

**KLASIFIKASI KUALITAS UDARA MENGGUNAKAN
METODE NAIVE BAYES DAN FUZZY NAIVE BAYES DENGAN
DATA HILANG DAN DATA IMPUTASI**

SKRIPSI

Oleh:

GUSNADI JULIANSAH PUTRA

NIM.08011181924008



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

2025

LEMBAR PENGESAHAN

**KLASIFIKASI KUALITAS UDARA MENGGUNAKAN
METODE NAIVE BAYES DAN FUZZY NAIVE BAYES DENGAN
DATA HILANG DAN DATA IMPUTASI**

SKRIPSI

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

Oleh:

GUSNADI JULIANSAH PUTRA

NIM. 08011181924008

Pembimbing Kedua



Des Alwine Zavanti, S.Si., M.Si
NIP. 197012041998022001

**Indralaya 11 Maret 2025
Pembimbing Utama**



Prof. Yulia Resti, S.Si., M.Si., Ph.D
NIP. 197307191997022001

Mengetahui,

Ketua Jurusan Matematika



Dr. Dian Cahyawati Sukanda, S.Si., M.Si
NIP. 197303212000122001

PERNYATAAN KEASLIAN KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama Mahasiswa : Gusnadi Juliansah Putra

NIM : 08011181924008

Fakultas/Jurusan : Matematika dan Ilmu Pengetahuan Alam/Matematika

Menyatakan bahwa skripsi ini adalah hasil karya ilmiah saya sendiri dan karya ilmiah ini belum pernah diajukan sebagai pemenuhan persyaratan untuk memperoleh gelar kesarjanaan strata satu (S1) dari Universitas Sriwijaya maupun perguruan tinggi lain.

Semua informasi yang dimuat didalam skripsi ini yang berasal dari penulis lain baik yang dipublikasi atau tidak telah diberikan penghargaan dengan mengutip nama sumber penulis baik yang secara benar. Semua isi dari skripsi ini sepenuhnya menjadi tanggung jawab saya sebagai penulis.

Demikianlah surat pernyataan ini saya buat dengan sebenarnya.

Indralaya, 11 Maret 2025
Penulis



Gusnadi Juliansah Putra
NIM. 08011181924008

LEMBAR PERSEMBAHAN

MOTTO

“Tetapi boleh jadi kamu tidak menyenangi sesuatu, padahal itu baik bagimu, dan boleh jadi kamu menyukai sesuatu, padahal itu tidak baik bagimu, Allah mengetahui sedang kamu tidak mengetahui.”

(QS. Al-Baqarah : 216)

**“Tidak ada jalan yang mulus,
Tanpa doa ibu yang sangat tulus”**

Skripsi ini saya persembahkan untuk :

- ✓ **Allah SWT**
- ✓ **Rasulullah SAW**
- ✓ **Kedua Orang Tua Saya**
- ✓ **Kakak Saya**
- ✓ **Keluarga Besar Saya**
- ✓ **Semua Guru dan Dosen Saya**
- ✓ **Sahabat dan Teman-teman Saya**
- ✓ **Almamater Saya**

KATA PENGANTAR

Assalamu'alaikum warahmatullahi wabarakatuh

Puji syukur kehadiran Allah Subhanahu Wa Ta'ala Yang Maha Pengasih lagi Maha Penyayang karena atas limpahan berkah rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “**Klasifikasi Kualitas Udara Menggunakan Metode Naive Bayes dan Fuzzy Naive Bayes Dengan Data Hilang Dan Data Imputasi**” dengan baik dan tepat pada waktunya. Shalawat serta salam semoga senantiasa selalu tercurahkan kepada Nabi Muhammad SAW beserta keluarga, para sahabat, serta seluruh pengikutnya hingga akhir zaman. Skripsi ini disusun sebagai salah satu syarat agar dapat memenuhi syarat untuk memperoleh gelar Sarjana Sains bidang studi Matematika di Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.

