

BAB I

PENDAHULUAN

1.1. Latar Belakang

Demensia merupakan jenis penyakit gangguan *neuropsikiatri* otak yang terjadi jika salah satu atau lebih dari sel otak berhenti bekerja sebagian ataupun seluruhnya. Seiring bertambahnya usia sel-sel yang berada di otak akan mengalami penurunan kinerja. Penyakit demensia sudah dikenal luas mendunia dan setiap tahunnya mengalami peningkatan terhadap gejala penyakit seiring bertambahnya usia. Sel-sel otak penderita demensia mengalami kerusakan parah pada fungsi otak. Penyakit ini akan merusak penilaian seseorang secara permanen, dapat merusak pemikiran ingatan serta menyebabkan penurunan kemampuan kognitif. Adapun salah satu penyakit demensia yang paling umum yaitu penyakit *Alzheimer Disease* (AD).

Penyakit *Alzheimer* salah satu penyebab paling utama dari demensia di seluruh dunia[1]. Penyakit ini kebanyakan dimiliki oleh orang yang sudah berumur diatas 65 tahun. Sampai saat ini penyakit *Alzheimer* masih belum ditemukan obatnya, Adapun para ahli medis menemukan ragam pengobatan hanya untuk memperlambat perkembangan kondisi serta meredakan gejala dari penyakit tersebut. Oleh karena itu, diperlukan deteksi untuk mendiagnosis penyakit *Alzheimer* sejak dini untuk memperlambat gejala yang ditimbulkan. Kebanyakan orang yang didiagnosis dengan *Alzheimer's* mengalami kesulitan dalam ingatan jangka pendek, kesulitan dalam berbicara atau menemukan kata yang tepat, kesulitan melakukan tugas-tugas sehari-hari, kesulitan dalam mengatur dan merencanakan aktivitas sehari-hari, perubahan perilaku, dan kemampuan kognitif yang menurun secara bertahap[2]. Menurut laporan *Alzheimer's Disease International* (ADI) setiap 3 tiga detik, seseorang di dunia mengalami demensia. Pada tahun 2020, sudah lebih dari 55 juta orang menderita demensia, dan jumlah ini mengalami kenaikan kurang lebih dua kali lipat setiap 20 tahun, pada tahun 2030 mencapai 78 juta dan 139 juta pada tahun 2050. Mayoritas peningkatan terjadi di negara-negara berpenghasilan rendah, di mana 60% pasien demensia tinggal. Pada tahun 2050, diperkirakan angka ini akan meningkat menjadi 71%. Populasi lanjut

usia tumbuh paling cepat di Cina, India, serta tetangga mereka di Asia Selatan dan Pasifik Barat [3].

Penyakit demensia dapat dideteksi dengan menggunakan pemindaian MRI (*Magnetic Resonance Imaging*). Pemindaian MRI dapat memberikan informasi detail tentang organ-organ yang berada dalam tubuh, termasuk organ otak. MRI menggunakan gelombang elektromagnetik dan medan dalam membuat citra otak secara terperinci untuk membantu mengidentifikasi ukuran otak dan perubahan struktural[4]. Oleh karena itu, MRI seringkali digunakan sebagai diagnosis berbagai penyakit organ tubuh, termasuk penyakit *Alzheimer* . Namun karena kompleksitas penyakit maka dibutuhkan keahlian dan pengalaman khusus untuk membaca MRI dari dokter radiologi. Adapun human error atau kesalahan manusia pada dokter atau radiolog dalam mendiagnosis pemindaian MRI scan penyakit *Alzheimer* 's dapat terjadi. Faktor kesalahan manusia dalam mendiagnosis terjadi karena kurangnya pengalaman ataupun kurangnya pengetahuan mendalam terkait penyakit yang dideteksi. Untuk mengurangi kesalahan dalam mendiagnosis penyakit *Alzheimer* , teknologi Machine learning dapat digunakan sebagai alat bantu dalam interpretasi hasil pemindaian MRI sehingga dapat diketahui sejauh mana AD telah menyebar.

