

ANALISIS PREDIKSI PENYAKIT DEMENSIA ALZHEIMER MENGUNAKAN METODE DECISION TREE C4.5 DAN NAIVE BAYES

^{1*}Okky Prasetya, ²Syaeful Machfud
Universitas Pamulang, Tangerang Selatan, Banten, Indonesia
e-mail: dosen02837@unpam.ac.id

Abstrak

Demensia adalah suatu sindrom akibat penyakit otak, biasanya bersifat kronik atau progresif serta terdapat gangguan fungsi luhur. Jenis demensia yang paling sering dijumpai yaitu demensia tipe Alzheimer, termasuk daya ingat, daya pikir, daya orientasi, daya pemahaman, berhitung, kemampuan belajar, berbahasa, dan daya kemampuan menilai. Organisasi Kesehatan Dunia (WHO) telah mengungkapkan bahwa 55 juta orang di seluruh dunia mengalami penurunan keterampilan dalam memori dan berpikir, yang sekarang dapat mempengaruhi individu berusia 30 sampai 40 tahun. Maka dari itu, penelitian ini memberikan solusi dengan menerapkan algoritma Decision Tree C4.5 di RapidMiner untuk memprediksi penyakit demensia alzheimer. Permasalahan yang saat ini terjadi adalah belum adanya prediksi penyakit demensia alzheimer serta belum adanya cara mengukur angka prediksi penyakit demensia alzheimer dengan metode Algoritma Decision Tree C4.5 dan Naive Bayes. Metodologi penelitian yang digunakan adalah alur proses dari mulai pengumpulan data, pengolahan awal data, model atau metode yang diusulkan atau dikembangkan, eksperimen dan pengujian model atau metode dan yang terakhir evaluasi dan validasi hasil. Tujuan Penelitian ini adalah untuk memprediksi penyakit demensia alzheimer menggunakan dataset yang digunakan dari kagel. Model yang dihasilkan diuji untuk mendapatkan nilai accuracy, precision, recall dan AUC dari setiap algoritma. Hasil yang diharapkan dari pengujian Alzheimer dengan menggunakan Metode Decision Tree C4.5 Dan Naive Bayes dapat memberikan pemecahan untuk permasalahan Alzheimer lebih akurat.

Kata kunci: Demensia Alzheimer, Decision Tree C4.5, Naive Bayes, Rapid Miner

Abstract

Diabetes mellitus is a group of metabolic diseases characterized by hyperglycemia that occurs due to abnormalities in insulin secretion, insulin action or both. The number of people affected by diabetes is increasing every year. The problem currently occurring is that there is no level of accuracy in predicting diabetes in applying the neural network algorithm using the backpropagation method. Therefore, this research provides a solution by applying the Backpropagation algorithm in RapidMiner to predict diabetes. The research methodology used is a process flow starting from data collection, initial data processing, proposed or developed models or methods, experimentation and testing of models or methods and finally evaluation and validation of the results. The aim of this research is to predict diabetes using the dataset used from Kagel. In this research, the hardware requirements consist of an ACER ASPIRE 4739 Laptop with Intel CORE I3 M380 2.53 GHz Processor specifications, 500 GB HDD, and 4 GB RAM and the software required is the Windows 7 Ultimate 64-Bit Operating System, RapidMiner along with the Dataset that will be tested. The resulting model was tested to obtain accuracy, precision, recall and AUC values for each algorithm. The expected results from testing the Diabetes Mellitus model using a neural network based on Backpropagation can provide a more accurate solution to the Diabetes Mellitus problem.

Keywords : Diabetes Mellitus, Backpropagation, RapidMiner, Artificial Neural Network

PENDAHULUAN

Penyakit alzheimer adalah kondisi neurologis yang menyerang sistem saraf pada otak. Penyakit alzheimer menyebabkan hilangnya neuron dan membuat penderita sulit untuk melakukan aktivitas sehari-hari dikarenakan kebingungan, ingatan yang tidak terorganisir, dan pada akhirnya akan menghilangnya kemampuan untuk mengingat (Z. Breijyeh and R. Karaman, 2020). Faktor utama yang memengaruhi perkembangan penyakit alzheimer adalah usia. Penyakit ini biasanya didiagnosis pada orang yang berusia 65 tahun ke atas, tetapi tidak menutup kemungkinan terjadi pada orang yang lebih muda. Selain itu, keturunan garis genetik dan cedera otak traumatis yang cukup berat juga terkait dengan peningkatan risiko. Organisasi Kesehatan Dunia (WHO) telah mengungkapkan bahwa 55 juta orang di seluruh dunia mengalami penurunan keterampilan dalam memori dan berpikir, yang sekarang dapat mempengaruhi individu berusia 30 sampai 40 tahun.

Penyakit ini menjadi salah satu tantangan terbesar dalam dunia kesehatan di abad ke-21 ini (P. Scheltens et al,2020). 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 .