Dengan segala hormat dan penuh kerendahan hati, penulis mengucapkan terima kasih yang tak terhingga dan terkhusus kepada orang tua tercinta, yaitu **Bapak Agus Perianto** dan **Ibu Suadesi Marleni** yang telah merawat, membimbing, mengarahkan dan selalu memberikan yang terbaik untuk penulis, baik untuk pendidikan, nasehat, perhatian, kasih sayang, motivasi, dan tak lupa doa yang selalu diberikan untuk penulis. Terimakasih untuk semua cinta yang selalu diberikan untuk penulis sehingga rela mengorbankan kebahagiaan ibu dan bapak hanya untuk kebahagiaan penulis. Dan tak lupa penulis juga mengucapkan terima kasih kepada semua yang sudah terlibat dan membantu dalam menyelesaikan skripsi ini, baik secara langsung maupun tidak langsung.

Untuk itu penulis menyampaikan ucapan terimakasih yang tak terhingga dan penghargaan kepada :

1. Bapak **Prof. Hermansyah, S.Si., M.Si., Ph. D** selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
2. Bapak **Dr. Dian Cahyawati Sukanda, S.Si., M.Si** selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
3. Ibu **Des Alwine Zayanti, S.Si., M.Si** selaku Sekretaris Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
4. Ibu **Prof. Dr. Yulia Resti, S.Si., M.Si., Ph. D** selaku Dosen Pembimbing Utama yang telah bersedia meluangkan waktu, pikiran, tenaga, dan nasehat dalam membimbing serta mengarahkan penulis dengan penuh kesabaran, pengertian dan perhatian baik selama perkuliahan maupun dalam penyelesaian skripsi sehingga penulis dapat menyelesaikan skripsi ini.
5. Ibu **Des Alwine Zayanti, S.Si., M.Si.**, selaku Dosen Pembimbing Kedua yang telah bersedia meluangkan waktu, pikiran, tenaga, dan nasehat dalam membimbing serta mengarahkan penulis dengan penuh kesabaran, pengertian dan perhatian baik selama perkuliahan maupun dalam penyelesaian skripsi sehingga penulis dapat menyelesaikan skripsi ini.
6. Ibu **Dr. Eka Susanti, M.Sc.**, dan bapak **Drs. Endro Setyo Cahyono, M.Si.**, selaku Dosen Pembahas dan penguji yang telah bersedia meluangkan waktu dalam memberikan tanggapan, kritik dan saran yang bermanfaat dalam perbaikan dan penyelesaian skripsi ini.

7. Bapak **Drs. Ali Amran, M.T.**, selaku Ketua Seminar skripsi yang telah memberikan arahan, saran, dan mengatur jalannya seminar sehingga dapat berjalan dengan baik.
8. Ibu **Dr. Dian Cahyawati Sukanda, S.Si., M.Si.**, selaku Dosen Pembimbing Akademik dan Sekretaris Seminar skripsi yang telah membantu, memberikan arahan, dan memberikan catatan serta masukan dan saran yang sangat berguna bagi penulis.
9. **Seluruh Dosen di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya** yang telah memberikan ilmu yang bermanfaat, arahan, nasehat serta bimbingan selama penulis menjalani penedidikan perkuliahan.
10. Bapak **Irwansyah** selaku admin dan Ibu **Hamidah** selaku pegawai tata usaha Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya yang telah membantu penulis selama masa perkuliahan terutama dalam proses administrasi.
11. Kakekku **Alm. Jubli** dan **Musijan**, Nenekku **Nurhayati** dan **Alm. Nur Hayati** Kakakku tersayang **Siti Maryam Ayunda Sari Dan Arman Zohri** Dan Adekku tercinta **Pandra Perianto, Lehha, M. Hatim Romadhon Dan Hanif Arkana Shidqi** yang selalu menjaga serta selalu memberikan kasih sayang, doa, motivasi, bantuan, dan nasihatnya.
12. Sahabatku **Raficko Deni Marantika** Terimakasih untuk canda tawa, suka duka, ketulusan, motivasi, pikiran, nasehat dan doanya.

