

Metode Naïve Bayes Untuk Memprediksi Penyakit Stroke

Haris

Program Studi Informatika, Universitas Nusa Mandiri
Jl. Kramat Raya No.18, RW.7, Daerah Khusus Ibukota Jakarta 10450
haris12180247@nusamandiri.ac.id

Diterima : 22 Agustus 2022

Disetujui : 01 Oktober 2022

Abstrak— Teknologi sangat amat berkembang pesat, sampai-sampai manusia sudah hampir terikat sepenuhnya oleh teknologi. Bahkan dibidang Kesehatan juga menggunakan teknologi-teknologi yang canggih sehingga dapat menyelesaikan pekerjaan di bidang kesehatan. Melihat pesatnya perkembangan di dunia Kesehatan, maka telah dilakukan penelitian untuk memprediksi penyakit stroke berdasarkan gaya hidup serta lingkungan masyarakat. Penelitian ini bertujuan untuk mengklasifikasi dan memprediksi orang-orang tentang apakah orang tersebut akan berpotensi terkena penyakit stroke atau tidak. Peneliti menggunakan metode klasifikasi Naive Bayes yang merupakan salah-satu metode sederhana untuk mengklasifikasi dan memprediksi sebuah data. Peneliti mengambil data orang-orang seperti nama, umur, beberapa gejala, lingkungan, dan sebagainya. Peneliti juga menggunakan aplikasi Python 3 untuk melakukan prediksi dan juga Rapidminer untuk melihat tingkat akurasi prediksi dari metode Naive Bayes ini. Maka dapat disimpulkan bahwa metode naïve bayes ini bisa digunakan untuk memprediksi sebuah penyakit seseorang, contohnya penyakit stroke.

Kata Kunci —*klasifikasi, stroke, naïve bayes, prediksi, rapidminer*

I. PENDAHULUAN

Penyakit Stroke adalah penyakit yang menyebabkan kematian tertinggi nomor tiga di Amerika, dan diperkirakan ada 4.5 juta kematian setiap tahun di seluruh dunia yang diakibatkan oleh penyakit ini[1]. Di Indonesia, pengidap penyakit stroke ini mencapai 2.137.941 orang di tahun 2013[2]. WHO atau Organisasi Kesehatan Dunia mendefinisikan penyakit stroke sebagai gangguan fungsi pada otak secara tiba-tiba. Tingginya jumlah kematian akibat penyakit ini dikarenakan ketidak-tahuan orang-orang terhadap penyakit stroke dan akhirnya membuat orang-orang jadi mengabaikan gejala-gejala penyakit ini, padahal jika kita sudah melihat gejalanya dan segera ditangani sedini atau secepat mungkin maka masih ada kemungkinan para pengidap penyakit stroke dapat diobati dan berangsur pulih.

Meski perkembangan di dunia medis begitu pesat, namun masih banyak orang-orang atau masyarakat yang tidak bisa menikmati

perkembangannya. Mereka atau masyarakat yang tidak beruntung ini tidak bisa menikmati perkembangan di dunia kesehatan dikarenakan kurangnya ekonomi mereka. Banyak orang-orang yang tidak bisa berobat atau berkonsultasi tentang penyakit mereka karena tidak memiliki biaya untuk berkonsultasi atau berobat, misalnya jika mereka ingin berkonsultasi tentang apakah mereka berkemungkinan besar terkena penyakit stroke dimasa mendatang atau tidak. Maka dari itu saya berkeinginan untuk membantu mereka agar dapat mendapatkan jawaban tentang keresahan mereka dengan cara menerapkan metode Naïve Bayes untuk memprediksi penyakit stroke pada masyarakat setempat.

Dengan adanya penelitian ini, saya harap bisa membantu mempermudah kita, khususnya masyarakat setempat yang saya survey agar lebih mudah untuk mengetahui apakah mereka berdampak atau berkemungkinan terkena penyakit stroke jika masih menerapkan gaya hidupnya yang sekarang atau tidak. Saya juga berharap setelah

adanya penelitian ini bisa membuat para masyarakat supaya lebih mawas diri dan mengubah gaya hidupnya kearah yang lebih sehat. Selain itu, saya juga menghitung tingkat keberhasilan persentase metode Naïve Bayes ini menggunakan aplikasi Rapid Miner untuk melihat seberapa tinggi akurasi prediksi dari metode Naïve Bayes pada dataset ini.

