



## Penerapan *Decision Tree* dan *Neural Network* untuk Prediksi *Severity Level* Pada Kasus Hipertensi di RSUD Khidmat Sehat Afiat (KISA) Depok

Arlie Rona Rafidah<sup>1</sup>, Mieke Nurmalasari<sup>1\*</sup>, Hosizah<sup>1</sup>, Dhiar Niken Larasati<sup>2</sup>

<sup>1</sup>Program Studi Manajemen Informasi Kesehatan, Universitas Esa Unggul  
*arlien.rona.r@gmail.com, mieke@esaunggul.ac.id, hosizah@esaunggul.ac.id*

<sup>2</sup>Badan Pusat Statistik  
*dhiarniken@bps.go.id*

### Keywords:

*Decision Tree, Classification, Neural Network, Severity Level*

### ABSTRACT

*Severity Level* is a component of the INA-CBGs code that indicates the severity of a case during treatment, influencing the INA-CBGs tariff rate. The aim of this study is to predict the severity level by implementing decision tree and neural network algorithms using Orange Data Mining. This research was conducted at Khidmat Sehat Afiat Regional Public Hospital in Depok City, utilizing 162 inpatient claim records with primary diagnoses of Hypertension and Hypertensive Heart Disease and secondary diagnoses of CHF, CKD, or both. Prediction was carried out on 114 testing data and 48 training data. Claim data were analyzed using Decision Tree and Neural Network, with testing results showing the highest score in neural network performance with an AUC of 62.5%, CA of 57%, F1 of 56.2%, and precision of 57.7%. Based on calculations from the confusion matrix, the neural network demonstrated better performance, with accuracy at 57.89%, precision at 65.6%, and recall at 80.76%. These results suggest that the neural network is recommended for predicting the severity level of hypertension cases at Khidmat Sehat Afiat Hospital, as it achieves higher accuracy than the Decision Tree.

### Kata Kunci

*Decision Tree, Klasifikasi, Neural Network, Severity Level*

### ABSTRAK

*Severity Level* adalah salah satu komponen kode INA CBGs yang menggambarkan tingkat keparahan kasus pada masa perawatan yang mempengaruhi besarnya tarif INA CBGs. Tujuan dari penelitian ini adalah melakukan prediksi *severity level* dengan implementasi *decision tree* dan *neural network* menggunakan *Orange Data Mining*. Penelitian ini dilakukan di RSUD Khidmat Sehat Afiat Kota Depok menggunakan 162 data dari laporan klaim rawat inap dengan diagnosa utama Hipertensi dan *Hypertensive Heart Disease* dan diagnosa sekunder CHF, CKD, dan keduanya. Prediksi dilakukan pada 114 data testing dan 48 data training. Data klaim dianalisa dengan *decision tree* dan *neural network* dihasilkan nilai testing dengan skor tertinggi pada *neural network* dengan AUC 62,5%, CA 57%, F1 56,2%, dan precision 57,7%. Berdasarkan perhitungan dari *confusion matrix* didapatkan *neural network* memiliki kinerja lebih baik dengan nilai akurasi 57,89%, *precision* 65,6%, *recall* 80,76%. Berdasarkan hasil tersebut maka *neural network* direkomendasikan untuk prediksi *severity level* kasus hipertensi di RSUD Khidmat Sehat Afiat karena memiliki akurasi lebih tinggi dibandingkan *Decision Tree*.

### Korespondensi Penulis:

Mieke Nurmalasari,

Submitted : 31-05-2024; Accepted : 10-11-2024;  
Published : 08-12-2024

Copyright (c) 2024 The Author (s)

Program Studi Manajemen Informasi Kesehatan,  
Universitas Esa Unggul, Indonesia,  
Jl. Arjuna Utara No.9, Duri Kupa, Jakarta Barat 11510  
Telepon: +6281297987035  
Email: mieke@esaunggul.ac.id

*This article is distributed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0)*

## 1. PENDAHULUAN

Sistem JKN yang diimplementasikan di Indonesia saat ini adalah BPJS Kesehatan sebagai badan hukum jaminan sosial. Jaminan sosial bertujuan untuk menjamin seluruh rakyat agar dapat memenuhi kebutuhan dasar hidupnya yang layak [1]. Pembiayaan program BPJS untuk pembayaran rumah sakit menggunakan sistem pembayaran *Indonesian Case Base Groups* (INA CBGs) dengan skema pembayaran *casemix*. *Casemix* merupakan skema pembayaran dengan pengklasifikasian penyakit berdasarkan kombinasi ciri klinis yang sama atau mirip dengan biaya perawatan disuatu rumah sakit (RS) [2]. Proses *grouping* dari data klinis (diagnosa utama, diagnosa sekunder, prosedur) dan data pasien (identitas pasien, umur, jenis kelamin, lama dirawat, dan lainnya) yang menentukan besarnya tarif INA-CBGs [3].

*Grouping* dalam software INA-CBGs akan menghasilkan kode INA-CBGs diantaranya digit I (*Case-Mix Main Groups*) sebagai klasifikasi tahap pertama dengan label huruf alphabet (A sampai Z), digit 2 (*Case Groups*) adalah sub-group kedua yang menunjukkan spesifikasi atau tipe kelompok kasus, digit 3 (*Case Type*) adalah sub-group ketiga sebagai spesifikasi CBGs dengan numerik, digit 4 (*Severity Level*) sebagai sub-group keempat yang menggambarkan tingkat keparahan kasus yang dipengaruhi adanya komorbiditas ataupun komplikasi dalam masa perawatan [4]. *Severity level* pada INA-CBGs terbagi menjadi 4 kode romawi yaitu "0" menunjukkan kasus rawat jalan, "I" menunjukkan ringan untuk rawat inap tingkat keparahan 1 untuk kasus tanpa komplikasi atau komorbid. "II" untuk kasus sedang rawat inap dengan tingkat keparahan 2 (mild komplikasi dan komorbiditi), "III" untuk kasus berat rawat inap dengan tingkat keparahan 3 (major komplikasi dan komorbiditi) [4]. Berdasarkan penjelasan di atas, diketahui bahwa *severity level* merupakan salah satu aspek yang menentukan besarnya tarif INA-CBGs pada setiap kasus CMG (*Case Mix Groups*), yang dilihat dari tingkat keparahan penyakit dan adanya komplikasi diagnosis sekunder.