13. Saudara seperjuangan satu kosan bernama **Pegi Karmiaji, Iqbal Bakti Effendi, Riski Mahardika, Rian Saputra, M. Nuralim, Hifzir Ridho, Rahmat, Alham, Yopan, kak Agus, Jenifer, Erwaini, Indah dan mbak Desi** yang sering membantu, memberikan semangat, memberikan motivasi serta memberikan dukungan kepada penulis.
14. My bro **Ahmad Budiman** selaku teman SMA sampai Kuliah yang telah memberikan semangat, dukungan serta motivasi kepada penulis
15. Ibu angkat **Puspa Ningsih** dan nenek **Rohaya** terimakasih untuk semua bantuannya selama menempuh perkuliahan. Terimakasih untuk canda tawa, suka duka, ketulusan, motivasi, pikiran, nasehat dan doanya.
16. Keluarga seperjuangan kuliah “team Serendipity” **Septa Lestari, Novi Fajaryanti, Anisa Safitri, Tia Sriyaningsih, Unsy Warzukni , Khairunnisa, Fauzi Darmawan, Ahmad Budiman, Putra Ramadhan, Muhammad Suedarmin, Kristiadi Manaek Silaban, Jimmy, Meigia Purnama Sari, Riandino Febriansyah** terimakasih atas semua bantuannya selama menempuh pendidikan didunia perkuliahan dan terimakasih untuk rasa sayang, peduli, cerita, canda tawa, suka duka, motivasi, nasehat dan doanya.
17. Kakak tingkat **2016, 2017, 2018** beserta adik tingkat angkatan **2020, 2021, 2022** yang telah membantu memberikan arahan dan bantuan selama perkuliahan serta dalam penulisan skripsi ini.

18. **Seluruh Mahasiswa Matematika Universitas Sriwijaya Angkatan 2019** yang telah banyak membantu penulis selama menempuh perkuliahan maupun dalam pembuatan skripsi.

19. **Semua pihak** yang telah membantu penulis dalam menyelesaikan skripsi ini yang tidak dapat disebutkan satu per satu, penulis hanya bisa mengucapkan terima kasih kepada semuanya yang telah membantu penulis, dan semoga segala kebbaikannya mendapat pahala dari Allah SWT.

Penulis berharap agar skripsi ini semoga bermanfaat bagi seluruh pihak yang membutuhkan terutama untuk mahasiswa-mahasiswi jurusan matematika serta terkhusus untuk yang membacanya.

Wassalamu'alaikum warahmatullahi wabarakatuh

Indralaya, 11 Maret 2025

Gusnadi Juliansah Putra

**USING AIR QUALITY CLASSIFICATION
NAIVE BAYES AND FUZZY NAIVE BAYES METHOD WITH
LOST DATA AND IMPUTATION OF DATA**

By:

Gusnadi Juliansah Putra

08011181924008

ABSTRACT

Healthy air plays a very important role in the survival of life, whereas polluted air will have a negative impact on living things. Therefore, it is necessary to classify air quality. The aim of this research is to carry out classification to determine good air quality using the Naive Bayes and Fuzzy Naive Bayes methods with missing data and imputed data. The research results show that using both the Naive Bayes and Fuzzy Naive Bayes methods by removing missing data obtains better results than imputing missing data. Classification of air quality data using the Naive Bayes method by removing missing data obtained an accuracy of 85.92%, precision of 8.21%, recall of 36.01% and Fscore of 31.45%. Meanwhile, classification of air quality data using the Naive Bayes method by imputing any missing data obtained an accuracy of 82.71%, precision of 23.78%, recall of 24.29% and Fscore of 20.72%. Classification of air quality data using the fuzzy Naive Bayes method by removing missing data obtained an accuracy of 85.6%, precision of 44.03%, recall of 46.04% and Fscore of 43.08%. Meanwhile, classification of air quality data using the fuzzy Naive Bayes method by imputing any missing data obtained an accuracy of 84.89%, precision of 27.2%, recall of 26.47% and Fscore of 24.97%.

