

BAB I

PENDAHULUAN

1.1 Latar Belakang

Bakteri *Mycobacterium tuberculosis* merupakan bakteri penyebab penyakit tuberkulosis yang dapat menular (Hardin, 2021). Pada tanggal 2 maret 1882 Robert Koch mengemukakan penemuan dari penyebab panyakit *tuberculosis* (TBC) sejak saat itu telah membuka jalan untuk diagnosa dan pengobatan penyakit ini (Kemenkes RI, 2018). 90% pasien yang terkonfirmasi TBC paru-paru tidak memberikan gejala (*asymptomatic*) (Sekitar et al., 2020). Menurut Kementerian Kesehatan (2020), diperkirakan kasus *tuberculosis* di Indonesia mencapai 845.000, dengan sekitar enam puluh sembilan persen atau atau diperkirakan sebanyak lima ratus empat puluh ribu yang terdeteksi. Nilai kematian akibat *tuberculosis* juga sangat tinggi, dengan tiga belas orang meninggal karena *tuberculosis* setiap jam (Mei et al., 2021). Penularan penyakit tuberkulosis bisa terjadi melalui udara, baik melalui tetesan lendir (*droplet nuclei*) maupun oleh lendir penderita *tuberculosis* yang tertinggal di alat makan pasien, sehingga infeksi melalui udara sangat mudah menular kepada orang lain. (Nomor & Demonstrasi, 2020). Penyakit ini biasanya muncul dengan gejala seperti batuk, demam, nyeri dada, sesak napas dan kelelahan (Bangun et al., 2019). Hal ini membuktikan bahwa penyakit ini memiliki gejala yang sangat beragam, sehingga diperlukan pemeriksaan radiologis untuk mengetahui apakah seorang pasien pernah terkonfirmasi *tuberculosis*. Bakteri tuberkulosis bisa juga menyerang organ lain yang terdapat di dalam tubuh manusia seperti meningen, ginjal, tulang, dan kelenjar getah bening (Rahmaniati & Apriyani, 2013).

Cara untuk mendiagnosis dan menganalisis penyakit yang diderita pasien dengan gambar dua dimensi atau sampel yang disebut dengan citra medis. Terdapat beberapa metode untuk memperoleh citra diantaranya adalah *X-ray*, *MRI* (*Magnetic resonance Imaging*), *USG* (*Ultrasonography*), *CT-Scan* (*Computed Tomography*), *Endoscopy*, dan *Nuclear Medicine* adalah metode yang dipakai untuk memperoleh sebuah citra. (Nafisah et al., 2021).

Kendala yang dihadapi adalah belum adanya label untuk mengetahui tingkat keparahan kerusakan paru-paru yang diderita pasien berdasarkan dari banyaknya data citra *rontgen*. Semakin banyaknya masyarakat yang terindikasi penyakit *tuberculosis* dan sedikitnya informasi mengenai kerusakan paru-paru yang diakibatkan dari bakteri penyebab *tuberculosis* tersebut. Tujuan pelabelan dari data citra rontgen paru-paru tersebut berguna untuk mengetahui letak dari kelas apakah berada pada indeks kanan atau indeks kiri atau indeks lainnya (Dyan; et al., 2018). Apabila diimplementasikan pada penelitian ini akan terbagi pada 3 kelas tingkat lesi aktif, lesi inaktif, dan luluh paru.

Salah satu algoritma yang dipergunakan untuk menggolongkan data atau membuat label dari data tersebut adalah algoritma *clustering K-Medoids*. Algoritma *K-Medoids* memiliki kemiripan dengan algoritma *K-Means*, namun terdapat perbedaan yang cukup penting, nilai tengah dihitung dengan rata-rata (mean) dan perhitungan jarak dihitung dari data pada masing-masing mean, tetapi pada algoritma ini, data akan digunakan sebagai nilai tengah / disebut dengan medoid, dan jumlah baris dihitung dari baris pertama data (Asmiatun, 2019).

Pada penelitian *clustering* tingkat keparahan TBC berdasarkan citra *rontgen* menggunakan algoritma *k-medoids* memiliki tujuan untuk mengelompokkan data citra rontgen paru-paru TBC dan menghasilkan label dari pengelompokan tersebut, agar dapat mengetahui tingkat keparahan kerusakan paru-paru yang diakibatkan oleh penyakit *tuberculosis* (TBC) yang diderita pasien. Alasan penggunaan algoritma *k-medoids* pada penelitian ini adalah karena algoritma *k-medoids* menggunakan sebuah objek sebagai pusat cluster yang mewakili setiap cluster. Algoritma *K-Medoids* dirancang untuk mengatasi kelemahan algoritma *K-Means* yang sensitif terhadap outlier, karena objek dengan nilai yang lebih besar menyimpang secara signifikan dari distribusi data. (Muslihudin et al., 2019). Kelebihan metode *K-Medoids* ialah waktu yang digunakan untuk komputasi sangat cepat, serta metode *K-Medoids* bisa menangani seluruh jenis data matriks. Tujuan metode *K-Medoids* ialah untuk membagi dataset menjadi beberapa kelompok. Metode *K-Medoids* sendiri dikenal menjadi salah satu

metode untuk mencapai akurasi cukup tinggi. Seperti pada penelitian Ni Nyoman yang menghasilkan akurasi 83% dengan menggunakan metode K-Medoids (Asmiatun, 2019).

