JURNAL TEKNOINFO
Volume 17, Nomor 2, Juli 2023, Page 351-357
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

PENERAPAN ALGORITMA NAÏVE BAYES DAN FORWARD SELECTION UNTUK PREDIKSI PENYAKIT STROKE

Tetra Praja Utama[1)], M. Said Haibuan[2)]

[1,2]Fakultas Ilmu Komputer, IIB Darmajaya
[1,2]Jl. ZA. Pagar Alam No.93, Gedong Meneng, Kec. Rajabasa, Kota Bandar Lampung, Lampung
Email: [1] Tetra.praja@gmail.com, [2]M.said@gmail.com

Abstract

Stroke is a major health problem in today's elite society. Currently, stroke is a serious problem that can occur in almost all parts of the world. A sudden stroke can cause death, physical and mental disability in both able-bodied and elderly people. In order to get batch data, one has to use data mining methods for classification, like Classification is the process of defining a concept or data class that describes or differentiates instances to evaluate an unknown class of objects. When classifying multiple attributes, a set of records is also defined, also known as an array. In continuous or categorical form, one of the attributes indicates the category of the record. With the dash problem above, of course, you have 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 Naïve Bayes Algorithm to classify the most important factors for this disease. Testing resulted in a fairly high accuracy of the Naïve Bayes Algorithm, which is equal to 95.13, but the results of this accuracy can still be improved by conducting further research to produce higher accuracy.
Keywords: *Accuracy, Stroke, Naïve Bayes Algorithm*

Abstrak

Stroke merupakan masalah kesehatan utama dalam masyarakat elit saat ini. Saat ini, stroke merupakan masalah serius yang dapat terjadi hampir di seluruh dunia. *Stroke* yang tiba-tiba dapat menyebabkan kematian, cacat fisik dan mental baik pada orang yang berbadan sehat maupun lanjut usia. Untuk mendapatkan data batch, seseorang harus menggunakan metode penambangan data untuk klasifikasi, seperti Klasifikasi adalah proses mendefinisikan konsep atau kelas data yang menjelaskan atau membedakan instance untuk mengevaluasi kelas objek yang tidak diketahui. Saat mengklasifikasikan beberapa atribut, sekumpulan *record* juga ditentukan, juga dikenal sebagai *array*. Dalam bentuk kontinyu atau kategorikal, salah satu atribut menunjukkan kategori rekaman. Dengan permasalahan dash diatas tentunya harus bisa mengatasi masalah tersebut banyak dilakukan penelitian dalam bidang ilmu komputer diantaranya adalah Klasifikasi Penderita Penyakit *Stroke* Menggunakan Algoritma *Naïve Bayes* untuk mengklasifikasikan faktor *withering* penting untuk penyakit ini. Pengujian menghasilkan akurasi Algoritma *Naïve Bayes* yang cukup tinggi yaitu sebesar 95.13 namun hasil akurasi tersebut masih dapat ditingkatkan lagi dengan melakukan penelitian lanjutan untuk menghasilkan akurasi lebih tinggi
Kata kunci: Akurasi, *Stroke*, *Algoritma Naïve Bayes*