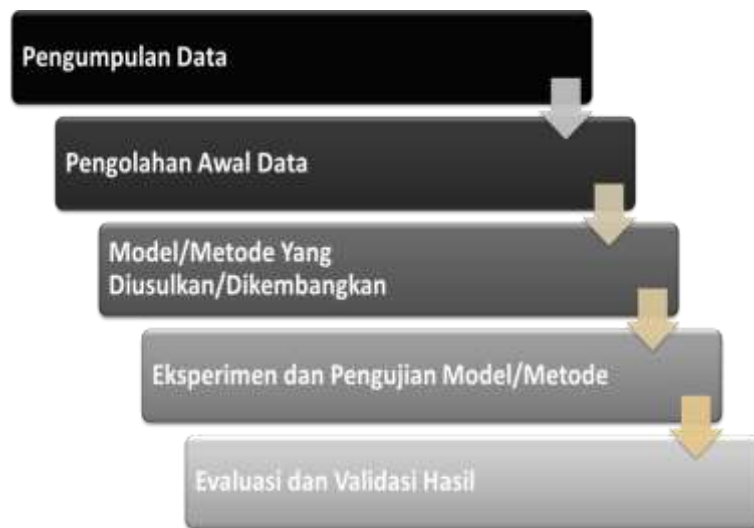
Data mining adalah proses untuk menemukan korelasi, pola, dan tren baru yang bermakna dengan memilah-milah data dalam jumlah besar yang disimpan di dalam repositori, menggunakan teknologi pengenalan pola serta teknik statistik dan matematika. Data mining adalah proses menemukan pola dan pengetahuan menarik dari data dalam jumlah yang besar . Istilah data mining memiliki beberapa pandangan, seperti knowledge discovery ataupun pattern recognition. Istilah knowledge discovery atau penemuan pengetahuan tepat digunakan karena tujuan utama dari data mining memang untuk mendapatkan pengetahuan yang masih tersembunyi di dalam bongkahan data. Sedangkan istilah untuk pattern recognition atau pengenalan pola tepat untuk digunakan karena guna menemukan pola yang tersembunyi di dalam bongkahan data. Neural Network (NN) atau Jaringan syaraf tiruan (JST) adalah suatu usaha untuk meniru fungsi otak manusia. Otak manusia diyakini terdiri dari jutaan unit pengolahan kecil, yang disebut neuron, yang bekerja secara paralel. Neuron saling terhubung satu sama lain melalui koneksi neuron. setiap individu neuron mengambil input dari satu set neuron. Ini kemudian memproses input tersebut dan melewati output untuk satu set neuron. Keluaran dikumpulkan oleh neuron lain untuk diproses lebih lanjut. Para Otak manusia adalah jaringan kompleks neuron di mana koneksi tetap melanggar dan membentuk. Banyak model mirip dengan otak manusia telah diusulkan. Kemampuan otak manusia seperti : mengingat, menghitung, menggeneralisasi, adaptasi, diharapkan neural network dapat meniru kemampuan otak manusia. Neural network berusaha meniru struktur/arsitektur dan cara kerja otak manusia sehingga diharapkan bisa dan mampu menggantikan beberapa pekerjaan manusia. Pekerjaan seperti mengenali pola (pattern recognition), prediksi, klasifikasi, pendekatan fungsi, optimisasi.

Decision Tree atau Pohon Keputusan adalah salah satu algoritma pembelajaran mesin yang paling populer dan mudah dimengerti. Algoritma ini membangun sebuah model prediksi dengan cara membuat serangkaian keputusan berdasarkan pada fitur-fitur yang dimiliki oleh data. Setiap keputusan dalam Decision Tree menghasilkan percabangan, sehingga membentuk struktur seperti pohon. Keputusan dipilih berdasarkan metrik seperti entropi atau keuntungan informasi untuk mencapai ketidakpastian yang lebih rendah. Decision Tree sangat berguna dalam klasifikasi, terutama ketika terdapat banyak variabel yang berpengaruh pada hasil klasifikasi. Algoritma ini juga dapat memberikan pemahaman visual dalam bentuk pohon yang mudah dipahami oleh manusia.

RapidMiner merupakan perangkat lunak yang bersifat terbuka (open source). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. RapidMiner memiliki kurang lebih 500 operator data mining, termasuk operator untuk input, output, data preprocessing dan visualisasi. RapidMiner merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri.

METODE

Tahapan ini adalah tahap dimana dibuatkan suatu rancangan untuk memproses data mulai dari input data sampai mendapatkan nilai akurasi yang terbaik dari algoritma *Decision Tree* C4.5 dan *Naive Bayes* dengan menggunakan aplikasi *RapidMiner* Studio.



