



Klasifikasi Penyakit Alzheimer Menggunakan Data Mining Decision Tree

Yulina Sari^{1*}, Zaehol Fatah²

¹ Sistem Informasi, Universitas Ibrahimy, Indonesia

² Sistem Informasi, Universitas Ibrahimy, Indonesia

^{1*}yulinasari653@gmail.com, ²zaeholfatah@gmail.com

Abstrak

Penyakit Alzheimer adalah penyakit neurodegeneratif yang mempengaruhi otak dan secara bertahap mengurangi kemampuan seseorang untuk berpikir, mengingat, dan berfungsi secara mandiri. Ini merupakan jenis demensia yang paling umum, dengan gejala awal yang biasanya terjadi pada individu di atas 65 tahun, meskipun dapat juga terjadi pada usia yang lebih muda. Beberapa gejala awal Alzheimer termasuk lupa akan peristiwa baru-baru ini, kesulitan menyelesaikan tugas sehari-hari, dan kebingungan tentang waktu atau tempat. Gejala penyakit dapat memburuk seiring berjalannya waktu, seperti kehilangan kemampuan berkomunikasi, perubahan suasana hati yang signifikan, dan penurunan fungsi fisik. Diagnosis tahap awal untuk penyakit Alzheimer dapat meningkatkan efisiensi terapi. Deteksi dini penyakit Alzheimer dapat memanfaatkan pendekatan matematis pada data mining untuk menganalisis data. Klasifikasi adalah salah satu metode data mining yang dapat diterapkan untuk pendeteksian dini atau prediksi penyakit Alzheimer. Penelitian yang berkenaan dengan klasifikasi penyakit Alzheimer belum banyak dilakukan. Oleh karena itu, penelitian ini akan membandingkan performa algoritma C4.5 dan algoritma AdaBoost dalam klasifikasi penyakit Alzheimer menggunakan teknik pengujian percentage split dan k-fold cross validation. Selain itu, penelitian ini akan menggunakan metode Algoritma Decision Tree karena dapat ditafsirkan dengan mudah oleh manusia dan memiliki akurasi yang tinggi sebesar 26.3%. Dapat mengubah data yang ada menjadi pohon keputusan dan aturan-aturan, sehingga dapat membantu dalam proses pengambilan keputusan yang kompleks menjadi lebih sederhana.

Kata Kunci: Penyakit Alzheimer, Data mining, Klasifikasi, Algoritma decision tree

PENDAHULUAN

Penyakit Alzheimer adalah penyakit neurodegeneratif yang mempengaruhi otak dan secara bertahap mengurangi kemampuan seseorang untuk berpikir, mengingat, dan berfungsi sendiri. Ini adalah jenis paling umum dari demensia, dan kasus awal biasanya terjadi pada orang di atas 65 tahun, tetapi juga bisa terjadi pada orang yang lebih muda. Beberapa gejala awal Alzheimer termasuk lupa akan peristiwa baru-baru ini, kesulitan menyelesaikan tugas sehari-hari, dan kebingungan tentang waktu atau tempat. Gejala penyakit dapat memburuk seiring berjalannya waktu, seperti kehilangan kemampuan berkomunikasi, perubahan suasana hati yang signifikan, dan penurunan fungsi fisik. Usia, faktor genetik, dan riwayat keluarga adalah faktor risiko utama, meskipun penyebab pastinya belum sepenuhnya dipahami. Selain itu, kerusakan saraf yang terjadi pada penderita Alzheimer disebabkan oleh perubahan kimiawi yang terjadi di otak, seperti pembentukan tangles tau dan plak beta-amyloid. Menurut World Alzheimer Report 2021, Alzheimer's Disease International (ADI) memperkirakan bahwa 75% penderita demensia tidak terdiagnosis secara global, dengan angka tersebut diyakini meningkat hingga 90% di beberapa negara berpenghasilan rendah dan menengah. Saat ini ada dua jenis demensia yaitu vaskular dan non-vaskular. Demensia vaskular, juga dikenal sebagai Alzheimer, adalah Perubahan yang terjadi pada tubuh yang disebabkan oleh penyumbatan pembuluh darah di otak dapat menyebabkan melemahnya ingatan manusia.[1]