1. PENDAHULUAN

Cerebral palsy adalah gangguan otak yang bermanifestasi sebagai gangguan lokal dan/atau global pada Aktivitas sistem saraf dan terjadi secara tiba-tiba, Lambat dan penurunan fungsi saraf yang cepat pada *stroke* yang disebabkan oleh penyakit serebrovaskular non-trauma. Penyakit *Hermos* menimbulkan gejala seperti: Kelumpuhan wajah atau anggota badan, bicara cadel, bicara cadel dan cepat. Gangguan fungsi saraf pada *stroke* disebabkan oleh gangguan *serebrovaskular* non traumatik. Penyakit *Hermos* menimbulkan gejala seperti:kelumpuhan wajah atau anggota badan, bicara cadel,l[1] kemungkinan perubahan kesadaran, gangguan penglihatan dan lain-lain. Anda mengalami *stroke* jika Anda pernah didiagnosis *stroke* oleh ahli kesehatan (dokter/perawat/bidan) atau belum pernah didiagnosis *stroke* oleh ahli kesehatan tetapi tiba-tiba mengalami kelumpuhan moncong tanpa otot mata atau bicara, atau bicara cadel.Kesulitan bahasa/komunikasi dan/atau ketidak mampuan memahami bahasa, jumlah korban *stroke* di Indonesia pada tahun 2013 diperkirakan sebanyak 1.236.825 orang (7,0‰). Kemungkinan adanya perubahan kesadaran, gangguan penglihatan dan lain-lain akan ditentukan oleh tenaga medis. Anda telah mengalami *stroke* jika *stroke* telah didiagnosis dokter (dokter/perawat/bidan) atau jika dokter pernah mendiagnosa anda terkena *stroke* tetapi anda tiba-tiba terkena *stroke* tanpa otot mata atau ucapan atau ucapan yang tidak jelas /kesulitan komunikasi dan/atau Ketidak mampuan memahami bahasa akan mempengaruhi sekitar 1.236.825 korban *stroke* (7,0 ‰) di Indonesia pada tahun 2013. Seperti yang didefinisikan

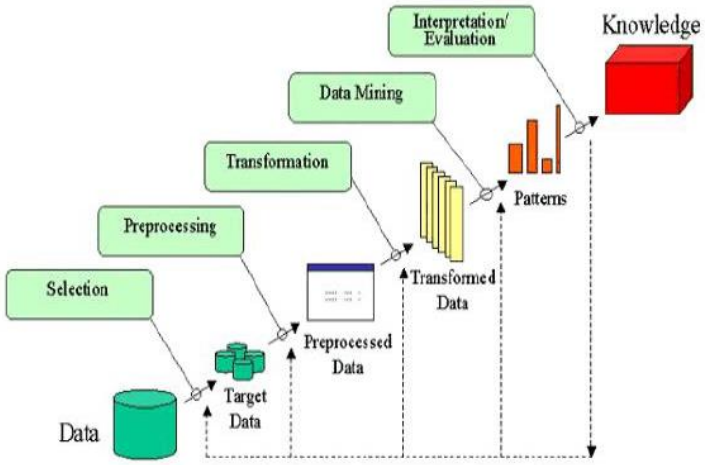
JURNAL TEKNOINFO
Volume 17, Nomor 2, Juli 2023, Page 351-357
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