Gambar 1 Tahapan Penelitian

Pada tahapan penelitian dibagi menjadi 5 tahapan antara lain pengumpulan data, pengolahan awal data, model atau metode yang diusulkan atau dikembangkan, eksperimen dan pengujian model atau metode dan yang terakhir evaluasi dan validasi hasil.

a. Pengumpulan Data

Pengumpulan data ditentukan atas data yang akan diproses dan mencari data yang yang tersedia serta memperoleh data tambahan yang dibutuhkan, kemudian menintegrasikan data tersebut didalam dataset.

b. Pengolahan awal data.

Dilakukan penyeleksian data, data dibersihkan dan ditransformasikan bentuk yang diinginkan sebelum dilakukan pembuatan model

c. Model atau metode yang diusulkan atau dikembangkan.

Data yang diteliti dianalisa kemudian dikelompokkan variabel mana yang berhubungan dengan satu sama lainnya, lalu dibuatkan model yang sesuai dengan jenis data. Pembagian data kedalam data latihan (training data) dan data uji (testing data) juga diperlukan untuk pembuatan model.

d. Eksperimen dan pengujian model atau metode

Pengujian model diusulkan pada model yang akan diuji untuk melihat hasil berupa rule yang akan dimanfaatkan dalam mengambil keputusan hasil penelitian.

e. Evaluasi dan validasi hasil

Pada penelitian ini dilakukan evaluasi terhadap model yang ditetapkan untuk mengetahui tingkat keakurasian model.

Dalam penelitian ini, pengumpulan data adalah data pasien yang memeriksakan penyakit demensia alzheimer dengan hasil yang didapat sebanyak 336 orang yang diperiksa. Dengan atribut dari setiap penyakit demensia alzheimer yang diperiksa adalah *Male-Female* (jenis kelamin), *Age* (umur), *Education* (pendidikan terakhir), *Social Economi Status* (status ekonomi sosial), *Mini Mental State Examination* (pemeriksaan keadaan mental), *Clinical Dementia Rating* (pemeriksaan demensia klinis), *Estimated total intracranial volume* (perkiraan total volume intrakranial), *Normalize Whole Brain Volume* (menormalkan seluruh otak), *Atlas Scaling Factor* (faktor perskalaan). Data pasien yang memeriksakan penyakit demensia alzheimer bisa di lihat pada tabel berikut.

Tabel 1 Dataset Keseluruhan

M/F	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF	Group
M	87	14	2	27	0	1987	0.696	0.88	Nondemented
M	88	14	2	30	0	2004	0.681	0.88	Nondemented
M	75	12	1	23	0.5	1678	0.736	1.05	Demented
M	76	12	1	28	0.5	1738	0.713	1.01	Demented

M	80	12	1	22	0.5	1698	0.701	1.03	Demented
F	88	18	3	28	0	1215	0.71	1.44	Nondemented
F	90	18	3	27	0	1200	0.718	1.46	Nondemented
M	80	12	4	28	0	1689	0.712	1.04	Nondemented
M	83	12	4	29	0.5	1701	0.711	1.03	Nondemented
M	85	12	4	30	0	1699	0.705	1.03	Nondemented
M	71	16	1	28	0.5	1357	0.748	1.29	Demented
M	73	16	1	27	1	1365	0.727	1.29	Demented
M	75	16	1	27	1	1372	0.71	1.28	Demented
F	93	14	2	30	0	1272	0.698	1.38	Nondemented

a. Algoritma C4.5

Tahapan Algoritma *Decision Tree* C4.5 adalah sebagai berikut :

1. Menyiapkan data *training*.
2. Menentukan akar dari pohon.
3. Hitung nilai Gain :

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i \quad (3.1)$$

Keterangan :

S : Himpunan kasus

n : Jumlah partisi S

pi: Proporsi Si terhadap S

4. Ulangi langkah ke-2 hingga semua tupel terpartisi

$$Gain(S, A) = S - \sum_{i=1}^n \frac{|S_i|}{|S|} * S_i \quad (3.2)$$

Keterangan :

S : Himpunan kasus

A: Atribut

n : Jumlah partisi atribut A

|Si| : Jumlah kasus pada partisi ke-i

|S| : Jumlah kasus dalam S

5. Proses partisi pohon keputusan akan berhenti saat semua tupel dalam node N mendapat kelas yang sama dan atau tidak ada atribut di dalam tupel yang dipartisi lagi dan atau tidak ada tupel di dalam cabang yang kosong.

b. Algoritma Naïve Bayes

Alur metode *Naïve Bayes* adalah sebagai berikut :

1. Baca data *training*
2. Hitung Jumlah dan probabilitas, namun apabila data numerik maka :
3. Cari nilai mean dan standar deviasi dari masing-masing parameter yang merupakan data numerik.