Dengan tahapan proses menggunakan data rontgen yang diambil dari situs *kaggle* sekitar tahun 2020 – tahun 2022 kemudian menginput data citra untuk melakukan proses *pre-processing*, *Pre-processing* data ialah langkah penting dalam data mining dengan memperbaiki kualitas citra (Yani, 2019). Pada penelitian kali ini *pre-processing* dilakukan dengan metode *Histogram Equalization* (HE), alasan digunakannya metode *Histogram Equalization* adalah Pada penelitian “Analisis Perbandingan *He* Dan *Clahe* Pada *Image Enhancement* Dalam Proses Segmenasi Citra Untuk Deteksi Fertilitas Telur” Metode *Histogram Equalization* (HE) mampu memberikan akurasi deteksi dengan persentase 96%(Saifullah, 2020). Data yang akan di *pre-processing* berguna untuk memastikan kualitas data yang digunakan baik sebelum dilakukan proses analisis data (Yani, 2019), selanjutnya data yang telah di *pre-processing* akan diekstraksi menggunakan ekstraksi fitur tekstur (GLCM) yang memiliki citra berskala keabuan (grayscale) pada penelitian “Evaluasi Ekstraksi Fitur GLCM dan LBP Menggunakan Multikernel SVM untuk Klasifikasi Batik” ekstraksi fitur Gray Level CoOccurrence Matrix (GLCM) memperoleh akurasi tertinggi pada kernel polynomial yaitu sebanyak 100% (Andono & Rachmawanto, 2021). Tujuan dari ekstraksi fitur *GLCM* ialah buat mendapatkan karakteristik berasal dari masing-masing citra. Sesudah berhasil diekstraksi dilakukanlah proses normalisasi data dengan menggunakan normalisasi data Min-Max seperti pada penelitian “Perbandingan Normalisasi Data Untuk Klasifikasi Wine Menggunakan Algoritma K-Nn” menghasilkan rata-rata akurasi sebesar 59,68% (Nasution et al., 2019). Setelah normalisasi selesai dilakukan proses selanjutnya adalah *clustering* dan yang akan terjadi selanjutnya adalah membuat aplikasi berbasis web untuk sistem informasi *cluster* dengan fitur validasi. Penelitian ini diharapkan sangat membantu dalam mengetahui label dari tingkat kerusakan paru-paru yang diderita pasien untuk memberikan informasi tentang jenis kerusakan paru-paru yang terindikasi tbc dan dapat memberikan obat yang sesuai agar tidak semakin

mempercepat kerusakan yang terdapat di paru-paru.

1.2 Identifikasi Masalah

Berdasarkan latar belakang diatas, maka identifikasi masalah pada penelitian ini adalah :

1. Banyak nya data citra rontgen paru-paru yang belum memiliki label untuk mengetahui tingkat keparahan kerusakan paru-paru yang diderita
2. Belum ada nya sistem informasi berbasis *web* dengan fitur validasi untuk mengetahui hasil dari *clustering* mengenai kerusakan paru-paru TBC.

1.3 Rumusan Masalah

Berdasarkan latar belakang diatas, maka rumusan masalah pada penelitian ini adalah :

1. Bagaimana membangun model *clustering* citra *rontgen* paru-paru menggunakan metode *K-Medoids*?
2. Bagaimana cara membangun aplikasi sistem informasi berbasis *web* untuk mengetahui label hasil dari *clustering* tingkat keparahan kerusakan paru-paru serta dapat memvalidasi label yang telah di hasilkan?

1.4 Tujuan Penelitian

Berdasarkan latar belakang serta rumusan masalah diatas, maka ada beberapa tujuan yang ingin dicapai diantaranya yaitu :

1. Untuk membangun model *clustering* dari citra *rontgen* paru-paru dengan memanfaatkan metode *K-Medoids*
2. Untuk membangun sebuah aplikasi berbasis *web* dalam mengetahui hasil label *clustering* dan selanjutnya dapat dilakukan proses validasi

1.5 Batasan Masalah

Agar penelitian ini tidak keluar dari rumusan masalah, maka peneliti membuat batasan masalah penelitian sebagai berikut:

1. Data citra yang digunakan merupakan hasil digitalisasi foto *rontgen* dada

format warna *grayscale*

2. Ekstraksi fitur menggunakan metode GLCM
3. Data yang diambil dari situs *Kaggle* pada tahun 2020 – tahun 2022

1.6 Manfaat Penelitian

Berdasarkan permasalahan dan juga tujuan penelitian, maka pada penelitian ini ada beberapa manfaat yang diharapkan, yaitu:

1. Menghasilkan sistem informasi berbasis web untuk mengetahui label dari tingkat keparahan kerusakan paru-paru yang diderita pasien dan dapat dimanfaatkan oleh tim medis dalam menentukan label kerusakan paru-paru
2. Meningkatkan pengetahuan dan pemahaman tentang membangun model clustering sistem berbasis komputer dan penggunaan Metode *K-Medoids Clustering*.
3. Dapat mengelompokkan data *rontgen* penyakit *tuberculosis* (TBC), sehingga dapat mengetahui tingkat keparahan kerusakan paru-paru yang diderita pasien berdasarkan label yang telah dihasilkan dari proses *clustering*.