

Komparasi Akurasi Algoritme CART dan *Neural Network* Untuk Diagnosis Penyakit *Diabetic Retinopathy*

Comparison of CART Algorithm Accuracy and Neural Network for Diagnosing Diabetic Retinopathy Disease

Pungkas Subarkah¹, Muhammad Marshal Abdallah², Septi Oktaviani Nur Hidayah³

^{1,2,3}Fakultas Ilmu Komputer, Universitas Amikom Purwokerto

e-mail: ¹subarkah@amikompurwokerto.ac.id, ²marshal.abdallah@gmail.com,

³nurhidayahsepti006@gmail.com

Abstrak

Diabetic retinopati atau DR ialah bagian dari suatu komplikasi penyakit diabetes yang dapat menimbulkan kematian untuk pengidapnya. Komplikasi tersebut berbentuk kehancuran bagian pada retina mata. Kandungan glukosa yang tinggi dalam darah merupakan pemicu pembuluh darah kapiler kecil jadi rusak serta bisa menimbulkan kebutaan. Penelitian ini menguji diagnosis penyakit diabetes retinopathy dengan melakukan klasifikasi menggunakan metode data mining. Pada penelitian ini diterapkan algoritme CART dan Algoritme Neural Network untuk mendiagnosis penyakit diabetes retinopati menggunakan dataset diperoleh dari UCI Repository Learning dari Universitas Debreen, Hongaria. Metode validasi dan evaluasi yang digunakan dalam penelitian ini yaitu 10-cross validation dan confusion matrix. Dari hasil pengujian yang didapatkan dengan algoritme CART didapati nilai akurasi yaitu 63.4231% precision 0.64%, recall 0.634%, dan f-measure 0.634% sedangkan pada algoritme Neural Network didapatkan hasil nilai akurasi 72.285% , precision 0.723%, recall 0.723%, dan F-Measure 0.723%. Dari hasil tersebut dapat disimpulkan bahwa algoritme Neural Network lebih tepat guna mendiagnosis penyakit diabetes retinopathy.

Kata kunci— Klasifikasi, Diabetes Retinopathy, Algoritme, CART, Neural Network

Abstract

Diabetic retinopathy or DR is part of a complication of diabetes that can cause death for sufferers. This complication takes the form of damage to the retina of the eye. High glucose content in the blood triggers small capillaries to become damaged and can cause blindness. This study examines the diagnosis of diabetes retinopathy by classifying it using the data mining method. In this study, the CART algorithm and the Neural Network algorithm were applied to diagnose diabetes retinopathy using a dataset obtained from the UCI Repository Learning from Debreen University, Hungary. The validation and evaluation methods used in this study are 10-cross validation and confusion matrix. From the test results obtained with the CART algorithm, the accuracy values are 63.4231%, 0.64% precision, 0.634% recall, and 0.634% f-measure, while the Neural Network algorithm results in an accuracy value of 72.285%, 0.723% precision, 0.723% recall, and F -Measure 0.723%. From these results it can be concluded that the Neural Network algorithm is more appropriate for diagnosing diabetes retinopathy.

Keywords— Classification, Diabetes Retinopathy. Algorithm, CART, Neural Network

1. PENDAHULUAN

Penyakit *diabetic retinopathy* atau DR ialah bagian dari suatu komplikasi penyakit diabetes yang dapat menimbulkan kematian untuk pengidapnya. Komplikasi tersebut berbentuk kehancuran pada bagian retina mata. Kadar glukosa yang tinggi dalam darah adalah penyebab pembuluh darah kapiler kecil menjadi pecah dan dapat menyebabkan kebutaan. Penyakit ini dimulai dengan atau pecahnya pembuluh darah kecil di retina, tumpahan darah yang pada saat itu menyebabkan jaringan menebal, mengembang, dan sekarat [1]. DR biasanya hendak berakibat secara langsung pada terganggunya penglihatan pengidap serta apabila terlambat ditangani hingga hendak menimbulkan pengidap hadapi kebutaan secara permanen. Kebanyakan pengidap melaksanakan pengecekan di dikala keadaan penyakit sudah merambah sesi beresiko, hal ini menyebabkan penyakit ini susah dipulihkan hingga saat ini. Perihal ini disebabkan watak dari penyakit tersebut ini yang tidak menampilkan indikasi yang nampak apabila masih pada sesi dini pada pengidap penderita[2].

DR ialah bagian dari komplikasi mikrovaskular diabetes mellitus mempunyai angka prevalensi yang cukup tinggi. Bersumber pada informasi riset yang dicoba di sebagian wilayah di Indonesia, diperkirakan prevalensi retinopathy sebesar 42,6%. Paling tidak hendak ditemui 24.600 orang dengan retinopathy serta dekat 10% dari jumlah tersebut alami kebutaan. Jumlah ini diperkirakan hendak bertambah pada tahun 2030 dengan ditaksir 98.400 orang mengidap diabetes serta dekat 11.000 orang di antara lain hadapi kebutaan. Mengingat sebagian besar kebutaan akibat diabetik retinopati ialah keadaan yang permanen serta sukar buat disembuhkan, hingga penangkalan serta penyembuhan jadi perihal yang berguna dalam proses pengobatan penyakit ini [3]. Dampak visual yang timbul dari suatu kebocoran retina salah satunya yaitu epistaksis [4].