Diagnosis tahap awal untuk penyakit alzheimer dapat meningkatkan efisiensi terapi untuk penyakit ini. Diagnosa penyakit alzheimer dapat dibantu dengan pendeteksian dini. Deteksi dini penyakit alzheimer dapat memanfaatkan pendekatan matematis pada data mining untuk menganalisis data. Data mining memiliki berbagai metode, salah satunya klasifikasi yang dapat diterapkan untuk pendeteksian dini atau prediksi penyakit alzheimer. Klasifikasi adalah proses untuk mengidentifikasi dan mengelompokkan objek yang memiliki karakteristik yang sama ke dalam beberapa kelas yang telah ditentukan.[2]

Terapi yang ada hingga saat ini hanya menargetkan untuk membantu menurunkan gejala dengan menggunakan berbagai obat dan psikoterapi, akan tetapi tidak menyembuhkan penyakit. Beberapa obat kimiawi yang dikembangkan untuk aplikasi klinis sejauh ini hanya mencapai efek terapeutik untuk tujuan pencegahan. Pengobatan tradisional herbal selama ribuan tahun telah mengumpulkan banyak pengalaman dalam terapi demensia. Studi farmakologi modern mengonfirmasi efek terapeutik dari banyak komponen aktif obat herbal. Efek terapeutik dari obat herbal ini dinilai sangat potensial untuk berbagai penyakit, termasuk gangguan neurodegeneratif.[3]

Penelitian yang berkenaan dengan klasifikasi penyakit alzheimer belum banyak dilakukan. Dataset yang digunakan pada penelitian ini memiliki atribut-atribut yang saling relevan dan informatif sehingga dapat memberikan hasil deteksi dan klasifikasi yang lebih baik. Pada penelitian ini akan dibandingkan performa algoritma C4.5 dan algoritma AdaBoost dalam klasifikasi penyakit alzheimer. Dalam pengujian kedua algoritma digunakan teknik pengujian percentage split dan k-fold cross validation. Pada percentage split dipilih ukuran split sebesar 80% untuk data latih dan 20% sebagai data uji. Pada teknik pengujian k-fold cross validation dipilih nilai k sebesar 10 dimana data akan dibagi menjadi sepuluh kelompok dan dapat bergantian sebagai data latih serta data uji sebanyak sepuluh kali. Penelitian ini akan menentukan algoritma dan teknik pengujian mana yang paling optimal dalam mendeteksi penyakit alzheimer berdasarkan kriteria akurasi, presisi, dan recall.[2]

Studi ini menggunakan metode Algoritma Decision Tree karena dapat ditafsirkan dengan mudah oleh manusia. Decision Tree, atau pohon keputusan, adalah model prediksi yang menggunakan struktur pohon. Idenya adalah mengubah data yang ada menjadi pohon keputusan dan aturan-aturannya. Salah satu manfaat utama penggunaan metode ini, kata Asmaul dalam jurnalnya adalah membuat proses pengambilan keputusan yang kompleks menjadi lebih sederhana, sehingga lebih mudah untuk menemukan solusi masalah saat membuat keputusan. [4]

Didapatkan hasil bahwa tingkat akurasi metode decision tree dapat dipengaruhi oleh beberapa faktor, diantaranya : jumlah data training, data testing, dan atribut yang digunakan. Penelitian lainnya yang dilakukan, didapatkan hasil bahwa aturan rules yang dihasilkan dari perhitungan metode decision tree sangat bermanfaat dan membantu dalam meningkatkan kualitas serta akurasi dari hasil analisa yang dilakukan kedepannya.[5]