Dalam beberapa tahun terakhir, pembelajaran mesin telah menjadi alat yang berkembang, kuat, dan memungkinkan di bidang medis. Model algoritma pembelajaran mesin juga dapat membantu mengotomatisasi prediksi penyakit *Alzheimer* dengan mengidentifikasi pola tersembunyi pada data menggunakan metode optimal sehingga hasil prediksi dapat menjadi tolak ukur[5]. Menggunakan metode *machine learning* karena metode ini lebih simple. Beberapa peneliti telah melakukan prediksi pada penyakit *Alzheimer* dengan menggunakan berbagai model *machine learning*. *Machine learning* merupakan aplikasi komputer yang diperoleh dengan cara pembelajaran yang berasal dari data sehingga menghasilkan prediksi yang diperoleh. Teknologi *machine learning* yang dapat digunakan dalam mendeteksi penyakit *Alzheimer* yaitu klasifikasi. Klasifikasi penyakit *Alzheimer* dapat menggunakan *dataset* berupa gambar MRI organ otak 2D atau 3D dari penderita penyakit *Alzheimer* . Gambar organ otak yang digunakan yaitu 2D karena

jika kedalam 3D akan ada 3 tampilan yang dapat terlihat diantaranya yaitu *sagittal*, *coronal* dan *axial*. Dari *dataset* citra MRI tersebut akan dibuat sebuah kerangka penelitian yang meliputi *preprocessing data*, *feature extraction*, *feature selection*, *training* dan *validasi* model serta interpretasi hasil.

Berdasarkan penelitian sebelumnya yang dilakukan oleh Minseok Song, dkk, yang berjudul "*Diagnostic Classification and Biomarker Identification of Alzheimer's Disease with Random Forest Algorithm*" *dataset* yang digunakan pada penelitian ini diperoleh dari database *Alzheimer's Disease Neuroimaging Initiative* (ADNI). Pada penelitian ini menguji beberapa model pembelajaran mesin dengan volume regional dari 2.250 MRI otak: 687 kontrol normal (NC), 1.094 gangguan kognitif ringan (MCI), dan 469 AD yang disediakan oleh ADNI (database *Inisiatif Neuroimaging Penyakit Alzheimer*). Tiga jenis kumpulan fitur (63, 29, dan 22 fitur) dipilih, Adapun beberapa metode yang digunakan untuk mengukur akurasi yaitu terdiri dari RF, *Support vector machine* (SVM), *Multi-layer perceptron* (MLP), dan *Convolutional neural network* (CNN). Hasil dari metode RF, MLP, dan CNN menunjukkan performa tinggi sebesar 90,2%, 89,6%, dan 90,5% dengan 63 fitur[6].

Penelitian selanjutnya dilakukan oleh Qian Zhang, dkk, yang berjudul "*Classification of Alzheimer's disease progression based on sMRI using gray matter volume and lateralization index*" eksperimen yang digunakan dalam penelitian ini diperoleh dari database ADNI (ADNI, <http://adni.loni.usc.edu/>). Dalam penelitian ini, sampel data yang digunakan berisi 203 gambar MRI struktural dari 94 subjek yang berusia antara 55 hingga 90 tahun dari Kumpulan data ADNI. Di antara gambar struktural tersebut, 67 gambar untuk 27 subjek NC (12 laki-laki dan 15 perempuan), 58 gambar untuk 27 subjek EMCI (18 laki-laki dan 9 perempuan), 39 gambar untuk 18 subjek LMCI (10 laki-laki dan 8 perempuan) dan 39 gambar untuk 22 subjek AD (12 laki-laki dan 10 perempuan). Pekerjaan ini berfokus pada diagnosis AD dengan bantuan komputer dan tahap prodromalnya termasuk EMCI dan LMCI. Fitur klasifikasi yang berisi volume GM di 64 wilayah otak berbeda dan LI untuk 15 wilayah otak berbeda pertama kali diekstraksi dengan ANOVA dan uji Kruskal-Wallis. Adapun beberapa metode yang digunakan untuk mengukur akurasi

yaitu terdiri dari RBF, Linear, Linear kernel, RF, DT dan KNN menghasilkan akurasi yaitu 84,8714%, 85,2648%, 88,109%, 84,2595%, 84,2595% dan 71,3448% [7].

