

**KLASIFIKASI PENYAKIT STROKE MENGGUNAKAN METODE
NAÏVE BAYES DAN *FUZZY NAÏVE BAYES* BERDASARKAN
*RESAMPLING REPEATED SPLIT VALIDATION***

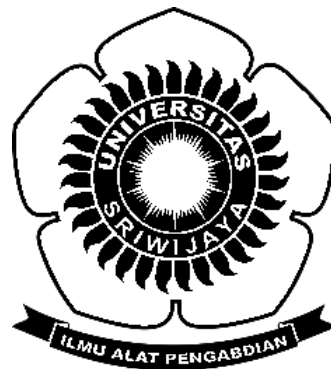
SKRIPSI

**Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana
di Jurusan Matematika pada Fakultas MIPA**

Oleh:

FEBRIANTI DWI NUR'AINI

08011381924062



**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS SRIWIJAYA**

2023

LEMBAR PENGESAHAN

**KLASIFIKASI PENYAKIT STROKE MENGGUNAKAN METODE
NAÏVE BAYES DAN FUZZY NAÏVE BAYES BERDASARKAN
RESAMPLING REPEATED SPLIT VALIDATION**

SKRIPSI

**Sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Matematika**

Oleh:

FEBRIANTI DWI NUR'AINI

NIM. 08011381924062

Indralaya, November 2023

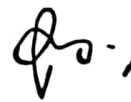
Pembimbing Kedua

Pembimbing Utama



Novi Rustiana Dewi, S.Si., M.Si

NIP. 19701113 199603 2 002



Prof. Dr. Yulia Resti, S.Si., M.Si

NIP. 19730719 199702 2 001

Mengetahui

Ketua Jurusan Matematika



Dr. Dian Cahyawati S, S.Si., M.Si

NIP. 19730321 200012 2 001

PERNYATAAN KEASLIAN KARYA ILMIAH

Sayaa yang bertanda tangan dibawah ini :

Nama : Febrianti Dwi Nur'aini

NIM : 08011381924062

Fakultas/Jurusan : Matematika dan Ilmu Pengetahuan Alam/Matematika

Menyatakan bahwa skripsi yang berjudul “Klasifikasi Penyakit Stroke Menggunakan Metode *Naïve Bayes* dan *Fuzzy Naïve Bayes* Berdasarkan *Resampling Repeated Split Validation*” adalah hasil karya saya sendiri dan karya ilmiah ini belum pernah diajukan sebagai pemenuhan persyaratan untuk memperoleh gelar kesarjanaan strata (S1) dari Universitas Sriwijaya maupun perguruan tinggi lainnya. Semua informasi yang dimuat dalam skripsi ini yang berasal dari penulis lain baik yang dipublikasikan atau tidak telah diberikan penghargaan dengan mengutip nama sumber penulis secara benar. Semua isi dari skripsi ini sepenuhnya menjadi tanggung jawab saya sebagai penulis.

Demikianlah pernyataan ini saya buat dengan sebenarnya.

Indralaya, 23 November 2023
Yang membuat pernyataan,



Febrianti Dwi Nur'aini
NIM. 08011381924062

HALAMAN PERSEMBAHAN

MOTTO

“Dan barang siapa yang bertakwa kepada Allah, niscaya Allah menjadikan baginya kemudahan dalam urusannya”

(QS. At-Talaq 65: Ayat 4)

“To achieve what you want, you must continue to pursue and strive to make it happen. Then at the same time take care of yourself and your health”

(Park Chanyeol)

Skripsi ini dipersembahkan untuk:

- Allah SWT
- Orang Tuaku Tercinta
- Adik - Adikku Tersayang
- Dosenku
- Teman Seperjuangan
- Almamaterku

KATA PENGANTAR

Assalamu 'alaikum warahmatullahi wabarakatuh

Puji syukur atas kehadiran ALLAH SWT yang telah memberikan rahmat, taufiq, hidayah dan innayah-Nya, sehingga penulis bisa menyelesaikan skripsi yang berjudul “**Klasifikasi Penyakit Stroke menggunakan Metode *Naïve Bayes* dan *Fuzzy Naïve Bayes* Berdasarkan *Resampling Repeated Split Validation*” dengan baik dan tepat waktu. Shalawat serta salam semoga selalu tercurahkan kepada Nabi Muhammad Shallallahu'alaihi wa sallam.**