METODE

Pengambilan data

Pada bagian ini dijelaskan bagaimana penulis mengumpulkan data untuk penelitian ini. Dengan menggunakan metode studi literatur untuk mengumpulkan data, dan jenis data sekunder yang digunakan adalah Informasi Kesehatan Komprehensif untuk Penyakit Alzheimer, yang diakses melalui situs web Kaggle, Kumpulan data ini berisi informasi kesehatan yang luas untuk 2.149 pasien, masing-masing diidentifikasi secara unik dengan ID mulai dari 4751 hingga 6900. Kumpulan data mencakup detail demografis, faktor gaya hidup, riwayat medis, pengukuran klinis, penilaian kognitif dan fungsional, gejala, dan diagnosis Penyakit Alzheimer. Data ini ideal untuk peneliti dan ilmuwan data yang ingin mengeksplorasi faktor-faktor yang terkait dengan Alzheimer, mengembangkan model prediktif, dan melakukan analisis statistik. Berikut ini *dataset* beserta penjelasan dari atribut-atributnya disajikan pada gambar 1 dibawah ini.

# PatientID	# Age	# Gender	# Ethnicity	# Education...	# BMI	# Smoking	# AlcoholCo...	# PhysicalA...	# DietQuality	# SleepQual...	# FamilyHis...
4751	73	0	0	2	22.927749238993864	0	13.29721772827684	6.327112473553353	1.3472143859081076	9.025678665766115	0
4752	89	0	0	0	26.82768119159682	0	4.5425238177221905	7.619884540163832	0.5187671386597053	7.151292743851223	0
4753	73	0	3	1	17.795882442817113	0	19.55588452555359	7.844987799974517	1.826334664579784	9.673574157961111	1
4754	74	1	0	1	33.88881784413547	1	12.289265546283783	8.428801358491492	7.43568414808302	8.392553685358862	0
4755	89	0	0	0	28.716973826446887	0	18.454356898619612	6.318468689368432	0.7954975889177474	5.597237677578526	0
4756	86	1	1	1	38.626885546278938	0	4.148143784276235	8.21186163866682868	1.584922818674888	7.261952584681128	0
4757	68	0	3	2	38.387621858169126	1	0.6468472785489217	9.257694989794447	5.897387927498878	5.477685593697583	0
4758	75	0	0	1	18.776889489162835	0	13.723825785512622	4.649458668217812	8.341983191582784	4.213289925183894	0
4759	72	1	1	0	27.833188388832352	0	12.167847629684712	1.531359787981249	6.736882843694737	5.748232868874191	0
4768	87	0	0	0	35.45638172642629	1	16.82868823659191	6.448772687368699	8.886819121177685	7.55177344379122	0
4761	89	0	3	1	39.46383422186146	0	9.811292128986842	8.819958351478192	0.43482827782982697	7.644897388851816	0
4762	78	0	0	2	22.463382649418685	1	19.3808182977973385	3.8346393816458988	8.279189583998566	8.312325536865838	0
4763	84	1	0	1	26.778945999187713	0	18.978821644757177	3.9788788724668274	7.8244173889347	8.253884551876643	1

Gambar 1. Dataset

Keterangan :

- 1 *age* : Usia pasien
- 2 *Gender* : Jenis kelamin pasien
- 3 *Etnicity* : Etnisitas pasien
- 4 *Education level* : Tingkat Pendidikan Pasien
- 5 *BMI* : Indeks Massa Tubuh pasien, mulai dari 15 hingga 40.
- 6 *Smoking* : Status merokok, di mana 0 menunjukkan Tidak dan 1 menunjukkan Ya.
- 7 *Alcohol consumption* : Konsumsi alkohol mingguan dalam satuan, mulai dari 0 hingga 20.
- 8 *Physical Activity* : Aktivitas fisik mingguan dalam jam, mulai dari 0 hingga 10.
- 9 *Diet quality* : Skor kualitas diet, mulai dari 0 hingga 10.
- 10 *Sleep quality* : Skor kualitas tidur, mulai dari 4 hingga 10.
- 11 *Family History Alzheimers* : Riwayat keluarga Penyakit Alzheimer, di mana 0 menunjukkan Tidak dan 1 menunjukkan Ya.

