Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD
Volume 6 ; Nomor 2 ; Juli 2023 ; Page 619-625
E-ISSN : 2615-5133 ; P-ISSN : 2621-8976
https://ojs.trigunadharma.ac.id/index.php/jsk/index

Penerapan Data Mining dengan Algoritma C5.0 Untuk Prediksi Penyakit Stroke

Fazrin Meila Azzahra Sofyan[1], Affani Putri Riyandoro[2], Devi Fitriani Maulana[3], Jajam Haerul Jaman[4]
[1,2,3,4] Sistem Informasi, Universitas Singaperbangsa Karawang
Email: [1]2010631250044@student.unsika.ac.id, [2] 2010631250026@student.unsika.ac.id, [3]2010631250037@student.unsika.ac.id, [4]jajam.haeruljaman@staff.unsika.ac.id
Email Penulis Korespondensi: 2010631250044@student.unsika.ac.id

Article History:
Received Jul 21th, 2023
Revised Jul 26th, 2023
Accepted Jul 30th, 2023

Abstrak

Penyakit stroke merupakan kondisi yang mempengaruhi sistem saraf dan dapat menyebabkan dampak yang serius pada kesehatan seseorang. WHO menyatakan sebanyak 13,7 juta kasus setiap tahunnya dan 5,5 juta orang diantaranya meninggal dunia akibat penyakit ini. Tujuan dari penelitian ini adalah untuk mengembangkan model prediksi yang dapat membantu dalam identifikasi dini risiko terjadinya stroke. Metode yang digunakan dalam penelitian ini adalah Knowledge Discovery in Databases (KDD) dengan menerapkan algoritma C5.0, yang merupakan salah satu algoritma klasifikasi yang efektif dalam mengolah data dengan atribut numerik maupun kategorikal. Pada metode Knowledge Discovery in Databases (KDD) terdiri dari beberapa tahap yang perlu dilakukan untuk penelitian ini, yaitu selection, preprocessing, transformation, data mining, dan evaluation. Untuk Algoritma C5.0 sendiri merupakan sebuah algoritma klasifikasi dalam bidang data mining yang secara khusus digunakan dalam teknik decision tree. Data yang digunakan dalam penelitian ini adalah dataset yang berisi informasi medis dan faktor risiko yang terkait dengan stroke. Hasil dari penelitian ini berupa Decision Tree (pohon keputusan) dengan nilai accuracy, recall, dan precision dengan melakukan split data 80% (data training) - 20% (data testing) hasil nilai Accuracy yang diperoleh sebesar 95%, Recall = 96%, dan Precision = 99%.
Kata Kunci : Data Mining, Prediksi, Stroke, Algoritma C5.0, Decision Tree

Abstract

Stroke is a condition that affects the nervous system and can have a serious impact on a person's health. WHO states that there are 13.7 million cases each year and 5.5 million people die from this disease. The purpose of this research is to develop a prediction model that can help in early identification of the risk of stroke. The method used in this research is Knowledge Discovery in Databases (KDD) by applying the C5.0 algorithm, which is one of the classification algorithms that is effective in processing data with numerical and categorical attributes. The Knowledge Discovery in Databases (KDD) method consists of several stages that need to be carried out for this research, namely selection, preprocessing, transformation, data mining, and evaluation. The C5.0 algorithm itself is a classification algorithm in the field of data mining which is specifically used in decision tree techniques. The data used in this research is a dataset containing medical information and risk factors associated with stroke. The results of this study are in the form of a Decision Tree with accuracy, recall, and precision values by splitting the data 80% (training data) - 20% (testing data) the results of the Accuracy value obtained are 95%, Recall = 96%, and Precision = 99%.
Keyword : *Data Mining, Prediction, Stroke, C5.0 Algorithm, Decision Tree*

1. PENDAHULUAN

Stroke adalah penyakit yang diakibatkan oleh aliran darah otak yang menyerang fungsi otak baik fokal maupun global dan berlangsung selama 24 jam atau lebih dan dapat menyebabkan kematian [1]. Stroke dapat terjadi ketika sel atau jaringan mati dikarenakan sebagian otak tidak mendapatkan aliran darah, hal itu dapat terjadi saat pembuluh darah pada otak tersumbat ataupun pecah yang menyebabkan kerusakan pada jaringan otak [2]. Gejala yang sering terjadi adalah mati rasa pada anggota tubuh, gejala lain yang sering dialami adalah kesulitan dalam berbicara sehingga sulit untuk

Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD
Volume 6 ; Nomor 2 ; Juli 2023 ; Page 619-625
E-ISSN : 2615-5133 ; P-ISSN : 2621-8976
https://ojs.trigunadharma.ac.id/index.php/jsk/index

memahami pembicaraan, kesulitan dalam melihat, kesulitan untuk berjalan, sakit kepala parah, dan kehilangan keseimbangan [3].