Dengan segala hormat dan ketulusan hati penulis ingin mengucapkan terima kasih dan mempersembahkan untuk cinta pertama dan panutan penulis, Ayahanda **Alm. Supardi**. Beliau memang tidak sempat melihat semua anaknya menyelesaikan bangku perkuliahan, namun beliau mampu mendidik penulis serta memberikan semangat hingga penulis mampu menyelesaikan studinya hingga sarjana. Kemudian untuk pintu surgaku, Ibunda **Devi Yarti** terima kasih sebesar-besarnya penulis berikan kepada beliau atas segala bentuk bantuan, doa dan semangat yang diberikan selama ini. Terima kasih atas kesabaran dan kebesaran hati dalam menghadapi penulis yang keras kepala, ibu menjadi penguat dan pengingat yang paling hebat. Terima kasih sudah menjadi tempatku untuk pulang, bu.

Dalam penyusunan skripsi ini, tentunya penulis mendapatkan begitu banyak bantuan dari beberapa pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada:

1. Bapak **Prof. Hermansyah, S.Si., M.Si., Ph.D.**, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
2. Ibu **Dr. Dian Cahyawati Sukanda, M.si.**, selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
3. Ibu **Prof. Dr. Yulia Resti, S.Si., M.Si.**, selaku Dosen Pembimbing utama yang telah memberikan bimbingan, arahan, saran, dan ilmu serta pembelajaran sehingga skripsi dapat diselesaikan.
4. Ibu **Novi Rustiana Dewi, S.Si., M.Si.**, selaku Dosen Pembimbing kedua yang telah memberikan bimbingan, arahan, saran, dan ilmu serta pembelajaran sehingga skripsi dapat diselesaikan.
5. Bapak **Drs. Putra Bahtera Jaya Bangun, M.Si.**, dan Ibu **Dr. Yuli Andriani, S.Si., M.Si.**, selaku Dosen Pembahas yang telah memberikan tanggapan dan saran yang sangat bermanfaat untuk perbaikan dan penyelesaian skripsi ini.
6. Ibu **Sri Indra Maiyanti, S.Si., M.Si.**, selaku Dosen Pembimbing Akademik yang telah membimbing dan memberikan arahan kepada penulis selama masa perkuliahan.
7. **Seluruh Bapak/Ibu Dosen dan Staff di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya** atas ilmu dan nasehat yang telah diberikan selama masa perkuliahan.
8. Bapak **Irwansyah** dan Ibu **Hamidah** yang telah membantu dalam proses administrasi.

9. Adik-adik penulis yaitu **Nilam Anggraini** dan **Dhimas Abimanyu** terima kasih sudah ikut serta dalam proses penulis menempuh pendidikan selama ini, terima kasih atas semangat, doa, dan cinta yang selalu diberikan kepada penulis. Tumbuhlah menjadi versi paling hebat, adik-adikku.
10. Sahabat di Jurusan Matematika, **Hutvina Latifatussolehah**, **Rayhannul Suci** dan **Deka Sephira Agatha** yang telah berjuang bersama, menemani di suka maupun duka, memberikan dorongan, semangat dan doa kepada penulis selama masa perkuliahan ini.
11. Sahabat penulis, **Angelia**, **Bella**, **Dedeh**, **Della**, **Meski**, **Nadia** dan **Puput** yang telah memberikan support, doa dan perhatian kepada penulis dalam menemani pembuatan skripsi ini.
12. Sahabat selama perantauan, **April**, **Arima**, **Puspa**, **Yessi**, **Delia**, **Yupita**, **Ari**, **Saribi** dan yang tidak bisa disebutkan satu persatu yang selalu memberikan dukungan dan motivasi kepada penulis.
13. Teman-teman seperjuangan skripsi, **Maulidiya**, **Afhriza**, **Laurensia**, **Risma**, **Dinda**, **Dieta**, dan **Rifat** yang memberikan bantuan selama perkuliahan skripsi ini.
14. Semua pihak yang tidak dapat disebutkan satu persatu, terimakasih atas dukungan dan bantuan selama penulis menyelesaikan skripsi.
15. Kepada penulis, **Febrianti Dwi Nur'aini** terimakasih atas segala kerja keras dan semangatnya sehingga tidak pernah menyerah salam mengerjakan tugas akhir skripsi ini. Semoga tetap selalu rendah diri, karena ini baru awal dari

semuanya, masih banyak tahap yang harus dilewati. Semoga segala kebaikan dan pertolongan semuanya mendapat berkah dari Allah SWT.