Data mining

Metode data mining memungkinkan pengguna mengakses jumlah data yang besar dalam waktu yang relatif cepat. Dengan kata lain, data mining adalah suatu alat dan aplikasi yang menggunakan analisis statistik pada data dalam proses ekstraksi atau penggalian data dan informasi yang belum diketahui sebelumnya. Secara sederhana, proses penggalian data menghasilkan penemuan informasi terbaru dengan menemukan pola atau aturan tertentu dalam sejumlah besar data dikutip dari buku berjudul DATA MINING DAN APLIKASINYA.[6]

Sebelum menghasilkan informasi baru, proses data mining melibatkan sejumlah proses, termasuk pembersihan data, proses penghilangan suara dan data yang tidak konsisten, proses integrasi data, yang menggabungkan data dari berbagai sumber, dan proses pemilihan dan seleksi. Semua proses ini dilakukan untuk memilih data yang akan digunakan dalam proses mining, sehingga meninggalkan data yang tidak digunakan. Mengubah data menjadi format yang dapat digunakan untuk perhitungan algoritmik dan evaluasi pola dikenal sebagai konversi data. merupakan proses analisis hasil penambangan menggunakan satuan pengukuran dan penyampaian pengetahuan, yang menunjukkan hasil penambangan.[7]

Decition tree

Decision tree, juga disebut sebagai "pohon keputusan", adalah model dari persoalan yang terdiri dari penyampaian keputusan dan solusi, serta sejumlah keputusan yang mengarah pada solusi. Konsep dasar dari pohon keputusan adalah mengubah sekumpulan data menjadi pohon keputusan yang menampilkan aturan keputusan. Pohon keputusan terdiri dari berbagai node untuk membentuk pohon berakar. Dalam pohon keputusan, node dengan satu masukan yang tepat dan dua atau lebih keluaran disebut internal node, node terakhir dan hanya memiliki masukan disebut leafnode, dan node terakhir dan hanya memiliki masukan disebut decisionnode atau terminal node. Akar adalah node yang tidak memiliki masukan tetapi memiliki dua atau lebih keluaran. Decision Tree adalah struktur pohon yang terdiri dari node-node.[5]

Proses decision tree diawali dengan mengubah bentuk data dari berbentuk tabel menjadi suatu tree kemudian dibentuk rule yang selanjutnya rule tersebut disederhanakan. Persyaratan yang harus dipenuhi dalam penerapan algoritma decision tree adalah membutuhkan atribut kelas karena decision tree merupakan representasi dari supervised Learning, training data harus berjumlah banyak, dan bervariasi. [8]. Model Decision Tree dipilih karena kemampuannya dalam menangani data dengan berbagai variabel dan memberikan interpretasi yang mudah dipahami[9]

Klasifikasi

Klasifikasi merupakan metode dalam data mining yang memisahkan data ke dalam kelompok atau kategori yang telah ditetapkan sebelumnya. Metode klasifikasi menggunakan pembelajaran terpantau, di mana data pelatihan yang sudah diberi label digunakan untuk menghasilkan aturan-aturan yang dapat mengklasifikasikan data uji ke dalam kelompok atau kelas yang telah ditentukan sebelumnya.[10]

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu berpendapatan tinggi, berpendapatan menengah, dan berpendapatan rendah. Contoh lain klasifikasi dalam bisnis dan penelitian adalah: Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau bukan. Memeriksa apakah suatu pengajuan hipotek oleh nasabah merupakan suatu kredit yang baik atau buruk. Mendiagnosis penyakit seorang pasien untuk mendapatkan termasuk kategori penyakit apa. [6]