Prevalensi diabetes mellitus di Indonesia cukup tinggi, saat ini penderita retinopati diabetik 7,6 juta penderita diabetes, namun hanya 39% yang menerima perawatan dari 1% yang berhasil mencapai target pengobatan[5]. Jumlah penderita diabetes di Indonesia diperkirakan akan mencapai 14,1 Juta orang padad tahun 2035[6]. *The DiabCare Asia 2008 Study* yang melibatkan 1785 penderita diabetes mellitus pada 18 pusat kesehatan primer dan sekunder di Indonesia melaporkan 42% penderita diabetes mellitus mengalami komplikasi retinopati[7]. Faktor resiko yang penderita diabetik retinopati antara lain yaitu lama menderita diabetes, kemampuan mengontrol kadar gula darah, ibu hamil yang menderita diabetes, penderita anemia, hiperlipidemia dan perokok[8]. Kendala yang dialami saat ini yaitu keterbatasan dokter spesialis mata untuk mendeteksi dini penyakit diabetes retinopathy sehingga dapat segera diobati sebagai upaya pencegahan.

Analisis dini penyakit DR harus dapat dilakukan, salah satunya dengan memanfaatkan strategi information mining, dimana interaksi teknik *data mining* bermaksud untuk menggambarkan penemuan baru pada diabetes retinopathy atau *dataset* DR. Strategi penambangan informasi menggambarkan siklus yang menggunakan prosedur AI yang dapat diukur, numerik, buatan manusia, dan AI untuk memisahkan dan membedakan data berharga serta informasi terkait dari kumpulan data.[9].

Peran *data mining* diantaranya proses klasifikasi. Dalam siklus pengaturan, umumnya digunakan untuk memutuskan pilihan seperti yang ditunjukkan oleh informasi baru yang diperoleh dari menegaskan informasi masa lalu menggunakan komputasi algoritmik. Teknik klasifikasi dapat diterapkan dalam semua bidang misalnya dalam bidang kesehatan[10], pada bidang kesehatan, metode klasifikasi bisa diterapkan. Dalam *dataset* karakterisasi terdapat satu karakteristik target. Ciri ini akan dicari dari informasi baru berdasarkan kualitas yang berbeda pada informasi yang lalu. Kuantitas karakteristik dapat mempengaruhi eksekusi suatu algoritme[11].

Pada penelitian terdahulu, telah dilakukan diagnosis menggunakan algoritme CART maupun algoritme *neural network* dalam berbagai bidang tidak hanya di bidang kesehatan.. Penelitian ini dilakukan oleh Bahram Choubin, dkk pada tahun 2018, dengan menerapkan metode CART untuk mengetahui sedimen sungai atau SSL adalah masalah penting dalam manajemen

sumber daya air dan lingkungan yang terintegrasi, karena sedimen mempengaruhi kualitas air dan habitat perairan. Tujuan dalam penelitian ini adalah dengan penggunaan metode algoritme CART untuk memperkirakan sedimen sungai (SSL) berdasarkan data hidro-meteorologi dengan nilai akurasi sebesar 77%[12]. Penelitian yang dilakukan dalam bidang kesehatan untuk diagnosis penyakit diabetes mellitus retinopathy, dengan menggunakan algoritme CART menghasilkan nilai akurasi sebesar 63%[9]. Penelitian yang dilakukan Hokino, dkk memperdiksi penyakit tuberkolosis dengan menggunakan algoritme CART dengan nilai akurasi sebesar 76%[13]. Sementara penelitian menggunakan metode algoritme *neural network* digunakan pada penelitian Derisma mengaplikasikan metode random forest, *neural network* dan naïve bayes untuk dibandingkan. Tujuannya untuk melakukan diagnosis penyakit kanker payudara berbasis komputasi yang dapat menghasilkan kondisi kanker seseorang berdasarkan akurasi algoritme. Hasil pengujian didapatkan nilai akurasi algoritme *neural network* sebesar 97.9%[14]. Penelitian selanjutnya adalah penelitian yang dilakukan oleh Jafar, dkk, pada penelitian ini bertujuan untuk mengklasifikasikan pasien dengan atribut yang dipilih yang meningkat akurasi diagnosis penyakit kanker. Untuk mengurangi kesalahan (yaitu rasio positif palsu) dan meningkatkan akurasi diagnosis penyakit dengan efisiensi klasifikasi yang lebih tinggi dan waktu klasifikasi yang lebih rendah, algoritma klasifikasi *ensemble neural network* teroptimisasi terbobot. Hasil yang diperoleh dalam penelitian ini menggunakan algoritme *neural network* sebesar 65%[15].

Pemilihan kedua algoritme dalam mendiagnosis penyakit retinopathy, karena dari kelebihan kedua algoritme ini sesuai dengan *dataset* yang digunakan. Kelebihan algoritme CART adalah suatu metode yang mempunyai sifat nonparametrik yang tepat untuk data yang berjenis *numeric*, untuk transformasi variabel yang independen, algoritme CART tidak memerlukan variabel yang akan terlebih dahulu karena sudah menghasilkan invariant guna transformasi variabel[16]. Pada klasifikasi algoritme CART (*Classification and Regression Trees*), sebuah record akan diklasifikasikan ke dalam salah satu dari sekian klasifikasi yang tersedia pada variabel tujuan berdasarkan nilai-nilai variabel prediktornya [17]. Sedangkan pada algoritme *neural network* atau jaringan syaraf tiruan merupakan algoritma klasifikasi yang mampu menukar struktur algoritmenya agar mampu memecahkan suatu masalah berdasarkan hasil informasi yang berasal dari eksternal maupun internal yang melewati jaringan tersebut. *Neural Network* dapat memodelkan hubungan yang kompleks antara input dan output untuk menemukan pola dari data[18]. Selain itu juga algoritme *neural network* merupakan strategi yang sangat mudah dan sukses untuk menangani karakterisasi[19].