II. TINJAUAN PUSTAKA

A. Penelitian Terkait

Pada penelitian ini, terdapat beberapa penelitian terkait yang menjadi referensi bagi peneliti untuk menulis penelitian ini, diantaranya ada penelitian dari Nofi Deffia Sari pada tahun 2018, pada penelitian ini disimpulkan bahwa klasifikasi prediksi kepuasan pelanggan dengan metode naïve bayes ini dapat membantu admin untuk menentukan klasifikasi kepuasan masyarakat terhadap layanannya[3].

Ada juga penelitian dari Felisita Klaralisa pada tahun 2019 yang berjudul “Penerapan Algoritma Naïve Bayes untuk Memprediksi Penyakit Malaria Pada Puskesmas Hanura”, di penelitian itu disimpulkan bahwa metode naïve bayes dapat untuk memprediksi penyakit malaria dapat memberikan kemudahan bagi pihak puskesmas dalam meprediksi penyakit malaria pada pasien[4].

Selain itu ada juga penelitian dari Muhammad Zulfiansyah Pratama pada tahun 2020 yang berjudul “Penentuan Kriteria dan Penerima Bantuan Pangan Non tunai di Desa Parungkuda Dengan Menggunakan Algoritma Naïve Bayes”, disitu disimpulkan bahwa naïve bayes sangat membantu petugas pembagian bantuan agar syarat penerima bantuan sesuai dengan kriteria penerima bantuan[5]. Penelitian-penelitian ini adalah penelitian terkait yang menjadi bahan referensi saya dalam melakukan penelitian ini.

B. Stroke

Stroke diperkirakan terus-menerus membuat perhatian para media, pasien, dokter,

dan juga peneliti. Diperkirakan ada 4.5 juta kematian setiap tahun diseluruh dunia yang diakibatkan oleh stroke ini, dan juga ada 9 juta orang lebih yang sedang berjuang melawan stroke. Hampir satu dari empat laki-laki dan satu dari lima perempuan diatas umur empat puluh lima yang diperkirakan terkena stroke. Diperkirakan ditahun 2023 ada peningkatan jumlah orang-orang yang terkena stroke[1].

Stroke atau biasa dipanggil ‘struk’ adalah kondisi dimana aliran darah ke otak terhambat dikarenakan penyumbatan pembuluh darah. Stroke ini bisa terjadi dikarenakan penyumbatan pembuluh darah dan untuk penyumbatannya ini ada dua jenis, yang pertama adalah penyumbatan pembuluh darah atau disebut *Stroke Iskemik*, lalu yang kedua adalah pecahnya pembuluh darah atau disebut *Stroke Hemoragik* dan hampir 70% kasus *Stroke Hemoragik* terjadi pada penderita *Hipertensi* (darah tinggi).

C. Naïve Bayes

Naïve Bayes Classifier adalah sebuah pengklasifikasi yang statis, metode ini dapat memprediksi sebuah probabilitas. Metode Naïve Bayes adalah metode pengklasifikasi yang berdasarkan teorema Bayes[6].

Keuntungan menggunakan metode Naive Bayes salah satunya adalah kita hanya butuh data training yang sedikit untuk memprediksi sebuah dataset. Naïve Bayes adalah penghitungan secara statistik untuk memprediksi peluang yang akan datang berdasarkan pengalaman atau permasalahan yang dihadapi sebelumnya sehingga dikenal sebagai Teorema Bayes[7]. Untuk persamaan dari klasifikasi Naive Bayes adalah sebagai berikut :

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)}$$

- H = Hipotesis data
- X = Data yang tidak diketahui classnya
- P(H) = Kemungkinan hipotesis H
- P(X) = Kemungkinan X

- $P(H|X)$ = Kemungkinan hipotesis H, berdasarkan kondisi X
- $P(X|H)$ = Kemungkinan X, berdasarkan kondisi hipotesis H

D. Data Mining

Data Mining adalah sebuah *Artificial Intelligent* yang sangat powerful, data mining bisa menemukan informasi-informasi yang sangat penting dengan cara menganalisa data-data dari manapun. Data-data yang didapat akan digunakan pada proses selanjutnya yang akan sangat membantu penelitian dalam mendapatkan jawaban atau hasil yang diinginkan[8]. Dengan kata lain, data mining adalah proses yang menggunakan teknik *statistic*, *AI*, *math*, dan juga *machine learning* untuk mendapatkan data-data terkait dari berbagai database.

