

PERBANDINGAN DECISION TREE C4.5 DAN SUPPORT VECTOR MACHINE (SVM) DALAM KLASIFIKASI PENDERITA STROKE BERBASIS PSO

Chindu Lintang Bhuana^{1*}, Handoyo Widi Nugroho²

¹Program Pascasarjana Magister IIB Darmajaya, ²IIB Darmajaya

chindulintangbuana@gmail.com¹, handoyo.wn@darmajaya.ac.id

Abstract

Stroke is a major health problem for today's modern society. At present, stroke is increasingly becoming a serious problem that is faced almost all over the world. This is because a sudden stroke can result in death, physical and mental disability in both productive and old age. To obtain stroke information data, it is necessary to carry out data mining processes such as classification. Classification is a process for determining a model that explains or distinguishes concepts or data classes, with the aim of being able to estimate the class of an object whose class is unknown, in classification a number of records are also given which are called training sets, which consist of several attributes, attributes can be continuous or categorical, one of the attributes specifies the class for the record. In the above problems regarding stroke, to be able to overcome this problem, a lot of research has been carried out in the field of computer science, including the Classification of Stroke Patients Using the C4.5 Decision Tree Algorithm and the Support Vector Machine (SVM) Algorithm to classify the most important factors for this disease. The test resulted in a fairly high accuracy of the C4.5 Decision Tree Algorithm, which was 96.11%. Referring to the results of the accuracy of the research, it can be seen that the Decision Tree produces high accuracy, but the results of this accuracy can still be improved by conducting further research to produce higher accuracy by adding Optimization Feature Weighting PSO.

Keywords: Accuracy; Stroke; PSO; C4.5 Decision Tree Algorithm; Support Vector Machine (SVM) Algorithm.

Abstrak

Stroke merupakan masalah kesehatan yang utama bagi masyarakat modern saat ini. Pada dewasa ini, stroke semakin menjadi masalah serius yang dihadapi hampir diseluruh dunia. Hal tersebut dikarenakan serangan stroke yang mendadak dapat mengakibatkan kematian, kecacatan fisik dan mental baik pada usia produktif maupun usia lanjut. Untuk mendapatkan data informasi stroke perlu dilakukan proses data mining seperti klasifikasi. Klasifikasi merupakan sebuah proses untuk menentukan model yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang kelasnya tidak diketahui, di dalam klasifikasi juga diberikan sejumlah record yang dinamakan training set, yang terdiri dari beberapa atribut, atribut dapat berupa kontinyu ataupun kategoris, salah satu atribut menunjukkan kelas untuk record. Pada permasalahan diatas mengenai penyakit stroke maka untuk dapat mengatasi masalah tersebut banyak dilakukan penelitian dalam bidang ilmu komputer diantaranya adalah Klasifikasi Penderita Penyakit Stroke Menggunakan Algoritma *Decision Tree* C4.5 dan Algoritma Support Vector Machine (SVM) untuk mengklasifikasikan faktor paling penting untuk penyakit ini. Pengujian menghasilkan akurasi Algoritma *Decision Tree* C4.5 yang cukup tinggi yaitu sebesar 96,11 %. Merujuk hasil akurasi dari penelitian tersebut dapat dilihat bahwa *DecisionTree* menghasilkan akurasi yang tinggi namun hasil akurasi tersebut masih dapat ditingkatkan lagi dengan melakukan penelitian lanjutan untuk menghasilkan akurasi lebih tinggi dengan menambahkan *Optimization Feature Weighting* PSO.

Kata Kunci: Akurasi; Stroke, PSO, Algoritma *Decision Tree* c4.5; Algoritma Support Vector Machine (SVM)

1. PENDAHULUAN

Stroke adalah penyakit pada otak berupa gangguan fungsi syaraf lokal dan/atau global, munculnya mendadak, progresif, dan cepat. Gangguan fungsi syaraf pada stroke disebabkan oleh gangguan peredaran darah otak non traumatik. Gangguan syaraf tersebut menimbulkan gejala antara lain: kelumpuhan wajah atau anggota badan, bicara tidak lancar, bicara tidak jelas (pelo), (Rahim et al., 2022) mungkin perubahan kesadaran, gangguan penglihatan, dan lain-lain. Didefinisikan sebagai stroke jika pernah didiagnosis menderita penyakit stroke oleh tenaga kesehatan (dokter/perawat/bidan) atau belum pernah didiagnosis menderita penyakit stroke oleh nakes tetapi pernah mengalami