Peneliti ini menguji keberhasilan *dataset* diabetes retinopathy dengan cara melakukan klasifikasi diagnosis pasien dengan menerapkan *machine learning* sehingga luarannya mendapatkan sebuah informasi dari pola yang ada dan mengekstraksi informasi, algoritme klasifikasi yang digunakan yaitu algoritme *Classification and Regression Trees* (CART) dan algoritme *neural network* dengan metode algoritme sebagai yang menghasilkan diagnosis penyakit diabetes retinopathy. Data yang digunakan dalam penelitian ini yaitu data primer dan sekunder yang diperoleh dari data public UCI WEKA yang terdiri dari 1151 data dan memiliki 19 atribut. Hasil akhir dari pemeriksaan ini diandalkan untuk memberikan referensi mengenai ketepatan simpulan retinopati diabetes dengan menggunakan perhitungan algoritme CART dan algoritme *neural network*.

2. METODE PENELITIAN

2.1 Dataset Diabetes Retinopathy

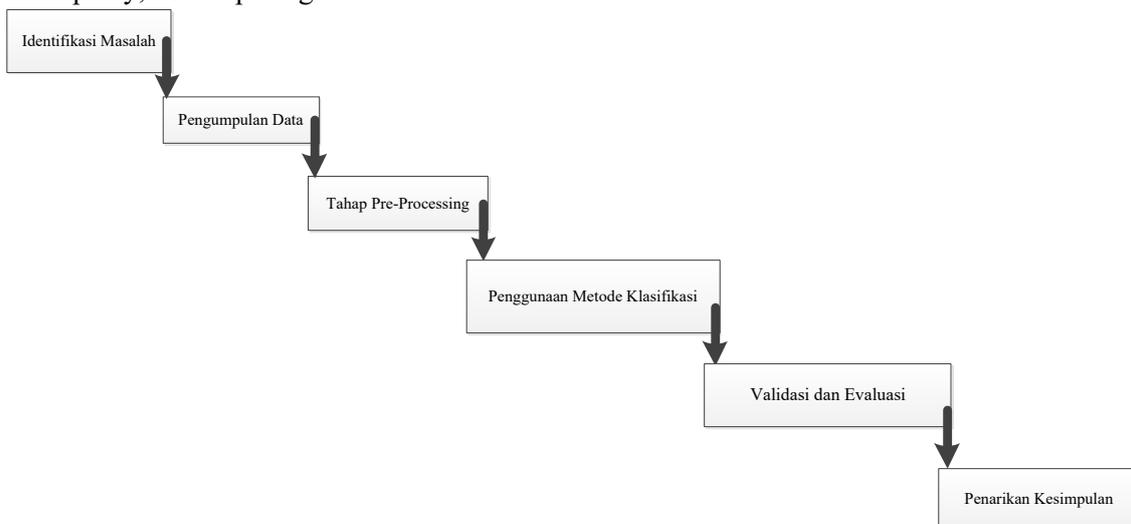
Data penelitian bersumber dari UCI repository *machine learning*. Data diabetic retinopathy serta diperoleh Universitas Debrecen, Hongaria. *Dataset* terdiri dari 1151 *record*, yang mempunyai 19 atribut (18 atribut serta 1 atribut sasaran). Atribut sasaran terdiri 2 ialah wajar diabet retinopati serta terindikasi diabet retinopati.

Tabel 1 *Attribute* pada *dataset* Diabetes Retinopathy

Nama	Informasi
<i>Attribute 0</i>	Hasil biner pada mutu evaluasi
<i>Attribute 1</i>	Hasil berpasangan <i>pra-skrining</i>
<i>Attribute 2 - 7</i>	Hasil identifikasi yang disampaikan MA
<i>Attribute 8 - 15</i>	Hasil normalisasi pada sel rusak yang keluar dari pembuluh darah ke dalam jaringan mata
<i>Attribute 16</i>	Jarak Euclidean ke titik fokus yang lebih tua dan titik fokus dari pelat optik
<i>Attribute 17</i>	Jarak lingkaran optik
<i>Attribute 18</i>	Hasil biner kategorisasi berbasis AM/FM
<i>Target_Attribute 19</i>	Label kelas. 1 = mengandung terindikasi DR (label akumulatif untuk kelas 1, 2, 3), kelas. 0 = tidak ada gejala DR.

2. 2 Desain Penelitian

Gambar 1. Yaitu proses tahapan untuk mendiagnosis penyakit diabetes retinopathy. Hal ini bertujuan untuk memudahkan dalam proses penelitian. Desain penelitian ini menggunakan algoritme CART dan algoritme Neural Network untuk mendiagnosis penyakit diabetes retinopathy, dilihat pada gambar dibawah ini:



Gambar 1 Tahapan Riset

2. 2.1 Identifikasi Masalah

Tahapan ini dilakukan untuk menentukan masalah dan teknik yang tepat agar fokus yang tepat dapat diselesaikan untuk tujuan antisipasi penyakit retinopati diabetes.

2.2.2 Pengumpulan Data

Tahapan pengumpulan memakai informasi sekunder yang digunakan, diperoleh dari *database UCI Repository*. *Dataset* terdiri dari 1151 *record* yang memiliki 19 atribut (18 atribut dan 1 atribut sasaan). Atribut sasaran terdiri 2 *output* yaitu wajar diabetes retinopathy dan terindikasi diabet retinopathy[20]. Pada 2 *output* pada *class* atribut ini akan digunakan sebagai *output* data penyakit diabet retinopati. Berikut jumlah dataset berdasarkan class, dapat dilihat pada tabel 2.