E. Data Mining

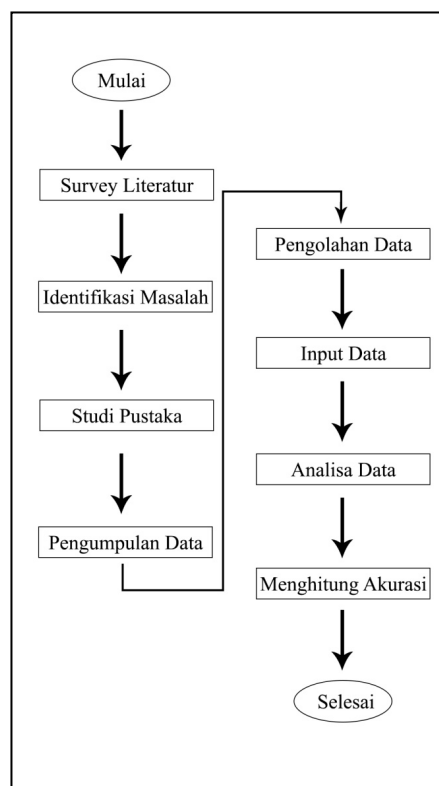
Klasifikasi merupakan teknik dalam *data mining* untuk mengelompokkan data berdasarkan keterikatan data terhadap data sampel[9]. Klasifikasi merupakan proses yang digunakan mencari model atau fungsi yang membedakan kelas sebuah data, tujuannya adalah untuk memprediksi suatu kelas yang labelnya belum diketahui. Model yang ada itu berasal dari kumpulan data training yang sudah dianalisis.

III. METODE PENELITIAN

Metode penelitian merupakan aktifitas untuk mencari dan merumuskan sebuah masalah hingga tersusun sebuah laporan[10]. Tahapan penelitian ditunjukkan dalam gambar 1. Berikut penjelasan dari tahapan penelitian pada Gambar 1 diatas :

- Survey Literatur - Pada tahapan ini, peneliti melakukan survei atau mengumpulkan berbagai informasi atau ilmu-ilmu yang berhubungan tentang judul penelitian.
- Identifikasi Masalah - Pada tahapan ini, peneliti melakukan identifikasi masalah yang berkaitan dengan permasalahan yang akan diteliti, pada penelitian ini

permasalahannya adalah tentang penyakit stroke, cara mencegah penyakit stroke, gejala-gejalanya, dan sebagainya.



Gambar 1. Tahapan penelitian.

- Studi Pustaka – Tahapan dimana peneliti mempelajari, mendalami, dan mengkaji informasi-informasi literatur.
- Pengumpulan Data – Tahapan dimana peneliti mengumpulkan data dengan cara mewawancarai masyarakat dan menyebar kuisioner secara online.
- Pengolahan Data – Tahapan dimana peneliti mengolah data yang dikumpulkan, misalnya membuang data-data yang tidak perlu, mengkategorikan jawaban, dan lain-lain.
- Input Data – Menginput data yang telah diolah ke dalam klasifikasi naïve bayes.
- Analisa Data – Melihat hasil dari proses klasifikasi naïve bayes.
- Menghitung Akurasi – Setelah data dianalisa, selanjutnya adalah menghitung tingkat akurasi dari proses klasifikasi ini.

A. Metodologi Penelitian

Jenis penelitian yang dipakai adalah metode kuantitatif. Penelitian dilakukan pada Jl. Gang Gatep, Mangga Dua Selatan. Peneliti melakukan wawancara kepada masyarakat di Mangga Dua untuk mendapatkan data yang dibutuhkan, selain itu peneliti juga menyebarkan kuisioner secara online ke beberapa sosial media untuk menjangkau lebih banyak orang.

B. Pra-proses Data

Pada Gambar 2 bisa dilihat adalah hasil dari pengumpulan dataset yang telah dilakukan, terdapat beberapa kelas diantaranya adalah Kelamin, Umur, Hipertensi, Jantung, Menikah, Lingkungan, BMI, Rokok, Alkohol, dan Stroke.