Di Indonesia, hipertensi merupakan salah satu penyakit yang menjadi masalah kesehatan dan merupakan penyakit yang memerlukan jangka panjang sehingga biaya yang dikeluarkan untuk pengobatan hipertensi sangat tinggi [5]. Penyakit hipertensi dapat disertai dengan berbagai komplikasi seperti stroke, gagal jantung, gagal ginjal dan kebutaan [6]. Dalam penelitian Agiwahyunto (2018) dalam palupi (2023) menyebutkan bahwa diagnosis sekunder dan komplikasi pasien dapat berpengaruh terhadap besaran tarif INA CBGs karena dapat menaikkan *severity level* atau tingkat keparahan pasien [7]. Prediksi *severity level* dapat dilakukan dengan menerapkan *data mining*.

*Data mining* adalah analisis data dalam jumlah besar yang berguna untuk menyelesaikan masalah, salah satunya dengan metode prediksi klasifikasi. Metode *decision tree* mengubah sejumlah besar fakta menjadi *decision tree* yang menggambarkan aturan. *Decision tree* juga bermanfaat untuk mengeksplorasi data dan menemukan hubungan tersembunyi antara berbagai calon variabel input dengan variabel target. *Decision tree* yang dapat digunakan untuk prediksi *severity level* pada model klasifikasi pohon keputusan menggunakan algoritma C4.5. Algoritma C4.5 berguna untuk klasifikasi data dengan menggunakan teknik *decision tree*. Algoritma ini mampu mengolah data numerik (kontinu) dan diskrit, menangani atribut dengan nilai yang hilang, serta menghasilkan aturan yang mudah diinterpretasikan. Selain itu, C4.5 dikenal sebagai salah satu algoritma yang paling cepat dibandingkan dengan algoritma klasifikasi lainnya [8].

Penelitian ini menggunakan algoritma dalam metode klasifikasi yaitu *Decision Tree*, dan *Neural Network* dengan tools *Orange Data Mining*. Kelebihan metode *decision tree* adalah sifatnya yang fleksibel sehingga dapat meningkatkan kualitas keputusan yang dihasilkan [9]. Sedangkan kelebihan metode *neural network* yaitu mampu untuk mengidentifikasi pola dalam data yang terkait dengan hasil tertentu dari populasi besar dan menerapkan temuan tersebut pada individu [10]. Di RSUD Khidmat Sehat Afiat diagnosa hipertensi menjadi salah satu diagnosa yang sering ditemui disertai berbagai macam diagnosa komplikasi. Berdasarkan latar belakang tersebut peneliti bertujuan untuk melakukan prediksi dengan implementasi pohon keputusan (*decision tree*) dan *neural network* memanfaatkan data hasil *grouping* penyakit hipertensi pada software INA CBGs pada pasien JKN di RSUD Khidmat Sehat Afiat pada tahun 2021 dan 2022 dengan diagnosis utama hipertensi dan HHD (*Hypertensive Heart Disease*).

## 2. METODE PENELITIAN

Jenis penelitian ini adalah penelitian analisis prediktif. Penelitian prediktif adalah penelitian yang dapat menjelaskan hasil peramalan atau hasil perkiraan suatu nilai berdasarkan data yang dimiliki. Penelitian ini bertujuan memprediksi *severity level* pada kasus hipertensi pasien rawat inap, melalui metode pengujian model klasifikasi pada *software orange data mining* sehingga dapat dihasilkan hasil prediksi *severity level* dari data klaim kasus hipertensi pada rawat inap.

Data yang digunakan pada penelitian ini adalah data sekunder yaitu laporan klaim JKN rawat inap tahun 2021 dan 2022. Pengumpulan data dilakukan di bagian *casemix* RSUD Khidmat Sehat Afiat (KISA) Kota Depok pada bulan Maret sampai dengan Juli 2023. Populasi pada penelitian ini adalah seluruh data pada laporan klaim JKN rawat inap pada tahun 2021 dan 2022. Sampel penelitian ini adalah seluruh data pasien rawat inap dengan kasus hipertensi pada laporan klaim rawat inap pada tahun 2021 dan 2022 sebanyak 162 data. Teknik pengambilan sampel yang digunakan dalam penelitian ini menggunakan teknik *purposive sampling* karena sampel penelitian ditentukan dengan kriteria diagnosa utama hipertensi baik hipertensi primer maupun sekunder.

Pengolahan data dilakukan melalui beberapa langkah yang sistematis untuk memudahkan proses analisis dan prediksi. Berikut adalah langkah-langkahnya:

### 1. Identifikasi Data yang Dibutuhkan

Langkah pertama adalah membuat checklist dokumen untuk menentukan data yang dibutuhkan. Pada prediksi *severity level* pada laporan klaim BPJS rawat inap tahun 2021 dan 2022. Data penting yang dikumpulkan meliputi: usia pasien, diagnosa utama, diagnosa sekunder, lama dirawat (*Length of Stay/LOS*), *severity level* yang dihasilkan dari grouping INA-CBGs.

### 2. Pengumpulan dan Pembersihan Data

Untuk memastikan tidak ada data yang hilang atau tidak sesuai.

### 3. Transformasi data agar variabel seperti usia pasien, lama dirawat, diagnosa utama, dan diagnosa sekunder siap diolah oleh model.

### 4. Penerapan Model

Model yang digunakan adalah *decision tree* dan *neural network* dengan *Orange Data Mining*. Data kemudian dibagi menjadi dua bagian, yakni *training data* dan *testing data*. *Data training* digunakan untuk melatih model, sedangkan *data testing* dipakai untuk mengukur kinerja model.