secara mendadak keluhan kelumpuhan pada satu sisi tubuh atau kelumpuhan pada satu sisi tubuh yang disertai kesemutan atau baal satu sisi tubuh atau mulut menjadi mencong tanpa kelumpuhan otot mata atau bicara pelo atau sulit bicara/komunikasi dan atau tidak mengerti pembicaraan. Jumlah penderita penyakit stroke di Indonesia tahun 2013 berdasarkan diagnosis tenaga kesehatan (Nakes) diperkirakan sebanyak 1.236.825 orang (7,0‰), (Reza et al., 2022) sedangkan berdasarkan diagnosis Nakes/gejala di-perkirakan sebanyak 2.137.941 orang (12,1‰). Berdasarkan diagnosis Nakes maupun diagnosis/gejala, Provinsi Jawa Barat memiliki estimasi jumlah pen-derita terbanyak yaitu sebanyak 238.001 orang (7,4‰) dan 533.895 orang (16,6‰), se-dangkan Provinsi Papua Barat memiliki jumlah penderita paling sedikit yaitu sebanyak 2.007 orang (3,6‰) dan 2.955 orang (5,3‰) (Rahim et al., 2022)

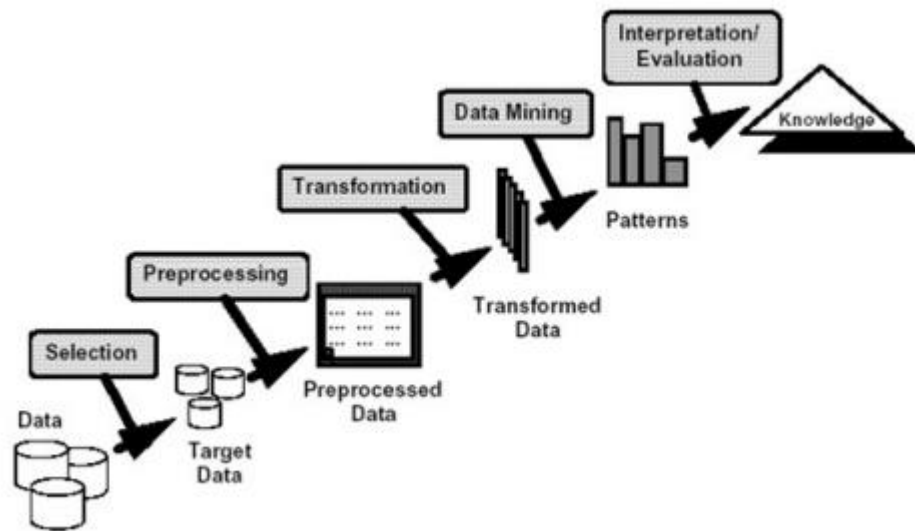
Berdasarkan hasil pengujian metode Naive bayes dan Support Vector Machine (SVM) terhadap prediksi keberhasilan pengobatan imunoterapi untuk penyakit kutil menggunakan bahasa pemrograman R pada RStudio, maka dapat diambil kesimpulan bahwa hasil pengujian menggunakan dapat memprediksi class untuk semua instance secara tepat dengan nilai akurasi 1, sedangkan metode SVM masih terdapat prediksi yang tidak tepat dengan tingkat akurasi 0.8. Hasil pengujian tersebut menunjukkan bahwa metode Naive bayes merupakan metode yang lebih unggul dibandingkan dengan metode Support Vector Machine (SVM). (Nurlia & Enri, 2021)

Dari metode klasifikasi data mining dengan algoritma C4.5 dan pengaplikasian pohon keputusan yang membentuk aturan tersebut terdapat akurasi pada data training yang berjumlah 130 dari 156 data pasien sebesar 82,31% sedangkan akurasi pada data testing yang berjumlah 26 dari 156 data pasien sebesar 76,92%. Perhitungan keduanya menggunakan confusion matrix. (Rahim et al., 2022) (Sulaeman et al., 2022) Berdasarkan penelitian dan pengujian Analisis Algoritma Support Vector Machine (SVM) Dalam Klasifikasi Penyakit Stroke yang telah penulis lakukan dapat disimpulkan bahwa, dalam penelitian menggunakan dataset dari Kaggle yang diterbitkan pada tahun 2021, menggunakan Machine Learning menggunakan metode Support Vector Machine untuk mengklasifikasi data. Penulis melakukan pengujian data dengan dua cara, yaitu menggunakan teknik Unbalanced Data untuk melakukan penelitian dengan data yang tidak sebanding, dan cara klasifikasi pada umumnya dengan data sebanding (Balanced). Untuk metode Support Vector Machine penulis menggunakan dua jenis kernel yaitu kernel Linear dan Polynomial, untuk data unbalanced pada kernel linear mendapatkan akurasi terbesar 76% dan untuk polynomial mendapatkan hasil akurasi 80%. Untuk data yang balanced dengan kernel linear penulis mendapatkan hasil akurasi tertinggi 77%, sedangkan pada kernel polynomial penulis mendapatkan hasil tertinggi 76%. (Akbar et al., 2022) (Saputri et al., 2022) Berdasarkan beberapa penelitian terdahulu, Pada penelitian ini peneliti akan melakukan penelitian untuk menganalisis dan membandingkan Decision Tree C4.5 dan Support Vector Machine (SVM) Berbasis PSO untuk melihat perbandingan tingkat akurasi, presisi, dalam klasifikasi penyakit stroke. Selain itu, hasil pengujian juga harus dianalisis untuk melihat seberapa efektif kedua algoritma tersebut (Ibrahim Irawan, Gani Hilmansyah, 2022)