Data yang terkumpul berjumlah 109 data, terdiri dari 32 data stroke dan 77 non-stroke. Untuk lebih jelasnya bisa dilihat pada Gambar 2 dibawah ini :

Kelamin	Umur	Hipertensi	Jantung	Menikah	Lingkungan	BMI	Rokok	Alkohol	Stroke
Wanita	22	Tidak	Tidak	Tidak	Perkotaan (Urban)	24.7	Tidak	Tidak	Tidak
Pria	23	Tidak	Tidak	Tidak	Perkotaan (Urban)	17.9	Tidak	Tidak	Tidak
Pria	39	Ya	Tidak	Ya	Perkotaan (Urban)	29.1	Ya	Tidak	Ya
Pria	22	Tidak	Tidak	Tidak	Perkotaan (Urban)	18.4	Tidak	Tidak	Tidak
Pria	44	Ya	Ya	Ya	Perkotaan (Urban)	27.5	Ya	Ya	Ya
Pria	37	Tidak	Tidak	Ya	Perdesaan (Rural)	23.1	Tidak	Tidak	Tidak
Wanita	23	Tidak	Tidak	Tidak	Perdesaan (Rural)	22.3	Tidak	Tidak	Tidak
Wanita	21	Tidak	Tidak	Tidak	Perkotaan (Urban)	26.7	Tidak	Tidak	Tidak
Pria	23	Tidak	Tidak	Tidak	Perkotaan (Urban)	27.1	Tidak	Ya	Tidak
Wanita	23	Tidak	Tidak	Tidak	Perkotaan (Urban)	17.5	Tidak	Tidak	Tidak
Pria	24	Tidak	Tidak	Tidak	Perkotaan (Urban)	22.9	Ya	Ya	Tidak
Wanita	22	Tidak	Tidak	Tidak	Perkotaan (Urban)	21.2	Tidak	Tidak	Tidak
Pria	41	Ya	Tidak	Ya	Perkotaan (Urban)	28.6	Ya	Ya	Ya
Wanita	34	Ya	Ya	Ya	Perdesaan (Rural)	24.4	Tidak	Tidak	Ya
Wanita	31	Ya	Tidak	Ya	Perdesaan (Rural)	23.5	Tidak	Tidak	Ya
Pria	47	Ya	Ya	Ya	Perkotaan (Urban)	26.4	Ya	Tidak	Ya
Wanita	23	Tidak	Tidak	Tidak	Perdesaan (Rural)	24.2	Tidak	Tidak	Tidak
Pria	38	Ya	Ya	Ya	Perkotaan (Urban)	25.4	Ya	Tidak	Ya
Pria	28	Ya	Tidak	Tidak	Perkotaan (Urban)	25.6	Ya	Tidak	Ya
Wanita	23	Tidak	Tidak	Tidak	Perkotaan (Urban)	16.6	Tidak	Tidak	Tidak
Wanita	36	Ya	Tidak	Ya	Perkotaan (Urban)	27.5	Tidak	Tidak	Ya
Wanita	23	Tidak	Tidak	Tidak	Perdesaan (Rural)	17.2	Tidak	Tidak	Tidak
Pria	27	Ya	Tidak	Tidak	Perkotaan (Urban)	23.2	Ya	Tidak	Tidak
Wanita	22	Tidak	Tidak	Tidak	Perkotaan (Urban)	16.2	Tidak	Tidak	Tidak
Pria	47	Ya	Ya	Ya	Perkotaan (Urban)	28.3	Ya	Tidak	Ya
Pria	23	Tidak	Tidak	Tidak	Perkotaan (Urban)	24.4	Tidak	Tidak	Tidak
Wanita	23	Tidak	Tidak	Tidak	Perkotaan (Urban)	19.4	Tidak	Tidak	Tidak
Pria	22	Tidak	Tidak	Tidak	Perkotaan (Urban)	27.8	Tidak	Tidak	Tidak
Wanita	44	Ya	Tidak	Ya	Perkotaan (Urban)	25.6	Tidak	Tidak	Tidak
Pria	42	Ya	Ya	Ya	Perdesaan (Rural)	26.6	Ya	Tidak	Ya
Wanita	40	Ya	Tidak	Ya	Perdesaan (Rural)	24.1	Tidak	Tidak	Tidak
Pria	25	Tidak	Tidak	Tidak	Perdesaan (Rural)	22.3	Ya	Tidak	Tidak
Pria	23	Tidak	Tidak	Tidak	Perdesaan (Rural)	21.5	Tidak	Tidak	Tidak
Pria	24	Tidak	Tidak	Tidak	Perdesaan (Rural)	21.9	Ya	Tidak	Tidak
Pria	25	Ya	Tidak	Tidak	Perkotaan (Urban)	20.4	Ya	Tidak	Tidak
Pria	47	Ya	Tidak	Ya	Perkotaan (Urban)	23.7	Ya	Tidak	Tidak
Wanita	45	Ya	Tidak	Ya	Perkotaan (Urban)	26.2	Tidak	Tidak	Ya
Wanita	19	Tidak	Tidak	Tidak	Perkotaan (Urban)	17.9	Tidak	Tidak	Tidak
Wanita	24	Tidak	Tidak	Tidak	Perkotaan (Urban)	17.2	Tidak	Tidak	Tidak
Wanita	30	Ya	Tidak	Ya	Perkotaan (Urban)	27.1	Tidak	Tidak	Tidak
Pria	27	Tidak	Tidak	Tidak	Perkotaan (Urban)	22.1	Ya	Tidak	Tidak

