

# KLASIFIKASI POLA PEMINJAM BUKU BERDASARKAN PROFESI MENGGUNAKAN ALGORITMA NAÏVE BAYES

Oleh:

Febri Rosita Dewi

Ade Eviyanti

Progam Studi Informatika

Universitas Muhammadiyah Sidoarjo

Januari, 2025

# Pendahuluan

Perpustakaan memiliki peran penting dalam mendukung literasi dan pendidikan masyarakat. Dinas Perpustakaan Sidoarjo memanfaatkan data peminjaman buku untuk memahami kebutuhan peminjam dari berbagai profesi seperti pelajar, mahasiswa, pegawai, dan wirausaha. Namun, klasifikasi peminjam secara manual sering kali tidak efisien dan cenderung bias.

Penelitian ini menerapkan algoritma Naive Bayes, yaitu Gaussian dan Multinomial Naive Bayes, untuk mengklasifikasikan peminjam berdasarkan profesi mereka. Gaussian Naive Bayes digunakan untuk atribut numerik, sedangkan Multinomial Naive Bayes lebih cocok untuk atribut kategori. Evaluasi hasil klasifikasi dilakukan untuk memastikan akurasi dan efisiensi algoritma dalam mengelompokkan data.

Tujuan penelitian ini adalah memberikan wawasan tentang kebiasaan membaca peminjam berdasarkan profesi, seperti minat guru pada buku pendidikan atau profesional teknologi pada buku sains. Hasil klasifikasi ini diharapkan membantu perpustakaan dalam mengoptimalkan koleksi, meningkatkan layanan, dan merancang program literasi yang lebih efektif.

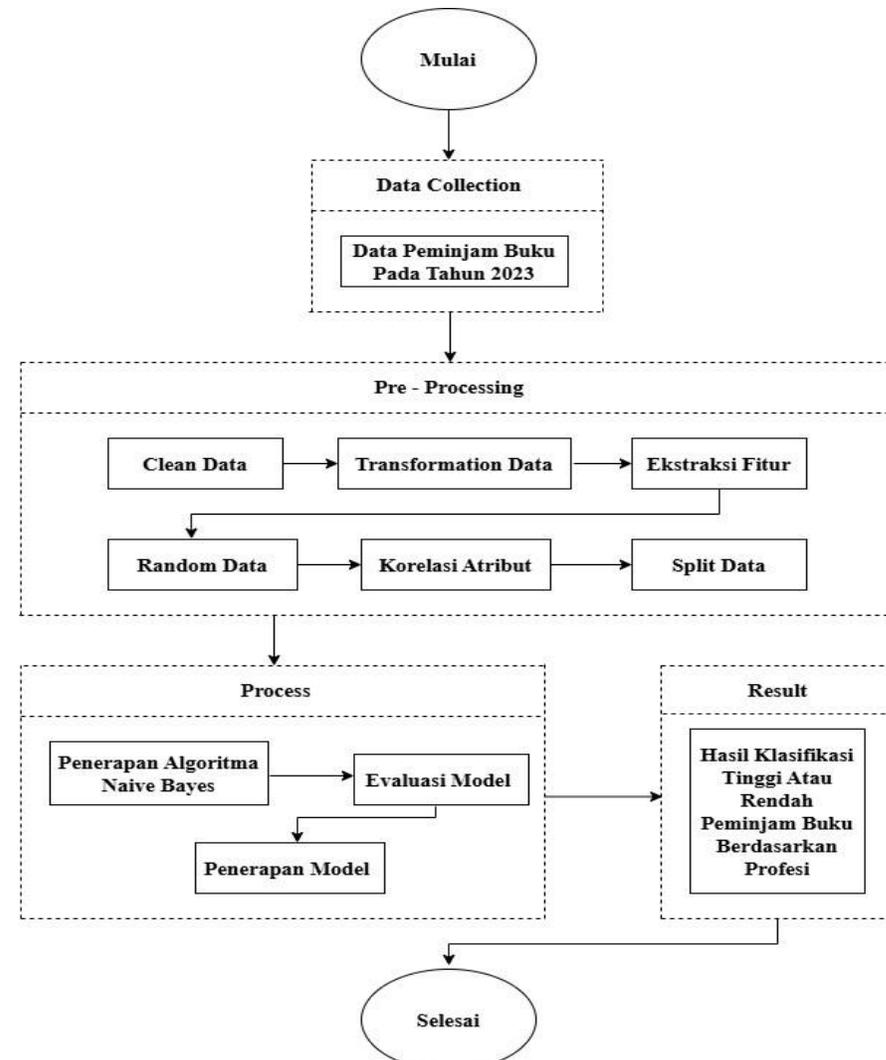
# Pertanyaan Penelitian (Rumusan Masalah)