oleh otoritas kesehatan)[2], Berdasarkan definisi tenaga medis/gejala, diperkirakan sebanyak 2.137.941 orang (12,1‰). Berdasarkan identifikasi dan diagnosis/gejala oleh petugas kesehatan, provinsi Jawa Barat memiliki pasien terbanyak yaitu 238.001 orang (7,4‰) dan 533.895 orang (16,6‰), sedangkan provinsi Papua Barat memiliki jumlah pasien paling sedikit. jumlah pasien bertambah 111. 2.007 orang (3,6‰) dan 2.955 orang (5,3%)[3]
Penelitian ini dilakukan untuk prediksi *Stroke* dengan menggunakan algoritma *Naïve Bayes.* Salah satu manfaat algoritma *Naïve Bayes* dapat digunakan untuk membantu pengambilan keputusan Prediksi Penyakit *Stroke* dan menganalisa penyakit *Stroke*. Dan digunakan untuk mencari nilai probabilitas masing-masing kriteria. Sedangkan *Forward selection* merupakan salah satu metode yang sering dipakai dalam *feature selection* untuk mencari nilai tengah *Aprecici. F -Measure dihitung* dengan menggabungkan *Precision* dan *Recall* menjadi satu nilai. *F-measure = 2 * (precision * recall) / (precision + recall) precision* adalah rasio dari jumlah prediksi yang benar dalam kelas positif terhadap jumlah total prediksi dalam kelas positif. *recall* adalah rasio dari jumlah prediksi yang benar dalam kelas positif terhadap jumlah total item dalam kelas positif. *Nilai F-measure* berkisar antara 0 dan 1, di mana nilai yang lebih tinggi menunjukkan kinerja yang lebih baik dari sistem klasifikasi.
Perkembangan teknologi sangat mempercepat perkembangan ilmu pengetahuan dan teknologi informasi, bahkan telah merambah banyak bidang, dan salah satu kelebihannya adalah terciptanya sistem peramalan. Berbagai metode klasifikasi banyak digunakan untuk memprediksi penyakit [4][5]. Sistem prediksi sangat berguna untuk memprediksi sesuatu yang kemungkinan akan terjadi berdasarkan informasi yang ada, untuk mengatasi kemungkinan kerugian atau untuk memaksimalkan kemungkinan keuntungan.
Berdasarkan hasil pengujian metode *Naive Bayes* dan *SVM* (*Support Vector Machine*) memprediksi keberhasilan pengobatan imunoterapi kutil menggunakan bahasa pemrograman R *Studio*, dapat disimpulkan bahwa hasil pengujian dapat memprediksi kelas untuk semua orang. Kasus benar pada tingkat akurasi 1, sedangkan metode *SVM* masih memberikan prediksi yang salah pada tingkat akurasi 0,8. Hasil pengujian menunjukkan bahwa metode *Naive Bayes* merupakan metode yang lebih baik dibandingkan dengan *Support Vector Machine* (SVM)[6]. Klasifikasi adalah teknik pengolahan data yang membagi objek menjadi beberapa kelas sesuai dengan jumlah kelas yang diinginkan. Salah satu algoritma metode klasifikasi adalah *algoritma Naïve Bayes. Forward selection* merupakan salah satu metode yang sering dipakai dalam *feature selection* mampu meningkatkan nilai akurasi.

2. METODE PENELITIAN

2.1 Alur Penelitian

Penelitian ini menggunakan *Knowledge Discovery* in *Database* (*KDD*)[7], tahapan model proses seperti yang ditunjukkan pada Gambar 1 Penjelasan dari setiap langkah penelitian seperti pada gambar 1 dibawah ini:



Gambar 1. Proses *Data Mining*[8]

Gambar 1 mengilustrasikan proses penambangan data, di mana data dipilih, dibersihkan, dan diproses sebelumnya sesuai dengan panduan dan pengetahuan para ahli di lapangan, yang mengumpulkan data internal dan eksternal dan mengintegrasikannya ke dalam gambaran keseluruhan organisasi. Mengunakan algoritma penambangan data yang

JURNAL TEKNOINFO
Volume 17, Nomor 2, Juli 2023, Page 351-357
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

diimplementasikan pada langkah ini untuk memeriksa data terintegrasi dan mengidentifikasi informasi berharga dengan mudah. Hasil penambangan data dievaluasi untuk menentukan apakah domain data dapat ditemukan dalam bentuk aturan yang diekstraksi dari jaringan. KDD (Penemuan Pengetahuan di *Database*)[9] bekerja sebagai berikut:

1. Data *Selelection*
   Pemilihan informasi dari informasi fungsional harus dilakukan pada tahap ekstraksi data sebelum memulai KDD.
2. *Preporcessing*
   Sebelum data mining dapat dilakukan, harus dilakukan proses pembersihan untuk menghapus duplikat data, memeriksa data yang tidak konsisten, dan kesalahan pada data seperti *error*. kesalahan diperbaiki. Proses pengayaan juga dilakukanyaitu, untuk mengkonfirmasi informasi yang ada dengan informasi lain yang relevan dan diperlukan untuk KDD, seperti informasi eksternal.
3. *Transformation*
   Kode data yang dipilih dengan cara yang membuatnya cocok untuk proses penambangan data. Proses pengkodean KDD adalah proses kreatif dan sangat bergantung pada jenis atau model data yang diambil dari *database*.
4. Penambangan data
   Proses menemukan pola atau informasi yang menarik dari data terpilih dengan menggunakan teknik atau metode tertentu
5. Interpretasi/Evaluasi
   Model data yang dibentuk dalam proses data mining wajib tersaji pada bentuk yang gampang dipahami pihak yang berkepentingan. langkah ini adalah bagian berdasarkan proses KDD yang dikenal menjadi rendering. Pada termin ini, kami menyelidiki apakah pola atau inputan yg ditemukan bertentangan menggunakan kabar atau hipotesis yg terdapat sebelumnya.

Evaluasi memainkan kiprah kunci pada membentuk pelaksanaan berbasis *datamining*. Ada beberapa cara buat melakukan evaluasi. Evaluasi nir semudah yg kita bayangkan. Ketika kami mempunyai data yang kami pakai pada pelatihan, kami nir dan merta menggunakannya menjadi indikator keberhasilan pelaksanaan kami. Oleh lantaran itu, diharapkan metode spesifik buat prediksi kinerja dari eksperimen menggunakan tipe data selain data pelatihan.[10]

Biasanya data yang cukup dapat digunakan untuk pengujian. Masalah umum adalah data. Oleh karena itu, kami perlu memastikan bahwa data yang kami gunakan untuk pelatihan dan pengujian berkualitas tinggi. [11]

2.2 *Algoritma Naïve Bayes*

Algoritma Naive Bayes merupakan salah satu metode yang dapat digunakan untuk mengklasifikasikan suatu kumpulan data. Algoritma ini menggunakan metode probabilitas dan statistik yang diusulkan oleh ilmuwan Inggris *Thomas Bayes* untuk memprediksi probabilitas masa depan berdasarkan pengalaman masa lalu. [8]

Naive Bayes adalah pembelajaran mesin yang menggunakan perhitungan probabilitas menggunakan konsep pendekatan Bayesian. Kata merendahkan naif berasal dari asumsi bahwa pengaruh skor atribut tidak bergantung pada kemungkinan kelas tertentu dari skor atribut lainnya. Menggunakan *teorema Bayes* dalam *algoritme naive Bayes* terdiri dari menggabungkan probabilitas sebelumnya dan probabilitas bersyarat dalam sebuah formula yang bisa dipakai buat menghitung probabilitas berdasarkan setiap kemungkinan klasifikasi.[12]

Rumus *Naïve Bayes* adalah:$P(H|X) = \frac{P(H)P(X|H)}{P(X)}$ (2.2)

Informasi:

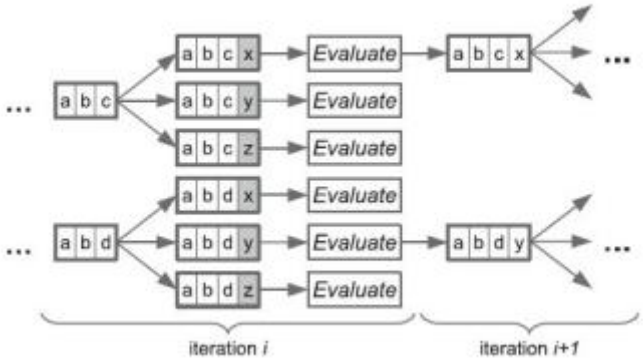
X = data kelas tidak diketahui
H = hipotesis data X, merupakan kategori tertentu
P(H|X) = probabilitas hipotesis H untuk kondisi X (probabilitas posterior)

2,3 *Forward selection*

Forward selection adalah prosedur langkah demi langkah yang tujuannya adalah untuk menambahkan variabel yang