Tabel 2 Jumlah *Dataset* berdasarkan *Class*

No	Klasifikasi	Jumlah <i>Record Dataset</i>
1	<i>Contains Signs</i>	611
2	<i>No Signs</i>	540
Jumlah		1151

2.2.3 Tahap *Pre-Processing*

Tahapan ini dilakukan suatu proses seleksi data yang bertujuan untuk mendapatkan data yang bersih dan siap digunakan sebagai bahan penelitian. Tahap *pre-processing* data meliputi identifikasi dan pemilihan atribut (*attribute identification and selection*), penanganan nilai pada atribut yang hilang atau atribut yang tidak lengkap (*handling missing values*) dan proses diskretisasi nilai.

2.2.4 Penggunaan Metode Klasifikasi

Metode yang digunakan dalam penelitian ini yaitu algoritme CART dan algoritme Neural Network, dari hasil *confusion matrix* dapat dihitung nilai *precision*, *recall* dan *F-Measure*. Berikut ini merupakan rincian perhitungan *confusion matrix* dapat dilihat pada gambar 2., sebagai berikut[21] :

Tabel 3 *Confusion Matrix*

<i>Correct classification</i>	<i>Classification as</i>	
	<i>Yes</i>	<i>No</i>
<i>Yes</i>	<i>True Positives (TP)</i>	<i>False Negatives (FN)</i>
<i>No</i>	<i>False Positives (FP)</i>	<i>True Negatives (TN)</i>

Adapun rumus penghitungan dari sebuah *precision*, *recall* dan *F-Measure* dari sebuah nilai *accuracy*[21], dapat dilihat dari persamaan dibawah ini :

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$F-Measure = \frac{2 \times precision \times recall}{precision + recall} \quad (3)$$

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (4)$$

2.2.5 Validasi dan Evaluasi

Dalam tahapan ini dilakukan validasi dan evaluasi serta pengukuran keakuratan hasil yang dicapai oleh model menggunakan teknik yang terdapat dalam aplikasi weka yaitu *confusion matrix* dan *10-fold cross validation*.

2.2.6 Penarikan Kesimpulan

Pada tahapan ini yaitu menyimpulkan hasil yang diperoleh dari penelitian dengan menggunakan algoritme CART dan algoritme Neural Network yang memberikan hasil akurasi terbaik untuk mendiagnosis penyakit diabetes retinopathy yang berdasarkan nilai *precision*, *recall*, dan *F-Measure* dari tiap masing-masing algoritme dengan tingkat klasifikasi [22] sebagai berikut :

- Excellent classification* = 0.90 – 1.00
- Good classification* = 0.80 – 0.90
- Fair classification* = 0.70 – 0.80
- Poor classification* = 0.60 – 0.70

e. $Failure = 0.50 - 0.60$

3. HASIL DAN PEMBAHASAN

3.1 Identifikasi Masalah

Tahapan ini dilakukan untuk menentukan masalah dan teknik yang tepat agar fokus yang tepat dapat diselesaikan untuk tujuanantisipasi penyakit retinopati diabetes. Dari hasil tersebut penulis melakukan penelitian dengan menggunakan algoritme CART dan *Neural Network* untuk menganalisis *dataset*.

3.2 Pengumpulan Data

Dalam penelitian penulis ini, data yang akan digunakan dalam penelitian ialah dengan mengambil dari *repository* diambil dari *database UCI Repository*. *Dataset* terdiri dari 1151 *record*, yang memiliki 19 *record*, yang memiliki 19 *attribute* (18 *attribute* dan 1 *attribute target*). *Attribute target* terdiri dari dua *output* yaitu normal diabetes retinopathy dan terindikasi diabetes retinopathy[20]. Kedua *output class attribute* ini akan digunakan sebagai *output* data penyakit diabetes retinopathy. Data ini harus diolah terlebih dahulu yakni melalui tahapan *pre-processing*, dimana tahapan ini bertujuan untuk menyesuaikan atribut-atribut yang akan digunakan dalam mengolah *dataset* ini akan memberikan hasil yang berbeda-beda pada hasil nilai akurasi dalam mendiagnosis penyakit diabetes retinopathy serta hal ini berkaitan dengan ada atau tidaknya kelengkapan nilai pada setiap atribut. Atribut-atribut pada *dataset* sebagai berikut :

- a. Atribut 0
Hasil biner pada mutu evaluasi.
- b. Atribut 1
Hasil berpasangan dari *pre-screening*.
- c. Atribut 2 – 7
Hasil identifikasi yang disampaikan MA.
- d. Atribut 8 – 15
Sel rusak yang keluar dari pembuluh darah ke dalam jaringan mata yang diwakili oleh seperangkat, yang menghasilkan jumlah piksel yang membentuk lesi pada bola mata
- e. Atribut 16
Jarak euclidean dari pusat makula dan pusat cakram optik untuk memberikan informasi penting kondisi pasien.
- f. Atribut 17
Jarak lingkaran pada optik
- g. Atribut 18
Hasil biner dari klasifikasi berbasis AM/FM.
- h. Atribut Kelas
Label kelas. 1 = mengandung tanda-tanda DR (label akumulatif untuk kelas Messidor 1, 2, 3), 0 = tidak ada tanda-tanda DR.

Atribut diatas yang akan digunakan oleh peneliti sebagai bahan eksplorasi dan diubah terlebih dahulu disesuaikan dengan jenis data yang dapat dibaca oleh aplikasi Weka.