Menurut World Health Organization, stroke menempati posisi kedua penyebab kematian dan penyebab kecacatan pada posisi ketiga [4]. Berdasarkan data pada tahun 2018 dari hasil diagnosis dokter, prevalensi stroke di Indonesia pada penduduk dengan rentang umur ≥ 15 tahun diperkirakan sebesar 10,9% atau sebanyak 2.120.362 orang [5]. Data dari World Stroke Organization mengungkapkan bahwa setiap tahunnya terdapat sekitar 13,7 juta kasus baru stroke dan sekitar 5,5 juta orang meninggal dunia akibat penyakit stroke. Lebih kurang 70% kasus stroke dan 87% kematian serta kecacatan akibat stroke terjadi di negara-negara dengan tingkat pendapatan rendah dan menengah [6].

Sangat penting untuk melakukan deteksi pada penyakit ini yang diharapkan dapat mengurangi risiko terkena stroke, oleh sebab itu dilakukan suatu prediksi dengan menerapkan algoritma C5.0 dimana algoritma tersebut merupakan penyempurnaan dari algoritma sebelumnya yaitu ID3 dan C4.5 yang dibentuk oleh Ross Quinlan pada tahun 1987 [7].

Beberapa penelitian terdahulu tentang algoritma C5.0, salah satunya yang dilakukan oleh Rizkya pada 108 data mahasiswa menghasilkan akurasi 91% yang mengajukan banding UKT, sedangkan akurasi 80% diperoleh dari sistem penentu UKT [8]. Penelitian lain yang dilakukan oleh Devi melakukan komparasi prediksi banjir menggunakan algoritma C5.0, SVM, dan Naïve Bayes. Dalam konteks ini, algoritma SVM dan C5.0 memiliki tingkat akurasi yang sama, yakni sebesar 93,75%, sementara algoritma Naive Bayes memiliki tingkat akurasi sebesar 81,25%. Oleh karena itu, dapat disimpulkan bahwa algoritma SVM dan C5.0 lebih akurat dalam melakukan prediksi [9]. Penelitian yang dilakukan oleh Natanael Benediktus menjelaskan tentang performa akademik siswa, data yang digunakan diambil dari LMS yang memiliki variable numerik dan kategorik. Algoritma C5.0 digunakan pada penelitian ini yang menghasilkan nilai akurasi sebesar 71,667% dengan data training sebesar 75% dan data testing 25% [10]. Menurut Nurnia Zamasi, dalam jurnal yang berjudul "Implementasi Algoritma C 5.0 Pada Analisa Data Potensi Pertanian dan Perternakan" menunjukkan bahwa, penerapan algoritma C5.0 dapat memberi gambaran potensi pertanian dan peternakan di setiap wilayahnya dengan menggunakan aplikasi *RapidMiner Classification Decision Tree* [11].

Dengan mempertimbangkan latar belakang yang telah dijelaskan sebelumnya, diharapkan penelitian ini dapat memberikan hasil yang akurat dengan menggunakan algoritma C5.0 pada prediksi penyakit stroke yang dapat dijadikan sumber informasi lebih lanjut untuk waktu ke depan.

2. METODOLOGI PENELITIAN

2.1 Algoritma C5.0

Algoritma C5.0 merupakan sebuah algoritma klasifikasi dalam bidang data mining yang secara khusus digunakan dalam teknik decision tree. Algoritma ini merupakan pengembangan dari dua algoritma sebelumnya yang dikembangkan oleh *Ross Quinlan* pada tahun 1987, yaitu ID3 dan C4.5. Proses pembentukan pohon *(tree)* pada algoritma C5.0 hampir mirip dengan algoritma C4.5. Kemiripan ini terutama terlihat dalam perhitungan entropy dan information gain. Namun, perbedaan utama terletak pada langkah lanjutan yang dilakukan oleh algoritma C5.0 setelah perhitungan *information gain*. Pada algoritma C4.5, perhitungan berhenti setelah menghitung *information gain*, sedangkan pada algoritma C5.0, langkah selanjutnya adalah menghitung gain ratio [12]. Rumus untuk mencari nilai Entropy:

$$Entropy\ (S) = \sum_{i=1}^{n} - P_i * log_2 P_i \qquad (1)$$

Keterangan :
S : Himpunan Kasus
n : Jumlah partisi S
Pi : Proporsi dari Si terhadap S