### 5. Evaluasi Model

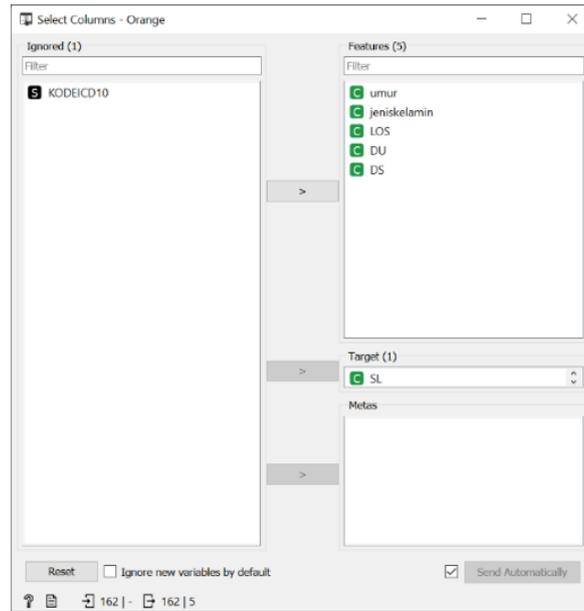
Hasil prediksi dievaluasi dengan menggunakan metrik akurasi, *precision*, *recall*, dan *ROC Curve*. Evaluasi ini membantu menentukan apakah model yang digunakan dapat memprediksi *severity level* secara akurat.

## 3. HASIL DAN ANALISIS

### 3.1 Pemilihan data (selection data) untuk penentuan *severity level* pada kasus hipertensi di RSUD Khidmat Sehat Afiat (KSA) Depok

Pemilihan data pasien (*selection data*) pada penelitian ini 162 data pasien rawat inap dengan diagnosis utama hipertensi sekunder yaitu tekanan darah tinggi yang disebabkan oleh penyakit tertentu dalam penelitian ini adalah HHD (*Hypertensive Heart Disease*), dan diagnosa sekunder CHF (*Congestive Heart Failure*) dan CKD (*Chronic Kidney Disease*) pada tahun 2021-2022. *Selection data* penelitian ini menggunakan beberapa atribut yaitu diagnosa utama, diagnosa sekunder, umur pasien, jenis kelamin, lama dirawat, dan *severity level* hasil *grouping* pada INA-CBGs. Seleksi data dilakukan untuk mengurangi kompleksitas atribut (variabel) yang akan dikelola pada *processing* dan analisis. Dengan cara ini akan membantu mempercepat proses pembelajaran/pemodelan [11].

Berikut pemilihan data yang dilakukan menggunakan *software orange data mining* dengan import file data laporan klaim.

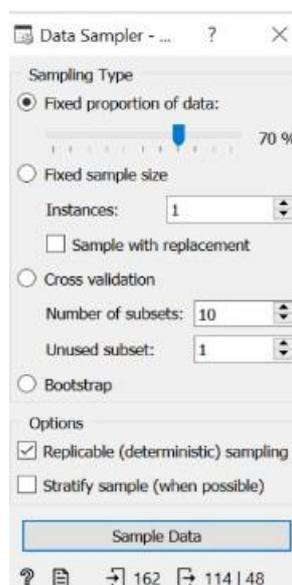


Gambar 1. Selection Data pada Orange Data Mining

Diketahui dari data terdapat 5 *features* yaitu diagnosa utama (DU), diagnosa sekunder (DS), umur, jenis kelamin, lama dirawat (LOS) dan 1 meta data yaitu kode ICD-10. Variabel target dalam dalam tabel yaitu *severity level* (SL). Dari pemilihan data untuk prediksi penentuan *severity level* pada kasus hipertensi di RSUD Khidmat Sehat Afiat (KISA) Depok pada penelitian ini sesuai dengan hasil penelitian Isma (2021) yang menyatakan bahwa *severity level* pada INA-CBGs dipengaruhi oleh diagnosa sekunder [12]. Diagnosis sekunder pasien dapat berpengaruh terhadap besaran tarif INA CBGs karena dapat menaikkan *severity level* atau tingkat keparahan pasien [7]. Kemudian pada Permenkes No 76 Tahun 2016 yang menyatakan bahwa data sosial dan data klinis pasien mempengaruhi hasil *grouping* INA-CBGs [4].

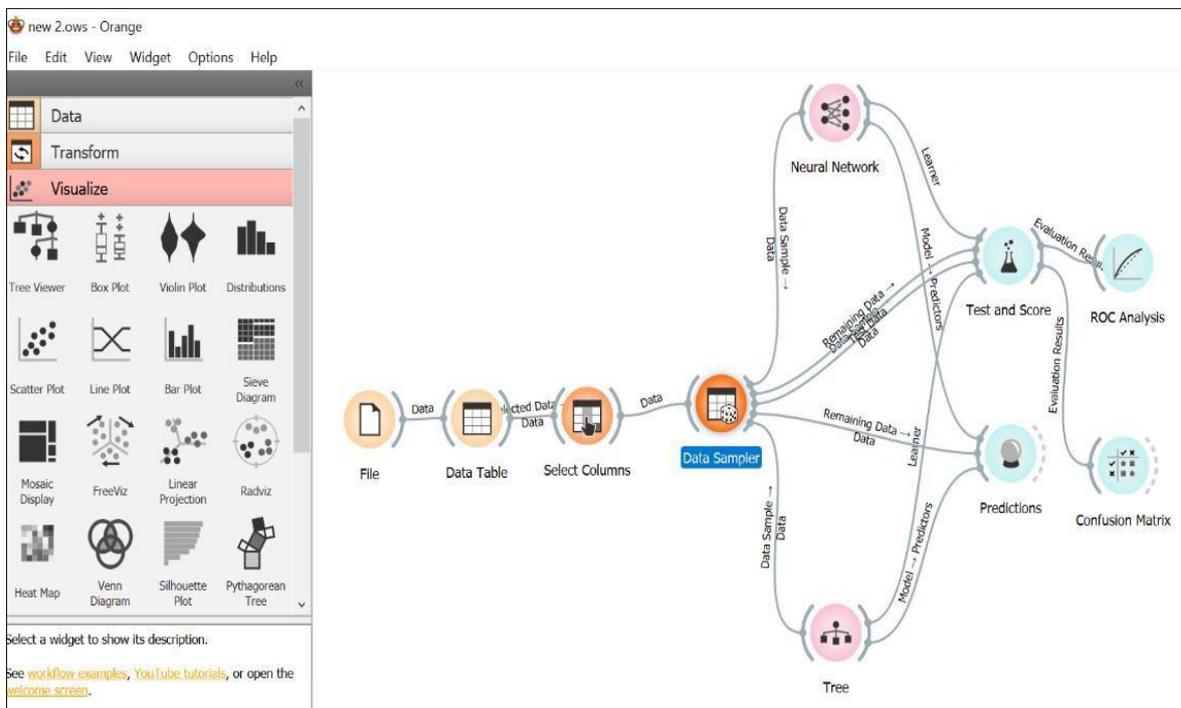
### 3.2 Implementasi permodelan *decision tree* dan *neural network* untuk penentuan *severity level* pada kasus hipertensi di RSUD Khidmat Sehat Afiat (KISA) Depok