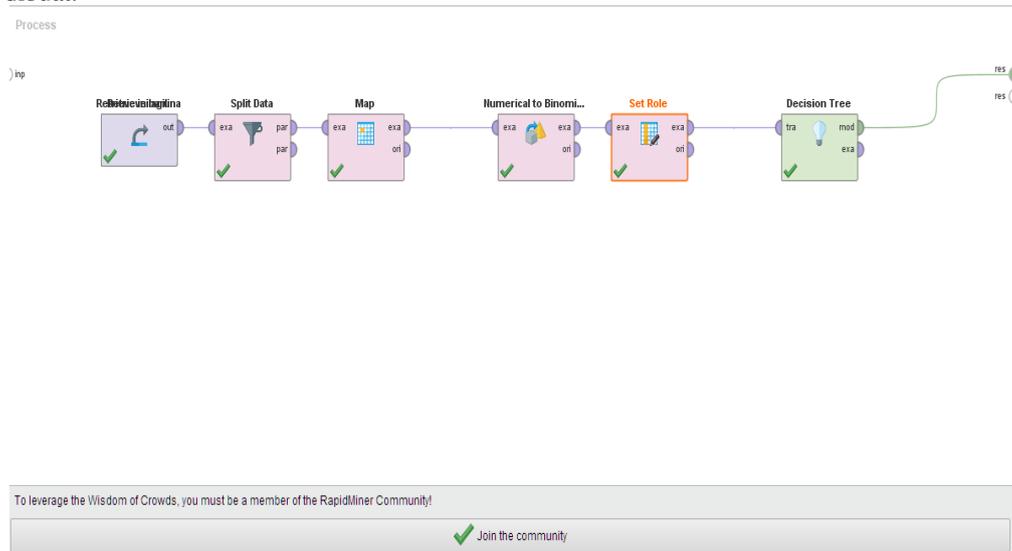
Rapid miner

Rapidminer adalah perangkat lunak open source untuk knowledge discovery dan data mining. Rapidminer merupakan Salah satu program untuk memproses penambahan data disebut RapidMiner. Rapidminer adalah sebuah Analisis teks yang fokus didalam pekerjaan yang dilakukan oleh RapidMiner text mining, yang melibatkan penggalian pola dari kumpulan data besar dan menggabungkannya dengan Teknik statistik, kecerdasan buatan, dan basis data. Tujuan dari analisis teks ini adalah untuk menghasilkan informasi yang akurat dan bermakna dari teks yang sangat besar atau kompleks. RapidMiner adalah perangkat lunak open source yang kuat untuk knowledge discovery dan data mining yang menawarkan berbagai fitur yang berguna. Berikut merupakan beberapa fitur Rapidminer yang meliputi: 1. Berlisensi gratis (open source): RapidMiner tersedia secara gratis untuk penggunaan nonkomersial dan komersial, dan tersedia dalam bentuk kode sumber yang dapat diakses oleh siapa saja. 2. Multiplatform karena deprogram dalam bahasa java: RapidMiner ditulis dalam bahasa pemrograman Java, yang memungkinkannya untuk berjalan pada berbagai platform, termasuk Windows, Mac OS X, dan Linux. 3. Internal data yang berbasis XML, yang memfasilitasi pertukaran data eksperimen: Data dalam RapidMiner disimpan dalam format XML yang fleksibel, sehingga memudahkan pertukaran data antara pengguna dan memungkinkan integrasi dengan alat lain. 4. Berisi scripting language untuk mengotomatisasi eksperimen: RapidMiner menyertakan bahasa scripting sendiri, yang memungkinkan pengguna untuk mengotomatisasi eksperimen dan memperluas fungsionalitas. 5. Memiliki GUI (Graphical User Interface), command line mode (batch mode), dan java API yang dapat dipanggil dari program lain: RapidMiner memiliki antarmuka pengguna grafis yang intuitif dan mudah digunakan, tetapi juga menyediakan mode baris perintah dan API Java untuk memungkinkan integrasi yang lebih kuat dengan alat lain. 6. Dapat diperluas dengan menyertakan ekstensi dan plugin: RapidMiner dapat diperluas dengan mudah melalui penyertaan ekstensi dan plugin, memungkinkan pengguna untuk menyesuaikan fungsionalitas sesuai dengan kebutuhan. 7. Alat plotting untuk model dan tampil data multidimensi: RapidMiner menyertakan alat plotting yang kuat untuk memvisualisasikan model dan data multidimensi, yang membantu pengguna untuk memahami dan menganalisis data dengan lebih baik.[11]