Adapun persamaan yang digunakan untuk menghitung nilai rata-rata hitung (mean) dapat dilihat sebagai berikut :

$$\mu = \frac{\sum_{i=1}^n x_i}{n} \quad \text{atau} \quad \mu = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

dimana :

μ : rata-rata hitung (mean)

X_i : nilai sample ke -i

n : jumlah sample

dan persamaan untuk menghitung nilai simpangan baku (standar deviasi) dapat dilihat sebagai berikut :

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1}}$$

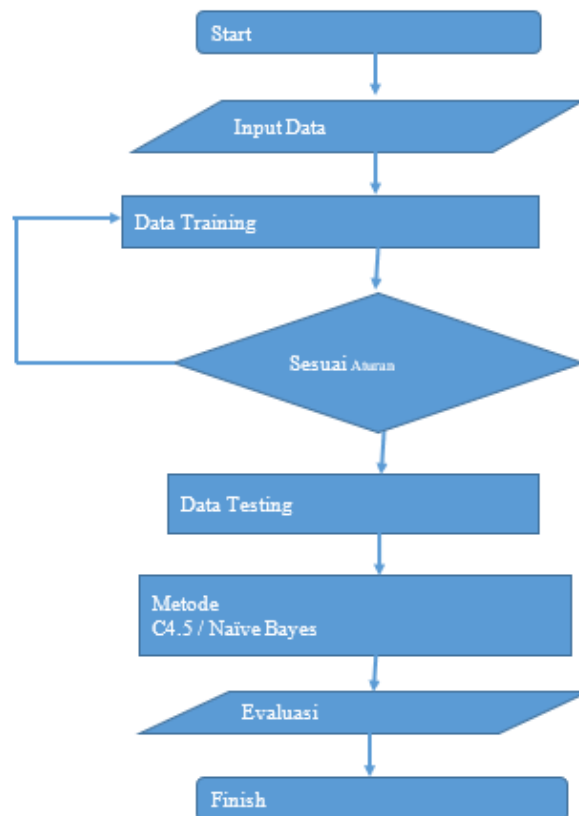
dimana :

σ : standar deviasi
Xi : nilai x ke-i
 μ : rata-rata hitung
n : jumlah sampel

4. Cari nilai probabilistik dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut.
5. Mendapatkan nilai dalam tabel mean, standard deviasi dan probabilitas
6. Solusi kemudian dihasilkan.
7. Evaluasi dan Validasi

Pada tahap evaluasi dan validasi ini menggunakan AUC, ROC dan Akurasi. Kemudian menganalisa akurasi dengan membandingkan kedua metode tersebut.

Pada tahap modeling ini dilakukan pemrosesan data traning sehingga akan membahas metode algoritma yang diujikan dengan memasukan data pasien penyakit demensia alzheimer kemudian di analisa dan dikomparasi.



Gambar 2 Model yang diusulkan.

HASIL DAN PEMBAHASAN

Algoritma *Decision Tree* C4.5

Algoritma Decision Tree C4.5 merupakan salah satu algoritma yang digunakan untuk membentuk pohon keputusan, metode pohon keputusan mengubah fakta menjadi pohon keputusan yang mempresentasikan sebuah aturan.

Data Training

Persiapan awal data (preparation data) Jumlah data diperoleh dari pengumpulan data yaitu sebanyak 236 record, dengan atribut male-female, Age, Education, Social Economic Status, Mini Mental State Examination, Clinical Dementia Rating, Estimated total intracranial volume, Normalize Whole Brain Volume, Atlas Scaling Factor dan hasil sebagai label Nondemented sebagai tidak demensia, Demented sebagai demensia. Sehingga menghasilkan data seperti terlihat pada tabel dibawah ini :

Tabel 2. Tabel Data Training

M/F	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF	Group
M	87	14	2	27	0	1987	0.696	0.88	Nondemented
M	88	14	2	30	0	2004	0.681	0.88	Nondemented
M	75	12	1	23	0.5	1678	0.736	1.05	Demented
M	76	12	1	28	0.5	1738	0.713	1.01	Demented
M	80	12	1	22	0.5	1698	0.701	1.03	Demented
F	88	18	3	28	0	1215	0.71	1.44	Nondemented
F	90	18	3	27	0	1200	0.718	1.46	Nondemented
M	80	12	4	28	0	1689	0.712	1.04	Nondemented
M	83	12	4	29	0.5	1701	0.711	1.03	Nondemented
M	85	12	4	30	0	1699	0.705	1.03	Nondemented
M	71	16	1	28	0.5	1357	0.748	1.29	Demented
M	73	16	1	27	1	1365	0.727	1.29	Demented
M	75	16	1	27	1	1372	0.71	1.28	Demented
F	93	14	2	30	0	1272	0.698	1.38	Nondemented

