitian yang dilakukan Rocky, Martini, dan Derwin (2012) yang mendiagnosa penyakit jantung dengan metode K-Means *clustering* pada model RBFNN menghasilkan nilai akurasi sebesar 85%. Idni, Endang, dan Tutik (2014) mengelompokkan *gender* berdasarkan nilai maksimum minimum amplitudo suara berbasis Fuzzy C-Means sebagai metode *clustering* menghasilkan akurasi sebesar 85%. Handayani dan Isye (2015), yang mengelompokkan eksudat keras pada penyakit *diabetic retinopathy* yaitu penyakit retina yang disebabkan oleh diabetes melitus dengan menggunakan metode K-Means *clustering* dalam tiga model *Neural Network* yang berbeda yaitu *Support Vector Machine* (SVM), *Multilayer Perceptron* (MLP), and *Radial Basis Function Neural Network* (RBFNN), dengan nilai akurasi masing-masing model secara berturut-turut yaitu 89,29%,

91,07%, dan 85,71%. Tita Suwening A. (2015) yang membandingkan klasifikasi kanker payudara menggunakan metode Fuzzy C-Means *cluster* pada model *Recurrent Neuro Fuzzy* menghasilkan nilai akurasi 95,83% pada *training*.

Hal-hal yang telah dijelaskan di atas dan cara pengelompokan yang berbeda ini melatarbelakangi penulis untuk melakukan penelitian menggunakan metode K-Means *clustering* dan Fuzzy C-Means *clustering* pada model *Radial Basis Function* untuk mengklasifikasi stadium kanker payudara. Penelitian ini bertujuan untuk mengetahui metode *clustering* yang lebih efektif digunakan pada model *Radial Basis Function* untuk mengklasifikasi stadium kanker payudara dengan menggunakan citra *mammogram*. Rancangan dan hasil penelitian dilakukan dengan menggunakan MATLAB R2013a. Penelitian dengan judul “Perbandingan K-Means dan Fuzzy C-Means *Clustering* pada Model *Radial Basis Function Neural Network* (RBFNN) untuk Klasifikasi Stadium Kanker Payudara” diharapkan dapat memberikan manfaat di bidang matematika dan bidang kesehatan.

## **B. Pembatasan Masalah**

Untuk menghindari pembahasan yang terlalu luas, batasan-batasan dalam penelitian ini adalah sebagai berikut.

1. Citra *mammogram* yang diolah dan digunakan sebagai input diperoleh dari *Mamograph Image Analysis Society* (MIAS).
2. Model yang digunakan untuk klasifikasi dalam penelitian ini adalah *Radial Basis Function Neural Network* dengan fungsi aktivasi *Gaussian*.



3. Pembelajaran RBFNN untuk mendapatkan bobot-bobot terbaik dilakukan dengan bantuan MATLAB R2013a.
4. Jarak dan pusat *cluster* pada metode K-Means *clustering* didapatkan dengan bantuan MINITAB
5. Pusat *cluster* pada metode Fuzzy C-Means *clustering* didapatkan dengan bantuan MATLAB R2013a
6. *Output* RBFNN diklasifikasikan menjadi tiga, yaitu normal, tumor (*benign*), dan kanker (*malignant*).

### **C. Rumusan Masalah**

Berdasarkan pada latar belakang maka rumusan masalah dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana proses penentuan stadium kanker payudara dengan menggunakan RBFNN ?
2. Bagaimana perbandingan ketepatan klasifikasi stadium kanker payudara dengan K-Means dan Fuzzy C-Means *clustering* pada model RBFNN ?

### **D. Tujuan Penelitian**

Sesuai dengan pokok permasalahan yang telah dirumuskan, maka tujuan penelitian ini adalah:

1. Mendiskripsikan proses penentuan stadium kanker payudara dengan menggunakan RBFNN.

2. Mendiskripsikan hasil perbandingan akurasi stadium kanker payudara dengan menggunakan K-Means dan Fuzzy C-Means *clustering* pada model RBFNN.

#### **E. Manfaat Penelitian**

Manfaat yang dapat diperoleh dari penelitian ini adalah:

1. Bagi penulis

Bagi penulis sendiri, penulisan skripsi ini dapat menambah pengetahuan dan wawasan tentang aplikasi model RBFNN dan pemrograman MATLAB dalam kehidupan sehari-hari, khususnya dalam bidang kesehatan.

2. Bagi para pembaca

Sebagai salah satu bahan dalam mempelajari model RBFNN dan MATLAB serta diharapkan penelitian ini dapat dijadikan sebagai referensi untuk penelitian selanjutnya.

3. Bagi perpustakaan Universitas Negeri Yogyakarta

Penulisan skripsi ini juga bermanfaat dalam menambah koleksi bahan pustaka yang bermanfaat bagi Universitas Negeri Yogyakarta pada umumnya, dan mahasiswa Fakultas Matematika dan Ilmu Pengetahuan Alam pada khususnya.