Rumus untuk mencari nilai *Information Gain* :

$$InformationGain(S, A) = Entropy\ (S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} * Entropy\ (Si) \qquad (2)$$

Keterangan :
S : Himpunan kasus
A : Atribut
n : Jmlah partisi atribut A
| Si | : Jumlah kasus pada partisi ke i
| S | : Jumlah kasus dalam S

Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD
Volume 6 ; Nomor 2 ; Juli 2023 ; Page 619-625
E-ISSN : 2615-5133 ; P-ISSN : 2621-8976
https://ojs.trigunadharma.ac.id/index.php/jsk/index

Rumus untuk mencari nilai *gain ratio* :

$$Gain\ Ratio = \frac{Infrmation\ Gain\ (S, A)}{\sum_{i=1}^{n} Entropy\ (S_i)} \quad (3)$$

2.2 Prediksi

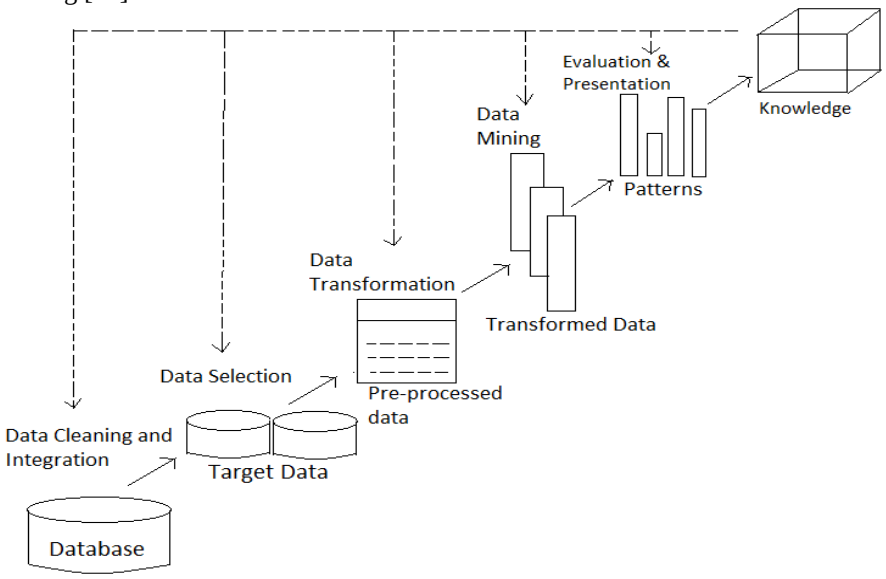
Prediksi merupakan proses untuk membuat perkiraan secara sistematis mengenai hal yang mungkin terjadi di masa yang akan datang berdasarkan informasi yang ada saat ini, sehingga dapat memperkecil kesalahan yang mungkin terjadi. Namun, prediksi tidak menghasilkan jawaban pasti terkait peristiwa yang akan terjadi, melainkan untuk memperkirakan jawaban yang paling mungkin terjadi [13].

2.3 Decision Tree

Decision Tree adalah metode klasifikasi yang menggambarkan pohon keputusan berdasarkan pendekatan *top-down*. Pendekatan ini melibatkan evaluasi semua atribut menggunakan ukuran statistik, seperti *information gain*, untuk menentukan sejauh mana suatu atribut efektif dalam mengklasifikasikan kumpulan sampel data. Algoritma *decision tree* secara otomatis menentukan atribut yang paling penting berdasarkan kemampuan dalam mengelompokkan data menjadi kelas yang benar. Keuntungan dari *decision tree* adalah representasi pohon yang dihasilkan dapat dipahami oleh manusia dan mampu menempatkan prediktor yang baik dalam urutan yang sesuai. Setelah model pohon terbentuk, komputasi untuk mengklasifikasikan data baru menggunakan model tersebut menjadi lebih cepat. Selain itu, *decision tree* juga dapat mengelola dengan baik data numerik dan data kategorikal [14].