Penulis berharap skripsi ini dapat bermanfaat dan menambah pengetahuan bagi semua pihak yang membutuhkan terutama mahasiswa/mahasiswi Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Indralaya, Oktober 2023

Penulis

***CLASSIFICATION OF STROKE DISEASE
USING NAÏVE BAYES METHOD AND FUZZY NAÏVE BAYES METHOD
BASED ON RESAMPLING REPEATED SPLIT VALIDATION***

By:

**FEBRIANTI DWI NUR'AINI
08011381924062**

ABSTRACT

A stroke is a nerve function that results from a blockage caused by blood clots and causes the oxygen and nutrients in the brain to suffer tissue damage. This study uses the Naïve Bayes and Fuzzy Naïve Bayes method based on Repeated Split Validation Resampling and variable selection use Chi square to classify stroke disease. The data used on this research is a Stroke Prediction Dataset from Kaggle's website. Classification of stroke disease using Naïve Bayes obtained an average Accuracy of 92,27%, Precision of 96,09%, Recall of 95.79%, and FScore of 95,94%, where here as using the naïve bayes method with chi square obtained an average Accuracy of 92,66%, Precision of 96,50%, Recall of 95.81%, and FScore of 96,16%. Classification of stroke disease using Fuzzy Naïve Bayes obtained an average Accuracy of 93,15%, Precision of 97,43%, Recall of 95.46%, and FScore of 96,44%, where as using Fuzzy Naïve Bayes method with chi square obtained an average Accuracy of 93,54%, Precision of 97,74%, Recall of 95,57%, and FScore of 96,64%. Based on factors that contribute to the prediction of stroke and calculations using the Naïve Bayes and Fuzzy Naïve Bayes methods, the results are 95% not a stroke and 5% of a stroke.

Keywords: Stroke, Naïve Bayes, Fuzzy Naïve Bayes, Repeated Split Validation.

**KLASIFIKASI PENYAKIT STROKE MENGGUNAKAN METODE
NAÏVE BAYES DAN FUZZY NAÏVE BAYES BERDASARKAN
RESAMPLING REPEATED SPLIT VALIDATION**

Oleh:

FEBRIANTI DWI NUR'AINI

08011381924062

ABSTRAK

Stroke merupakan gangguan fungsi saraf yang terjadi karena sumbatan yang disebabkan oleh gumpalan darah sehingga menyebabkan pasokan oksigen dan nutrisi yang ada di otak mengalami kerusakan jaringan otak. Penelitian ini menggunakan metode *Naïve Bayes* dan *Fuzzy Naïve Bayes* berdasarkan *Resampling Repeated Split Validation* dan seleksi variabel menggunakan uji *chi square* untuk mengklasifikasikan penyakit stroke. Data yang digunakan pada penelitian ini adalah *Stroke Prediction Dataset* dari website Kaggle. Hasil klasifikasi menggunakan metode *Naïve Bayes* mendapatkan *accuracy* sebesar 92,27%, *precision* sebesar 96,09%, *recall* sebesar 95,79%, dan *fscore* sebesar 95,94%, sedangkan menggunakan metode *Naïve Bayes* dengan uji *chi square* mendapatkan *accuracy* sebesar 92,66%, *precision* sebesar 96,50%, *recall* sebesar 95,81%, dan *fscore* sebesar 96,16%. Hasil klasifikasi menggunakan metode *Fuzzy Naïve Bayes* mendapatkan *accuracy* sebesar 93,15%, *precision* sebesar 97,43%, *recall* sebesar 95,46%, dan *fscore* sebesar 96,44%, sedangkan menggunakan metode *Fuzzy Naïve Bayes* dengan uji *chi square* mendapatkan *accuracy* sebesar 93,54%, *precision* sebesar 97,74%, *recall* sebesar 95,57%, dan *fscore* sebesar 96,64%. Berdasarkan faktor-faktor yang berkontribusi terhadap prediksi penyakit stroke dan perhitungan menggunakan metode *Naïve Bayes* dan *Fuzzy Naïve Bayes* maka hasil prediksi tidak terkena stroke sebesar 95% dan terkena stroke sebesar 5%.