2. KERANGKA TEORI

2.1. Datamining

Data mining adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam database. Menurut data mining merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat yang tersimpan di dalam database besar. (Rahim et al., 2022) (Reza et al., 2022)



Gambar 1. Proses KDD (Sri Diantika et al., 2021)

1. Pemilihan Data

dari sekumpulan data operasional yang perlu dilakukan sebelum penggalan informasi dalam KDD dimulai. Data dari hasil seleksi digunakan untuk proses data mining, disimpan pada suatu berkas, terpisah dari basis data operasional.

2. *Pre-processing/Cleaning*

Tahap cleaning ini mencakup seperti membuang data duplikasi, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan dalam menulis (tipografi).

3. *Transformation*

Tahap ini dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data. Data diubah atau digabung ke dalam format khusus sebelum bisa diaplikasikan. Sebagai contoh, beberapa metode standar seperti analisis asosiasi dan cluterung hanya bisa menerima input data kategorikal. Karenanya data berupa numerik yang berlanjut perlu dibagi menjadbeberapa interval, seperti pada atribut data umur dapat ditranformasikan ke dalam rentang umur.

4. *Data Mining*

Merupakan proses mencari pola atau informasi yang menarik dalam data yang terpilih dengan menggunakan teknik atau metode tertentu. Pemilihan metode atau algoritma yang tepat sangatbergantung pada tujuan dan proses KDD.

5. *Interpretation/Evaluasi*

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya

2.2. Algoritma C4.5

Information gain diciptakan oleh Claude Shannon dengan mempelajari nilai informasi dari data, dan menggunakan nilai tersebut sebagai acuan dalam menentukan atribut yang akan digunakan dalam menyusun pohon

keputusan. *Atribut* yang dipilih akan menghasilkan partisi dengan data yang lebih seragam, dan dapat menghasilkan pohon keputusan yang sederhana mungkin dengan perulangan yang sedikit. Berikut persamaan data dalam *tuple* D.

$$Info(D) = \sum_{i=1}^n -p_i \log_2(p_i) \quad (2.4)$$

Dimana p_i merupakan probabilitas tuple dalam D yang menjadi kelas C_i dengan asumsi $|C_{i,D}|/|D|$. $Info(D)$ atau disebut juga *entropy* dari D merupakan rata rata informasi yang diperlukan untuk identifikasi tuple dalam D.

Jika Nilai A adalah nilai diskrit maka data D akan dipisahkan sejumlah nilai data A sehingga nilai setiap cabang akan murni dan sejenis. Setelah percabangan pertama, jumlah percabangan yang mungkin terjadi diukur dengan persamaan.

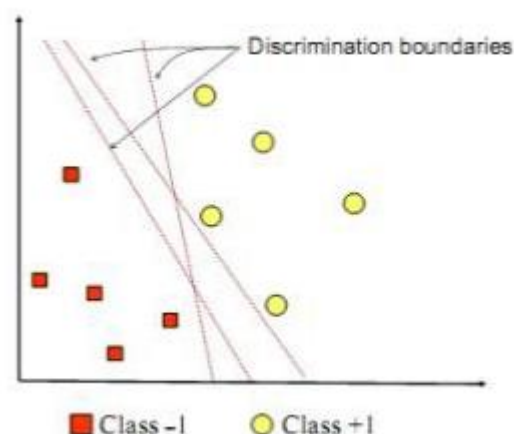
$$Info_A(D) = \sum_{j=1}^v \frac{|D_j|}{|D|} \times Info(D_j) \quad (2.5)$$