2.4 Knowledge Discovery in Database

Knowledge Discovery in Database (KDD) adalah metode untuk mengumpulkan data historis untuk menentukan pola dan hubungan pada dataset. Hasil dari data mining tersebut dapat digunakan sebagai acuan dalam pengambilan keputusan di masa yang akan datang [13].



Gambar 1 Metode KDD

a. *Selection*
Tahap selection merupakan tahap pemilihan data yang diterapkan sebelum masuk ke tahap penggalian informasi dalam metode Knowledge Discovery in Databases (KDD). Hasil data selection tersebut akan digunakan dalam proses data mining selanjutnya. Pada tahap ini dilakukan pemilihan atribut yang akan digunakan pada dataset penyakit Stroke yang diperoleh dari situs penyedia data yaitu Kaggle.

b. *Preprocessing*
Tahap Preprocessing melibatkan proses pembersihan data yang menjadi fokus KDD sebelum melakukan data mining. Pada tahap ini dilakukan pembersihan data pada kolom "bmi" dan "smoking_status" untuk menghilangkan data yang memiliki nilai *NaN* dan *Unknown*.

c. *Transformation*

Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD
Volume 6 ; Nomor 2 ; Juli 2023 ; Page 619-625
E-ISSN : 2615-5133 ; P-ISSN : 2621-8976
https://ojs.trigunadharma.ac.id/index.php/jsk/index

Tahap Transformation melibatkan proses transformasi data yang telah dipilih agar sesuai untuk proses data mining. Pada tahap ini dilakukan transformasi tipe data pada atribut "gender" dan "smoking_status" yang menggunakan tipe data string menjadi integer.

d. *Data Mining*
Data Mining, merupakan proses untuk mencari pola atau informasi menarik dalam data terpilih menggunakan teknik atau metode tertentu. Pada penelitian ini menggunakan algoritma C5.0 untuk menentukan klasifikasi dari prediksi penyakit Stroke.

e. *Evaluation*
Hasil yang berupa pola informasi dari proses data mining yang telah ditampilkan, dapat disajikan ke dalam bentuk yang mudah dimengerti. Hal ini penting agar informasi yang dihasilkan dapat dimanfaatkan dengan baik oleh pihak yang membutuhkan.

2.5 Confusion Matrix

Confusion matrix adalah metode evaluasi yang menggunakan tabel matrix. Evaluasi dengan menggunakan confusion matrix menghasilkan akurasi, recall, dan presisi. Akurasi mengindikasikan jumlah data yang diklasifikasikan dengan benar setelah pengujian dilakukan. TP (True Positive) adalah jumlah data positif yang diklasifikasikan dengan benar oleh sistem, TN (True Negative) adalah jumlah data negatif yang diklasifikasikan dengan benar oleh sistem, FN (False Negative) adalah jumlah data negatif yang salah diklasifikasikan oleh sistem, dan FP (False Positive) adalah jumlah data positif yang salah diklasifikasikan oleh sistem [15]

3. HASIL DAN PEMBAHASAN

Dataset yang digunakan diambil dari situs penyedia data yaitu Kaggle dengan link berikut https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset. Dataset ini memiliki 12 atribut yang dapat dilihat pada tabel 1.

Tabel 1 Dataset Penyakit Stroke

Nama Atribut	Keterangan
id	Pengidentifikasi unik
gender	Jenis kelamin pada pasien
age	Usia pada pasien
hypertension	Status riwayat hipertensi yang dimiliki oleh pasien
heart_disease	Status riwayat penyakit jantung yang dimiliki oleh pasien
ever_married	Status pernikahan
work_type	Status pekerjaan pasien
Residence_type	Tipe tempat tinggal
avg_glucose_level	Rata-rata glukosa dalam darah
bmi	Body Mass Index (Indeks Massa Tubuh)
smoking_status	Status merokok
stroke	Status pasien mengidap stroke

Pada penelitian untuk memprediksi penyakit stroke menggunakan algoritma C5.0, peneliti menggunakan 8 atribut saja yaitu age, hypertension, heart_disease, avg_glucose_level, bmi, smoking_status dan stroke setelah melewati tahap selection. Menangani missing value pada atribut bmi dan smoking_status pada tahan preprocessing hingga tahap transformation untuk mengubah tipe data yang dimiliki pada atribut gender dan smoking_status menjadi tipe data integer.