Penelitian selanjutnya dilakukan oleh Xiaomu Tang, dkk, yaitu berjudul “*Comparing different algorithms for the course of Alzheimer ’s disease using machine learning*“ dalam penelitian ini, 560 subjek yang memenuhi syarat berjumlah 0–15.000 dalam database AD *Neuroimaging Initiative* (ADNI) yang dipilih secara acak. Pada penelitian ini menggunakan 4 kelas yaitu *cognitive normal* (CN) berjumlah 140 data, 230 data pada *Early Mild Cognitive Impairment* (EMCI). 110 data pada *Late Mild Cognitive Impairment* (LMCI). Dan 80 pada *Alzheimer ’s disease*(AD). Algoritma yang digunakan yaitu RF, DT serta SVM untuk mengklasifikasikan dan memprediksi perkembangan penyakit AD yang berbeda. Penelitian ini menemukan bahwa CN-AD memiliki akurasi klasifikasi tertinggi, disusul EMCI-AD, CN-LMCI, LMCI-AD, EMCI-LMCI, dan CN-EMCI. Pada prediksi CN-AD, AUC pengklasifikasi RF sebesar 0,92 lebih tinggi dibandingkan AUC pengklasifikasi SVM dan DT[8].

Berdasarkan ketiga contoh penelitian terdahulu, maka dapat dilihat bahwa metode *Random Forest* (RF) memang dapat menghasilkan hasil akurasi yang cukup tinggi pada penelitian sebelumnya. Pada penelitian sebelumnya kebanyakan menggunakan 2 kelas. Oleh karena itu, pada penelitian ini saya akan membandingkan 4 kelas dan menggunakan metode *Random Forest* (RF) dengan *extraction feature Local Binary Pattern*(LBP). Pada penelitian ini akan menggunakan metode CRISP-DM (*Cross-Industry Standard Process Model for Data Mining*) yang mana metode ini menjabarkan tentang proses data mining dalam enam tahap yaitu *Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, Deployment*.

Melalui penelitian yang berjudul “*Klasifikasi Penyakit Alzheimer ’s Pada Citra Magnetic Resonance Imaging (Mri) Menggunakan Algoritma Random Forest (Rf)*” diharapkan dapat memberikan hasil kalsifikasi yang optimal sehingga dapat membantu tenaga kesehatan dalam melakukan diagnosis AD secara optimal.

1.2. Perumusan Masalah

Berikut ini merupakan uraian dari perumusan masalah yang meliputi permasalahan, metode usulan, dan pertanyaan penelitian:

1.2.1. Permasalahan

Semakin tinggi kompleksitas penyakit AD maka membutuhkan tenaga medis khusus yang dapat mengidentifikasi hasil MRI, adanya faktor-faktor yang mempengaruhi terjadinya kesalahan pada mengidentifikasi yang dilakukan oleh tenaga medis sehingga mengakibatkan diagnosis penyakit AD selama ini kurang optimal. Metode klasifikasi citra MRI yang akurat dan otomatis sangat dibutuhkan dalam mengatasi permasalahan tersebut. Pada penelitian ini menggunakan ekstraksi fitur *Local Binary Pattern*(LBP) yang berfungsi untuk mengidentifikasi tekstur citra MRI pada data ADNI. Penggunaan ekstraksi fitur diharapkan dapat menghasilkan nilai akurasi yang akurat dalam mengklasifikasi citra MRI menggunakan *Random Forest*, Selain itu, kekurangan pada penelitian sebelumnya hanya terdapat perbedaan jumlah data yang digunakan dan pada variasi data yang digunakan. Pada penelitian ini akan menggunakan 800 data citra dan menggunakan 4 tingkat kelas keparahan penyakit sebagai *output* klasifikasi.

1.2.2. Solusi Permasalahan

Pada penelitian ini, peneliti mengusulkan metode *Random Forest* (RF) dalam meningkatkan akurasi diagnosis penyakit *Alzheimer* melalui *dataset Alzheimer Disease Neuroimaging Initiative* (ADNI). Pada teknik ekstraksi fiturnya menggunakan teknik *Local Binary Pattern*(LBP) yang berperan untuk mengidentifikasi tekstur pada data citra MRI. Hasil dari ekstraksi fitur tersebut yang nanti digunakan untuk data input pada algoritma *Random Forest* (RF).

1.2.3. Pertanyaan Peneliti

1. Bagaimana performa (*Accuracy, Precision, Recall, F1-score*) yang didapatkan dari proses klasifikasi penyakit *Alzheimer* dengan menggunakan metode *Random Forest* (RF)?