3.3 Tahap *Pre-Processing*

Tahapan ini, dilakukan identifikasi atribut serta penyeleksian dari *dataset* diabetes retinopathy agar data yang diperoleh ialah data yang benar-benar siap untuk digunakan. Tabel 4., ialah *dataset* diabetes reitnopathy yang telah dilakukan penyesuaian atribut untuk digunakan pada aplikasi Weka.

Tabel 4 Data *Pre-processing*

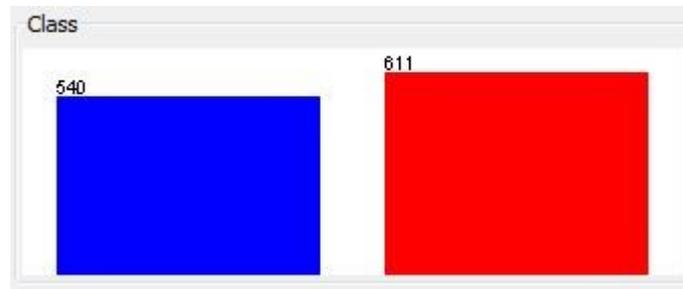
Data Orisinal	Data Hasil <i>Pre-Processing</i>	Deskripsi
1	1	Hasil biner dari mutu evaluasi
1	1	Hasil berpasangan dari <i>pre-screening</i>
56	56	Hasil dari deteksi MA
55	55	
46	46	
41	41	
37	37	
28	28	
6.960789	6.960789	
2.747318	2.747318	
1.544124	1.544124	
0.631409	0.631409	
0.201806	0.201806	
0.207558	0.207558	
0.160528	0.160528	
0.133009	0.133009	
0.11772	0.11772	
0.522871	0.522871	Jarak euclidean dari pusat makula dan pusat cakram optik untuk memberikan informasi penting mengenai kondisi pasien.
1	1	Hasil biner dari klasifikasi berbasis AM / FM
1	1	Label kelas. 1 = mengandung tanda-tanda DR (label akumulatif untuk kelas Messidor 1, 2, 3), 0 = tidak ada tanda-tanda DR
<i>Contains Signs</i>	0	Class, <i>Contains Signs</i>
<i>Signs</i>	1	Class, <i>Signs</i>

Dari hasil analisis tabel 4., tentang hasil data *pre-processing dataset* diabetes retinopathy yang telah dilakukan, bahwa semua atribut yang digunakan nilainya lengkap atau tidak *missing value*.

3.4 Penggunaan Metode Klasifikasi

Setelah melalui tahap *pre-processing* selanjutnya *dataset* mulai diolah dengan menggunakan aplikasi Weka. Tahapan ini digunakan dan bertujuan untuk menghasilkan *confusion matrix* berdasarkan metode evaluasi *10-fold cross validation*, dimana *dataset* tersebut dibagi menjadi 10 *subsets* (9 *subsets* sebagai *training sets* dan 1 *subsets* sebagai *testing set*) dengan 10 kali iterasi / pengulangan. Adapun *classifier* yang digunakan untuk uji coba pada dataset adalah SimpleCart (*Classification And Regression Trees*) dan MultiLayerPerceptron (Neural Network). Pada gambar 2 merupakan visualisasi data secara keseluruhan yang telah dijabarkan sesuai dengan atribut yang digunakan serta sudah melewati tahapan *pre-processing*. Selanjutnya data-data tersebut diklasifikasi menggunakan algoritme CART (*SimpleCart*) dan algoritme Neural Network (*MultiLayerPerceptron*) dalam penggunaan pada aplikasi Weka yang

menghasilkan pola klasifikasi. Berikut gambar 2 tentang visualisasi informasi *dataset*, sebagai berikut :



Gambar 2 Visualisasi Informasi *dataset*

Keterangan dari gambar 2 diatas, sebagai berikut:

- Grafik dengan warna biru merupakan *class* terindikasi diabetes retinopathy.
- Grafik dengan warna merah merupakan *class* normal atau tidak terindikasi diabetes retinopathy.

Pada algoritme CART menghasilkan nilai akurasi sebesar 63.4231%. Nilai akurasi tersebut didapatkan dari nilai *precision*, *recall*, dan *F-Measure*. Berikut proses perhitungannya dapat dilihat pada tabel 5. *Confusion* Kelas “0”.

Tabel 5 *Confusion* Kelas “0”

368 (<i>True Positive</i>)	172 (<i>False Negative</i>)
249 (<i>False Positive</i>)	362 (<i>True Negative</i>)

Pada tabel 5. *Confusion* Kelas “0”, merupakan penjabaran dari nilai *confusion matrix* yang dapat dilihat pada tabel tabel 5. *Confusion* Kelas “0”.