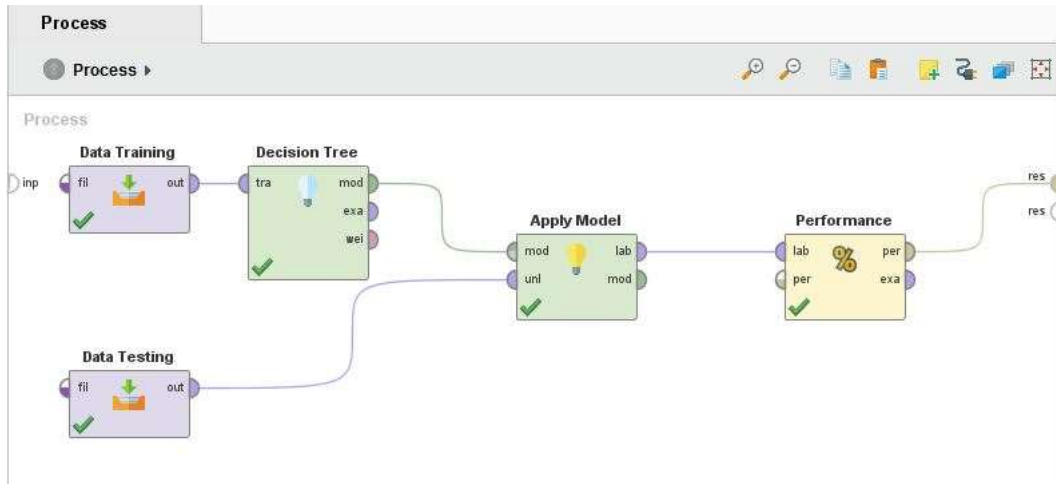
Setelah melakukan persiapan data *training* berjumlah 236 data record, langkah selanjutnya adalah melakukan persiapan data *testing* sebanyak 100 record. Data seperti terlihat pada tabel dibawah ini :

Tabel 3. Tabel Data Testing

M/F	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF	Group
M	71	16	3	27	0.5	1479	0.75	1.19	Demented
F	74	12	3	29	0	1344	0.739	1.31	Nondemented
F	75	12	3	29	0	1338	0.747	1.31	Nondemented
F	77	12	3	29	0	1344	0.74	1.31	Nondemented
F	76	16	1	28	0	1346	0.762	1.3	Nondemented
F	79	16	1	29	0	1354	0.739	1.3	Nondemented
F	78	18	1	30	0	1440	0.666	1.22	Nondemented
F	80	18	1	30	0	1436	0.663	1.22	Nondemented
F	82	18	1	29	0	1442	0.644	1.22	Nondemented
F	70	11	4	29	0.5	1295	0.748	1.36	Demented
F	71	11	4	28	0.5	1284	0.741	1.37	Demented
M	74	18	2	30	0	1636	0.68	1.07	Nondemented
M	78	18	2	27	0	1645	0.663	1.07	Nondemented
M	74	18	2	28	0.5	1659	0.739	1.06	Demented

Pengujian Decision Tree

Pengujian data dilakukan dengan menggunakan aplikasi *RapidMiner Studio*, berikut gambar alur pengujian algoritma *Decision Tree C4.5*.



Gambar 4. 1 Alur Pengujian *Decision Tree*

Pada gambar 4.1 Alur Pengujian Decision Tree C4.5 yaitu dengan menginputkan Data Training dan Data Testing kedalam aplikasi RapidMiner kemudian inputkan metode Decision Tree, Apply Model dan Performance pada menu operator untuk mendapatkan hasil Akurasi dari metode Decision Tree.

PerformanceVector (Performance)

Table View Plot View

accuracy: 94.00%

	true Demented	true Nondemented	class precision
pred Demented	37	0	100.00%
pred Nondemented	6	57	90.48%
class recall	86.05%	100.00%	

Gambar 4. 2 Hasil Akurasi *Decision Tree* C4.5

Hasil pengujian yang telah dilakukan dengan menggunakan model *Decision Tree C4.5* memperoleh akurasi sebesar 94.00% seperti pada gambar 4.2.

PerformanceVector (Performance) X

Criterion:
accuracy
AUC
precision
recall

Table View Plot View

precision: 100.00% (positive class: Demented)

	true Nondemented	true Demented	class precision
pred. Nondemented	57	6	90.48%
pred. Demented	0	37	100.00%
class recall	100.00%	96.95%	

Gambar 4. 3 Hasil Precision *Decision Tree* C4.5

Hasil pengujian yang telah dilakukan dengan menggunakan model *Decision Tree* memperoleh precision sebesar 100.00 %.