Dari latar belakang yang telah disampaikan, rumusan masalah dalam penelitian ini adalah:

- 1) Bagaimana penerapan algoritma Naive Bayes untuk mengklasifikasikan peminjam buku berdasarkan profesi?
- 2) Seberapa efektif algoritma Naive Bayes (Gaussian dan Multinomial) dalam mengklasifikasikan peminjam buku berdasarkan profesi?

# Metode

Tahapan penelitian ini dimulai dari pengumpulan data (Data Collection), dilanjutkan dengan preprocessing data, pemrosesan, dan diakhiri dengan analisis hasil (Result).



# Metode

## 1. Data Collection

Data didapatkan setelah melakukan wawancara langsung dengan pihak Dinas Perpustakaan dan Kearsipan Kabupaten Sidoarjo sehingga diketahui data transaksi peminjam buku. Data yang digunakan berasal dari Dinas Perpustakaan dan Kearsipan Kabupaten Sidoarjo, yang memiliki jumlah 4476 pada tahun 2023.

## 2. Pre – Processing

Preprocessing adalah langkah penting dalam analisis data untuk menyiapkan data sebelum pemodelan. Pada penelitian ini, preprocessing meliputi pembersihan data, transformasi, ekstraksi fitur, pengacakan data, analisis korelasi atribut, dan pembagian data.

- **Clean Data** = proses menghapus kesalahan, inkonsistensi, informasi tidak relevan, serta menangani data hilang atau duplikat dengan imputasi median atau penghapusan
- **Transformation Data** = Merubah data category menjadi format numerik
- **Ekstraksi Fitur** = memilih fitur-fitur yang relevan untuk model klasifikasi.

# Metode

- **Random Data** = Yang elemen-elemen datanya disusun secara acak
- **Korelasi Atribut** = Correlation attribute menunjukkan hubungan antara variabel dalam dataset untuk menemukan pola penting guna mendukung keputusan, prediksi, atau pengelompokan.
- **Split Data** = Split data awal 80:20, kemudian ditambahkan rasio 90:10, 75:25, dan 50:50. Pengujian dilakukan 2 model menggunakan Gaussian dan Multinomial Naïve Bayes untuk membandingkan akurasi

# Metode

## 3. Process

Process data dalam konteks pemodelan menggunakan algoritma naïve bayes yang mencakup serangkaian tahapan yang sistematis untuk mengolah data, menerapkan algoritma, dan mengevaluasi hasil. Berikut adalah rincian dari tahapan tersebut.

- **Penerapan Algoritma Naïve Bayes** = Teknik statistik untuk mengklasifikasikan data dengan mengestimasi peluang keanggotaan kelas menggunakan Teorema Bayes
- **Evaluasi Model** = Mengukur performa menggunakan confusion matrix dan metrik seperti akurasi, presisi, recall, dan F1-score. Penelitian ini mengevaluasi Gaussian Naïve Bayes dan Multinomial Naïve Bayes.
- **Penerapan Model** = Model yang telah dilatih diterapkan pada data baru untuk mengklasifikasikan profesi peminjam dan merekomendasikan buku sesuai profesi.