Hasil yang diperoleh pada pengujian menggunakan algoritma C5.0 adalah berupa pohon keputusan (*decision tree*), nilai *accuracy, recall,* dan *precision*.

a. *Nilai Accuracy, Recall dan Precision*
Nilai yang diperoleh dari hasil *testing* data sebesar 20% dari dataset, sehingga menghasilkan nilai accuracy, recall dan precision seperti tabel 2 berikut.

Tabel 2 Tabel Confusion Matrix

	TRUE	FALSE
Berpotensi Stroke	651	4
Tidak Berpotensi Stroke	30	1

Untuk mencari nilai *accuracy, recall dan precision* menggunakan rumus dari *Confusion Matrix* sebagai berikut:

Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD
Volume 6 ; Nomor 2 ; Juli 2023 ; Page 619-625
E-ISSN : 2615-5133 ; P-ISSN : 2621-8976
https://ojs.trigunadharma.ac.id/index.php/jsk/index

- Accuracy

$$\frac{TP + TN}{TP + TN + FP + FN} \; x \; 100\%$$

$$\frac{651 + 1}{651 + 1 + 4 + 30} \; x \; 100\% = 95\%$$

- Recall

$$\frac{TP + FN}{TP} \; x \; 100\%$$

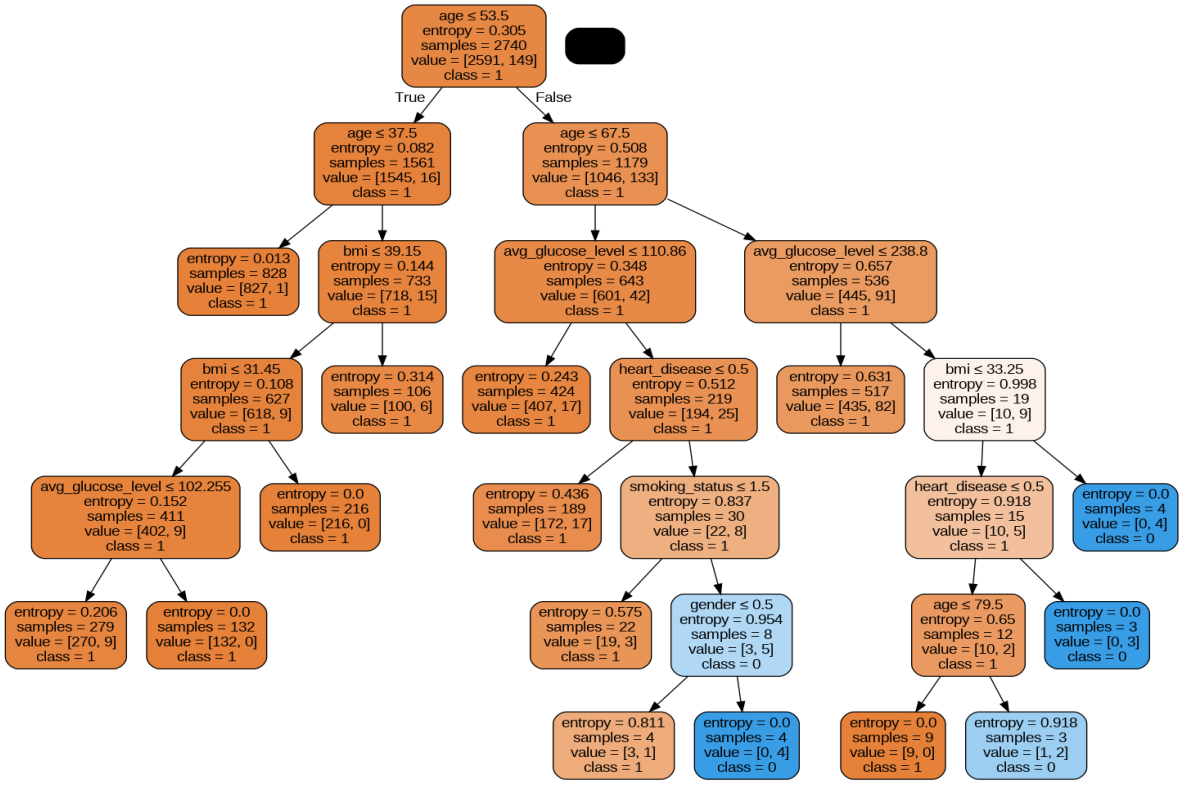
$$\frac{651 + 1}{30 + 651} \; x \; 100\% = 96\%$$

- Precision

$$\frac{TP}{FP + TP} \; x \; 100\%$$

$$\frac{651}{4 + 651} \; x \; 100\% = 99\%$$

b. *Decision Tree*
Decision Tree yang diperoleh dari pengolahan data Python menggunakan testing data sebanyak 20% untuk memprediksi pasien terkena penyakit stroke.