Gambar 2. Dataset terkumpul.

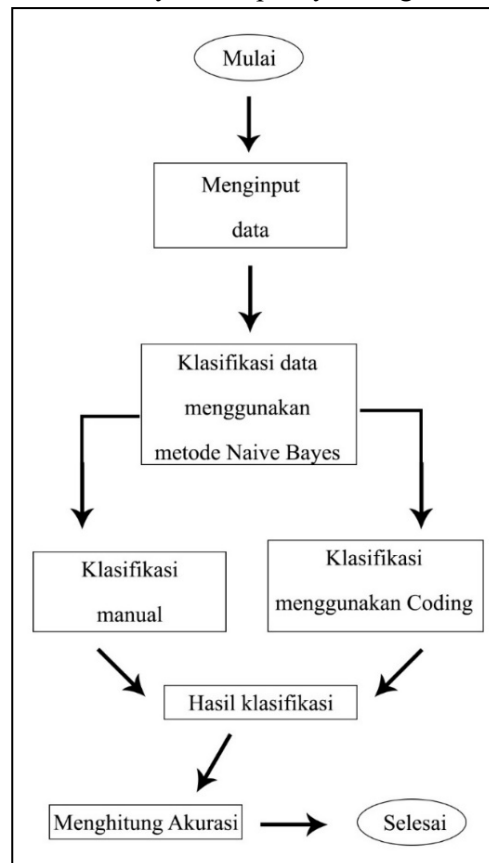
Setelah dataset terkumpul, selanjutnya dilakukan *data reduction* yaitu pengurangan atau pembuangan kelas yang tidak penting, lalu juga

dilakukan *data transformation* yaitu mengubah data atau mengkategorikan data kebeberapa kategori. Untuk *data transformation* ini misalnya pada bagian kelamin, pria diubah menjadi 0, dan Wanita diubah menjadi 1. Untuk jelasnya bisa dilihat pada Gambar 3 dibawah ini :

Kelamin	Umur	Hipertensi	Jantung	Menikah	Lingkungan	BMI	Rokok	Alkohol	Stroke
1	1	0	0	0	1	2	0	0	0
0	1	0	0	0	1	0	0	0	0
0	4	1	0	1	1	3	1	0	1
0	1	0	0	0	1	0	0	0	0
0	5	1	1	1	1	3	1	1	1
0	4	0	0	1	0	2	0	0	0
1	1	0	0	0	0	1	0	0	0
1	1	0	0	0	1	3	0	0	0
0	1	0	0	0	1	3	0	1	0
1	1	0	0	0	1	0	0	0	0
0	1	0	0	0	1	1	1	1	0
1	1	0	0	0	1	1	0	0	0
0	5	1	0	1	1	3	1	1	1
1	3	1	1	1	0	2	0	0	1
1	3	1	0	1	0	2	0	0	1
0	6	1	1	1	1	3	1	0	1
1	1	0	0	0	0	2	0	0	0
0	4	1	1	1	1	3	1	0	1
0	2	1	0	0	1	3	1	0	1
1	1	0	0	0	1	0	0	0	0
1	4	1	0	1	1	3	0	0	1
1	1	0	0	0	0	0	0	0	0
0	2	1	0	0	1	2	1	0	0
1	1	0	0	0	1	0	0	0	0
0	6	1	1	1	1	3	1	0	1
0	1	0	0	0	1	2	0	0	0
1	1	0	0	0	1	1	0	0	0