Implementasi pohon keputusan (*decision tree*) dan *neural network* untuk penentuan *severity level* dilakukan dengan tools *orange data mining* dengan membagi data menggunakan data sampler menjadi 2 yaitu *data testing* dan *data training*. Dalam penelitian ini dihasilkan *data testing* 70 % yaitu 114 dari 162 dan *data training* 48 dari 162 data.



Gambar 2. Data Sampler

Permodelan pohon keputusan (*decision tree*) dan *neural network* dilakukan dengan desain seperti pada Gambar 3.



Gambar 3. Model Klasifikasi pada *Orange Data Mining*

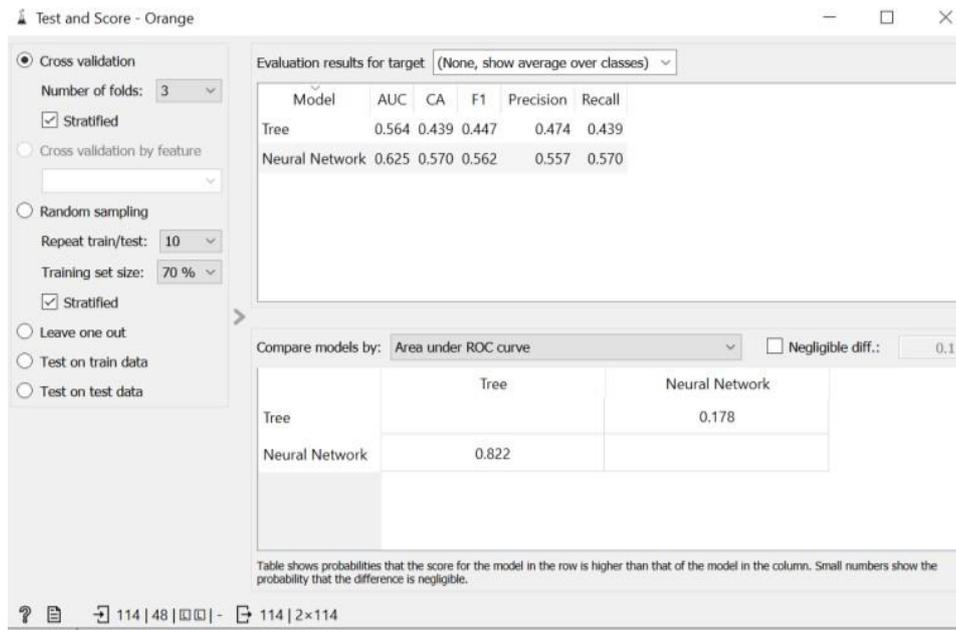
Tabel 1. Tabel Hasil Prediksi *Orange Data Mining*

Hasil Prediksi	Sesuai	Tidak Sesuai
<i>Neural Network</i>	16 (33,33%)	32 (66,67%)
<i>Decision Tree</i>	17 (35,41%)	31 (64,5%)

Dari permodelan *decision tree* dan *neural network* didapatkan hasil prediksi *severity level* pada kasus hipertensi seperti pada tabel 3.1. Implementasi *decision tree* dan *neural network* untuk penentuan *severity level* dilakukan dengan *tools orange data mining* sesuai dengan penelitian El-Jerjawi (2018) [10]. Pada aplikasi ini dapat dilakukan berbagai proses pengolahan *data mining* seperti transformasi data, visualisasi, pemodelan supervised dan unsupervised, serta evaluasi [10]. Proses klasifikasi menggunakan software *orange data mining* dimulai dengan menyiapkan *dataset training* untuk memulai pelatihan model dan memproses gambar yang ada dengan *Image Embedding*. Kemudian akan dilanjutkan dengan pengolahan data yang akan dilakukan oleh algoritma yang telah dipilih [13].