2. Bagaimana pengaruh nilai radius ekstraksi fitur *Local Binary Pattern*(LBP) dan jumlah pohon pada metode *Random Forest* (RF) terhadap hasil akurasi penyakit *Alzheimer* 's ?

1.3. Tujuan dan Manfaat

Dalam penelitian yang dilakukan terdapat tujuan dan manfaat yaitu:

1.3.1. Tujuan

Tujuan yang ingin dicapai dalam penelitian ini adalah :

1. Mengetahui nilai nilai performa (*Accuracy, Precision, Recall, F1-score*) dari proses klasifikasi penyakit *Alzheimer* menggunakan metode *Random Forest* (RF)
2. Mendapatkan pengukuran akurat terkait dengan pengaruh penggunaan nilai radius ekstraksi fitur *Local Binary Pattern*(LBP) dan jumlah pohon pada metode *Random Forest* terhadap akurasi klasifikasi penyakit *Alzheimer* 's.

1.3.2. Manfaat

Penelitian ini memiliki manfaat untuk mengetahui hasil dari klasifikasi penyakit *Alzheimer* yang mana dapat membantu tenaga medis atau dokter dalam mendiagnosis penyakit *Alzheimer*, serta memberikan kontribusi riset tentang klasifikasi penyakit *Alzheimer* di bidang machine learning.

1.4. Batasan

Terdapat batasan masalah dalam penelitian ini:

1. Penelitian ini menggunakan metode *Random Forest* (RF) sebagai *classifier*. Serta untuk *feature extraction Local Binary Pattern*(LBP).
2. Data yang digunakan berasal dari data publik *Alzheimer* 's *Disease Neuroimaging Initiative* (ADNI) yang dikumpulkan oleh peneliti untuk menentukan perkembangan penyakit *Alzheimer* 's, data tersebut dapat diakses pada <https://ida.loni.usc.edu/>.
3. Data yang digunakan sebanyak 800 citra yang memiliki 4 kategori kelas yaitu *Cognitively Normal* (CN), *Early Mild Cognitive Impairment* (EMCI), *Late Mild Cognitive Impairment* (LMCI), dan *Alzheimer* 's *Disease* (AD). Pada

setiap kelasnya memiliki 200 data dari 100 subjek yang berformat DCM (*.dicom) dan rata-rata ukurannya 256 x 256 *pixel* dan memiliki 170 *slice*. Data tersebut nantinya akan dihilangkan terlebih dahulu bagian non otak dan akan disimpan dengan format NIfTI-GZ (*.nii.gz).

4. *Output* yang dihasilkan berupa hasil klasifikasi pada MRI Scan otak pasien dengan kategori *Cognitively Normal* (CN), *Early Mild Cognitive Impairment* (EMCI), *Late Mild Cognitive Impairment* (LMCI), dan *Alzheimer 's Disease* (AD).

1.5. Sistematika Skripsi

Sistematika skripsi yang digunakan untuk menyusun proposal pada penelitian ini yakni terdiri dari 3 bab, diantaranya :

BAB I PENDAHULUAN

Menjelaskan mengenai latar belakang, perumusan masalah, permasalahan, metode usulan, pertanyaan penelitian, tujuan dan manfaat penelitian, batasan masalah, serta sistematika penulisan terkait penelitian yang dilakukan.

BAB II KAJIAN PUSTAKA

Kajian pustaka berisi landasan teori yang berkaitan dengan penyakit *Alzheimer 's*, Citra MRI Scan penyakit *Alzheimer 's*, *Machine learning*, metode yang akan digunakan yaitu *Random Forest* (RF), serta proses yang akan dilakukan seperti segmentasi citra, pemilihan *slice* sampai ke evaluasi model.

BAB III METODE PENELITIAN

Membahas terkait metode perancangan sistem yang mencakup jenis penelitian, tahapan penelitian, analisa kebutuhan sistem serta skenario pengujian program klasifikasi penyakit *Alzheimer 's* menggunakan *Random Forest* (RF).

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi hasil dan pembahasan implementasi dari metode yang diusulkan dan analisis terhadap hasil implementasi.

BAB V KESIMPULAN DAN SARAN

Bab ini merupakan bab yang memuat kesimpulan dari hasil penelitian yang dilakukan serta saran untuk pengembangan pada penelitian selanjutnya berdasarkan hasil analisis yang diperoleh pada bab IV.