RapidMiner merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. RapidMiner ditulis dengan menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi.[12]

HASIL DAN PEMBAHASAN

Gambar dibawah ini menunjukkan Antarmuka pengguna perangkat lunak RapidMiner Studio, sebuah alat yang populer untuk pemrosesan data dan pembelajaran mesin, digambarkan di sini. Tampilan ini menunjukkan sebuah proses yang sedang dibangun. Proses ini terdiri dari beberapa bagian atau operator yang saling terhubung, yang membentuk alur kerja. Komponen-komponen ini menunjukkan berbagai tahapan dalam proses analisis data, mulai dari membaca data (menggunakan Microsoft Excel), melakukan transformasi data (menggunakan numerik ke nominal), dan membangun model klasifikasi (menggunakan klasifikasi). Koneksi antar komponen menunjukkan aliran data dari satu tahap ke tahap berikutnya. Selain itu, komponen evaluasi, juga dikenal sebagai cross-validation, digunakan untuk mengevaluasi kinerja model yang dibuat.



Gambar 2. Model Klasifikasi



Gambar 3. Split Data

Split data digunakan untuk beberapa tujuan penting dalam proses analisis data dan machine learning:

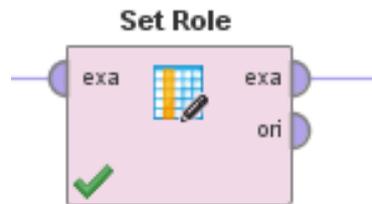
1. Dengan membagi dataset menjadi bagian pelatihan dan pengujian, dan membantu mengukur kinerja model secara lebih akurat.
2. Pembagian data membantu mengurangi risiko overfitting, di mana model terlalu cocok dengan data pelatihan dan tidak dapat generalisasi dengan baik pada data baru.
3. Dengan menggunakan data pengujian yang terpisah dapat menilai kinerja model menggunakan metrik evaluasi seperti akurasi, presisi, dan recall. Ini memberikan gambaran yang lebih jelas tentang seberapa baik model akan bekerja di dunia nyata.
4. Selain itu juga dapat menggunakan teknik lain seperti cross-validation untuk lebih lanjut mengoptimalkan dan menilai model dengan cara menyeluruh.



Gambar 4. Map

MAP (Mean Average Precision) biasanya digunakan untuk mengukur kinerja model dalam konteks klasifikasi atau rekomendasi, terutama ketika ada beberapa kategori atau label yang dipertimbangkan. Berikut adalah beberapa fungsi dan kegunaan MAP dalam menghitung data:

1. MAP memberikan gambaran tentang seberapa baik model dalam memberikan prediksi yang relevan, terutama dalam masalah multi-label atau multi-class.
2. MAP mengukur rata-rata presisi pada setiap tingkat recall, memberikan informasi yang lebih detail dibandingkan hanya menggunakan akurasi. Ini penting dalam situasi di mana ada ketidakseimbangan kelas.
3. Dalam konteks sistem rekomendasi, MAP dapat digunakan untuk mengevaluasi seberapa baik model dalam meranking item berdasarkan relevansi.
4. Dengan menggunakan MAP sebagai metrik evaluasi yang dapat melakukan tuning hyperparameter untuk meningkatkan kinerja model.