Keywords: *Air Quality, Fuzzy Sets, Naive Bayes, Fuzzy Naive Bayes.*

**KLASIFIKASI KUALITAS UDARA MENGGUNAKAN
METODE NAIVE BAYES DAN FUZZY NAIVE BAYES DENGAN
DATA HILANG DAN DATA IMPUTASI**

Oleh:

Gusnadi Juliansah Putra

08011181924008

ABSTRAK

Udara yang sehat sangat berperan penting dalam keberlangsungan hidup sebaliknya udara yang tercemar akan mendatangkan dampak buruk bagi makhluk hidup. Oleh karena itu perlu melakukan pengklasifikasian terhadap kualitas udara. Tujuan penelitian ini adalah melakukan klasifikasi untuk menentukan kualitas udara yang baik menggunakan metode Naive Bayes dan Fuzzy Naive Bayes dengan data hilang dan data imputasi. Hasil penelitian menunjukkan bahwa, baik menggunakan metode Naive Bayes maupun Fuzzy Naive Bayes dengan menghapus data yang hilang memperoleh hasil lebih baik dibandingkan mengimputasi data yang hilang. Klasifikasi data kualitas udara menggunakan metode Naive Bayes dengan menghapus data hilang memperoleh akurasi sebesar 85,92%, Presisi sebesar 8,21%, *recall* sebesar 36,01% dan *Fscore* sebesar 31,45%. Sedangkan klasifikasi data kualitas udara dengan menggunakan metode Naive Bayes dengan mengimputasi setiap data yang hilang memperoleh akurasi sebesar 82,71%, presisi sebesar 23,78%, *recall* sebesar 24,29% dan *Fscore* sebesar 20,72%. Klasifikasi data kualitas udara menggunakan metode fuzzy Naive Bayes dengan menghapus data hilang memperoleh akurasi sebesar 85,6%, presisi sebesar 44,03%, *recall* sebesar 46,04% dan *Fscore* sebesar 43,08%. Sedangkan klasifikasi data kualitas udara dengan menggunakan metode fuzzy Naive Bayes dengan mengimputasi setiap data yang hilang memperoleh akurasi sebesar 84,89%, presisi sebesar 27,2%, *recall* sebesar 26,47% dan *Fscore* sebesar 24,97%.

Kata Kunci: Kualitas Udara, Himpunan Fuzzy, Naive Bayes, Fuzzy Naive Bayes.

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PENGESAHAN	ii
LEMBAR PERSEMBAHAN	iii
KATA PENGANTAR.....	iv
ABSRTRACT	ix
ABSTRAK	x
DAFTAR ISI.....	xi
DAFTAR TABEL	xiv
DAFTAR GAMBAR.....	xviii
DAFTAR LAMPIRAN	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Udara	5
2.2 Pencemaran Udara.....	5

2.3 Parameter Pencemaran Udara.....	6
2.4 Data Mining.....	9
2.5 Klasifikasi	9
2.6 Dikritisasi.....	10
2.7 Himpunan Fuzzy	10
2.8 Fungsi Keanggotaan.....	11
2.9 Peluang	15
2.10 Metode Naive Bayes	15
2.11 Fuzzy Naive Bayes.....	19
2.10 <i>Confusion Matrix</i>	19
BAB III METODOLOGI PENELITIAN	22
3.1 Tempat.....	22
3.2 Waktu	22
3.3 Data Penelitian.....	22
3.4 Metode Penelitian	23
BAB IV HASIL DAN PEMBAHASAN	28
4.1 Deskripsi Data	28
4.2 Klasifikasi Dengan Metode NB.....	29
4.2.1 Klasifikasi Metode NB Dengan Data Hilang	29
4.2.2 Klasifikasi Metode NB Dengan Data Imputasi	39
4.3 Klasifikasi Dengan Metode FNB	49
4.3.1 Metode FNB Dengan Menghapus Data Hilang.....	49
I. Fungsi Keanggotaan Fuzzy.....	51
II. FNB	53
4.3.2 Metode FNB Dengan Menginputasi Data Hilang	60
I. Fungsi Keanggotaan Fuzzy.....	62
II. FNB	64
4.4 Analisis Hasil.....	71

BAB V KESIMPULAN DAN SARAN	72
5.1 Kesimpulan.....	72
5.2 Saran.....	73
DAFTAR PUSTAKA	74
LAMPIRAN.....	76