Gambar 3. Data display.

Setelah dilakukan pengolahan data, selanjutnya adalah penginputan data kedalam metode naïve bayes, tahapannya sebagai berikut :



Gambar 4. Tahapan proses klasifikasi.

Selanjutnya adalah proses klasifikasi menggunakan hitung manual, pertama dengan mencari probabilitas priornya dengan rumusnya sebagai berikut :

$$P(H|X) = P(X|H) / P(H)$$

- $P(\text{Stroke}) = 32/109 = 0.293$
- $P(\text{Non-stroke}) = 77/109 = 0.706$

Sudah didapat nilai probabilitas prior untuk stroke adalah 0.293, sedangkan probabilitas prior non-stroke 0.706.

IV. HASIL DAN PEMBAHASAN

Setelah melakukan analisa pada data prediksi penyakit stroke dengan metode Naïve Bayes, maka hasil yang dicapai oleh peneliti adalah guna mengetahui prediksi penyakit stroke seseorang berdasarkan gaya hidup dan lingkungannya misalnya umur, riwayat penyakit, menikah, dan atribut-atribut lainnya. Untuk mempermudah proses klasifikasi naïve bayes, selain menggunakan perhitungan manual, peneliti juga menggunakan coding naïve bayes untuk mempercepat proses perhitungan dan juga sebagai alat komparasi hasil klasifikasi. Berdasarkan dataset yang terkumpul, saya melakukan proses perhitungan naïve bayes dengan dataset yang terkumpul secara manual dengan langkah-langkah sebagai berikut :

- Menyiapkan data yang akan digunakan dalam perhitungan, untuk jelasnya bisa dilihat pada table 1 dibawah ini :

Tabel 1. Sample data.

Klm	Umr	Hpt	Jtg	Mnk	Lkg	BMI	Rkk	Alk	Strk
Pria	44	Ya	No	Ya	U	23.7	Ya	No	Ya
Pria	33	No	No	Ya	U	24.4	Ya	No	No
Pria	22	No	No	No	U	19.4	No	No	No
Pria	35	No	No	Ya	U	22.9	Ya	No	No
Pria	42	Ya	Ya	Ya	U	26.5	Ya	Ya	Ya
Pria	42	Ya	No	Ya	U	23.6	Ya	No	Ya

- Selanjutnya membuat atribut kelas yaitu Stroke = Ya dan Tidak.
- Setelah membuat kelas status, selanjutnya menghitung probabilitas kemunculan dari masing-masing atribut mulai dari kelamin hingga alkohol.
- Setelah mendapatkan hasil nilai probabilitas kemunculan dari masing-masing attribute, selanjutnya adalah menghitung prior nya. Seperti yagn telah dibahas tadi probabilitas prior untuk stroke adalah 0.293, sementara non-stroke adalah 0.706.
- Setelah mendapatkan prior, selanjutnya adalah mencari nilai likelihood dari tiap kelas.
- Setelah mendapatkan semua nilai, selanjutnya tinggal mengkalikan semua nilai dari stroke dan non-stroke, lalu mengkomparasi hasil dari perkaliannya.

Untuk lebih jelasnya bisa dilihat menggunakan data uji dibawah ini:

- Kelamin Pria
- Umur 47
- Hipertensi Ya
- Jantung Ya
- Menikah Ya
- Lingkungan Urban
- BMI 26,4
- Rokok Ya
- Alkohol Ya
- Stroke ???