PerformanceVector (Performance) X

Criterion: accuracy, AUC, precision, recall

Table View Plot View

recall: 86.05% (positive class: Demented)

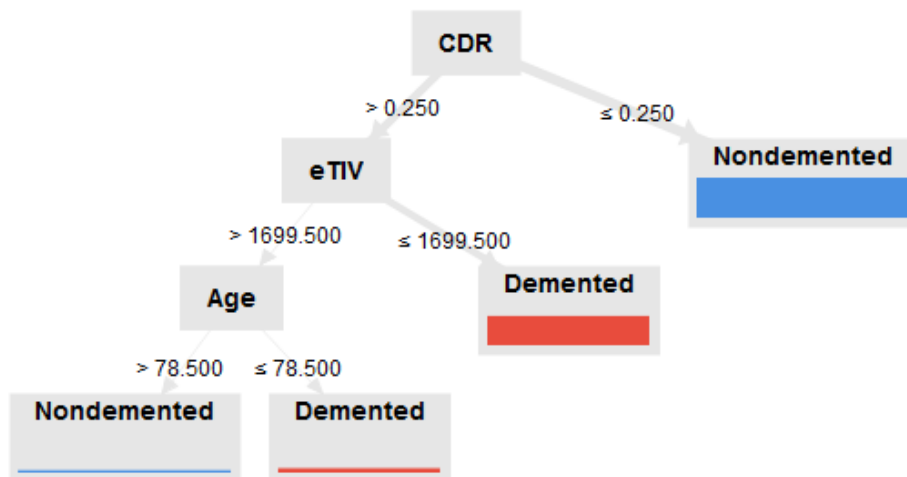
	true Nondemented	true Demented	class precision
pred Nondemented	57	6	90.48%
pred Demented	0	37	100.00%
class recall	100.00%	86.05%	

Gambar 4. 4 Hasil recall Decision Tree C4.5

Hasil pengujian yang telah dilakukan dengan menggunakan model *Decision Tree* memperoleh recall sebesar 86.05 %.

Pohon Keputusan Decision Tree C4.5

Pohon keputusan dapat dilihat pada Gambar 4.4 dibawah ini :



Gambar 4. 5 Pohon Keputusan Decision Tree C4.5

Entropy dan Gain

Tabel 4. 1 Tabel Hitung Entropy dan Gain

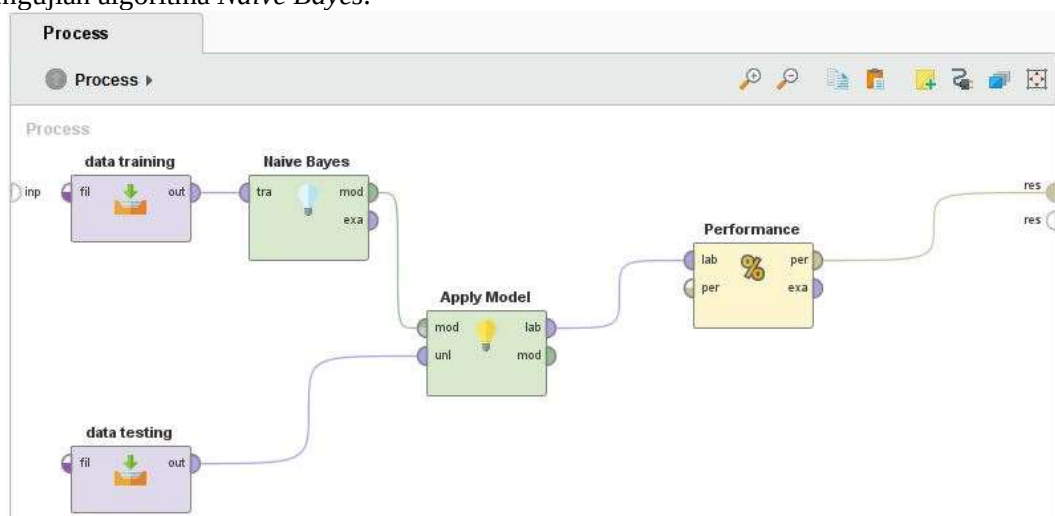
Atribut		Jumlah Kasus	Demented	Non demented	Entropy	Gain
Total		15	6	9	0,970950594	
M/F						0,253844614
	Male	11	4	2	0,977871792	
	Female	4	4	0	0	
Age						0
	71	1	1	0	0	
	73	1	1	0	0	
	75	2	2	0	0	
	76	1	1	0	0	
	80	2	1	1	1	
	83	1	0	1	0	
	85	1	0	1	0	

	87	1	0	1	0	
	88	2	0	2	0	
	90	1	0	1	0	
	93	1	0	1	0	
	95	1	0	1	0	
Education						-6
	12	6	3	3	1	
	14	4	0	4	0	
	16	3	3	0	0	
	18	2	0	2	0	
SES						- 1,169,925,001
	1	6	6	4	0,389975	
	2	4	0	2	0	
	3	2	0	3	0	
	4	3	0	0	0	
	5	0	0	0	0	
MMSE						0
	22	1	1	0	0	
	23	1	1	0	0	
	27	4	2	2	1	
	28	4	2	2	1	
	29	2	0	2	0	
	30	3	0	3	0	
CDR						0
	0	8	0	8	0	
	0.5	5	1	4	0,72192809 5	
	1	2	2	0	0	
eTIV						0
	1987	1	0	1	0	
	2004	1	0	1	0	
	1678	1	1	0	0	
	1738	1	1	0	0	
	1698	1	1	0	0	
	1215	1	0	1	0	
	1200	1	0	1	0	
	1689	1	0	1	0	
	1701	1	0	1	0	
	1699	1	0	1	0	
	1357	1	1	0	0	
	1365	1	1	0	0	
	1372	1	1	0	0	
	1272	1	0	1	0	
	1257	1	0	1	0	
nWBV						0
	0.696	1	0	1	0	
	0.681	1	0	1	0	
	0.736	1	1	0	0	
	0.713	1	1	0	0	