DAFTAR TABEL

Tabel 2. 1 <i>Multiclass Confusion Matrix</i>	20
Tabel 3. 1 Variabel Respon.....	22
Tabel 3. 2 Variabel Prediktor	23
Tabel 4. 1 Deskripsi Data.....	28
Tabel 4. 2 Data Penelitian	29
Tabel 4. 3 Data Setelah Menghapus Data Hilang NB.....	30
Tabel 4. 4 Nilai Maksimum, Minimum Dan <i>Range</i> Data Hilang NB.....	31
Tabel 4. 5 Batasan Semua Kategori Data Hilang NB	32
Tabel 4. 6 Data Hasil Dikritisasi Data Hilang NB	32
Tabel 4. 7 Data <i>Training</i> Data Hilang NB	33
Tabel 4. 8 Data <i>Testing</i> Data Hilang NB	33
Tabel 4. 9 Nilai Peluang <i>Likelihood</i> Data Hilang NB.....	35
Tabel 4. 10 Hasil Peluang <i>Posterior</i> Data Hilang NB	36
Tabel 4. 11 <i>Confusion Matrix</i> Data Hilang NB	37
Tabel 4. 12 Nilai <i>Confusion Matrix</i> Data Hilang NB	37
Tabel 4. 13 Akurasi, Presisi, <i>Recall</i> dan <i>Fscore</i> Data Hilang NB	39
Tabel 4. 14 Data Setelah Menginputasi Data Hilang NB	39
Tabel 4. 15 Nilai Maksimum, Minimum Dan <i>Range</i> Data Imputasi NB	40
Tabel 4. 16 Batasan Semua Kategori Data Imputasi NB	41
Tabel 4. 17 Data Hasil Dikritisasi Data Imputasi NB	42
Tabel 4. 18 Data <i>Training</i> Data Imputasi NB	42
Tabel 4. 19 Data <i>Testing</i> Data Imputasi NB	43

Tabel 4. 20 Nilai Peluang <i>Likelihood</i> Data Imputasi NB.....	45
Tabel 4. 21 Hasil Peluang <i>Posterior</i> Data Imputasi NB	46
Tabel 4. 22 <i>Confusion Matrix</i> Data Imputasi NB	47
Tabel 4. 23 Nilai <i>Confusion Matrix</i> Data imputasi NB	47
Tabel 4. 24 Akurasi, Presisi, <i>Recall</i> dan <i>Fscore</i> Imputasi NB.....	48
Tabel 4. 25 Interval Himpunan Fuzzy Variabel X_1 Data Hilang FNB	49
Tabel 4. 26 Himpunan Fuzzy Variabel X_1 Data Hilang FNB	50
Tabel 4. 27 Himpunan Fuzzy Data Hilang FNB.....	51
Tabel 4. 28 Nilai Keanggotaan Tertinggi Variabel X_1 Data Hilang FNB	52
Tabel 4. 29 Data <i>Training</i> Data Hilang FNB.....	53
Tabel 4. 30 Data <i>Testing</i> Data Hilang FNB	53
Tabel 4. 31 Nilai Peluang <i>Likelihood</i> Data Hilang FNB	55
Tabel 4. 32 Hasil Peluang <i>Posterior</i> Data Hilang FNB.....	57
Tabel 4. 33 <i>Confusion Matrix</i> Data Hilang FNB	57
Tabel 4. 34 Nilai <i>Confusion Matrix</i> Data Hilang FNB	58
Tabel 4. 35 Akurasi, Presisi, <i>Recall</i> dan <i>Fscore</i> Data Hilang FNB.....	59
Tabel 4. 36 Interval Himpunan Fuzzy Variabel X_1 Data imputasi FNB.....	60
Tabel 4. 37 Himpunan Fuzzy Variabel X_1 Data Imputasi FNB.....	61
Tabel 4. 38 Himpunan Fuzzy Data Hilang FNB.....	62
Tabel 4. 39 Nilai Keanggotaan Tertinggi Variabel X_1 Data Imputasi FNB	63
Tabel 4. 40 Data <i>Training</i> Data Imputasi FNB.....	64
Tabel 4. 41 Data <i>Testing</i> Data Imputasi FNB	64
Tabel 4. 42 Nilai Peluang <i>Likelihood</i> Data Imputasi FNB	67