Dimana $\frac{|D_j|}{|D|}$ merupakan bobot dari partisi j. $Info_A(D)$ merupakan informasi yang diperlukan untuk mengklasifikasikan tuple dari D pada partisi A. Semakin kecil hasil persamaan ini, semakin baik pula partisi yang dihasilkan. Nilai dari sebuah atribut menentukan penting tidaknya atribut tersebut dalam penyusunan pohon keputusan. Semakin besar nilai atribut berarti semakin baik pula klasifikasi yang akan terbentuk jika cabang dipecah menurut atribut tersebut, untuk melihat nilai dari atribut A digunakan persamaan. (Azizah, 2021)

$$Gain(A) = Info(D) - Info_A(D) \quad (2.6)$$

2.3. Support Vector Machine (SVM)

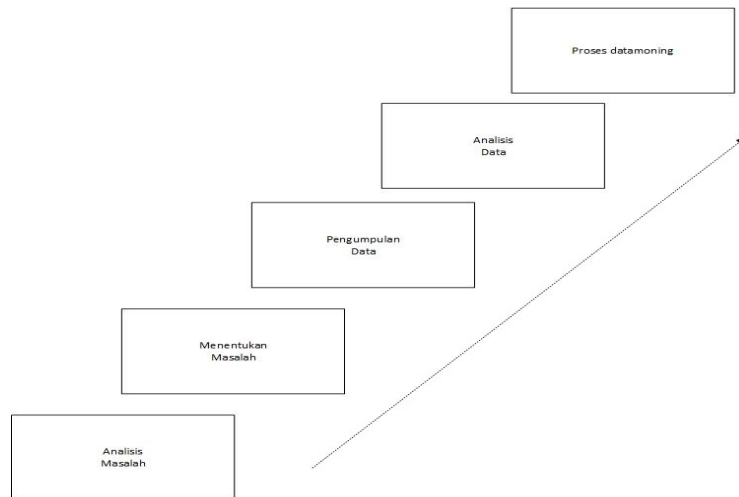
Metode pembelajaran mesin yang telah menjadi sangat populer untuk analisis neuroimaging dalam beberapa tahun terakhir. Karena kesederhanaan dan fleksibilitasnya yang relatif untuk mengatasi berbagai masalah klasifikasi, SVM secara khusus memberikan kinerja prediksi yang seimbang, bahkan dalam studi di mana ukuran sampel mungkin terbatas. Dalam penelitian gangguan otak, SVM biasanya digunakan menggunakan analisis pola multivoxel (MVPA) karena kesederhanaan relatifnya membawa risiko overfitting yang lebih rendah bahkan menggunakan data pencitraan dimensi tinggi. Baru-baru ini, SVM telah digunakan dalam konteks psikiatri presisi, terutama untuk aplikasi yang melibatkan prediksi diagnosis dan prognosis penyakit otak seperti penyakit Alzheimer, skizofrenia, dan depresi. Di bagian terakhir bab ini, kami meninjau sejumlah studi terbaru yang menggunakan SVM untuk aplikasi tersebut. (Putra et al., 2021)(Rohman et al., 2020)(Saputra et al., 2021)



Gambar 2. SVM berusaha untuk menemukan hyperplane terbaik yang memisahkan kedua kelas

3. METODOLOGI

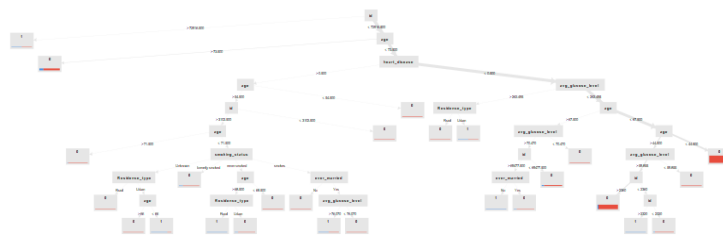
Metodologi penelitian ini dilakukan secara sistematis yang dapat digunakan sebagai pedoman untuk peneliti dalam melaksanakan penelitian agar hasil yang dicapai tidak menyimpang dan tujuan yang diinginkan dapat terlaksana dengan baik dan sesuai dengan tujuan yang telah ditetapkan sebelumnya, dengan kerangka kerja sebagai berikut



Gambar 3. Kerangka kerja yang diterapkan.