Gambar 2 Decision Tree Prediksi Pasien Terkena Penyakit Stroke

Berdasarkan decision tree mengenai prediksi penyakit stroke pada Gambar 2 di atas menghasilkan beberapa aturan atau rules yang terbentuk, diantaranya:

a. Jika age ≤ 53.5 dan age ≤ 37.5 maka hasil = Berpotensi Stroke
b. Jika age > 37.5, bmi ≤ 39.15, bmi ≤ 31.45 dan avg_glucose_level ≤ 102.26 maka hasil = Berpotensi Stroke
c. Jika age > 37.5, bmi ≤ 39.15, bmi ≤ 31.45 dan avg_glucose_level > 102.26 maka hasil = Berpotensi Stroke

d. Jika age > 53.5, age ≤ 67.5 dan avg_glucose_level ≤ 110.86 maka hasil = Berpotensi
e. Jika age > 53.5, age ≤ 67.5, avg_glucose_level > 110.86 dan heart_disease ≤ 0.5 maka hasil = Berpotensi Stroke
f. Jika age > 53.5, age ≤ 67.5, avg_glucose_level > 110.86, heart_disease > 0.5 dan smoking_status ≤ 1.5 maka hasil = Berpotensi Stroke
g. Jika age > 53.5, age ≤ 67.5, avg_glucose_level > 110.86, heart_disease > 0.5, smoking_status > 1.5 dan gender ≤ 0.5 maka hasil = Berpotensi Stroke
h. Jika age > 67.5 dan avg_glucose_level ≤ 238.80 maka hasil = Berpotensi Stroke
i. Jika age > 67.5, avg_glucose_level > 238.80, bmi ≤ 33.25, heart_disease ≤ 0.5 dan age ≤ 79.5 maka hasil = Berpotensi Stroke
j. Jika age > 67.5, avg_glucose_level ≤ 238.80, bmi > 33.25, heart_disease > 0.5 dan age > 79.5 maka hasil = Tidak Berpotensi Stroke

4. KESIMPULAN

Stroke meruakan penyakit yang menyerang fungsi otak yang diakibatkan oleh aliran darah tersumbat ataupun pecah hingga menyebabkan jaringan pada otak rusak, hal ini berlangsung selama 24 jam dan dapat menyebabkan kematian. WHO mengatakan sebanyak 13,7 juta kasus setiap tahunnya dan 5,5 juta orang diantaranya meninggal dunia akibat penyakit ini. Pada penelitian ini dataset diperoleh dari situs Kaggle yang memiliki 12 atribut serta data pasien sebanyak 4.000 data, hanya delapan atribut saja yang digunakan dengan alasan atribut tersebut sesuai dengan prediksi penyakit stroke. Algoritma C5.0 digunakan dalam proses penelitian ini sehingga menghasilkan sebuah pohon keputusan serta nilai Accuracy, Recall dan Precision dengan melakukan split data sebesar 80% (data training) – 20% (data testing), hasil nilai Accuracy yang diperoleh sebesar 95%, Recall = 96%, dan Precision = 99%. Hasil yang diperoleh dari algoritma C5.0 cukup baik untuk memprediksi penyakit stroke sehingga untuk penilitian yang dilakukan selanjutnya dapat menghasilkan perolehan yang diharapkan lebih baik lagi menggunakan algoritma lain.

UCAPAN TERIMA KASIH

Kami mengucapkan terima kasih kepada semua pihak yang telah mendukung penelitian ini dengan memberikan saran, masukan, dan bantuan teknis. Tanpa kontribusi mereka, penelitian ini tidak akan berhasil.