Kelas “0”

$$Precision = \frac{TP}{TP+FP} = \frac{368}{368+249} = 0,596$$

$$Recall = \frac{TP}{TP+FN} = \frac{368}{368+172} = 0,681$$

$$F-Measure = \frac{2 \times precision \times recall}{precision + recall} = \frac{2 \times 0,596 \times 0,681}{0,596 + 0,681} = \frac{2 \times 0,596 \times 0,681}{0,596 + 0,681} = \frac{2 \times 0,405876}{1,277} = \frac{0,811752}{1,277} = 0,636$$

Tabel 6 *Confusion* Kelas “1”

362 (<i>True Positive</i>)	249 (<i>False Negative</i>)
172 (<i>False Positive</i>)	368 (<i>True Negative</i>)

Kelas “1”

$$Precision = \frac{TP}{TP+FP} = \frac{362}{362+272} = 0,678$$

$$Recall = \frac{TP}{TP+FN} = \frac{362}{362+172} = 0,592$$

$$F\text{-Measure} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} = \frac{2 \times 0,678 \times 0,592}{0,678 + 0,592} = \frac{2 \times 0,401376}{1,27} = \frac{0,802752}{1,27} = 0,632$$

Setelah melakukan perhitungan nilai precision, recall dan F-measure dari masing-masing kelas, maka langkah selanjutnya ialah menghitung nilai rata-rata dari kelas tersebut. Rumus untuk menghitung nilai *Weighted Avg*. Sebagai Berikut:

$$\text{Weighted Avg (precision)} = \frac{0,596 \times A + 0,678 \times B}{1151} = \frac{0,596 \times 540 + 0,678 \times 611}{1151} = \frac{321,84 + 420,368}{1151} = \frac{742,208}{1151} = 0,64$$

$$\text{Weighted Avg (recall)} = \frac{0,681 \times A + 0,592 \times B}{1151} = \frac{0,681 \times 540 + 0,592 \times 611}{1151} = \frac{367,74 + 361,712}{1151} = \frac{729,452}{1151} = 0,634$$

$$\text{Weighted Avg (F-Measure)} = \frac{0,636 \times A + 0,632 \times B}{1151} = \frac{0,636 \times 540 + 0,632 \times 611}{1151} = \frac{343,44 + 386,152}{1151} = \frac{729,592}{1151} = 0,634$$

Hasil diatas adalah estimasi nilai ketepatan yang berasal dari *confusion matrix* hasil algoritme CART yang disajikan pada tabel 7. Berikut ini nilai presisi berdasarkan *confusion matrix*, dapat dilihat pada tabel 7.

Tabel 7 Nilai Akurasi Berdasarkan *Confusion Matrix* Algoritme CART

Class	Precision	Recall	F-Measure
0	0,596	0,681	0,636
1	0,678	0,592	0,632
Weighted Avg	0,64	0,634	0,634

Pengujian pada dataset diabetes retinopathy dengan algoritme CART, didapatkan hasil klasifikasi menggunakan aplikasi Weka 3.6 dan waktu yang diperlukan yaitu 0.63 *second*.

Selanjutnya algoritme Neural Network nilai akurasi yang dihasilkan sebesar 72.285%. Perhitungan *precision*, *recall* dan *F-Measure* tahap awal dari nilai hasil akurasi. Proses perhitungannya dapat dilihat pada tabel 8. *Confusion* Kelas "0".

Tabel 8 *Confusion* Kelas "0"

383 (<i>True Positive</i>)	157 (<i>False Negative</i>)
162 (<i>False Positive</i>)	449 (<i>True Negative</i>)

Tabel 8., diatas yaitu uraian *confusion matrix* yang dilihat pada tabel 8. *Confusion* kelas "0".

Kelas "0"

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{383}{383 + 162} = 0.703$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{383}{383 + 157} = 0.709$$

$$F\text{-Measure} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} = \frac{2 \times 0,703 \times 0,709}{0,703 + 0,709} = \frac{2 \times 0,49842}{1,412} = \frac{0,99684}{1,412} = 0.706$$

Tabel 9 *Confusion* Kelas “1”

449 (<i>True Positive</i>)	162 (<i>False Negative</i>)
157 (<i>False Positive</i>)	383 (<i>True Negative</i>)

Kelas “1”

$$Precision = \frac{TP}{TP+FP} = \frac{449}{449+157} = 0.741$$

$$Recall = \frac{TP}{TP+FN} = \frac{449}{449+162} = 0.735$$

$$F-Measure = \frac{2 \times precision \times recall}{precision + recall} = \frac{2 \times 0.741 \times 0.735}{0.741 + 0.735} = \frac{2 \times 0.544635}{1.476} = \frac{1.08927}{1.476} = 0.738$$

Dari hasil nilai *precision*, *recall* dan *F-Measure* yang dihasilkan dari masing-masing kelas, selanjutnya menghitung nilai rata-rata dari kelas tersebut (*Weighted Avg*). Menghitung *Weighted Avg* menggunakan rumus, sebagai berikut :

$$Weighted\ Avg\ (precision) = \frac{0.703 \times A + 0.741 \times B}{1151} = \frac{0.703 \times 540 + 0.741 \times 611}{1151} = \frac{379.62 + 453.751}{1151} = \frac{833.371}{1151} = 0.723$$

$$Weighted\ Avg\ (recall) = \frac{0.709 \times A + 0.735 \times B}{1151} = \frac{0.709 \times 540 + 0.735 \times 611}{1151} = \frac{382.86 + 499.085}{1151} = \frac{881.945}{1151} = 0.723$$

$$Weighted\ Avg\ (F-Measure) = \frac{0.706 \times A + 0.738 \times B}{1151} = \frac{0.706 \times 540 + 0.738 \times 611}{1151} = \frac{381.24 + 450.918}{1151} = \frac{832.158}{1151} = 0.723$$

Hasil diatas merupakan perhitungan nilai akurasi berdasarkan *confusion matrix* hasil algoritme neural network yang disajikan pada tabel 10. Nilai akurasi berdasarkan nilai *confusion matrix* algoritme neural network, sebagai berikut :

Tabel 10 Nilai Akurasi Berdasarkan *Confusion Matrix* Algoritme Neural Network

<i>Class</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
0	0.703	0.706	0.709
1	0.741	0.735	0.738
<i>Weighted Avg</i>	0.723	0.723	0.723

Pengklasifikasian untuk menguji *dataset* diabetes retinopathy dengan algoritme neural network menggunakan aplikasi weka 3.6 dan waktu yang dibutuhkan sejumlah 4.79 *second*. Dari hasil perhitungan pada algoritme CART dan algoritme neural network. Berikut perbedaan hasil dari kedua algoritme tersebut, yang disajikan pada tabel 11. Perbandingan hasil akurasi algoritme CART dan algoritme Neural Network.