	0.701	1	1	0	0	
	0.71	1	0	1	0	
	0.718	1	0	1	0	
	0.712	1	0	1	0	
	0.711	1	0	1	0	
	0.705	1	0	1	0	
	0.748	1	1	0	0	
	0.727	1	1	0	0	
	0.71	1	1	0	0	
	0.698	1	0	1	0	
	0.703	1	0	1	0	
ASF						0
	0.883	1	0	1	0	
	0.876	1	0	1	0	
	1.046	1	1	0	0	
	1.01	1	1	0	0	
	1.034	1	1	0	0	
	1.444	1	0	1	0	
	1.462	1	0	1	0	
	1.039	1	0	1	0	
	1.032	1	0	1	0	
	1.033	1	0	1	0	
	1.293	1	1	0	0	
	1.286	1	1	0	0	
	1.279	1	1	0	0	
	1.38	1	0	1	0	
	1.396	1	0	1	0	

Pengujian Algoritma Naive Bayes

Pengujian data dilakukan dengan menggunakan aplikasi *RapidMiner Studio*, berikut gambar alur pengujian algoritma *Naive Bayes*.



Gambar 4. 6 Alur Pengujian Metode Naïve Bayes

Tahapan pertama diawali dengan menginputkan Data *Training* melalui read excel, dan juga menginputkan Data *Testing* melalui read excel, kemudian hubungkan Data *Training* menggunakan model *Naive Bayes* dan insert Apply model, hubungkan Data *Testing* ke Apply model kemudian hubungkan dengan performance untuk mendapatkan hasil yang diinginkan.

Row No.	Group	M/F	Age	EDUC	SES	MMSE	CDR	eTV	mWBV	ASF
1	Nondem	M	87	14	2	27	0	1987	0.696	0.883
2	Nondem	M	88	14	2	30	0	2004	0.681	0.876
3	Demented	M	75	12	1	23	0.500	1578	0.738	1046
4	Demented	M	76	12	1	28	0.500	1738	0.713	1010
5	Demented	M	80	12	1	22	0.500	1688	0.701	1034
6	Nondem	F	88	18	3	28	0	1215	0.710	1444
7	Nondem	F	90	18	3	27	0	1200	0.718	1462
8	Nondem	M	80	12	4	28	0	1689	0.712	1039
9	Nondem	M	83	12	4	29	0.500	1701	0.711	1032
10	Nondem	M	85	12	4	30	0	1699	0.705	1033
11	Demented	M	71	16	1	28	0.500	1357	0.748	1293
12	Demented	M	73	16	1	27	1	1365	0.727	1286
13	Demented	M	76	16	1	27	1	1372	0.710	1279
14	Nondem	F	93	14	2	30	0	1272	0.698	1380
15	Nondem	F	95	14	2	29	0	1257	0.703	1396

ExampleSet (336 examples; 1 special attribute, 9 regular attributes)

Gambar 4. 7 Hasil Apply Model Naïve Bayes

Hasil prediksi demesia dan tidak demesia dengan menggunakan Metode Naïve Bayes yang berjumlah 336 record dengan 9 atribut.

	true Demented	true Nondemented	class precision
pred. Demented	43	1	97.73%
pred. Nondemented	9	56	100.00%
class recall	100.00%	98.25%	

Gambar 4. 8 Hasil akurasi pengujian Metode Naïve Bayes.

Hasil pengujian yang telah dilakukan dengan menggunakan model Naïve Bayes memperoleh akurasi sebesar 99.00 %.

	true Nondemented	true Demented	class precision
pred. Nondemented	56	0	100.00%
pred. Demented	1	43	97.73%
class recall	98.25%	100.00%	

Gambar 4. 9 Hasil precision pengujian Metode Naïve Bayes.

Hasil pengujian yang telah dilakukan dengan menggunakan model Naïve Bayes memperoleh precision sebesar 97.73%.

SIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan perhitungan menggunakan aplikasi RapidMiner dengan metode Decision Tree dan Naïve Bayes maka dapat disimpulkan bahwa dari hasil pengujian yang telah dilakukan dengan menggunakan 236 data training dan 100 data testing

menghasilkan nilai akurasi untuk analisis prediksi penyakit demensia alzheimer dengan menggunakan dua perbandingan algoritma klasifikasi data mining, dan menghasilkan nilai dari masing-masing algoritma yaitu untuk Decision Tree akurasinya 94.00% dan AUC=0.500. Sedangkan akurasi Naive Bayes = 99.00% dan AUC = 1.000.