Kata kunci: Stroke, *Naïve Bayes*, *Fuzzy Naïve Bayes*, *Repeated Split Validation*.

DAFTAR ISI

HALAMAN PENGESAHAN	ii
HALAMAN PERSEMBAHAN	iii
KATA PENGANTAR	iv
ABSTRACT	viii
ABSTRAK	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	4
1.4 Tujuan Masalah	4
1.5 Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA	6
2.1 Penyakit <i>Stroke</i>	6
2.2 <i>Statistical Machine Learning</i>	7
2.3 Klasifikasi	8
2.4 Diskritisasi	9
2.5 <i>Naïve Bayes</i>	9
2.6 <i>Fuzzy Naïve Bayes</i>	11
2.7 <i>Laplace Smoothing</i>	12
2.8 <i>Repeated Split Validation</i>	13
2.9 Peluang	14
2.10 Himpunan <i>Fuzzy</i>	15
2.11 Fungsi Keanggotaan	15
2.12 Uji <i>Chi Square</i> (UCS)	19
2.13 <i>Confusion Matriks</i>	21
BAB III METODOLOGI PENELITIAN	23

3.1	Tempat.....	23
3.2	Waktu	23
3.3	Metode Penelitian.....	23
3.3.1	Jenis dan Sumber Data.....	23
3.3.2	Variabel Penelitian.....	23
3.4	Teknik Analisis Data	24
BAB IV HASIL DAN PEMBAHASAN.....		29
4.1	Deskripsi Data	29
4.2	Diskritisasi Data	30
4.3	Klasifikasi Menggunakan Metode <i>Naïve Bayes</i> (NB).....	30
4.3.1	Resampling Menggunakan <i>Split Validation</i>	30
4.3.2	<i>Naïve Bayes</i>	31
4.3.3	<i>Confusion Matrix Naïve Bayes</i>	34
4.3.4	Metode <i>Naïve Bayes</i> dengan Uji <i>Chi Square</i> (UCS)	36
4.3.5	<i>Confusion Matrix Naïve Bayes</i> dengan Uji <i>Chi Square</i>	43
4.4	Klasifikasi Menggunakan Metode <i>Fuzzy Naïve Bayes</i> (FNB).....	44
4.4.1	Menentukan Himpunan Universal	44
4.4.2	Menentukan Nilai Keanggotaan <i>Fuzzy</i>	50
4.4.3	<i>Resampling</i> Menggunakan <i>Split Validation</i>	52
4.4.4	<i>Fuzzy Naïve Bayes</i>	53
4.4.5	<i>Confusion Matriks Fuzzy Naïve Bayes</i>	57
4.4.6	Metode <i>Fuzzy Naïve Bayes</i> dengan Uji <i>Chi Square</i> (UCS).....	58
4.4.7	<i>Confusion Matriks Fuzzy Naïve Bayes</i> dengan UCS	65
4.5	Analisis Hasil.....	66
BAB V KESIMPULAN DAN SARAN		68
5.1.	Kesimpulan.....	68
5.2.	Saran.....	69
DAFTAR PUSTAKA		70
LAMPIRAN		73