4. HASIL DAN PEMBAHASAN

Metode untuk meningkatkan tingkat akurasi dalam klasifikasi adalah penggunaan teknik optimasi. Sehingga peneliti menggunakan eksperimen lanjutan dengan menggunakan teknik optimasi seleksi fitur Particle Swarm Optimization (PSO) dengan algoritma Decision Tree C4.5. dan Support Vector Machine (SVM) Penerapan Metode dan algoritma pada rapidminer 9. 10 ditunjukkan pada gambar di bawah ini



Gambar 4. Pohon Keputusan

Tree

id > 72916.500: 1 {1=1, 0=1}

id ≤ 72916.500

| age > 73.500: 0 {1=111, 0=447}

| age ≤ 73.500

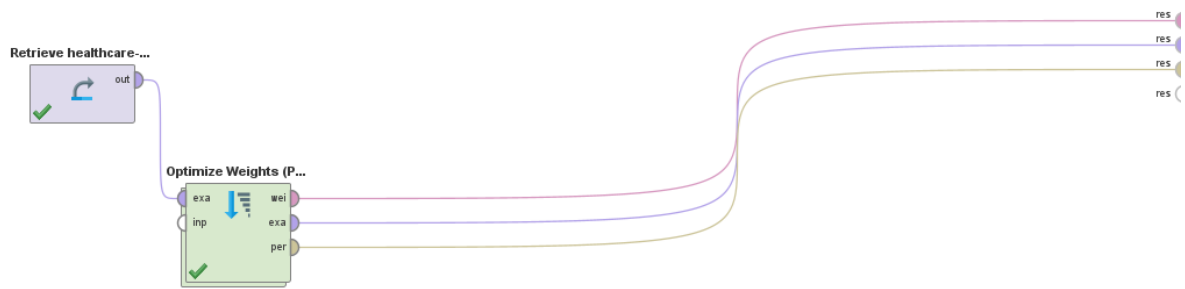
| | heart_disease > 0.500

| | | age > 54.500

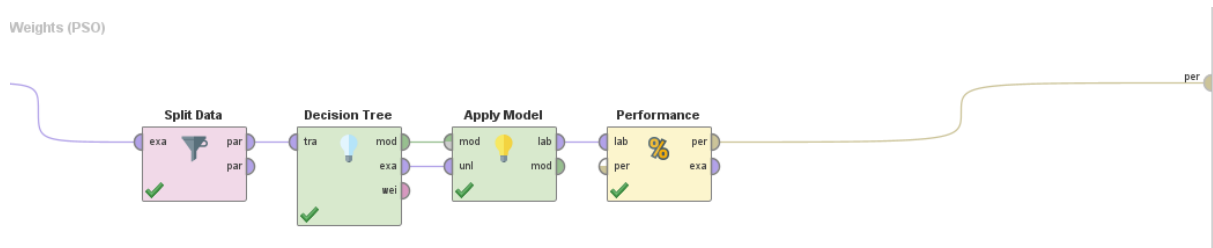
| | | | id > 3102.500

| | | | | age > 71.500: 0 {1=1, 0=20}
| | | | | age ≤ 71.500
| | | | | smoking_status = Unknown
| | | | | Residence_type = Rural: 0 {1=0, 0=10}
| | | | | Residence_type = Urban
| | | | | age > 66: 0 {1=0, 0=2}
| | | | | age ≤ 66: 1 {1=4, 0=1}
| | | | | smoking_status = formerly smoked: 0 {1=6, 0=24}
| | | | | smoking_status = never smoked
| | | | | age > 68.500
| | | | | Residence_type = Rural: 1 {1=2, 0=0}
| | | | | Residence_type = Urban: 0 {1=1, 0=5}
| | | | | age ≤ 68.500: 0 {1=0, 0=20}
| | | | | smoking_status = smokes
| | | | | ever_married = No: 0 {1=0, 0=3}
| | | | | ever_married = Yes
| | | | | avg_glucose_level > 76.070: 1 {1=10, 0=10}
| | | | | avg_glucose_level ≤ 76.070: 0 {1=0, 0=4}
| | | | id ≤ 3102.500: 0 {1=0, 0=9}
| | | age ≤ 54.500: 0 {1=0, 0=38}
| | heart_disease ≤ 0.500
| | | avg_glucose_level > 262.495
| | | Residence_type = Rural: 0 {1=0, 0=2}
| | | Residence_type = Urban: 1 {1=1, 0=1}
| | | avg_glucose_level ≤ 262.495
| | | age > 67.500
| | | | avg_glucose_level > 70.470
| | | | id > 69477.500
| | | | | ever_married = No: 1 {1=2, 0=0}
| | | | | ever_married = Yes: 0 {1=1, 0=6}
| | | | | id ≤ 69477.500: 0 {1=24, 0=179}
| | | | avg_glucose_level ≤ 70.470: 0 {1=0, 0=25}
| | | | age ≤ 67.500
| | | | age > 44.500
| | | | | avg_glucose_level > 59.855
| | | | | id > 2360: 0 {1=70, 0=1440}
| | | | | id ≤ 2360
| | | | | id > 2320: 1 {1=2, 0=0}
| | | | | id ≤ 2320: 0 {1=3, 0=35}
| | | | | avg_glucose_level ≤ 59.855: 0 {1=0, 0=64}