### 3.3 Pengujian data kesesuaian hasil prediksi *severity level* pada kasus hipertensi di RSUD Khidmat Sehat Afiat (KSA) Depok

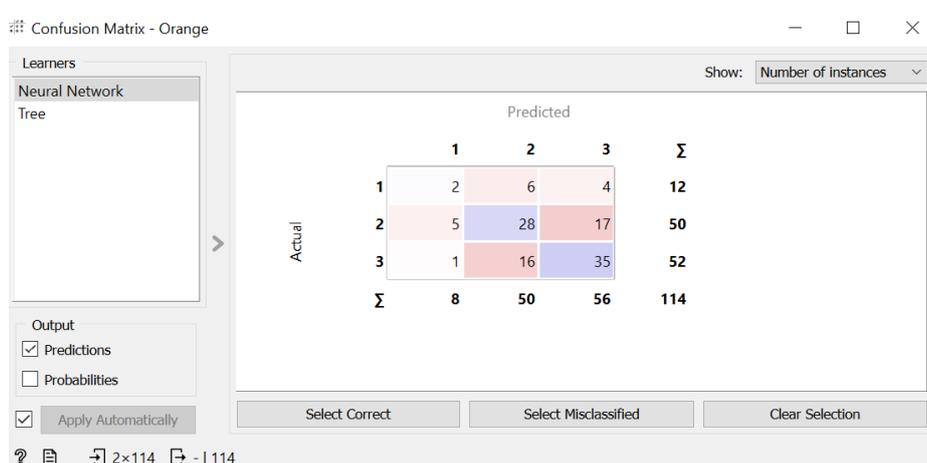
Pengujian data kesesuaian hasil prediksi *severity level* pada penelitian ini dilakukan dengan tiga metode yaitu *test and score*, *contusion matrix*, dan *ROC Curve*. Pada *test and score* perhitungan model dilihat dengan nilai *precision*, *recall*, dan *accuracy* pada Gambar 4.



Gambar 4. Hasil *Test and Score Orange Data Mining*

Berdasarkan hasil *test and score* pada Gambar 4, algoritma *neural network* menunjukkan performa terbaik dengan nilai AUC tertinggi, yaitu 0,625 (62,5%). Selain itu, *neural network* juga mencatat nilai CA tertinggi sebesar 0,570 (57,0%), nilai F1 tertinggi sebesar 0,562 (56,2%), dan nilai *precision* tertinggi sebesar 0,557 (57,7%). AUC menunjukkan tingkat prediksi, semakin tinggi hasilnya maka semakin tinggi juga tingkat prediksinya [14]. Sedangkan apabila F1 menunjukkan skor yang lebih baik maka model klasifikasi memiliki *precision* dan *recall* yang baik [15]. Dengan demikian, rasio prediksi benar lebih tinggi pada *neural network* dibandingkan algoritma lainnya. Hal ini sejalan dengan penelitian Hozairi (2021), yang menyatakan bahwa nilai *precision* menggambarkan perbandingan antara jumlah prediksi positif yang benar dengan keseluruhan data positif yang benar [16]. Selain itu, hasil *test and score* ini konsisten dengan penelitian terdahulu oleh Santoni (2020) yang menyimpulkan bahwa algoritma *neural network* memiliki kinerja yang lebih baik dibandingkan dengan algoritma *decision tree* [17].

*Confusion Matrix* merupakan Indikator kinerja untuk masalah klasifikasi *machine learning*. *Confusion Matrix tabel* yang berisi empat kemungkinan kombinasi angka prediksi dan aktual. Terdapat empat komponen utama yang merepresentasikan hasil proses klasifikasi, yaitu: *True Positive* (TP), yang merupakan jumlah data positif yang diprediksi benar; *True Negative* (TN), yang merupakan jumlah data negatif yang diprediksi benar; *False Positive* (FP), yang terjadi ketika data negatif diprediksi sebagai data positif; dan *False Negative* (FN), yang terjadi ketika data positif diprediksi sebagai data negative [19]. Hasil evaluasi performa *neural network* dapat dilihat pada Gambar 5, yang menampilkan jumlah prediksi benar dan salah untuk setiap tingkat *severity level* dari total 114 data yang diuji.



Gambar 5. *Confusion Matrix Neural Network*

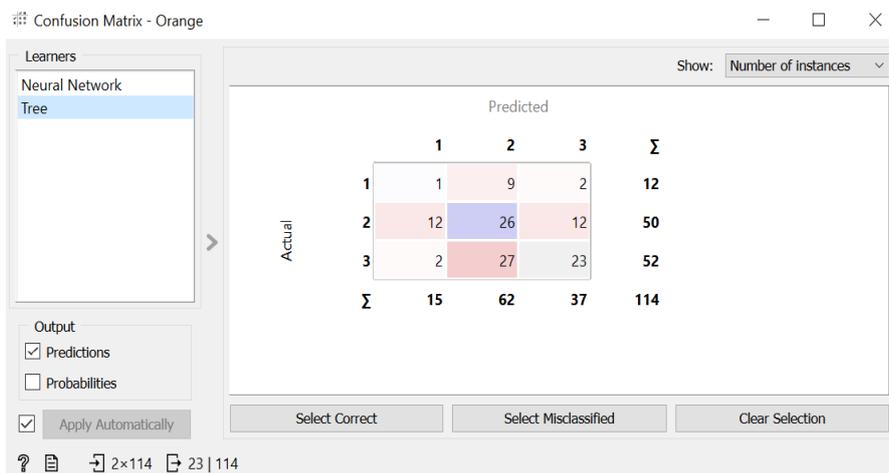
*Confusion Matrix* pada Gambar 5 menunjukkan hasil sebagai berikut: *True Positive* (TP) sebanyak 63, *True Negative* (TN) sebanyak 3, *False Positive* (FP) sebanyak 33, dan *False Negative* (FN) sebanyak 15. Berdasarkan penelitian Liang (2022), perhitungan *Accuracy*, *Precision*, dan *Recall* untuk metode *neural network* dilakukan menggunakan rumus-rumus sebagai berikut [18].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+TP} \times 100\% = \frac{63+3}{63+3+33+15} \times 100\% = 57,89\%$$

$$Precision = \frac{TP}{TP+FP} \times 100\% = \frac{63}{63+33} \times 100\% = 65,6\%$$

$$Recall = \frac{TP}{TP+FN} \times 100\% = \frac{63}{63+15} \times 100\% = 80,76\%$$

Dari perhitungan di atas *confusion matrix* pada *neural network* didapatkan *accuracy* 57,89%, *precision* 65,6% dan *recall* 80,76%. Sedangkan evaluasi *confusion matrix* pada *decision tree* dapat dilihat pada Gambar berikut.