DAFTAR TABEL

Tabel 2.1 Confusion Matrix	21
Tabel 2.2 Kategori Ketepatan Klasifikasi	22
Tabel 3. 1 Kategori tiap Variabel.....	23
Tabel 4. 1 Deskripsi Variabel.....	29
Tabel 4. 2 Diskritisasi Data.....	30
Tabel 4. 3 Data Latih.....	31
Tabel 4. 4 Data Uji.....	31
Tabel 4. 5 Hasil Peluang Likelihood.....	32
Tabel 4. 6 Hasil Prediksi Naïve Bayes.....	34
Tabel 4. 7 Confusion Matrix Naïve Bayes.....	35
Tabel 4. 8 Tabel Kontingensi variabel	36
Tabel 4. 9 Tabel Nilai O_i semua cell.....	37
Tabel 4. 10 Nilai $O_i - E_i$, $O_i - E_i^2$, dan nilai $O_i - E_i^2 E_i$	38
Tabel 4. 11 Hasil perhitungan UCS variabel pengaruh dan variabel respon	38
Tabel 4. 12 Data Latih.....	39
Tabel 4. 13 Data Uji.....	39
Tabel 4. 14 Hasil Peluang Likelihood.....	41
Tabel 4. 15 Hasil Prediksi NB dengan UCS	42
Tabel 4. 16 Confusion Matrix NB dengan UCS	43
Tabel 4. 17 Interval himpunan fuzzy variabel pengaruh X_2	44
Tabel 4. 18 Himpunan fuzzy variabel pengaruh X_2	45
Tabel 4. 19 Interval himpunan fuzzy variabel pengaruh X_8	46
Tabel 4. 20 Himpunan fuzzy variabel pengaruh X_8	47
Tabel 4. 21 Interval himpunan fuzzy variabel pengaruh X_9	48
Tabel 4. 22 Himpunan fuzzy variabel pengaruh X_9	49
Tabel 4. 23 Nilai Keanggotaan Terbesar variabel Pengaruh X_2	51
Tabel 4. 24 Nilai Keanggotaan Terbesar variabel Pengaruh X_8	51
Tabel 4. 25 Nilai Keanggotaan Terbesar variabel Pengaruh X_9	52
Tabel 4. 26 Data Latih.....	53
Tabel 4. 27 Data Uji.....	53
Tabel 4. 28 Hasil Peluang Likelihood.....	55
Tabel 4. 29 Hasil Prediksi Fuzzy Naïve Bayes	57
Tabel 4. 30 Confusion Matrix Fuzzy Naïve Bayes	57
Tabel 4. 31 Tabel Kontingensi variabel	58
Tabel 4. 32 Nilai O_i semua cell.....	59
Tabel 4. 33 Nilai $O_i - E_i$, $O_i - E_i^2$, dan nilai $O_i - E_i^2 E_i$	60
Tabel 4. 34 Hasil perhitungan UCS variabel pengaruh dan variabel respon	61
Tabel 4. 35 Data Latih.....	62
Tabel 4. 36 Data Uji.....	62
Tabel 4. 37 Hasil Peluang Likelihood.....	63

Tabel 4. 38 Hasil Prediksi FNB dengan UCS	65
Tabel 4. 39 Confusion Matrix FNB dengan UCS	65
Tabel 4. 40 Perbandingan hasil dari Metode NB dan FNB.....	66

DAFTAR GAMBAR

Gambar 2.1 Kurva-S Penyusutan.....	16
Gambar 2.2 Kurva-S Pertumbuhan.....	17
Gambar 2.3 Kurva Trapesium.....	18

DAFTAR LAMPIRAN

Lampiran 1. Nilai Ketepatan Metode Naïve Bayes	73
Lampiran 2. Nilai Ketepatan Metode Naïve Bayes dengan UCS	73
Lampiran 3 Nilai Ketepatan Metode Fuzzy Naïve Bayes.....	74
Lampiran 4. Nilai Ketepatan Metode FNB dengan UCS.....	75

BAB I

PENDAHULUAN

1.1 Latar Belakang

Stroke merupakan penyakit saraf yang disebabkan adanya gangguan peredaran darah pada otak yang terjadi secara mendadak, atau terdapat sumbatan yang disebabkan adanya gumpalan darah sehingga pasokan oksigen yang ada di otak mengalami kerusakan jaringan otak. *Stroke* dibagi menjadi dua tipe, yaitu *stroke* iskemik penyebabnya adalah kurangnya suplai oksigen ke dalam otak akibat adanya gumpalan darah yang menyumbat pembuluh darah. Yang kedua adalah *stroke* hemoragik penyebabnya adalah pecahnya pembuluh darah pada rongga antara otak dengan tengkorak sehingga mengakibatkan pasien kekurangan oksigen pada otak dan mengganggu sirkulasi normal darah pada otak (Puspitasari, 2020).