Tabel 11 Perbandingan Hasil Akurasi Algoritme CART dan algoritme Neural Network

Algoritme	Hasil akurasi	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	Waktu
CART	63.4231%.	0.64	0.634	0.634	0.63 <i>second</i>
Neural Network	72.285%.	0.723	0.723	0.723	4.79 <i>second</i>

Dari penjabaran pada Tabel 11., bahwa nilai akurasi algoritme neural network lebih baik dibandingkan dengan algoritme CART dalam mendiagnosis penyakit diabetes retinopathy.

3.5 Validasi dan Evaluasi

Untuk mengukur tingkat akurasi dari algoritme klasifikasi yang digunakan yaitu menggunakan *confusion matrix* yang disajikan pada tabel 12 dan tabel 13. Pada tabel 12 merupakan tabel yang dihasilkan oleh *confusion matrix* dari pengujian *dataset diabetes retinopath* menggunakan algoritme CART dengan metode *10-fold cross validation*

Tabel 12 *Confusion Matrix* Algoritme CART

Class	0	1
0	383	157
1	162	449
1151	545	606

Dari tabel 12 terlihat bahwa jumlah data hasil bentukan *rule* yang tidak terindikasi penyakit diabetes retinopathy yang sama dengan data *testing* yang juga tidak terindikasi penyakit diabetes retinopathy sebanyak 383. Kemudian, jumlah data hasil bentukan *rule* yang terindikasi penyakit diabetes retinopathy dengan data testing yang tidak terindikasi penyakit diabetes retinopathy sebanyak 157. Selanjutnya, jumlah data hasil bentukan *rule* yang terindikasi penyakit diabetes retinopathy dan data testing yang tidak terindikasi penyakit diabetes retinopathy sebanyak 162. Terakhir, jumlah data hasil bentukan *rule* yang terindikasi penyakit diabetes retinopathy yang sama dengan data testing bentukan *rule* yang juga terindikasi penyakit diabetes retinopathy sebanyak 449. Sedangkan pada tabel 13 merupakan tabel yang dihasilkan oleh *confusion matrix* dari pengujian *dataset diabetes retinopathy* menggunakan algoritme neural network dengan metode *10-fold cross validation*.

Tabel 13 *Confusion Matrix* Algoritme Neural Network

Class	0	1
0	368	172
1	249	362
1151	617	534

Dari tabel 13 terlihat bahwa jumlah data hasil bentukan *rule* yang tidak terindikasi penyakit diabetes retinopathy yang sama dengan data *testing* yang juga tidak terindikasi penyakit diabetes retinopathy sebanyak 368. Kemudian, jumlah data hasil bentukan *rule* yang terindikasi penyakit diabetes retinopathy dengan data testing yang tidak terindikasi penyakit diabetes retinopathy sebanyak 172. Selanjutnya, jumlah data hasil bentukan *rule* yang terindikasi penyakit diabetes retinopathy dan data testing yang tidak terindikasi penyakit diabetes retinopathy sebanyak 249. Terakhir, jumlah data hasil bentukan *rule* yang terindikasi penyakit diabetes retinopathy yang sama dengan data testing bentukan *rule* yang juga terindikasi penyakit diabetes retinopathy sebanyak 362.

3.6 Penarikan Kesimpulan

Dari hasil perhitungan yang telah dilakukan pada kedua algoritme, maka didapatkan hasil akurasi dari masing-masing algoritme yaitu 63.4231% dengan nilai *precision* 0.64%, *Recall* 0.634%, dan *F-Measure* 0.634%. pada algoritme CART dan 72.285% pada algoritme Neural Network dengan nilai *precision* 0.723%, *Recall* 0.723%, dan *F-Measure* 0.723% disajikan pada tabel 14.

Tabel 14 Komparasi Akurasi Algoritme CART dan algoritme Neural Network

Algoritme	Hasil akurasi	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	Waktu
CART	63.4231%.	0.64	0.634	0.634	0.63 <i>second</i>
Neural Network	72.285%.	0.723	0.723	0.723	4.79 <i>second</i>

4. KESIMPULAN

Berdasarkan hasil penelitian didapati bahwa ini hasil diagnosis menggunakan algoritme neural network memiliki akurasi sebesar 72.285% lebih baik daripada algoritme CART 63.4231%., pada dataset diabetes retinopathy, dengan nilai selisih 8.862%. Perihal ini sebagai referensi pihak rumah sakit dalam teknik pemilihan diagnosis penyakit diabetes retinopathy sehingga mendapatkan luaran informasi yang sesuai.

5. SARAN

Berdasarkan kesimpulan diatas, terdapat beberapa saran yang telah dirangkum untuk dapat dikembangkan pada penelitian selanjutnya, sebagai berikut :

- a. Meningkatkan hasil akurasi yang lebih baik pada diagnosis penyakit diabet retinopathy dengan menambahkan algoritme klasifikasi.
- b. Melakukan proses seleksi fitur *data mining* pada algoritme dalam klasifikasi penyakit diabet retinopathy.
- c. Melakukan optimasi pada algoritme *data mining* guna meningkatkan nilai akurasi yang lebih akurat.