Gambar 6. *Confusion Matrix Decision Tree*

Pada *confusion matrix decision tree* pada Gambar 6 menghasilkan nilai *True Positif* (TP) 26, *True Negatif* (TN) 24, *False Positif* (FP) 27, dan *False Negatif* (FN) 43. Perhitungan *Accuracy*, *Precision*, dan *Recall* dari metode *decision tree* adalah sebagai berikut.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+TP} \times 100\% = \frac{26+24}{26+24+27+37} \times 100\% = 43,85\%$$

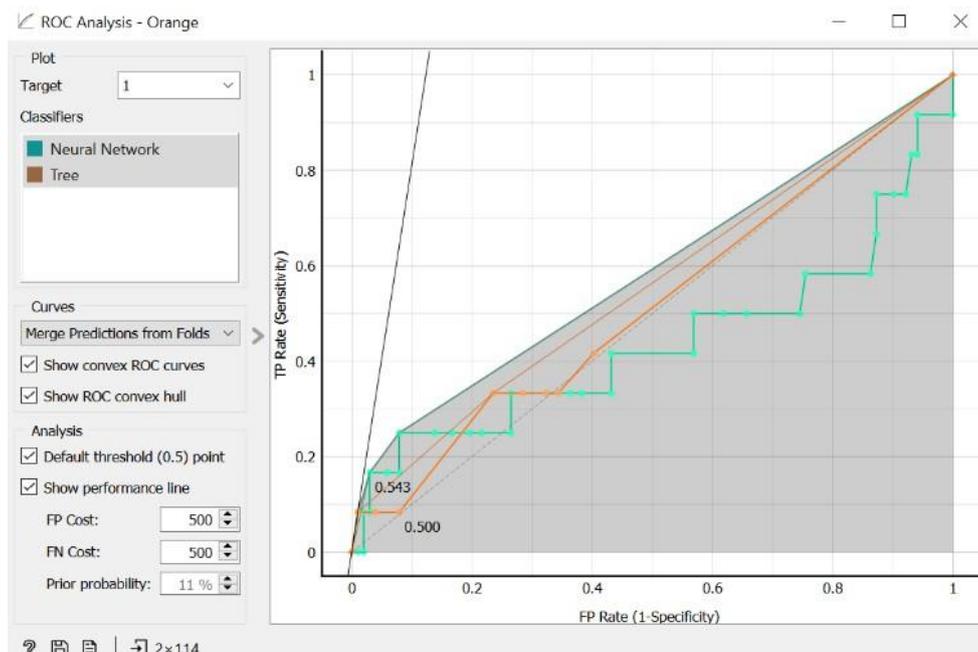
$$Precision = \frac{TP}{TP+FP} \times 100\% = \frac{26}{26+27} \times 100\% = 49,05\%$$

$$Recall = \frac{TP}{TP+FN} \times 100\% = \frac{26}{26+43} \times 100\% = 37,5\%$$

Dari perhitungan di atas, *confusion matrix* pada *decision tree* didapatkan *accuracy* 43,85%, *precision* 49,05%, dan *recall* 39%. Dari keempat perhitungan klasifikasi tersebut dapat diketahui bahwa kinerja dari model *neural network* memiliki *accuracy* lebih tinggi yaitu 57,89%. Pada *precision*, *neural network* memiliki nilai lebih tinggi yaitu 65,6%. Begitu pula pada *recall*, *neural network* memiliki nilai lebih tinggi yaitu 80,76%. Dari hasil *confusion matrix*, *neural network* memiliki kinerja yang baik dibanding *decision*

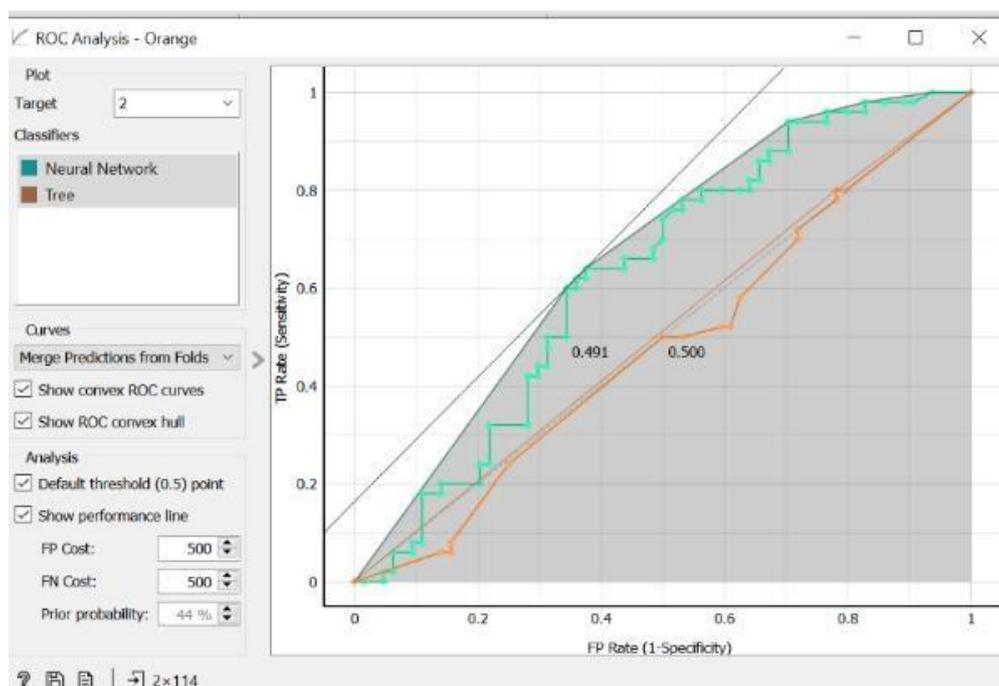
*tree*. *Confusion matrix* merupakan pengukuran performa untuk klasifikasi *machine learning* dengan indikator *accuracy* yang menggambarkan keakuratan model dapat mengklasifikasikan dengan benar, *precision* yang menggambarkan prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif, dan *recall* yang menggambarkan rasio prediksi benar positif dibandingkan dengan data yang benar positif [19].