Tabel 4. 43 Hasil Peluang <i>Posterior</i> Data Imputasi FNB	68
Tabel 4. 44 <i>Confusion Matrix</i> Data Imputasi FNB	69
Tabel 4. 45 Nilai <i>Confusion Matrix</i> Data Imputasi FNB	69
Tabel 4. 46 Akurasi, Presisi, <i>Recall</i> dan <i>Fscore</i> Data Hilang FNB	70
Tabel 4. 47 Perbandingan Hasil Metode NB dan FNB	71

DAFTAR GAMBAR

Gambar 2.1 Kurva Linier Naik	12
Gambar 2. 2 Kurva Linier Turun	13
Gambar 2. 3 Kurva Segitiga.....	14

DAFTAR LAMPIRAN

Lampiran 1 Dataset setelah di Fuzzy	76
Lampiran 2 Dataset setelah di Fuzzy	77

BAB I

PENDAHULUAN

1.1 Latar Belakang

Udara adalah salah satu bagian dari alam yang dibutuhkan makhluk hidup untuk bernafas karena udara mengandung oksigen yang dibutuhkan untuk proses metabolisme. Udara yang sehat sangat berperan penting dalam keberlangsungan hidup sebaliknya udara yang tercemar akan mendatangkan dampak buruk bagi makhluk hidup. Pencemaran udara adalah keadaan dimana udara mengalami penurunan kualitas yang disebabkan oleh transportasi, industri, dan domestik. Pencemaran udara secara tidak langsung juga disebabkan oleh pertumbuhan penduduk, pembangunan tata ruang yang tidak seimbang dan kurangnya tingkat kesadaran masyarakat terhadap pencemaran udara (Inaku dan Novianus, 2020).

Masalah kualitas udara menjadi salah satu masalah utama yang perlu diperhatikan karena berdampak pada kesehatan dan lingkungan. Pada tahun 2019, India mencatatkan angka kematian akibat polusi udara tertinggi di dunia, dengan sekitar 1,67 juta kematian setiap tahunnya. Faktor penyebab utamanya adalah polusi udara dari kendaraan, industri, pembakaran sampah, dan pola cuaca yang buruk. Tingginya populasi dan urbanisasi juga menjadi faktor yang memperparah masalah ini. Naive Bayes dapat digunakan untuk mengklasifikasikan data kualitas udara di India menjadi kategori yang berbeda berdasarkan faktor-faktor yang mempengaruhinya. Beberapa faktor yang dapat dijadikan variabel dalam klasifikasi

ini antara lain tingkat polusi udara, jenis polutan, lokasi geografis, kondisi cuaca, dan tingkat populasi.

Metode Naive Bayes dan fuzzy Naive Bayes telah banyak digunakan dalam berbagai penerapan. Keduanya menggunakan konsep teorema Bayes dalam menentukan probabilitas kelas, namun metode fuzzy Naive Bayes melibatkan konsep keanggotaan himpunan fuzzy (Resti *et al.* 2023). Metode Naive Bayes dan fuzzy Naive Bayes dalam mengklasifikasi seringkali terdapat data hilang yaitu data variabel bebas tidak lengkap. Data hilang sepatutnya diimputasi untuk menghindari hasil prediksi yang bias atau hasil prediksi yang tidak tepat. Beberapa metode yang digunakan untuk mengimputasi (mengisi) data yang hilang diantaranya adalah metode rata-rata (mean) dan median.