| | | | | $\text{age} \leq 44.500: 0 \{1=10, 0=2515\}$



Gambar 5. Proses 2 Particle Swarm Optimization (PSO) dan Decision Tree C4.5



Gambar 6. Proses 2 Particle Swarm Optimization (PSO) dan Decision Tree C4.5

accuracy: 96.11%

	true 1	true 0	class precision
pred. 1	51	16	76.12%
pred. 0	123	3387	96.50%
class recall	29.31%	99.53%	

Gambar 7. Nilai Accuracy Particle Swarm Optimization (PSO) dan Decision Tree C4.5

precision: 96.50% (positive class: 0)

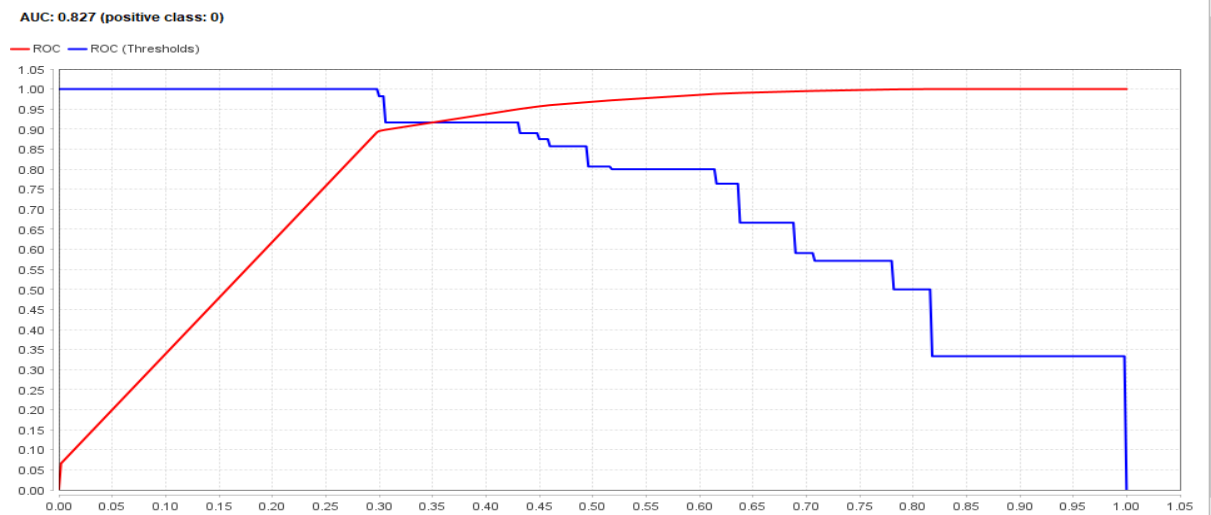
	true 1	true 0	class precision
pred. 1	51	16	76.12%
pred. 0	123	3387	96.50%
class recall	29.31%	99.53%	

Gambar 8. Nilai Precision Particle Swarm Optimization (PSO) dan Decision

recall: 99.53% (positive class: 0)

	true 1	true 0	class precision
pred. 1	51	16	76.12%
pred. 0	123	3387	96.50%
class recall	29.31%	99.53%	

Gambar 9. Nilai Confusion Matrix Recall Particle Swarm Optimization (PSO) dan Decision Tree C4.5



Gambar 10. Nilai dan Curve AUC Particle Swarm Optimization (PSO) dan Decision Tree C4.5

PerformanceVector:

accuracy: 96.11%

ConfusionMatrix:

True:	1	0
1:	51	16
0:	123	3387

precision: 96.50% (positive class: 0)

ConfusionMatrix:

True:	1	0
1:	51	16
0:	123	3387

recall: 99.53% (positive class: 0)

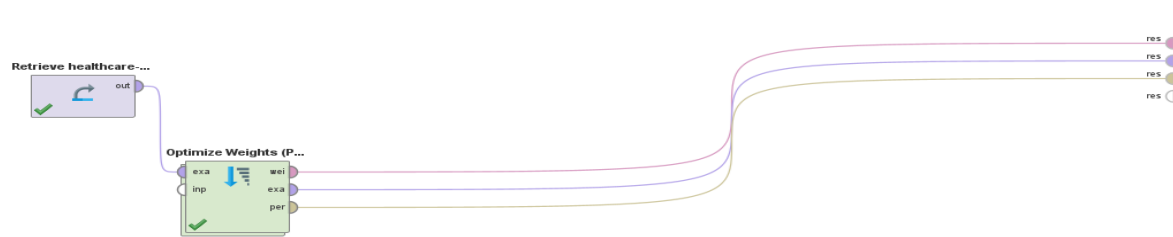
ConfusionMatrix:

True:	1	0
1:	51	16
0:	123	3387

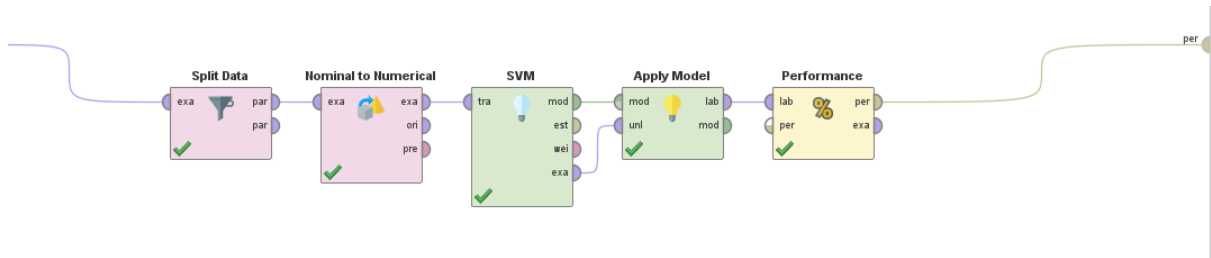
AUC (optimistic): 0.957 (positive class: 0)

AUC: 0.827 (positive class: 0)

AUC (pessimistic): 0.698 (positive class: 0)



Gambar 11. Proses 2 Particle Swarm Optimization (PSO) dan Decision Tree C4.5

**Gambar 12.** Proses 2 Particle Swarm Optimization (PSO) dan Decision Tree C4.5

accuracy: 95.33%

	true 1	true 0	class precision
pred. 1	7	0	100.00%
pred. 0	167	3403	95.32%
class recall	4.02%	100.00%	

Gambar 13. Nilai Accuracy Particle Swarm Optimization (PSO) dan SVN

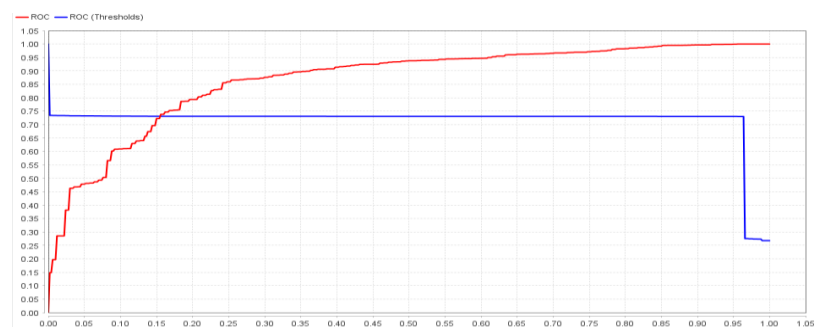
precision: 95.32% (positive class: 0)

	true 1	true 0	class precision
pred. 1	7	0	100.00%
pred. 0	167	3403	95.32%
class recall	4.02%	100.00%	

Gambar 14. Nilai precision Particle Swarm Optimization (PSO) dan SVN

recall: 100.00% (positive class: 0)

	true 1	true 0	class precision
pred. 1	7	0	100.00%
pred. 0	167	3403	95.32%
class recall	4.02%	100.00%	

Gambar 15. Nilai Confusion Matrix Recall Particle Swarm Optimization (PSO) dan SVN**Gambar 16.** Nilai dan Curve AUC Particle Swarm Optimization (PSO) dan SVN

PerformanceVector:

accuracy: 95.33%

ConfusionMatrix:

True:	1	0
1:	7	0
0:	167	3403

precision: 95.32% (positive class: 0)

ConfusionMatrix:

True:	1	0
1:	7	0
0:	167	3403

recall: 100.00% (positive class: 0)

ConfusionMatrix:

True:	1	0
1:	7	0
0:	167	3403

AUC (optimistic): 0.867 (positive class: 0)

AUC: 0.867 (positive class: 0)

AUC (pessimistic): 0.867 (positive class: 0)