JURNAL TEKNOINFO
Volume 17, Nomor 2, Juli 2023, Page 351-357
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

dimanipulasi ke persamaan satu persatu. Seleksi awal dimulai dengan seperangkat karakteristik kosong, kemudian ditambahkan karakteristik yang digunakan pada putaran pertama, semua karakteristik dievaluasi secara individual. Properti yang merupakan bagian dari properti yang sudah ada dan yang baru dibuat ditambahkan ke kumpulan properti dan dinilai kembali. Untuk mengurangi jumlah evaluasi, hanya himpunan bagian dari fitur terbaik yang disimpan, seperti yang ditunjukkan pada Gambar 2.[2]



Gambar 2. Metode *Forward Selection*

Data yang akan dilatih bervariasi dari satu variabel ke tingkat atau jumlah variabel yang menghasilkan nilai efisiensi, akurasi, atau kesalahan yang paling rendah. Misalnya pengujian data dengan dua variabel memberikan kesalahan yang lebih kecil, dan jika pengujian ulang dengan tiga variabel memberikan nilai kesalahan yang lebih besar untuk dua variabel, kesalahan terkecil masih ada pada variabel kedua, artinya variabel kedua penting. proses berhenti ketika semua variabel independen telah diuji. Algoritma pemilihan diuji pada setiap kombinasi data dari data variabel 1 periode hingga data variabel 10 periode untuk membandingkan data mana yang menunjukkan akurasi terbaik [10]

3. HASIL DAN PEMBAHASAN

3.1 Datasets

Pengolahan *dataset* Stroke sebanyak 5110 *records*. terdiri dari 12 atribut yang terdari dari. *Gender, Age, hypertension, heart_disease, ever_married work_type, Residence_type, avg_glucose_level, bmi dan smoking_statu s*akan dilakukan beberapa penyeleksian untuk menghasilkan data yang dibutuhkan.untuk pembentukan pola algoritma klasifikasi data mining dari metode *Naïve Bayes* dan *forward Selection*.

w No.	stroke	id	gender	age	hypertension	heart_disea...	ever_married	work_type	Residence_t...	avg_glucos...	I
97	0	41512	Male	57	0	0	Yes	Govt_job	Rural	76.620	2
98	0	64520	Male	68	0	0	Yes	Self-employed	Urban	91.680	4
99	0	579	Male	9	0	0	No	children	Urban	71.880	1
00	0	7293	Male	40	0	0	Yes	Private	Rural	83.940	1
01	0	68398	Male	82	1	0	Yes	Self-employed	Rural	71.970	2
02	0	36901	Female	45	0	0	Yes	Private	Urban	97.950	2
03	0	45010	Female	57	0	0	Yes	Private	Rural	77.930	2
04	0	22127	Female	18	0	0	No	Private	Urban	82.850	4
05	0	14180	Female	13	0	0	No	children	Rural	103.080	1
06	0	18234	Female	80	1	0	Yes	Private	Urban	83.750	1
07	0	44873	Female	81	0	0	Yes	Self-employed	Urban	125.200	4
08	0	19723	Female	35	0	0	Yes	Self-employed	Rural	82.990	3
09	0	37544	Male	51	0	0	Yes	Private	Rural	166.290	2
10	0	44679	Female	44	0	0	Yes	Govt_job	Urban	85.280	2