Gambar 5. Set Role

Fungsi "Set Role" di RapidMiner digunakan untuk menentukan peran atau atribut dari kolom dalam dataset. Ini sangat penting dalam konteks persiapan data dan pemodelan. Berikut adalah beberapa fungsi utama dari "Set Role":

1. Menentukan Target Variabel yaitu dapat menetapkan kolom tertentu sebagai target atau label yang ingin diprediksi oleh model. Ini penting dalam proses klasifikasi atau regresi.
2. Mengatur Atribut Fitur yaitu kolom-kolom lain dalam dataset dapat ditetapkan sebagai atribut input (features), yang akan digunakan oleh model untuk membuat prediksi.
3. Selain itu juga dapat menandai kolom tertentu sebagai ID (identifikasi) untuk menunjukkan bahwa kolom tersebut tidak perlu dipertimbangkan dalam proses pelatihan model. Kolom ini biasanya tidak memengaruhi prediksi, seperti ID pelanggan.
4. Dengan menetapkan peran yang sesuai untuk setiap kolom dapat lebih mudah menganalisis data dan menghasilkan visualisasi yang relevan.

Berikut gambar dibawah ini adalah hasil pohon decision tree dan persentase klasifikasi



Gambar 6. Pohon keputusan untuk Alcohol consumption

Tree

```
AlcoholConsumption > 19.687: true {false=0, true=5}
AlcoholConsumption <= 19.687: false {false=182, true=60}
```

Gambar 6. Description

Perhitungan :

- Jika $AlcoholConsumption > 19.687$:
Prediksi benar (True): 5
Prediksi salah (False): 0
- Jika $AlcoholConsumption \leq 19.687$:
Prediksi benar (True): 60
Prediksi salah (False): 182
- Hasil perhitungan :
Jumlah prediksi benar = $5+60 = 65$

Total data = $5+0+60+182 = 247$

Akurasi = $24765 \approx 0.263$ (26.3%)

Akurasi dari hasil ini adalah sekitar 26.3%.

KESIMPULAN

Dalam penelitian dapat disimpulkan bahwa pendekatan data mining, khususnya metode klasifikasi dapat berkontribusi dalam mendeteksi penyakit Alzheimer lebih awal. Dengan menggunakan dataset yang komprehensif, penelitian ini membandingkan kinerja algoritma dalam hal akurasi, presisi, dan recall, serta mengidentifikasi teknik pengujian yang paling optimal. Algoritma Decision Tree menjadi pilihan yang menarik karena kemudahan interpretasi hasilnya, yang dapat membantu dalam pengambilan keputusan terkait diagnosis dan perawatan.

Secara keseluruhan, penelitian ini memberikan wawasan berharga mengenai potensi penggunaan teknik klasifikasi dalam deteksi dini penyakit Alzheimer dan menggarisbawahi pentingnya penelitian lebih lanjut dalam mengembangkan metode diagnostik yang lebih efektif. Data yang dikumpulkan dari berbagai sumber juga menunjukkan relevansi dalam memahami faktor-faktor yang mempengaruhi penyakit ini, serta mengembangkan model prediktif untuk analisis lebih lanjut.

UCAPAN TERIMA KASIH

Ucapan terima kasih diberikan kepada Bapak Zaehol Fatah, M.Kom sebagai dosen pembimbing kami yang senantiasa memberikan kritikan yang memotivasi dan membangun demi keberlangsungan keberhasilan kami. Tidak lupa pula ucapan terima kasih yang paling utama di tujukan kepada seseorang di balik figur yaitu orang tua kami yang tidak pernah lelah dalam menasehati dan memberikan motivasi yang bersifat membangun juga tentunya, Kalian semua hebat *Jazakumullahu Khairon Absanal jazaa'*.

