

## ANALISIS ALGORITMA K-NEAREST NEIGHBOR UNTUK PREDIKSI KELULUSAN MAHASISWA TEPAT WAKTU (STUDI KASUS: ISB ATMA LUHUR)

Irsanda Hardimas, Delpiah Wahyuningsih, Rezky Yuranda

Teknik Informatika, Institut Sains dan Bisnis Atma Luhur

Jln. Selindung Kec Gabek Pangkalpinang, Indonesia

2111500017@mahasiswa.atmaluhur.ac.id

### ABSTRAK

Ketepatan waktu lulus mahasiswa merupakan indikator vital bagi kualitas dan akreditasi perguruan tinggi. Namun, ISB Atma Luhur menghadapi kendala dalam mengidentifikasi mahasiswa yang berisiko terlambat lulus secara dini akibat banyaknya variabel akademik dan keterbatasan pemantauan manual. Penelitian ini bertujuan untuk menganalisis performa algoritma *K-Nearest Neighbor* (KNN) dalam memprediksi kelulusan mahasiswa tepat waktu sebagai solusi sistem peringatan dini. Metode penelitian meliputi prapemrosesan data, standardisasi *Z-Score*, pembagian data latih dan uji (80:20), serta optimasi parameter *K* melalui *hyperparameter tuning*. Hasil pengujian menunjukkan bahwa model mencapai performa optimal pada nilai *K* = 3 dengan tingkat akurasi sebesar 98,31%. Metrik evaluasi lainnya menghasilkan *precision* 0,98, *recall* 1,00, dan *F1-score* 0,99. Temuan ini membuktikan bahwa algoritma KNN sangat efektif dalam mengklasifikasikan status kelulusan mahasiswa, sehingga dapat digunakan oleh manajemen institusi untuk melakukan intervensi akademik yang lebih tepat sasaran bagi mahasiswa yang berisiko.

**Kata kunci :** *K-Nearest Neighbor, kelulusan mahasiswa, klasifikasi, data mining, prediksi, machine learning*

### 1. PENDAHULUAN

Pendidikan tinggi berperan krusial dalam mencetak sumber daya manusia kompeten, di mana kualitasnya diukur melalui standar akreditasi Badan Akreditasi Nasional Perguruan Tinggi (BAN-PT). Salah satu indikator utamanya adalah persentase kelulusan mahasiswa tepat waktu [1], [2].

Kelulusan tepat waktu bukan hanya mencerminkan keberhasilan mahasiswa dalam meraih gelar sarjana, tetapi juga menunjukkan efisiensi manajemen akademik suatu institusi [3], [4].

Namun, pada realitanya, tingkat kelulusan tepat waktu seringkali belum mencapai target yang diharapkan, yang jika dibiarkan dapat berdampak pada penurunan nilai akreditasi program studi [2].

Permasalahan utama di ISB Atma Luhur, sebagai institusi yang berfokus pada teknologi dan bisnis, adalah besarnya tantangan dalam memantau masa studi mahasiswa di program studi Sistem Informasi, Teknik Informatika, dan Bisnis Digital. Banyaknya parameter yang mempengaruhi kelulusan, mulai dari faktor demografis hingga capaian akademik di setiap semester, membuat proses pemantauan secara manual menjadi tidak efisien [5], [6].

Ketidakmampuan mendeteksi risiko keterlambatan lulus secara dini dapat menghambat upaya intervensi yang seharusnya dilakukan oleh pihak program studi [7], [8].

Sebagai penyelesaian masalah, penelitian ini memanfaatkan teknik *data mining* dan *machine learning* yang telah terbukti efektif dalam menggali pengetahuan dari *database* akademik besar untuk melakukan klasifikasi [9].

Algoritma *K-Nearest Neighbor* (KNN) dipilih karena kesederhanaan konsepnya yang berbasis

kemiripan pola data (*lazy learning*) [10], [11]. KNN telah sukses diimplementasikan di berbagai bidang, seperti deteksi penyakit [12] hingga klasifikasi objek fisik [13], karena ketangguhannya dalam menangani data multidimensi. Meskipun algoritma lain seperti C4.5 dan Naive Bayes sering digunakan [3], [5], [14], [15], KNN menawarkan fleksibilitas unik melalui pengaturan parameter nilai tetangga (*K*) dan jenis perhitungan jarak yang mampu memberikan hasil kompetitif dan stabil pada data mahasiswa jika dilakukan pengujian parameter yang tepat [6], [9], [16], [17].