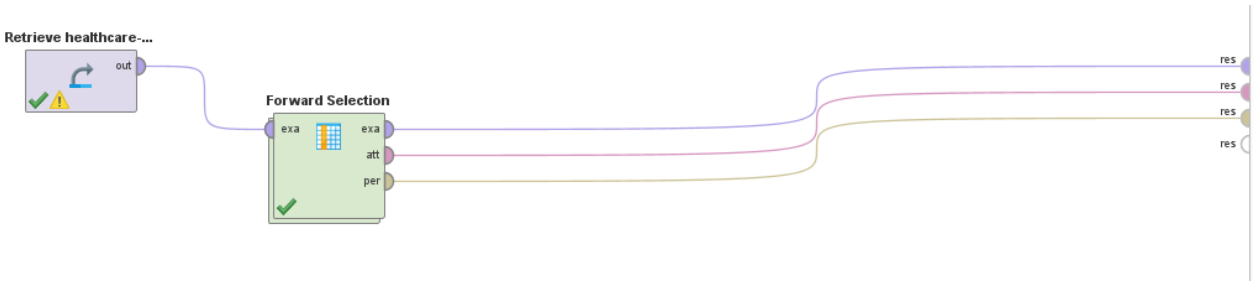
Gambar 3. Potongan Dataset

JURNAL TEKNOINFO
Volume 17, Nomor 2, Juli 2023, Page 351-357
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

3.2 Pembahasan

3.1.2 Eksprimen dan Pengujian

Dalam pengujian ini, penulis menggunakan *Cross Validation.* Pengelola validasi silang ini selanjutnya memisahkan data dengan cara memisahkan data agregat dari data awal dan sisanya untuk validasi data. Gambar 4 di bawah ini menunjukkan penggunaan data yang digunakan dalam klasifikasi *stroke* menggunakan *algoritma Naive Bayes.*
Proses untuk klasifikasi dengan menggunakan *Forward Selection* mencari Nilai *Accuracy, Precesion dan Reccall* Seperti gambar dibawah ini.

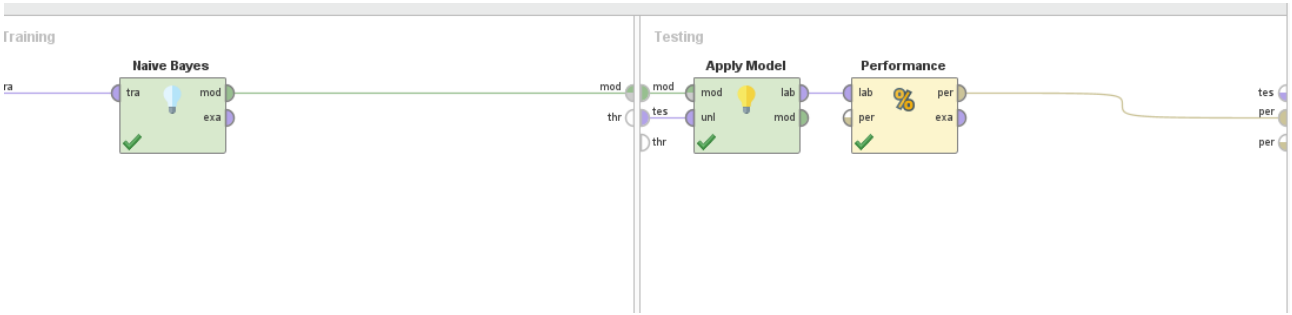


Gambar 4. Uji dengan Rapid Miner

Proses untuk klasifikasi dengan menggunakan *Forward Selection dan Cross Validation* mencari Nilai *Accuracy, Precesion dan Reccall* Seperti gambar dibawah ini.

Gambar 5. Uji Dengan Rapid Miner

Proses untuk klasifikasi dengan menggunakan *Forward Selection. Naive Bayes, Aply Model* dan *Performance Cross Validation* mencari Nilai *Accuracy, Precesion dan Reccall* Seperti gambar dibawah ini



Gambar 6. Uji Dengan Rapid Miner

3.2.2 Evaluasi dan Validasi

Menguji *Algoritma Naive Bayes* Pada langkah ini peneliti menggunakan metode *Algoritma Naive* Bayes untuk pemprosesan data atau penerapan data yang dibersihkan di *Rapidminer*. Berdasarkan pengujian yang dilakukan dengan aplikasi *Rapidminer*, diperoleh hasil sebagai berikut: *Algoritma Naive Bayes* mencapai akurasi sebesar 95,13%
Hasil dari proses *dataset* penyakit *stroke* didapatkan nilaii *Accuracy* 95,16% Seperti gambar dibawah ini