DAFTAR PUSTAKA

- [1] R. J. Alfirdausy and S. Bahri, "Implementasi Algoritma K-Nearest Neighbor untuk Klasifikasi Diagnosis Penyakit Alzheimer," *Techno.Com*, vol. 22, no. 3, pp. 635–642, 2023, doi: 10.33633/tc.v22i3.8393.
- [2] A. A. Mortara, M. Permatasari, A. Desiani, Y. Andriani, and M. Arhami, "Perbandingan Algoritma C4.5 dan Adaptive Boosting dalam Klasifikasi Penyakit Alzheimer," *J. Teknol. dan Inf.*, vol. 13, no. 2, pp. 196–207, 2023, doi: 10.34010/jati.v13i2.10525.
- [3] Jan Sudir Purba, "Potential Implication of Treatments for Alzheimer's Disease: Current and Future," *Medicinus*, vol. 36, no. 1, pp. 3–10, 2023, doi: 10.56951/medicinus.v36i1.112.
- [4] M. N. Baharudin and I. K. Dwi Nuryana, "Implementasi Algoritma Decision Tree untuk Klasifikasi Surat pada Aplikasi Mobile E-Surat Dinas Komunikasi dan Informatika Kota Kediri Berbasis Android," *J. Informatics Comput. Sci.*, vol. 4, no. 01, pp. 76–85, 2022, doi: 10.26740/jinacs.v4n01.p76-85.
- [5] F. D. Pratama and H. D. Bhakti, "Implementasi Aplikasi Prediksi Ketepatan Pembayaran Customer Perusahaan Dengan Metode Decision Tree," *Indexia*, vol. 5, no. 01, p. 72, 2023, doi: 10.30587/indexia.v5i01.5082.
- [6] Y. Ardilla *et al.*, *DATA MINING DAN APLIKASINYA*. Penerbit Widina, 2021.
- [7] Nurhidayah, "Peguruang: Conference Series," vol. 3, no. November, pp. 10–14, 2021.
- [8] M. Indah werdiningsih, Barry Nuqoba, *DATA MINING MENGGUNAKAN ANDROID WEKA dan SPSS*. Surabaya: Airlangga University Press, 2020. [Online]. Available: https://www.google.co.id/books/edition/Data_Mining_Menggunakan_Android_Weka_dan/xEwCEAAAQBAJ?hl=id&gbpv=1&dq=decision+tree+definisi%3F&pg=PA81&printsec=frontcover
- [9] M. F. R. Assauqi, "DECISION TREE BERDASARKAN GEJALA KLINIS DAN DATA LABORATORIUM," vol. 2, no. 4, pp. 42–53, 2024.
- [10] P. B. N. Setio, D. R. S. Saputro, and Bowo Winarno, "Klasifikasi Dengan Pohon Keputusan Berbasis Algoritme C4.5," *Prism. Pros. Semin. Nas. Mat.*, vol. 3, pp. 64–71, 2020.
- [11] M. Jundanuddin *et al.*, "Implementasi Data Mining Pada Penilaian Kinerja Guru Menggunakan Metode K-Nearest Neighbor (K-NN) Studi Kasus Pada Mi Almagiri Silo-Jember Implementation Of Data Mining On Teacher Performance Appraisal Using K-Nearest Neighbor (K-NN) Method Case Study," vol. 13, no. 105, pp. 869–880, 2024.
- [12] Aprilla Dennis, "Belajar Data Mining dengan RapidMiner," *Innov. Knowl. Manag. Bus. Glob. Theory Pract. Vols* 1 2, vol. 5, no. 4, pp. 1–5, 2013, [Online]. Available: http://esjournals.org/journaloftechnology/archive/vol1no6/vol1no6_6.pdf%5Cnhttp://www.aircse.org/journal/na/5413nsa02.pdf