Tujuan dari penelitian ini adalah melakukan analisis mendalam untuk menemukan model KNN yang paling optimal untuk konteks institusi ISB Atma Luhur [18].

Berbeda dengan penelitian sebelumnya yang hanya melakukan komparasi antar algoritma secara umum [4], [11], penelitian ini akan menganalisis pengaruh variasi nilai *K* terhadap tingkat akurasi, *precision*, dan *recall* pada dataset yang mencakup tiga program studi. Hasil analisis ini diharapkan dapat memberikan rekomendasi model prediksi terbaik bagi manajemen ISB Atma Luhur dalam meningkatkan persentase kelulusan tepat waktu melalui sistem peringatan dini yang akurat.

### 2. TINJAUAN PUSTAKA

#### 2.1. Data Mining dan Klasifikasi

Data mining merupakan proses penggalian pengetahuan dari dataset yang besar untuk menemukan pola atau informasi yang berguna [15].

Salah satu tugas utama dalam data mining adalah klasifikasi, yaitu proses menempatkan objek data ke dalam kategori atau kelas yang telah

ditentukan sebelumnya berdasarkan atribut yang dimiliki [6].

Dalam konteks pendidikan, teknik klasifikasi digunakan untuk memprediksi performa atau status kelulusan mahasiswa [14].

## 2.2. Algoritma K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) adalah algoritma pembelajaran berbasis contoh (*instance-based learning*) yang bekerja dengan cara mencari kedekatan antara data baru dengan data yang sudah ada (tetangga) dalam ruang fitur [10].

Prinsip utama KNN adalah objek yang memiliki karakteristik serupa akan cenderung berada di kelas yang sama [11].

Nilai K: Penentuan jumlah tetangga terdekat (nilai  $K$ ) sangat krusial; nilai  $K$  yang terlalu kecil dapat menyebabkan *overfitting*, sementara  $K$  yang terlalu besar dapat mengaburkan batas antar kelas [9].

Perhitungan Jarak: Biasanya menggunakan *Euclidean Distance* untuk mengukur kemiripan antar data [16].

## 2.3. Kelulusan Mahasiswa Tepat Waktu

Kelulusan tepat waktu didefinisikan sebagai penyelesaian masa studi sesuai dengan kurikulum yang ditetapkan (misalnya 8 semester untuk jenjang S1) [2].

Ketepatan waktu ini menjadi metrik utama bagi institusi seperti ISB Atma Luhur untuk menjaga standar akreditasi dan efisiensi manajemen akademik [1].

## 2.4. Penelitian Terkait

Penelitian mengenai prediksi kelulusan mahasiswa telah berkembang pesat dengan berbagai pendekatan algoritma. Sofiyanto menunjukkan bahwa KNN sangat efektif dalam mengidentifikasi pola kelulusan melalui klasifikasi biner [10].

Efektivitas KNN juga diperkuat oleh Mubarak yang membandingkannya dengan Naive Bayes menggunakan metodologi CRISP-DM, di mana KNN memberikan performa yang lebih stabil pada data akademik yang melibatkan durasi skripsi [9]. Selain itu, perbandingan antara KNN dan C4.5 dilakukan oleh Fakih pada data penerima beasiswa, yang membuktikan bahwa KNN memiliki akurasi kompetitif terhadap algoritma berbasis pohon keputusan [19].

Optimasi pada algoritma KNN juga menjadi fokus penting dalam literatur terbaru. Firmansyah melaporkan bahwa penggunaan teknik SMOTE untuk menangani ketidakseimbangan data (*imbalance class*) dapat meningkatkan performa KNN secara signifikan [18].

Di sisi lain, Yulihartati & Ramadhanu mengusulkan penggunaan *Principal Component Analysis* (PCA) untuk mereduksi dimensi data, yang terbukti mempercepat proses komputasi KNN tanpa mengurangi akurasi pada klasifikasi fitur [13].