5. KESIMPULAN

Dari penelitian yang telah dilakukan pembuatan menggunakan algoritma C4.5, dan Support Vector Machine dengan berbasis Particle Swarm Optimazation dan menggunakan data penderita penyakit stroke Model yang dihasilkan dikomparasi untuk dapat diketahui metode terbaik hasil pengujian dengan mengukur kinerja ketiga metode tersebut menggunakan confusion matrix, kurva ROC. Metode C4.5 berbasis PSO menghasilkan nilai akurasi 96,11 % dan nilai AUC 0,827. Metode Support Vector Machine berbasis PSO menghasilkan nilai akurasi 95,33 % dan nilai AUC 0,867. Nilai Akurasi Tertinggi yaitu pada algoritma C4.5 dan Nilai AUC tertinggi yaitu pada algoritma Support Vector Machine berbasis PSO. Dengan demikian metode Support Vector Machine dan algoritma C4.5 berbasis Particle Swarm Optimazation adalah metode yang terbaik.

DAFTAR PUSTAKA

- Akbar, F., Wira Saputra, H., Karel Maulaya, A., Fikri Hidayat, M., & Rahmadden. (2022). *Implementasi Algoritma Decision Tree C4.5 dan Support Vector Regression untuk Prediksi Penyakit Stroke*. 2(October), 61–67.
- Azizah, N. (2021). *Komparasi Metode Klasifikasi Decision Tree Algoritma C4.5 Dan Random Forest Untuk Prediksi Penyakit Stroke*.
- Ibrahim Irawan, Gani Hilmanasyah, L. R. H. Y. (2022). Perbandingan algoritma naïve bayes dan c4.5 untuk klasifikasi bantuan rumah sehat. *Juik (Jurnal Ilmu Komputer)*, 2.
- Nurlia, E., & Enri, U. (2021). Penerapan Fitur Seleksi Forward Selection Untuk Menentukan Kematian Akibat Gagal Jantung Menggunakan Algoritma C4.5. *Jurnal Teknik Informatika Musirawas*) Elin Nurlia, 6(1), 42.
- Putra, A., Haeirudin, D., Khairunnisa, H., & Latifah, R. (2021). Analisis Sentimen Masyarakat Terhadap Kebijakan PPKM Pada Media Sosial Twitter Menggunakan Algoritma Svm. *Seminar Nasional Sains Dan Teknologi 2021*,

November, 1–6.

- Rahim, A. M. A., Sunyoto, A., & Arief, M. R. (2022). Stroke Prediction Using Machine Learning Method with Extreme Gradient Boosting Algorithm. *MATRIK: Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 21(3), 595–606. <https://doi.org/10.30812/matrik.v21i3.1666>
- Reza, D. A. M., Siregar, A. M., & Rahmat. (2022). Penerapan Algoritma K-Nearest Neighbord Untuk Prediksi Kematian Akibat Penyakit Gagal Jantung. *Scientific Student Journal for Information, Technology and Science*, III(1), 105–112.
- Rohman, R. S., Saputra, R. A., & Firmansaha, D. A. (2020). Komparasi Algoritma C4.5 Berbasis PSO Dan GA Untuk Diagnosa Penyakit Stroke. *CESS (Journal of Computer Engineering, System and Science)*, 5(1), 155. <https://doi.org/10.24114/cess.v5i1.15225>
- Saputra, D., Irmayani, W., Purwaningtias, D., & Sidauruk, J. (2021). A Comparative Analysis of C4.5 Classification Algorithm, Naïve Bayes and Support Vector Machine Based on Particle Swarm Optimization (PSO) for Heart Disease Prediction. *International Journal of Advances in Data and Information Systems*, 2(2), 84–95. <https://doi.org/10.25008/ijadis.v2i2.1221>
- Saputri, N. D., Khalid, K., Rolliawati, D., Informasi, P. S., Islam, U., Sunan, N., Gresik, K., Timur, J., Disease, D. S., & Tree, D. (2022). *Komparasi Penerapan Metode Bagging dan Adaboost pada Algoritma C4 . 5 untuk Prediksi Penyakit Stroke*. 11(September), 567–577.
- Sri Diantika, Windu Gata, & Hiya Nalatissifa. (2021). Komparasi Algoritma SVM Dan Naive Bayes Untuk Klasifikasi Kestabilan Jaringan Listrik. *Elkom : Jurnal Elektronika Dan Komputer*, 14(1), 10–15. <https://doi.org/10.51903/elkom.v14i1.319>
- Sulaeman, K. R., Setianingsih, C., & Saputra, R. E. (2022). Analisis Algoritma Support Vector Machine Dalam Klasifikasi Penyakit Stroke. *EProceedings of Engineering*, 9(3), 922–928. <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/17909/17544%0Ahttps://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/17909>