# Metode

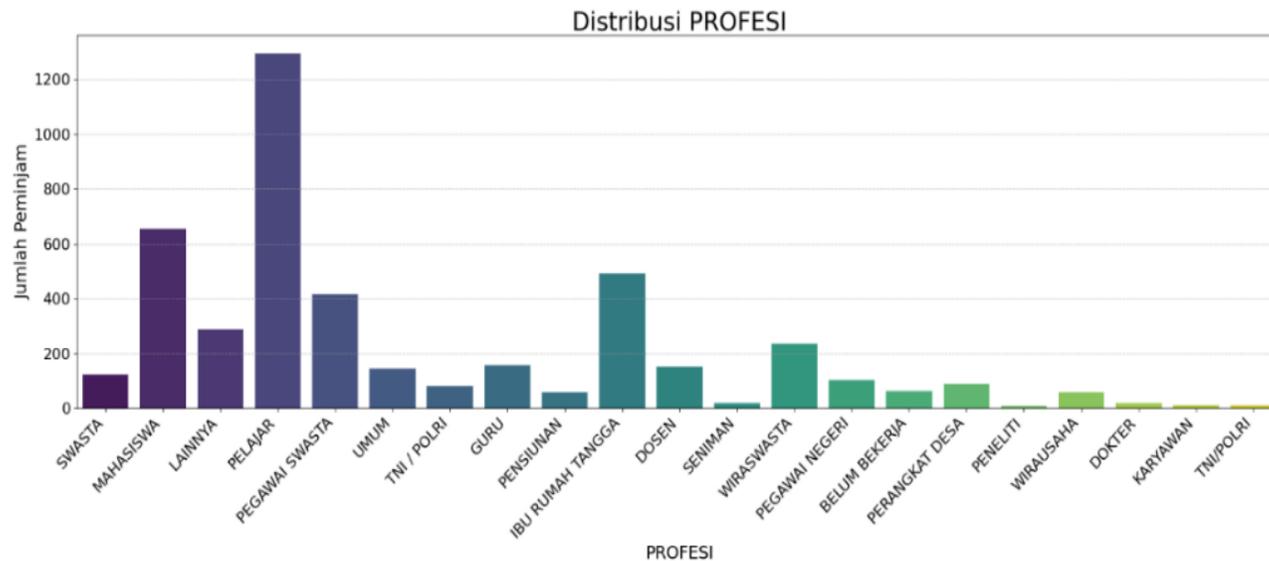
## 4. Result

Result adalah hasil klasifikasi peminjam buku berdasarkan waktu pengembalian yang dikelompokkan menjadi dua kategori yaitu tinggi dan rendah.

- **Tinggi** = Mengembalikan <10 hari, menunjukkan disiplin dan minat baca tinggi.
- **Rendah** = Mengembalikan >10 hari, menunjukkan minat rendah atau kurang disiplin.

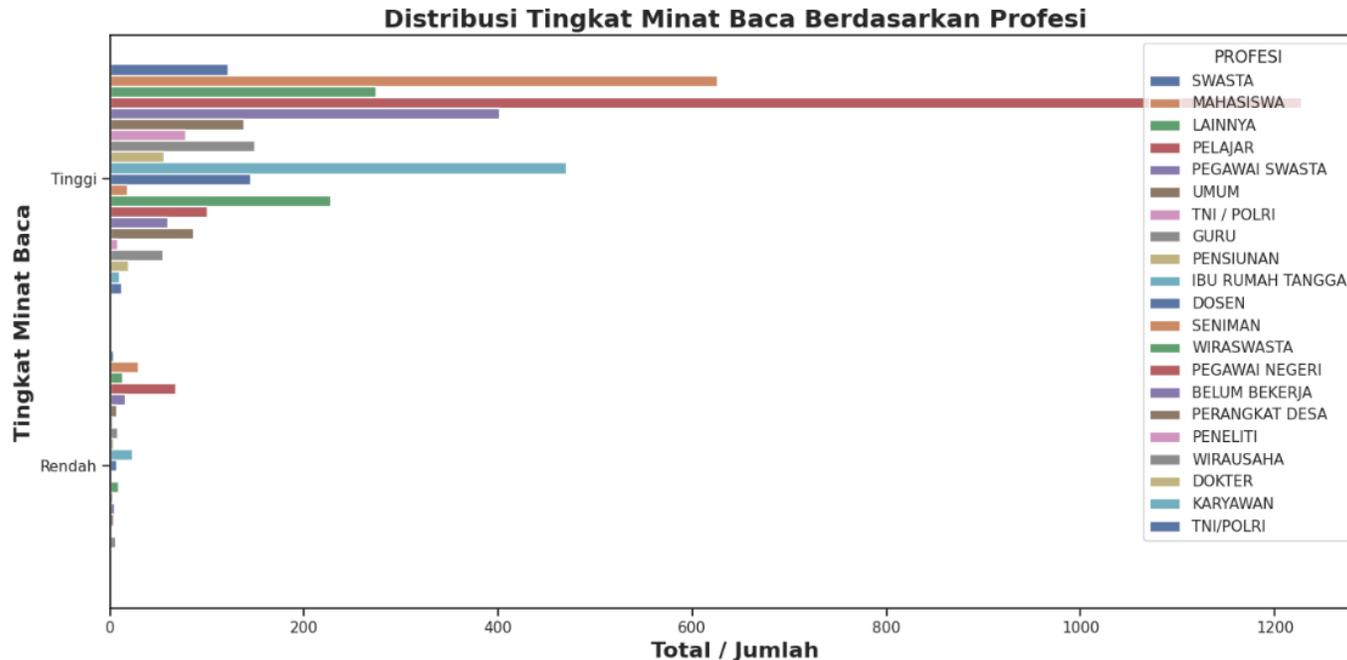
# Hasil

Penelitian ini mengungkapkan beragam profil peminjam buku di perpustakaan berdasarkan data yang disajikan dalam bentuk grafik. Berikut pembahasan hasil yang ditemukan:



Gambar disamping menunjukkan Pelajar sebagai peminjam terbanyak (1.294), diikuti Mahasiswa (655), Ibu Rumah Tangga (493), dan Pegawai Swasta (417). Hal ini menunjukkan perpustakaan menarik minat dari berbagai profesi, terutama akademik dan swasta

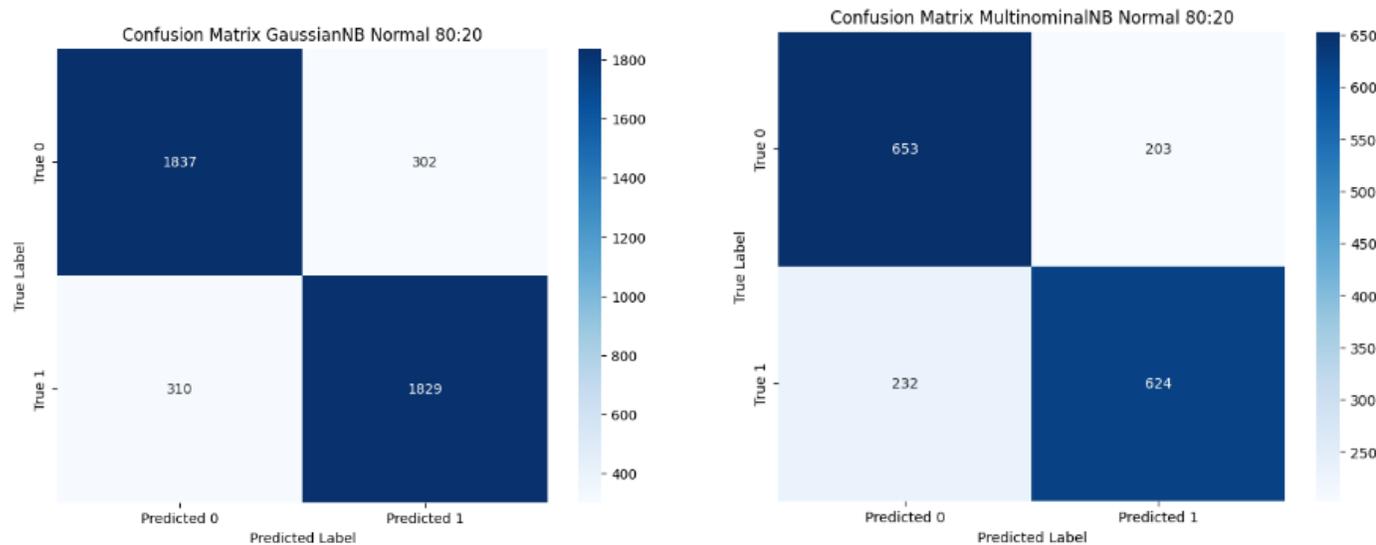
# Hasil



Gambar di samping menunjukkan distribusi minat baca berdasarkan profesi. Pelajar memiliki minat baca tertinggi, diikuti Mahasiswa, Ibu Rumah Tangga, dan Pegawai Swasta. Profesi seperti Dokter dan Peneliti memiliki minat baca yang lebih rendah.

# Hasil

## ❖ Hasil Pengujian Model Skenario Normal Rasio 80:20



## Tabel Perbandingan

Performa	Gaussian NB	Multinomial NB
Accuracy	0.85%	0.75%
Precision Negatif (0)	0.85	0.74
Precision Positif (1)	0.85	0.75
Recall Negariff(0)	0.87	0.76
Recall Positif (1)	0.85	0.73
F-1 Score Negatif (0)	0.86	0.75
F-1 Score Positif (1)	0.85	0.74
Support Negatif (0)	856	856
Support Positif (1)	856	856

Pada matrix evaluasi di Tabel di atas, terlihat bahwa nilai akurasi tertinggi diperoleh dari model Gaussian Naive Bayes, yaitu sekitar 85%, sedangkan model Multinomial Naive Bayes mencapai akurasi sekitar 75%.