Hal ini menunjukkan bahwa metode Naïve Bayes adalah metode terbaik untuk pemecahan masalah analisis prediksi penyakit demensia Alzheimer dibandingkan dengan metode Decision Tree.

DAFTAR PUSTAKA

- Aldjawad, M., Andryana, S., & Andrianingsih. (2021). Penerapan Metode Perbandingan Dempster-Shafer dengan Certainty . Jurnal JTIC (Jurnal Teknologi Informasi dan Komunikasi), 144-152.
- Alfirdausy, R. J., & Bahri, S. (2023). Implementasi Algoritma K-Nearest Neighbor untuk Klasifikasi Diagnosis Penyakit Alzheimer. Techno.COM, 635-642.
- Andrian, F., Martha, S., & Rahmayuda, S. (2020). Sistem Peramalan Jumlah Mahasiswa Baru menggunakan Metode Triple Exponential Smoothing. Jurnal Komputer Dan Aplikasi, 08(01), 112–121.
- Arrasyid, R. H., Ratna Kumala, E. S., & Afnan, N. H. (2020). Al-Iqra Medical Journal : Jurnal Berkala Ilmiah Kedokteran. “Potensi Kombinasi Memantin Dan Fitomelatonin Dari Ekstrak Biji Buncis (Phaseolus Vulgaris L) Terenkapsulasi Nano-Chelator Sebagai Inhibitor Agregasi B-Amiloid Dan Antioksidan : Modalitas Kuratif Mutakhir Pada Penderita Alzheimer”, e-ISSN : 2549-225X. Vol. 3 No. 2, Agustus 2020, Hal. 61-70.
- Arwin, L., & Pratiwi, J. N. (2020). Jurnal Ilmu Keperawatan Jiwa Persatuan Perawat Nasional Indonesia Jawa Tengah. “Peran Neuroprotektor Astaxanthin Dalam Pencegahan Penyakit Alzheimer”. e-ISSN 2621-2978 p-I Volume 3 No 1, Hal 47 – 52.
- Dewi, N. R., Desiani, A., & Salamah, F. (2023). Algoritma K-Nearest Neighbor (KNN) dan Single Layer Perceptron (SLP) Untuk Klasifikasi Alzheimer. JTT (Jurnal Teknologi Terapan), 92-101.
- Dhiya, S. S., Ramadhan, I. N., Qurrohman, M. T., Dewi, N., & Ariyanti. (2023). Pencegahan Kerusakan Hati dan Penyakit alzheimer akibat Konsumsi Alkohol Terselubung Dengan Pola Hidup sehat . Jurnal Peduli Masyarakat, 409-416.
- Ghoneim, S. (2019, April 2). Accuracy, Recall, Precision, F-Score & Specificity, which to optimize on? Retrieved Juli 14, 2022,
- Isa, I. G. (2022). Buku Ajar Sistem Pendukung Keputusan. Pekalongan: Penerbit NEM.
- Jihan, L. A., & Pratiwi, N. (2020). Jurnal Ilmu Keperawatan Jiwa. 2020 “Peran Neuroprotektor Astaxanthin Dalam Pencegahan Penyakit Alzheimer”. Volume 3 No 1, Hal 47 – 52, Februari 2020 Persatuan Perawat Nasional Indonesia Jawa Tengah.
- Mortara, A. A., Permatasari, M., Desiani, A., Andriani, Y., & Arhami, M. (2023). Perbandingan Algoritma C4.5 dan Adaptive Boosting dalam Klasifikasi Penyakit . Jurnal Teknologi dan Informasi (JATI), 196-207.
- Peraturan Pemerintah Republik Indonesia. (t.thn.). Peraturan Pemerintah Republik Indonesia Nomor 17.
- Tambunan, P. G., Hutagao, N. M., & Wirawan, S. A. (2020). Al-Iqra Medical Journal : Jurnal Berkala Ilmiah Kedokteran. “Potensi Penggunaan Cairan Sereb Rospinal Amyloid 42 Untuk Diagnosis Dini Penyakit Alzheimer”. e-ISSN : 2549-225X. Vol. 3 No. 2, Agustus 2020, Hal. 56-65.
- Wijaya, H. D., & Dwiasnati, S. (2020). JURNAL INFORMATIKA. “Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat”. Vol.7 No.1 April 2020, Halaman 1~7 ISS.
- Wildah, S. K., Agustiani, S., Ramadhan, M. R., Gata, W., & Nawawi, H. M. (2020). Deteksi Penyakit Alzheimer Menggunakan Algoritma Naïve Bayes dan Correlation Based Feature Selection. JURNAL INFORMATIKA, 166-173.
- Z. Breijyeh and R. Karaman, “Comprehensive Review on Alzheimer ’ s Disease .,” Molecules, vol. 25, p. 5789, 2020, doi: 10.3390/molecules25245789.