UCAPAN TERIMA KASIH

Terima kasih kepada LPPM Universitas Amikom Purwokerto yang telah mendanai penelitian ini dengan program Hibah Penelitian Amikom Purwokerto, dan kami ucapkan terima kasih juga kepada Fakultas Ilmu Komputer - Universitas Klabat yang telah memberikan kesempatan sehingga artikel ilmiah ini dapat diterbitkan pada Jurnal Cogito.

DAFTAR PUSTAKA

- [1] B. S. A. Ayu. Ratna Gitasari. Hidayat. "Klasifikasi Penyakit Diabetes Retinopati Berdasarkan Citra Digital dengan menggunakan Metode Wavelet dan Support Vector Machine.." Universitas Telkom. 2013.
- [2] I. M. Dewi. "Mengenal Lebih Jauh Retinopathy Diabetik." 2018. [Online]. Available: <https://www.columbiaasia.com/indonesia/health-articles/mengenal-lebih-jauhretinopathy-diabetik>.
- [3] "Diabetik Retina : Pergeseran Paradigma Kebutaan pada Era Milenial." [Online]. Available: <http://www.yankes.kemkes.go.id/read-retinopati-diabetik-pergeseran>

- paradigma-kebutaan-pada-era-milenial-5984.html. Diakses Hari Selasa. 03 Maret 2020. Pukul 14.57 WIB.
- [4] S. Faust. O.. Acharya. R. U.. Ng. E. Y. K.. Ng. K. H. and J.S. “Algorithms A for the Automated Detection of A Diabetic Retinopathy Using Digital Fundus A Image : A Review.” *JaMed Syst.* 2010.
- [5] Novonordisk. *The Blueprint for Change Programme. Where Economics and Health Meet : Changing Diabetes In Indonesia.* Jakarta. 2013.
- [6] IDF.. “Diabetes Melitus Atlas.” *Sixth Edition.* 2013. [Online]. Available: [http://www.idf.org/Diabetes Melitusatlas/download-book](http://www.idf.org/Diabetes_Melitusatlas/download-book). Di akses Pada Hari Selasa. 03 Maret 2020. Pukul 15.04 WIB.
- [7] P. Soewondo. S. Soegondo. K. Suastika. A. Pranoto. Soeatmadji. and A. D.W.. Tjokrorawiro. “The DiabCare Asia 2008 Study – Out Comes On Control and Complication of Type 2 Diabetic Patients in Indonesia.” *MedJ J Indones.* vol. 19. no. 4. pp. 235–44. 2010.
- [8] A. . Khurana. *Comprehensive Ophthalmology 4th ed.* New Delhi: New Age International. 2007.
- [9] P. Subarkah. “Penerapan Algoritma Klasifikasi Classification And Regression Trees (Cart) Untuk Diagnosis Penyakit Diabetes Retinopathy.” *MATRIK J. Manajemen. Tek. Inform. dan Rekayasa Komput.* vol. 19. no. 2. pp. 294–301. 2020.
- [10] and D. . S. Christobel. Angeline. “An Empirical Comparison of Data Mining Classification Methods.” 2011. pp. 24– 28.
- [11] E. Prasetyo. *Data Mining Konsep dan Aplikasi Menggunakan Matlab.* Yogyakarta: Andi Offset. 2012.
- [12] B. Choubin. H. Darabi. O. Rahmati. F. Sajedi-hosseini. and B. Kløve. “Science of the Total Environment River suspended sediment modelling using the CART model : A comparative study of machine learning techniques.” *Sci. Total Environ.* vol. 615. pp. 272–281. 2018.
- [13] V. Hokino *et al.*. “International Journal of Medical Informatics Development of CART model for prediction of tuberculosis treatment loss to follow up in the state of São Paulo . Brazil : A case – control study.” vol. 141. no. December 2019. 2020.
- [14] Derisma and F. Febrian. “Perbandingan Teknik Klasifikasi Neural Network . Support Vector Machine . dan Naive Bayes dalam Mendeteksi Kanker Payudara.” vol. 7. no. 1. pp. 53–62. 2020.
- [15] J. A. Alzubi. B. Bharathikannan. and S. Tanwar. “Boosted Neural Network Ensemble Classification For Lung Cancer Disease Diagnosis.” *Appl. Soft Comput. J.* vol. 80. pp. 579–591. 2019.
- [16] R. Timofeev. *Classification and Regression Trees (CART) Theory and Applications.* Berlin: Humboldt University. 2004.

-
- [17] D. Susanto. S., dan Suryadi. *Pengantar Data Mining*. Yogyakarta: Andi Offset. 2010.
- [18] A. A. Lorena. S., Ginting. B. R., & Permana. “Penerapan Data Mining Untuk Klasifikasi Kelayakan Nasabah Dalam Pengajuan Kredit Menggunakan Matlab.” 2016. pp. 1–10.
- [19] Imaniar Ramadhani, Jondri, and Rismila. “Prediction Of Multi Currency Exchange Rates Using Correlation Analysis And Backpropagation..” *E-Health Bioeng. Conf. (EHB)*.. 2017.
- [20] “Machine Learning Repository.” 2018. [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/Diabetic+Retinopathy+Debreccen+Data+Set#>.
- [21] M. Han. J., & Kamber. *Data Mining Concepts. Model and Techniques 2nd Edition*. San Fransisco: Elsevier. 2006.
- [22] F. Gorunescu. *Data mining Concepts. Models and Techniques*. Verlen Berlin: Springer. 2011.