JURNAL TEKNOINFO
Volume 17, Nomor 2, Juli 2023, Page 351-357
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

accuracy: 95.13% +/- 0.06% (micro average: 95.13%)

	true 1	true 0	class precision
pred. 1	0	0	0.00%
pred. 0	249	4861	95.13%
class recall	0.00%	100.00%	

Gambar 7. Nilai *Accuracy*

Hasil dari proses *dataset* penyakit *stroke* didapatkan nilai *Precisions* 95,13% Seperti gambar dibawah ini

precision: 95.13% +/- 0.06% (micro average: 95.13%) (positive class: 0)

	true 1	true 0	class precision
pred. 1	0	0	0.00%
pred. 0	249	4861	95.13%
class recall	0.00%	100.00%	

Gambar 8. Nilai Presisions

Hasil dari proses *dataset* penyakit *stroke* didapatkan nilai *Recall* 100,00 % Seperti gambar dibawah ini

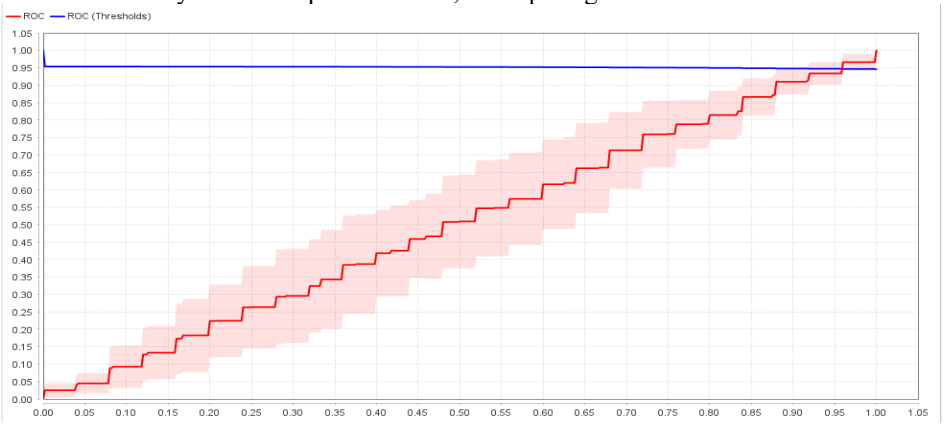
recall: 100.00% +/- 0.00% (micro average: 100.00%) (positive class: 0)

	true 1	true 0	class precision
pred. 1	0	0	0.00%
pred. 0	249	4861	95.13%
class recall	0.00%	100.00%	

Gambar 9. Nilai *Recall*

Perhitungan F=*Measure* nilai Tengah *recall aprecici*.
F1 Score = 2 * (Recall*Precission) / (*Recall* + *Precission*)
dalam kasus di atas, F1 *Score* = 2* (95,13%*100.00%) / (95,13% + 100,00%) = 9.513*2=19.02600/195.13=9,750%

Berikut adalah hasil *kurva AUC* yaitu mendapatkan nilai 0,089 seperti gambar dibawah in



Gambar 10. *Kurva AUC*

JURNAL TEKNOINFO
Volume 17, Nomor 2, Juli 2023, Page 351-357
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

Tabel 1. hasiil uji

Naïve Bayes	Akurasi	Presisi	Recall	AUC
	95.13%	95.13%	100%	0.089
Perhitungan *F=Masure Algoritma Nive Bayes*	F1 *Score* = 2* (95,13%*100.00%) / (95,13% + 100,00%) = 9.513*2=19.02600/195.13=9,750%			

4. KESIMPULAN

penelitian yang telah dilakukan menggunakan algortima *Naïve Bayes* dengan *Forward Determination*s dan menggunakan dataset Penyakit *Stoke*.untuk menentukan metode terbaik guna memperoleh hasil pengujian dengan mengukur keefektifan kedua metode tersebut menggunakan kurva *ROC*. *Metode algoritma Niave-Bayes* memberikan skor akurasi sebesar 95,13 dengan nilai *AUC* sebesar 0,087.