Pada metode *ROC Curve* diperoleh hasil evaluasi yang digunakan untuk melihat seberapa baik sebuah model dapat memisahkan contoh positif dan negatif dan untuk mengidentifikasi ambang batas terbaik untuk memisahkannya. Pada prediksi *severity level* ini dibagi menjadi 3 seperti pada gambar 7.



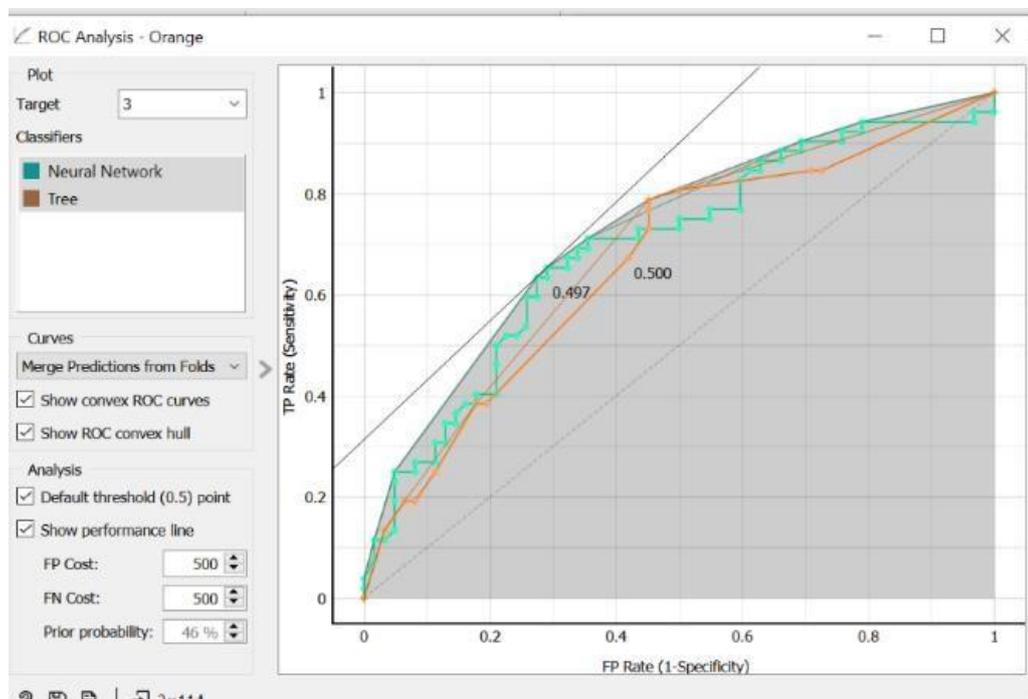
Gambar 7. ROC Curve Severity Level I

Kurva ROC di atas menunjukkan bahwa hasil analisis ROC pada prediksi *severity level I* pada *neural network* adalah 0,543 (54,3%), dan *decision tree* adalah 0,5 (50%). Oleh karena itu, *severity level I* yang dengan nilai akurasi terbaik adalah *neural network* karena kurvanya lebih mendekati titik 1.



Gambar 8. ROC Curve Severity Level II

Analisis ROC pada prediksi *severity level II* pada *neural network* adalah 0.491 dan *decision tree* adalah 0,5. Sehingga pada *severity level II* model dengan nilai akurasi terbaik adalah *decision tree* karena kurvanya mendekati titik 1.



Gambar 9. ROC Curve Severity Level III

Pada kurva ROC *severity level III* di atas menunjukkan bahwa hasil analisis ROC pada prediksi *severity level III* pada *neural network* adalah 0.497 dan *decision tree* adalah 0,5. Oleh karena itu, pada *severity level III* model yang memiliki nilai akurasi paling baik adalah *decision tree* karena kurvanya mendekati titik 1. Analisis ROC Curve ini sesuai dengan pernyataan Santoso dkk (2019) yang menyatakan bahwa ROC (Receiver Operating Characteristics) curve adalah kurva yang digunakan untuk melihat seberapa baik sebuah model dapat memisahkan contoh positif dan negatif dan untuk mengidentifikasi ambang batas terbaik untuk memisahkannya. Semakin mendekati titik 1,0 maka model semakin baik [19].

#### 4. KESIMPULAN

Implementasi *decision tree* dan *neural network* untuk penentuan *severity level* menghasilkan *data testing* 70% (114 data) dan *data training* 30% (48 data). Hasil prediksi tertinggi pada *decision tree* sebesar 35,41%, tetapi *neural network* menunjukkan performa terbaik dengan AUC 62,5%, CA 57%, F1 56,2%, dan precision 57,7%. Pada *confusion matrix*, *neural network* menunjukkan *accuracy* 57,89%, *precision* 65,6%, dan *recall* 80,76%. Pada ROC curve, model *neural network* lebih baik dalam memisahkan nilai positif dan negatif untuk *severity level I, II, dan III*. Disimpulkan bahwa *neural network* adalah metode klasifikasi terbaik. Rumah sakit disarankan dapat menerapkan melakukan prediksi *severity level* dengan menggunakan model *neural network* pada *orange data mining*. Penelitian selanjutnya disarankan untuk melakukan *balance sampling*, menambah variabel, dan mengimplementasikan metode *data mining* lainnya untuk membandingkan keakuratan dan menemukan metode terbaik dalam prediksi *severity level*.