Penelitian terdahulu dilakukan oleh Riyanah dan Fatmawati (2021) mengenai penerapan metode Naive Bayes untuk klasifikasi penerima bantuan surat keterangan tidak mampu, menggunakan 35 data *training* dan 8 data *testing*, setelah melakukan perhitungan dengan metode Naive Bayes memperoleh akurasi sebesar 62,86%. Penelitian lain juga dilakukan oleh Ali *et al* (2021) mengenai analisis sentimen tweet KRI Nanggala 402 di twitter menggunakan metode Naive Bayes, menggunakan 39 data *training* dan 14 data *testing*, setelah melakukan perhitungan dengan metode Naive Bayes memperoleh akurasi sebesar 73,00%. Setelah penjelasan masalah tersebut dan berdasarkan dari beberapa kajian penelitian terdahulu muncul ketertarikan terhadap peneliti untuk melakukan penelitian menggunakan metode Naive Bayes dan metode fuzzy Naive Bayes dengan data hilang dan data imputasi terhadap kualitas udara di India.

1.2 Rumusan Masalah

1. Bagaimana tingkat ketepatan dari pengklasifikasian kualitas udara menggunakan metode Naive Bayes dan fuzzy Naive Bayes dengan menghapus data hilang?.
2. Bagaimana tingkat ketepatan dari pengklasifikasian kualitas udara menggunakan metode Naive Bayes dan fuzzy Naive Bayes dengan mengimputasi data hilang?.
3. Bagaimana hasil perbandingan tingkat ketepatan dari pengklasifikasian kualitas udara menggunakan metode Naive Bayes dan fuzzy Naive Bayes dengan data hilang dan data imputasi.

1.3 Batasan Masalah

1. Data hilang diimputasi dengan rata-rata (mean).
2. Tingkat ketepatan yang digunakan adalah akurasi, presisi, *recall* dan *Fscore*.
3. Data yang digunakan adalah data kualitas udara India pada tahun 2015 hingga 2020.
4. Variabel yang digunakan berjumlah 9 variabel.

1.4 Tujuan

1. Memperoleh tingkat ketepatan dari pengklasifikasian kualitas udara menggunakan metode Naive Bayes dan fuzzy Naive Bayes dengan menghapus data yang hilang.

2. Memperoleh tingkat ketepatan dari pengklasifikasian kualitas udara menggunakan metode Naive Bayes dan fuzzy Naive Bayes dengan mengimputasi data yang hilang.
3. Memperoleh hasil perbandingan tingkat ketepatan dari pengklasifikasian kualitas udara menggunakan metode Naive Bayes dan fuzzy Naive Bayes dengan data hilang dan data imputasi.

1.5 Manfaat

Manfaat bagi ilmuwan adalah:

1. Penelitian ini diharapkan menjadi sumber bacaan dalam mengklasifikasi kualitas udara menggunakan metode Naive Bayes dan fuzzy Naive Bayes dengan data hilang dan data imputasi.
2. Penelitian ini dapat digunakan sebagai rujukan untuk penelitian selanjutnya dalam mengklasifikasi objek dengan menggunakan metode Naive Bayes fuzzy Naive Bayes.

Manfaat bagi masyarakat adalah:

1. Membantu memonitoring kualitas udara.
2. Membantu masyarakat yang tinggal di lingkungan pabrik mengetahui kondisi udara disekitar tempat tinggal.