# Pembahasan

## ❖ Perbandingan Model, Rasio Data dan Skenario Pengujian

Tabel Perbandingan Akurasi Model Gaussian Naïve Bayes

Splitting Data	Normal Data	Random Data	Korelasi Atribut Data
90:10	85%	97%	96%
80:20	85%	97%	96%
75:25	86%	97%	96%
50:50	86%	97%	97%

Pada matriks evaluasi di tabel di atas, terlihat bahwa nilai akurasi tertinggi diperoleh dari metode pengolahan data Random Data pada model Gaussian Naïve Bayes, yaitu sekitar 97%, sedangkan metode Normal Data hanya mencapai akurasi sekitar 85%-86%.

# Pembahasan

Tabel Perbandingan Akurasi Model Multinomial Naïve Bayes

Rasio Data	Normal Data	Random Data	Korelasi Atribut Data
90:10	73%	95%	88%
80:20	75%	76%	89%
75:25	74%	96%	89%
50:50	75%	95%	92%

Pada matriks evaluasi di tabel di atas, terlihat bahwa nilai akurasi tertinggi diperoleh dari metode Random Data pada model Multinomial Naïve Bayes, yaitu sekitar 96%, sedangkan metode Normal Data hanya mencapai akurasi sekitar 73%-75%.

# Manfaat Penelitian

- Manfaat penelitian ini adalah memberikan wawasan bagi perpustakaan untuk mengembangkan layanan berbasis data, mengoptimalkan koleksi buku sesuai kebutuhan profesi, serta meningkatkan program literasi. Hasil klasifikasi pola peminjaman berdasarkan profesi membantu perpustakaan memahami preferensi pengguna dan merancang layanan yang lebih responsif.
- Selain itu, perbandingan model, rasio data, dan skenario pengujian membantu mengidentifikasi model yang paling efektif, seperti Gaussian Naive Bayes dan Multinomial Naive Bayes. Pengujian dengan berbagai rasio data dan skenario memberikan wawasan tentang performa model, memungkinkan perpustakaan memilih metode yang tepat untuk mengelola data peminjam dan meningkatkan akurasi layanan.

# Referensi

- [1] W. Nasihuddin, “Peran Perpustakaan Sebagai Media Literasi Digital Masyarakat,” *Jurnal Politik dan Sosial Kemasyarakatan* 11, 2019.
- [2] N. S. N. Arimbi, “Pelayanan dan Fasilitas Perpustakaan Umum Daerah Sidoarjo (Perpusda),” *Jatim News*. Accessed: Jul. 03, 2024. [Online]. Available: <https://jatim.news/2024/05/04/pelayanan-dan-fasilitas-perpustakaan-umum-daerah-sidoarjo-perpusda/>
- [3] A. R. Abdillah and F. N. Hasan, “Analisis Sentimen Terhadap Kandidat Calon Presiden Berdasarkan Tweets Di Sosial Media Menggunakan Naive Bayes Classifier,” *SMATIKA JURNAL*, vol. 13, no. 01, pp. 117–130, Jul. 2023, doi: 10.32664/smatika.v13i01.750.
- [4] E. Martantoh and N. Yanih, “Implementasi Metode Naive Bayes Untuk Klasifikasi Karakteristik Kepribadian Siswa Di Sekolah MTS Darussa’adah Menggunakan PHP MySQL Implementation of Naive Bayes Method for Classification of Student’s Personality Characteristics at MTS Darussa’adah School Using PHP Mysql,” 2022.
- [5] R. A. Anggraini, G. Widagdo, A. S. Budi, and M. Qomaruddin, “Penerapan Data Mining Classification untuk Data Blogger Menggunakan Metode Naive Bayes,” *Jurnal Sistem dan Teknologi Informasi (JUSTIN)*, vol. 7, no. 1, p. 47, Jan. 2019, doi: 10.26418/justin.v7i1.30211.
- [6] T. Y. Yanto and F. Maedjaja, “Implementasi Data Mining Untuk Menganalisis Pola Peminjaman Buku Perpustakaan Dengan Menggunakan Metode Apriori,” *Jurnal InFact Sains dan Komputer*, vol. 8, no. 02, 2023
- [7] L. Utari and Y. Ulfah, “Penerapan Metode Naive Bayes untuk Prediksi Minat Baca Berdasarkan Usia,” vol. 11, pp. 67–74, 2021, doi: 10.36350/jbs.v11i1.
- [8] E. Irfiani, Y. Kusnadi, S. Sunarti, and F. Handayanna, “Implementasi Data Mining dalam Mengklasifikasi Minat Baca Pada Perpustakaan Daerah Menggunakan Algoritma C4.5,” *JOINS (Journal of Information System)*, vol. 8, no. 2, pp. 106–114, Nov. 2023, doi: 10.33633/joins.v8i2.8004.
- [9] R. N. Ainum, N. Hidayat, and A. A. Soebroto, “Klasifikasi Buku Perpustakaan menggunakan Metode Naive Bayes,” 2022. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [10] R. N. Ainum, N. Hidayat, and A. A. Soebroto, “Klasifikasi Buku Perpustakaan menggunakan Metode Naive Bayes,” 2022. [Online]. Available: <http://j-ptiik.ub.ac.id>