Metode *fuzzy naïve bayes* merupakan penggabungan dari teori himpunan *fuzzy* dengan klasifikasi *naïve bayes*. *Naïve bayes* adalah metode klasifikasi berdasarkan metode Teorema Bayes dengan asumsi independen yang kuat, dan *naive bayes* terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan kumpulan data yang besar (Lishania *et al.*, 2019). Naive Bayes memunculkan kebebasan sederhana, hubungan antar variabel di kelas sepenuhnya independen dari variabel lain. Asumsi sederhana ini mengakibatkan perkiraan suboptimal hasil yang diharapkan tidak akurat, khususnya untuk masalah domain teks (Zaidi *et al.* 2013). Himpunan *fuzzy* menjelaskan tentang faktor ketidakpastian pada proses pengelompokan dengan menggunakan himpunan *fuzzy* faktor ketidakpastian tersebut bisa diperhitungkan sehingga dapat mengurangi

kesalahan dalam proses klasifikasi tetapi dalam himpunan *fuzzy* memungkinkan seseorang untuk bekerja pada keadaan yang tidak pasti dan samar serta memecahkan masalah yang tidak tetapi dalam himpunan *fuzzy* memungkinkan seseorang untuk bekerja pada keadaan yang tidak pasti dan samar serta memecahkan masalah yang tidak. Metode *fuzzy naïve bayes* digunakan sebagai proses pengklasifikasian penyakit *stroke* pada penelitian ini.

Ada beberapa peneliti yang telah melakukan penelitian dengan *fuzzy naïve bayes* dan *naïve bayes* yang akan dilakukan pada penelitian ini seperti Resti *et al*, (2023) tentang “*Fuzzy Discretization on the Multinomial Naïve Bayes Method for Modeling Multiclass Classification of Corn Plant Diseases and Pests*”. Adapun penelitian yang dilakukan oleh Insani (2017) tentang “Sistem Pakar Deteksi Penyakit *Stroke* menggunakan Metode *Fuzzy Naïve Bayes*” dengan menggunakan metode gabungan *fuzzy* dan *naïve bayes* menghasilkan akurasi sebesar 85,6%. Ada juga yang melakukan penelitian yang dilakukan oleh Shabrina (2023) tentang “Klasifikasi Jenis Pensiun Pegawai PTPN II Tahun 2012-2019 Menggunakan Metode *Fuzzy Naïve Bayes* Berdasarkan Resampling Repeated Split Validation” menghasilkan akurasi sebesar 90,94% dan juga penelitian yang dilakukan oleh Antofa (2023) tentang “Komparasi Metode *Fuzzy Naïve Bayes* Dan Deep Neural Network Untuk Klasifikasi Indeks Harga Saham Gabungan (IHSG) Bursa Efek Indonesia Tahun 2012-2022” menghasilkan akurasi sebesar 57,32%.

Adapun juga penelitian yang dilakukan oleh Azizah (2021) dengan data yang sama melakukan penelitian tentang “Komparasi Metode Klasifikasi *Decision Tree* Algoritma C4.5 Dan *Random Forest* Untuk Prediksi Penyakit *Stroke*” dengan

variabel yang sama seperti jenis kelamin, umur, riwayat hipertensi, riwayat penyakit jantung, status menikah, tempat tinggal, jenis pekerjaan, kadar gula darah, BMI, status merokok, dan *stroke* dengan jumlah data sebanyak 5110 entri data dan mendapatkan rata-rata hasil akurasi untuk metode *decision tree* adalah 92,56% dan untuk metode *random forest* adalah 93,80%. Adapun penelitian yang dilakukan oleh Shobayo *et al* (2023) tentang “*Prediction of Stroke Disease with Demographic and Behavioural Data Using Random Forest Algorithm*” dengan rata-rata hasil akurasi untuk metode *random forest* sebesar 94,11% untuk metode *logistic regression* sebesar 91,43% dan metode *decision tree* sebesar 88,83%.