Pertama adalah menghitung class stroke nya seperti pada table 2 dibawah ini :

Tabel 2. Perhitungan manual Stroke

Non-stroke = $77/109 = 0.706$
Pria & Non-stroke = $46/77 = 0.597$
Umur 47 (6) & Non-stroke = $8/77 = 0.103$
Hipertensi (Ya) & Non-stroke = $16/77 = 0.207$
Jantung (Ya) & Non-stroke = $1/77 = 0.012$
Menikah (Ya) & Non-stroke = $30/77 = 0.389$
Lingkungan (Urban) & Non-stroke = $61/77 = 0.792$
BMI (3) & Non-stroke = $17/77 = 0.220$
Rokok (Ya) & Non-stroke = $31/77 = 0.402$
Alkohol (Tidak) & Non-stroke = $75/77 = 0.974$
Rokok (Ya) & Non-stroke = $31/77 = 0.402$
$0.706 * 0.597 * 0.103 * 0.207 * 0.012 * 0.389 * 0.792 * 0.220 * 0.402 * 0.974 = \mathbf{0.002}$

Hasil dari perkalian pada class stroke menunjukkan 0.004. Setelah itu kita mencari hasil dari keals non-stroke dengan cara yang sama. Bisa dilihat pada Tabel 3 di bawah ini;

Untuk perkalian pada kelas Non-stroke, didapatkan angka 0.002. Setelah mendapatkan kedua angka dari stroke dan non-stroke, selanjutnya adalah mengkomparasi antara hasil dari stroke dan non-stroke. Dsini hasil terbesarnya ada pada kelas Stroke, maka bisa diprediksi bahwa hasil dari data uji ini menghasilkan jawaban Stroke.

Selanjutnya adalah menghitung klasifikasi menggunakan coding naïve bayes, coding ini dibuat menggunakan aplikasi Jupyter Notebook dengan Bahasa pemrograman Python. Untuk jelasnya bisa dilihat pada gambar dibawah ini :

```

In [8]: from sklearn.naive_bayes import GaussianNB
        data=GaussianNB()
        data.fit(combain,stroke)

Out[8]: GaussianNB()

In [15]: result=data.predict([[1,4,0,0,1,0,3,0,0]])
        print("Prediksi Apakah Kamu Berpotensi Terkena Stroke")
        print("Hasil = ",result)
        print("(0 = Kamu tidak berpotensi terkena Stroke)")
        print("(1 = Kamu berpotensi terkena Stroke)")

Prediksi Apakah Kamu Berpotensi Terkena Stroke
Hasil = [0]
(0 = Kamu tidak berpotensi terkena Stroke)
(1 = Kamu berpotensi terkena Stroke)
    
```

Gambar 5. Coding Naïve Bayes

Stroke = $32/109 = 0.293$
Pria & Stroke = $24/32 = 0.75$
Umur 47 (6) & Stroke = $5/32 = 0.156$
Hipertensi (Ya) & Stroke = $31/32 = 0.968$
Jantung (Ya) & Stroke = $16/32 = 0.5$
Menikah (Ya) & Stroke = $30/32 = 0.937$
Lingkungan (Urban) & Stroke = $23/32 = 0.718$
BMI (3) & Stroke = $24/32 = 0.75$
Rokok (Ya) & Stroke = $24/32 = 0.75$
Alkohol (Tidak) & Stroke = $21/32 = 0.656$
Rokok (Ya) & Stroke = $24/32 = 0.75$
$0.293 * 0.75 * 0.156 * 0.968 * 0.5 * 0.937 * 0.718 * 0.75 * 0.75 * 0.656 = \mathbf{0.004}$

Pada gambar 5 diatas, terlihat hasil prediksi naïve bayes ini sudah bisa menampilkan hasil prediksi dari sebuah dataset yang ingin diuji, caranya hanya dengan memasukkan atau menginput data uji kedalam codingnya, lalu secara otomatis coding akan melakukan proses pengklasifikasian terhadap data tersebut dan akan mengeluarkan hasil dari pengklasifikasian ini.

Setelah melakukan pengklasifikasian pada dataset yang uji, selanjutnya adalah menghitung tingkat akurasi terhadap dataset yang diuji menggunakan aplikasi Rapidminer, bisa dilihat pada gambar 6 dibawah ini :

```

SimpleDistribution

Distribution model for label attribute Stroke

Class 0 (0.706)
9 distributions

Class 1 (0.294)
9 distributions
    
```

Gambar 6. Probabilitas Prior Rapidminer

Pada gambar 6 diatas, terlihat class 0 atau kelas Non-stroke adalah 0.706, sedangkan kelas 1 atau kelas Stroke adalah 0.294, maka dapat disimpulkan bahwa probabilitas prior pada perhitungan rapidminer dan manual sudah sesuai atau sama.