DAFTAR PUSTAKA

- Anggara, M.A. (2021). Pengelompokan Risiko Kredit Menggunakan Metode Naive Bayes. Skripsi Jurusan Matematika Fakultas Matematika Dan Ilmu Pengetahuan Alam Universitas Sriwijaya. Indralaya.
- Candana, E.W.H.,& Gunadu, I.G.A., & Divayana, D.G.H. (2021). Perbandingan *fuzzy tsukamoto*, *mamdini* dan *sugeno* dalam penentuan hari baik pernikahan berdasarkan wariga menggunakan confuse on matrix. *Jurnal Ilmu Komputer Indonesia (JIK)*, 6(2), 14–22.
- Djamaludin, M. A., Triayudi, A., & Mardiani, E. (2022). Analisis Sentimen Tweet KRI Nanggala 402 di Twitter menggunakan Metode Naive Bayes Classifier. *Jurnal JTIC (Jurnal Teknologi Informasi dan Komunikasi)*, 6(2), 161-166.
- Gandhi, B.S., & Megawaty, D.A. & Alita, D. (2021). Aplikasi monitoring dan penentuan peringkat kelas. *Jurnal Informatika dan Rekayasa Perangkat Lunak (JATIKA)*, 2(1), 54–63.
- Inaku, A.H.R., & Novianus, C. (2020). Pengaruh pencemaran udara PM 2,5 dan PM 10 terhadap keluhan Pernapasan Anak di Ruang Terbuka Anak di DKI Jakarta. *Jurnal ARKESMAS (Arsip Kesehatan Masyarakat)*, 5(2), 9–16.
- Kusuma, S. R., Hartati, R. S., & Sukerayasa, I. W. (2020). Pengaruh Jumlah Fungsi Keanggotaan pada Metode Fuzzy Logic Terhadap Hasil Peramalan Beban Listrik Jangka Panjang. *Jurnal SPEKTRUM Vol*, 7(1).
- Kusumaningsih, D., Wicahyo, A., & Pudoli, A. (2021). Penggunaan Algoritma Naive Bayes dalam klasifikasi Pengaruh Pencemaran Udara. *Jurnal ICT: Information Communication & Technology*, 21(1), 103-108.
- Megalina, Y. (2015). Pengaruh pencemaran udara di daerah terminal Amplas bagi kehidupan masyarakat. *Jurnal Pengabdian Kepada Masyarakat*, 21(79), 95-101.
- Muslim, A.M., & Prasetyo, B., & Harum, E.L, *et al.* (2019). *Data mining algoritma C4.5 disertai contoh kasus penerapannya dengan program komputer*. Unnes. Semarang. Indonesia.

- Norwati, D., & Prayogi S.A.(2021). Implementasi *naïve bayes classifier* dan *confusion matrix* pada analisis setimen berbasis teks pada twitter. *Jurnal Sains Komputer & Informatika*, 7(1), 214–219.
- Pujianto, U., & Widiyaningtyas, T., & Prasetya, D.D., & Romadhon, B. (2017). Penerapan algoritma *naïve bayes classifier* untuk klasifikasi judul skripsi dan tugas akhir berdasarkan kelompok bidang keahlian. *Jurnal Teknologi Elektro dan Kejuruan*, 27(1), 79–92.
- Rahman, M.F., & Darmawidjadja M.I., & Alamsyah, A. (2017). Klasifikasi untuk diagnosa diabetes menggunakan metode *Bayesian Regularization Neural Network* (RBNN), *Jurnal Informatika*, 11(1), 36-45.
- Resti, Y., Irsan, C., Neardiaty, A., Annabila, C., & Yani, I. (2023). *Fuzzy Discretization on the Multinomial Naïve Bayes Method for Modeling Multiclass Classification of Corn Plant Diseases and Pests. Mathematics*, 11(8), 1761.
- Ribeiro, M. S., Santos, M. F., & Farias, M. C. (2018). Kualitas Udara dan Emisi Gas Rumah Kaca di Kawasan Perkotaan: Studi Kasus di Bandung, Indonesia. *Journal of Environmental Science and Health, Part B*, 63(1), 15-26.
- Riyanah, N., & Fatmawati, F. (2021). Penerapan Algoritma Naive Bayes Untuk Klasifikasi Penerima Bantuan Surat Keterangan Tidak Mampu. *JTIM: Jurnal Teknologi Informasi Dan Multimedia*, 2(4), 206-213.
- Shirazi, M., & Asgari, F. (2017). *Investigating the relationship between air quality and health indicators in urban areas of Iran. Journal of Environmental Health*, 16(2), 145-154.
- Sugiyarto, (2021), *pengantar statistika matematika 1*, Magnum Pustaka Utama, DI Yogyakarta.
- Tangkelayuk. A ., Evangs. M. 2022. Klasifikasi kualitas air dengan metode KNN, *Naive Bayes* dan *Decision Tree*. *Jurnal Teknik Informatika dan Sistem Informasi*. Vol 9. No 2. Hal 1109-1119.
- Utami, D., & Devi, P.A.R. (2022). Klasifikasi kelayakan penerima bantuan program keluarga harapan (pkh) menggunakan metode *weighted naïve bayes* dengan *laplace smoothing*. *Jurnal Ilmiah Penelitian dan Pembelajaran Informatika*, 7(4),1373–1384.