Berdasarkan hasil penelitian terdahulu dan kelebihan dari beberapa metode yang telah dijelaskan, maka akan dikaji pengklasifikasian data penyakit *stroke* menggunakan metode *fuzzy naïve bayes* berdasarkan *repeated split validation*.

1.2 Rumusan Masalah

Berdasarkan pada latar belakang di atas, rumusan masalah pada penelitian ini adalah sebagai berikut :

1. Berapa tingkat ketepatan pada klasifikasi penyakit *stroke* menggunakan metode *naïve bayes*.
2. Berapa tingkat ketepatan pada klasifikasi penyakit *stroke* menggunakan metode *fuzzy naïve bayes*.
3. Manakah metode yang memiliki tingkat ketepatan lebih baik.

1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah :

1. Data yang digunakan berasal dari website kaggle.com, data ini terdiri dari 11 variabel yaitu jenis kelamin, umur, riwayat hipertensi, riwayat penyakit jantung, status menikah, tempat tinggal, jenis pekerjaan, kadar gula darah, BMI, status merokok, dan *stroke*.
2. Data berjumlah 5110 entri data dengan validasi data menggunakan *repeated split validation* dengan 30 kali pengulangan.
3. Data dipartisi menjadi 80% data *train* dan 20% data *ujit*.
4. Fungsi keanggotaan *fuzzy* yang digunakan ada 3 yaitu kurva sigmoid penyusutan, kurva sigmoid pertumbuhan, dan kurva trapesium.
5. Tingkat ketepatan klasifikasi penelitian ini dibatasi pada nilai *Accuracy*, *Precision*, *Recall*, dan *Fscore*.

1.4 Tujuan Masalah

Tujuan dari penelitian ini adalah sebagai berikut:

1. Memperoleh hasil dari klasifikasi penyakit *stroke* menggunakan metode *naïve bayes*.
2. Memperoleh hasil dari klasifikasi penyakit *stroke* menggunakan metode *fuzzy naïve bayes*.
3. Menentukan manakah metode yang memiliki tingkat ketepatan lebih baik.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah:

1. Diharapkan bisa memberikan manfaat kepada instansi kesehatan supaya mengetahui faktor-faktor apa saja yang mempengaruhi penyakit *stroke*, agar dapat digunakan untuk memperkecil resiko *stroke*.
2. Penelitian ini diharapkan dapat dijadikan sebagai bahan referensi penelitian selanjutnya tentang penyakit *stroke*.

DAFTAR PUSTAKA

- Anggraini, R. A., Widagdo, G., Budi, A.S., & Qomaruddin, M. (2019). Penerapan Data Mining Classification Untuk Data Blogger Menggunakan Metode Naïve Bayes. *Jurnal Sistem dan Teknologi Informasi (JUSTIN)* 7(1), pp. 47.
- Azizah, N. (2021). *Komparasi Metode Klasifikasi Decision Tree Algoritma C4.5 dan Random Forest untuk Prediksi Stroke*. Skripsi Jurusan Matematika FMIPA Universitas Sriwijaya. Dipublikasikan.
- Bustamante, C., Leonardo, G., & Rogelio, S. (2007). Fuzzy Naive Bayesian Classification in RoboSoccer 3D: A Hybrid Approach to Decision Making. *LNAI 4434*, pp. 507-515.
- Davvaz, B., Imam, M., & Soleha. (2021). Himpunan fuzzy dan rough sets. *Limits: Journal of Mathematics and Its Applications* 18(1), pp. 79.
- Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. *International Statistical Review*. 2, pp. 1-758.
- Haris. (2022) Metode Naïve Bayes untuk Memprediksi Penyakit Stroke. *Universitas Nusa Mandiri*, 6(1).
- Hariyani., & Sugiyarto, S. (2020) Diskritisasi Equal-Width Interval pada Naïve Bayes (Studi Kasus: Klasifikasi Pasien TBC). *AdMathEdu*, 10(2), pp. 91-102.
- Insani, S. A., (2017). *Sistem Pakar Deteksi Dini Penyakit Stroke menggunakan Metode Fuzzy Naïve Bayes*. Skripsi Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Brawijaya. Dipublikasikan.
- Kilimci, Z. H., & Murat, C. G. (2015). Evaluation of Classification Models for Language Processing. *International Symposium on Innovations in Intelligent Systems and Applications (INISTA)*, pp. 1-8.
- Lishania, I., Rito, G., & Yuki, N. N. (2019). Perbandingan Klasifikasi Metode Naive Bayes dan Metode Decision Tree Algoritma (J48) pada Pasien Penderita Penyakit Stroke di RSUD Abdul Wahab Sjahranie Samarinda. *Jurnal Ekspansional*. 10(2), pp. 136.
- Listiowarni, I., & Nilam, R., (2019). Implementasi Naïve Bayesian Dengan Laplacian Smoothing untuk Peminatan dan Lintas Minat Siswa SMAN 5 Pamekasan. *Jurnal SISFOKOM (Sistem Informasi dan Komputer)* 8(2), pp. 124–29.