#### REFERENSI

- [1] F. Z. Luthfiyah, "Hubungan Pelayanan Rawat Jalan Pasien Dengan Premi Badan Penyelenggara Jaminan Sosial (BPJS) Berdasarkan UU No 24 Tahun 2011," *Nusant. J. Pendidikan, Seni, Sains dan Sos. Humanioral*, vol. 1, no. 1, pp. 1–15, 2022.
- [2] S. Puspitorini, S. Kusumadewi, and L. Rosita, "SPK Penentuan Severity Level Kasus Penyakit Dengan Pohon Keputusan," in *Seminar Nasional Informatika Medis (SNIMed) 2016*, 2016, pp. 16–32.
- [3] I. Indriyanti, N. Ichsan, H. Fatah, T. Wahyuni, and E. Ermawati, "Implementasi Orange Data Mining Untuk Prediksi Harga Bitcoin," *J. Responsif Ris. Sains dan Inform.*, vol. 4, no. 2, pp. 118–125, Aug. 2022, doi:

- 10.51977/jti.v4i2.762.
- [4] Kementerian Kesehatan Republik Indonesia, *Peraturan Menteri Kesehatan Nomor 76 Tahun 2016 tentang Pedoman Indonesian Case Base Groups (INA-CBG) dalam Pelaksanaan Jaminan Kesehatan Nasional*. Jakarta: Kementerian Kesehatan Republik Indonesia, 2016.
- [5] B. Sedayu, S. Azmi, and R. Rahmatini, "Karakteristik Pasien Hipertensi di Bangsal Rawat Inap SMF Penyakit Dalam RSUP DR. M. Djamil Padang Tahun 2013," *J. Kesehat. Andalas*, vol. 4, no. 1, pp. 65–69, Jan. 2015, doi: 10.25077/jka.v4i1.192.
- [6] S. Regina, "Gambaran Indikasi Rawat Inap Pasien Hipertensi Di Rumah Sakit Sumber Waras Tahun 2017," Universitas Tarumanagara, 2018.
- [7] I. P. Palupi, S. Sugiarsi, and T. A. Sutrisno, "Analisis Hubungan Kelengkapan Diagnosis Sekunder Dengan Kenaikan Tarif INA-CBGs DiRSUD Pandan Arang Boyolali," *Indones. J. Heal. Inf. Manag.*, vol. 3, no. 2, pp. 1–7, 2023, doi: <https://doi.org/10.54877/ijhim.v3i2.126>.
- [8] R. H. Pambudi, B. D. Setiawan, and I. Indriati, "Penerapan Algoritma C4.5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 7, pp. 2637–2643, 2018, [Online]. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/1660>
- [9] A. P. Permana, K. Ainiyah, and K. F. H. Holle, "Analisis Perbandingan Algoritma Decision Tree, kNN, dan Naive Bayes untuk Prediksi Kesuksesan Start-up," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 6, no. 3, pp. 178–188, Sep. 2021, doi: 10.14421/jiska.2021.6.3.178-188.
- [10] N. S. El-Jerjawi and S. S. Abu-Naser, "Diabetes Prediction Using Artificial Neural Network," *Int. J. Adv. Sci. Technol.*, vol. 121, pp. 55–64, 2018, doi: <http://dx.doi.org/10.14257/ijast.2018.121.05>.
- [11] D. T. Wilujeng, M. Fatekurohman, and I. M. Tirta, "Analisis Risiko Kredit Perbankan Menggunakan Algoritma K-Nearest Neighbor dan Nearest Weighted K-Nearest Neighbor," *Indones. J. Appl. Stat.*, vol. 5, no. 2, pp. 142–148, Oct. 2023, doi: 10.13057/ijas.v5i2.58426.
- [12] I. Oktadiana, "Perbandingan Biaya Riil Pada Pasien Diabetes Mellitus Tipe 2 Dengan Tarif INA-CBG'S Di Rumah Sakit Umum Daerah," *J. Farm. Tinctura*, vol. 2, no. 2, pp. 42–51, 2021.
- [13] E. Pranadjaya, E. S. Pangestu, C. O. Sereati, S. Octaviani, and M. Darmawan, "Perbandingan Algoritma Machine Learning menggunakan Orange Data Mining untuk Klasifikasi Jenis Kendaraan pada Sistem Tilang Digital," *J. Elektro*, vol. 17, no. 1, pp. 41–47, Apr. 2024, doi: 10.25170/jurnalelektro.v17i1.5429.
- [14] E. Mardiani *et al.*, "Membandingkan Algoritma Data Mining Dengan Tools Orange untuk Social Economy," *Digit. Transform. Technol.*, vol. 3, no. 2, pp. 686–693, Nov. 2023, doi: 10.47709/digitech.v3i2.3256.
- [15] A. Hartanto and Thooyibah, "Implementasi Orange Data Mining Untuk Prediksi Penderita Diabetes," in *Prosiding Seminar Kecerdasan Artifisial, Sains Data, dan Pendidikan Masa Depan PROKASDADIK*, Program Studi Teknik Informatika, Program Pascasarjana, Universitas Pamulang, 2023, pp. 20–23.
- [16] H. Hozairi, A. Anwari, and S. Alim, "Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta Naive Bayes," *Netw. Eng. Res. Oper.*, vol. 6, no. 2, pp. 133–144, Nov. 2021, doi: 10.21107/nero.v6i2.237.
- [17] M. M. Santoni, N. Chamidah, and N. Matondang, "Prediksi Hipertensi menggunakan Decision Tree, Naïve Bayes dan Artificial Neural Network pada software KNIME," *Techno.Com*, vol. 19, no. 4, pp. 353–363, Nov. 2020, doi: 10.33633/tc.v19i4.3872.
- [18] J. Liang, "Confusion Matrix: Machine Learning," *POGIL Act. Clear. J.*, vol. 3, no. 4, 2022, [Online]. Available: <https://pac.pogil.org/index.php/pac/article/view/304>
- [19] S. Pramana, B. Yuniarto, S. Mariyah, I. Santoso, and R. Nooraeni, *Data Mining dengan R: Konsep Serta Implementasi*. Bogor: In Media, 2018.