DAFTAR PUSTAKA

[1] D. Prajarini, S. Tinggi, S. Rupa, D. Desain, and V. Indonesia, "Perbandingan Algoritma Klasifikasi Data Mining Untuk Prediksi Penyakit Kulit," *Informatics J.*, vol. 1, no. 3, p. 137, 2016.

[2] E. Nurlia and U. Enri, "Penerapan Fitur Seleksi Forward Selection Untuk Menentukan Kematian Akibat Gagal Jantung Menggunakan Algoritma C4.5," *J. Tek. Inform. Musirawas) Elin Nurlia*, vol. 6, no. 1, p. 42, 2021.

[3] D. Haryadi, D. Marini Umi Atmaja, A. Rahman Hakim, and N. Suwaryo, "Identifikasi Tingkat Resiko Penyakit Stroke Menggunakan Algoritma Regresi Linear Berganda," *Deny Haryadi, SNTEM*, vol. 1, no. November, pp. 1198–1207, 2021.

[4] N. Azizah, "Komparasi Metode Klasifikasi Decision Tree Algoritma C4.5 Dan Random Forest Untuk Prediksi Penyakit Stroke," 2021.

[5] D. A. M. Reza, A. M. Siregar, and Rahmat, "Penerapan Algoritma K-Nearest Neighbord Untuk Prediksi Kematian Akibat Penyakit Gagal Jantung," *Sci. Student J. Information, Technol. Sci.* , vol. III, no. 1, pp. 105–112, 2022.

[6] "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/engineeri."

[7] H. Bugis, "Metode Naïve Bayes Untuk Memprediksi Penyakit Stroke," *J. SISKOM-KB (Sistem Komput. dan Kecerdasan …*, 2022, [Online]. Available: https://www.jurnal.tau.ac.id/index.php/siskom-kb/article/view/317.

[8] I. Lishania, R. Goejantoro, and Y. N. Nasution, "Perbandingan Klasifikasi Metode Naive Bayes dan Metode Decision Tree Algoritma (J48) pada Pasien Penderita Penyakit Stroke di RSUD Abdul Wahab Sjahranie Samarinda," *J. Eksponensial*, vol. 10, no. 2, pp. 135–142, 2019, [Online]. Available: http://jurnal.fmipa.unmul.ac.id/index.php/exponensial/article/view/571.

[9] R. S. Rohman, R. A. Saputra, and D. A. Firmansaha, "Komparasi Algoritma C4.5 Berbasis PSO Dan GA Untuk Diagnosa Penyakit Stroke," *CESS (Journal Comput. Eng. Syst. Sci.*, vol. 5, no. 1, p. 155, 2020, doi: 10.24114/cess.v5i1.15225.

[10] A. F. Hermawan, F. R. Umbara, and F. Kasyidi, "Prediksi Awal Penyakit Stroke Berdasarkan Rekam Medis menggunakan Metode Algoritma CART (Classification and Regression Tree)," vol. 7, no. 2, pp. 151–164, 2022.

[11] N. D. Saputri *et al.*, "Komparasi Penerapan Metode Bagging dan Adaboost pada Algoritma C4 . 5 untuk Prediksi Penyakit Stroke," vol. 11, no. September, pp. 567–577, 2022.

[12] F. Akbar, H. Wira Saputra, A. Karel Maulaya, M. Fikri Hidayat, and Rahmaddeni, "Implementasi Algoritma Decision Tree C4.5 dan Support Vector Regression untuk Prediksi Penyakit Stroke," vol. 2, no. October, pp. 61–67, 2022.