DAFTAR PUSTAKA

[1] R. Saraswati, D and Khariri, “Transisi Epidemiologi Stroke Sebagai Penyebab Kematian Pada Semua Kelompok Usia Di Indonesia,” *J. Kedokt.*, vol. 2, no. 1, pp. 81–86, 2021, [Online]. Available: https://conference.upnvj.ac.id/index.php/sensorik/article/view/1001
[2] Kemenkes RI, “Apa itu Stroke ?,” 2018. https://p2ptm.kemkes.go.id/infographic-p2ptm/stroke/apa-itu-stroke (accessed May 29, 2023).
[3] D. Hisni, M. Evelianti Saputri, and Sujarni, “Stroke Iskemik Di Instalasi Fisioterapi Rumah Sakit Pluit Jakarta Utara Periode Tahun 2021,” *Penelit. Keperawatan Kontemporer*, vol. 2, no. 1, pp. 140–149, 2022.
[4] D. P. K. Singh, “World Stroke Day,” 2021. https://www.who.int/southeastasia/news/detail/28-10-2021-world-stroke-day (accessed May 29, 2023).
[5] Rokom, “Tingkatan Kualitas dan Layanan Stroke Lewat Transformasi Kesehatan,” 2022. https://sehatnegeriku.kemkes.go.id/baca/rilis-media/20221011/4641254/tingkatan-kualitas-dan-layanan-stroke-lewat-transformasi-kesehatan/ (accessed May 29, 2023).
[6] K. Suandari, “Gambaran Kemampuan Komunikasi Verbal Pada Pasien Stroke Di Rumah Sakit Umum Daerah Buleleng Bali Tahun 2021,” 2021.
[7] A. C. Wijaya, N. A. Hasibuan, and P. Ramadhani, “Implementasi Algoritma C5 . 0 Dalam Klasifikasi Pendapatan Masyarakat (Studi Kasus : Kelurahan Mesjid Kecamatan Medan Kota),” *Inf. dan Teknol. Ilm.*, vol. 13, pp. 192–198, 2018.
[8] R. N. Amalda, N. Millah, and I. Fitria, “Implementasi Algoritma C5.0 Dalam Menganalisa Kelayakan Penerima Keringanan Ukt Mahasiswa Itk,” *Teorema Teor. dan Ris. Mat.*, vol. 7, no. 1, p. 101, 2022, doi: 10.25157/teorema.v7i1.6692.
[9] D. Fitrianah, W. Gunawan, and A. P. Sari, “Studi Komparasi Algoritma Klasifikasi C5.0, SVM dan Naive Bayes dengan Studi Kasus Prediksi Banjir,” *Techno.Com*, vol. 21, no. 1, pp. 1–11, 2022, doi: 10.33633/tc.v21i1.5348.
[10] N. Benediktus and R. S. Oetama, “Algoritma Klasifikasi Decision Tree C5 . 0 untuk Memprediksi Performa Akademik Siswa,” *Ultimatics*, vol. XII, no. 1, pp. 14–19, 2020.
[11] N. Zamasi, “Implementasi Algoritma C 5 . 0 Pada Analisa Data Potensi Pertanian dan Perternakan,” *TIN Terap. Inform. Nusant.*, vol. 2, no. 4, pp. 184–190, 2021.
[12] A. Apriyadi, M. R. Lubis, and B. E. Damanik, “Penerapan Algoritma C5.0 Dalam Menentukan Tingkat Pemahaman Mahasiswa Terhadap Pembelajaran Daring,” *Komputa J. Ilm. Komput. dan Inform.*, vol. 11, no. 1, pp. 11–20, 2022, doi: 10.34010/komputa.v11i1.7386.
[13] D. S. O. Panggabean, E. Buulolo, and N. Silalahi, “Penerapan Data Mining Untuk Memprediksi Pemesanan Bibit Pohon Dengan Regresi Linear Berganda,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 1, p. 56, 2020, doi:

10.30865/jurikom.v7i1.1947.

[14] L. Karlitasari, I. W. Sriyasa, I. Wahyudi, and H. B. Santosi, "Prediksi Morfologi Jamur Menggunakan Algoritma C5.0," *J. Teknoinfo*, vol. 17, no. 1, p. 271, 2023, doi: 10.33365/jti.v17i1.2372.

[15] P. W. Kastawan, D. M. Wiharta, and M. Sudarma, "Implementasi Algoritma C5.0 pada Penilaian Kinerja Pegawai Negeri Sipil," *Maj. Ilm. Teknol. Elektro*, vol. 17, no. 3, p. 371, 2018, doi: 10.24843/mite.2018.v17i03.p11.