# Referensi

- [11] H. Latipa Sari and I. Yati Beti, "KLIK: Kajian Ilmiah Informatika dan Komputer Penerapan Data Mining Dalam Pengelompokan Buku Yang Dipinjam Menggunakan Algoritma K-Means," *Media Online*, vol. 3, no. 6, pp. 925–933, 2023, doi: 10.30865/klik.v3i6.826.
- [12] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *JURIKOM (Jurnal Riset Komputer)*, vol. 8, no. 6, p. 219, Dec. 2021, doi: 10.30865/jurikom.v8i6.3655.
- [13] I. K. A. Wiraguna, E. Setyati, and E. Pramana, "Prediksi Anak Stunting Berdasarkan Kondisi Orang Tua Dengan Metode Support Vector Machine Dengan Study Kasus Di Kabupaten Tabanan-Bali," *SMATIKA JURNAL*, vol. 12, no. 01, pp. 47–54, Jun. 2022, doi: 10.32664/smatika.v12i01.662.
- [14] A. Setiawan, R. Febrio Waleska, M. Adji Purnama, and L. Efrizoni, "Sidomulyo barat, Sidomulyo Bar," 2024. [Online]. Available: <http://e-journal.stmiklombok.ac.id/index.php/jireISSN.2620-6900>
- [15] M. Riyyan, H. Firdaus, J. H. Ronggo Waluyo, T. Timur, and J. Barat, "PERBANDINGAN ALGORITME NAÏVE BAYES DAN KNN TERHADAP DATA PENERIMAAN BEASISWA (Studi Kasus Lembaga Beasiswa Baznas Jabar)," 2022. [Online]. Available: <http://e-journal.stmiklombok.ac.id/index.php/jireISSN.2620-6900>
- [16] A. Felicia Watratan, A. B. Puspita, D. Moeis, S. Informasi, and S. Profesional Makassar, "Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia," 2020. [Online]. Available: <http://journal.isas.or.id/index.php/JACOST>
- [17] I. M. A. A. D. Putra, I. M. G. Sunarya, and I. G. A. Gunadi, "Perbandingan Algoritma Naive Bayes Berbasis Feature Selection Gain Ratio dengan Naive Bayes Kovenisional dalam Prediksi Komplikasi Hipertensi," *JTIM : Jurnal Teknologi Informasi dan Multimedia*, vol. 6, no. 1, pp. 37–49, Apr. 2024, doi: 10.35746/jtim.v6i1.488.
- [18] F. Rizki, M. P. Kharisma Putra, M. A. Assuja, and F. Ariany, "Implementasi Deep Learning Lenet Dengan Augmentasi Data Pada Identifikasi Anggrek," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 4, no. 3, pp. 357–366, Sep. 2023, doi: 10.33365/jatika.v4i3.3652.
- [19] R. Darmawan et al., "KLASIFIKASI DIAGNOSA PENYAKIT TIROID MENGGUNAKAN METODE RANDOM FOREST," 2024. [Online]. Available: <http://e-journal.stmiklombok.ac.id/index.php/jireISSN.2620-6900>
- [20] A. P. J. Dwitama, "DETEKSI UJARAN KEBENCIAN PADA TWITTER BAHASA INDONESIA MENGGUNAKAN MACHINE LEARNING: REVIU LITERATUR," *Jurnal Sains, Nalar, dan Aplikasi Teknologi Informasi*, vol. 1, no. 1, Aug. 2021, doi: 10.20885/snati.v1i1.5.