- Mishra, S., Vanli, O. A., Huffer, F. W., & Jung, S. (2016). Regularized Discriminant Analysis for Multi-sensor Decision Fusion and Damage Detection with Lamb-waves. *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace System*. pp. 728-741.
- Neardiaty, A. (2022). *Klasifikasi Hama dan Penyakit Tanaman Jagung menggunakan Metode Fuzzy Random Forest berdasarkan Resampling Repeated K-fold Cross Validation*. Skripsi Jurusan Matematika FMIPA Universitas Sriwijaya. Dipublikasikan.
- Pambudi, R. E., Sriyanto., & Firmansyah. (2022). Klasifikasi Penyakit Stroke menggunakan Algoritma Decision Tree C.45. *Jurnal Teknik*. 16(2), pp. 221-226.
- Puspitasari, P. N. (2020). Hubungan Hipertensi terhadap Kejadian Stroke. *Jurnal Ilmiah Kesehatan Sandi Husada*. 12(2), pp. 922-926.
- Resti, Y., Chandra, I., Adinda, N., Choirunnisa, A., & Irsyadi, Y. (2023). Fuzzy Discretization on the Multinomial Naïve Bayes Method for Modeling Multiclass Classification of Corn Plant Diseases and Pests. *Mathematics Journal MDPI*. 11, pp. 1761.
- Retnoningsih, E., & Rully, P. (2020). Mengenal Machine Learning dengan Teknik Supervised dan Unsupervised Learning Menggunakan Python. *Bina Insani Ict Journal*. 7(2), 156.
- Rohman, R. S., Rizal, A.S., & Dasya, A.F. (2020). Komparasi Algoritma C4.5 Berbasis PSO dan GA untuk Diagnosa Penyakit Stroke. *CESS (Journal of Computer Engineering, System and Science)*. 5(1), pp. 155-161.
- Shabrina, M., (2023). *Klasifikasi Jenis Pensiun Pegawai PTPN II Tahun 2012-2019 menggunakan Metode Fuzzy Naïve Bayes berdasarkan Resampling Repeated Split Validation*. Skripsi Jurusan Matematika FMIPA Universitas Sriwijaya. Dipublikasikan.
- Shobayo, O., Oluwafemi, Z., Modupe, O. O., & Bayode, O. (2023). Prediction of Stroke Disease with Demographic and Behavioural Data Using Random Forest Algorithm. *Analytics Journal MDPI*. 2, pp. 604-617.
- Sugiyarto. (2021). *Pengantar Statistika Matematika I*. Magnum Pustaka Utama.
- Surono, S., & Rizki, D. A. P. (2021). Optimization Of Fuzzy C-Means Clustering Algorithm with Combination of Minkowski and Chebyshev Distance Using Principal Component Analysis. *International Journal of Fuzzy Systems*. 23(1), pp. 139–144.

Zaidi, N.A., Jesus, C., Mark, J.C., & Geoffrey, I.W. (2013). Alleviating Naive Bayes Attribute Independence Assumption by Attribute Weighting. *Journal of Machine Learning Research*. 14, pp. 1947–1988.