Selanjutnya menghitung tingkat akurasinya menggunakan aplikasi rapidminer, bisa dilihat pada gambar 7 dibawah :

accuracy: 93.94%			
	true 0	true 1	class precision
pred 0	21	0	100.00%
pred 1	2	10	83.33%
class recall	91.30%	100.00%	

Dapat dilihat pada gambar 6 diatas, tingkat akurasi pada dataset ini lumayan tinggi yaitu 93.94%. Maka dapat disimpulkan bahwa pemrediksian dataset ini memiliki akurasi yang cukup baik.

V. PENUTUP

Kesimpulan yang didapat pada penelitian ini adalah bahwa metode naïve bayes dapat melakukan prediksi tidak hanya tentang penyakit stroke. Selain itu, bisa dilihat tingkat akurasi naïve bayes pada dataset ini sangat tinggi yaitu sekitar 93.94% dihitung menggunakan aplikasi rapidminer. Saran yang didapat pada penelitian ini adalah yaitu untuk menambah jumlah datasetnya agar proses prediksinya bisa lebih akurat lagi, misalnya menambah beberapa atribut dan kelas lainnya. Saran lainnya juga bisa dikembangkan lagi codingnya untuk dijadikan aplikasi sederhana yang UI-nya lebih mudah dipakai oleh banyak kalangan.

DAFTAR PUSTAKA

- [1] C. D. A. Wolfe, "The impact of stroke," *Br. Med. Bull.*, vol. 56, pp. 275–286, 2000, Accessed: Jun. 08, 2022. [Online]. Available: <https://academic.oup.com/bmb/article/56/2/275/303250>
- [2] O. A. Nastiti, "SISTEM PAKAR KLASIFIKASI STROKE DENGAN METODE NAIVE BAYES CLASSIFIER DAN CERTAINTY FACTOR SEBAGAI ALAT BANTU DIAGNOSIS," 2016.
- [3] N. D. Sari, "Penerapan klasifikasi kepuasan pelanggan go-jek menggunakan metode algoritma naïve bayes," p. 60, 2018.
- [4] H. Halimah, D. Linda, and F. Klaralia, "Penerapan Algoritma Naïve Bayes untuk Memprediksi Penyakit Malaria pada Puskesmas Hanura," *Teknika*, vol. 14, no. x, pp. 57–63, 2020.
- [5] M. Z. PRATAMA, "Penentuan Kriteria Dan Penerima Bantuan Pangan Non Tunai Di Desa Parungkuda Dengan Menggunakan Algoritma Naïve ...," 2020, [Online]. Available: <https://repository.bsi.ac.id/index.php/repo/viewitem/27910>
- [6] K. M. Leung, "Naive Bayesian Classifier," 2007.
- [7] A. T. Susilo, H. Setiawan, R. A. Saputro, T. Purwadi, and A. Saifudin, "Penggunaan Metode Naïve Bayes untuk Memprediksi Tingkat Kemenangan pada Game Mobile Legends," *J. Teknol. Sist. Inf. dan Apl.*, vol. 4, no. 1, p. 46, 2021, doi: 10.32493/jtsi.v4i1.7807.
- [8] A. Algarni, "Data Mining in Education," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, pp. 456–461, 2016, Accessed: Jun. 07, 2022. [Online]. Available: www.ijacsa.thesai.org
- [9] I. Oktanisa and A. A. Supianto, "Perbandingan Teknik Klasifikasi Dalam Data Mining Untuk Bank a Comparison of Classification Techniques in Data Mining for," *Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 5, pp. 567–576, 2018, doi: 10.25126/jtiik20185958.
- [10] S. Menengah Kejuruan Muhammadiyah, B. Bambang Supiyarto, B. Eka Purnama, and G. Kristianto Nugroho, "Pembuatan Media Pembelajaran Ketrampilan Komputer dan Pengelolaan Informasi Pada," Online.