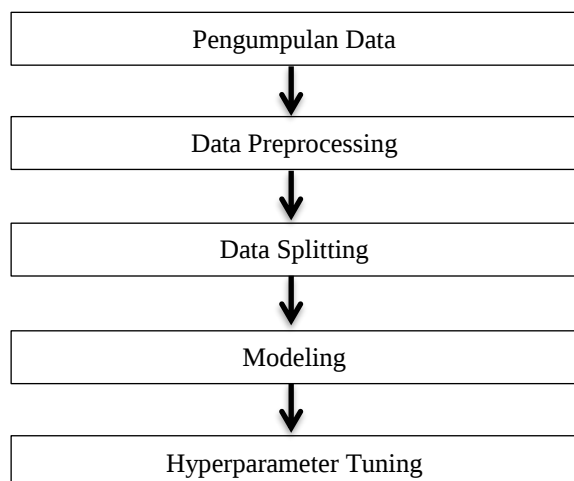
Selain di bidang akademik, ketangguhan KNN juga dibuktikan dalam klasifikasi medis untuk deteksi diabetes [20] dan analisis sentimen aplikasi berbasis AI [17].

Penelitian terdahulu menunjukkan bahwa meskipun algoritma lain seperti C4.5 sering memberikan akurasi tinggi pada data terstruktur [3], [6], KNN tetap unggul dalam menangani data yang memiliki ketergantungan fitur yang kuat dan pola kemiripan yang dinamis [11], [16].

Secara keseluruhan, beberapa studi perbandingan [1], [14] menekankan bahwa tidak ada satu algoritma yang selalu unggul di setiap dataset. Hal ini memperkuat alasan dilakukannya analisis mendalam terhadap algoritma KNN pada studi kasus spesifik di ISB Atma Luhur, guna menemukan parameter  $K$  yang paling optimal bagi karakteristik data mahasiswa di sana [8].

## 3. METODE PENELITIAN

Tahapan penelitian diimplementasikan langsung ke dalam alur kode Python dengan urutan sebagai berikut:



Gambar 1. Diagram Alur Penelitian

Dari gambar 1 menjelaskan sebagai berikut:

- Pengumpulan Data (*Data Loading*)** : Data akademik mahasiswa dari tiga program studi (Sistem Informasi, Teknik Informatika, dan Bisnis Digital) ISB Atma Luhur dimuat ke dalam dataframe Pandas melalui format .csv. Data ini memuat atribut-atribut prediktor dan kolom target kelas (Lulus Tepat Waktu atau Terlambat).
- Pra-pemrosesan dan Transformasi Data (*Data Preprocessing*)** : Pembersihan Data dimana Menghapus atribut yang tidak memiliki relevansi komputasi terhadap kelulusan, seperti Nama atau NIM, menggunakan fungsi `drop()`. Standardisasi (*Scaling*) dimana Algoritma KNN sangat sensitif terhadap skala data yang bervariasi (misalnya nilai umur puluhan, sedangkan nilai IPK hanya skala 0-4). Oleh karena itu, dilakukan standardisasi fitur menggunakan `StandardScaler` dari pustaka Sklearn agar semua fitur memiliki

- rentang nilai distribusi yang seimbang dengan rata-rata (mean) nol dan varians satu.
- Pembagian Dataset (Data Splitting): Dataset yang telah diskalakan dibagi menjadi dua bagian, yaitu data latih (training data) dan data uji (testing data). Pembagian ini menggunakan fungsi `train_test_split` (rasio 80% data latih dan 20% data uji) agar model dapat diuji secara objektif pada data yang belum pernah dilihat sebelumnya.
  - Pemodelan KNN (*KNN Classifier*): Data latih dimasukkan ke dalam model klasifikasi menggunakan modul *KNeighborsClassifier*. Proses ini menghitung kedekatan (jarak) antar titik data menggunakan metrik jarak (umumnya *Euclidean* atau *Minkowski* dengan parameter  $p=2$ ).
  - Pencarian Nilai  $K$  Optimal (*Hyperparameter Tuning*): Untuk menghindari tebakan nilai  $K$  secara acak, dilakukan perulangan (for-loop) pada rentang nilai  $K$  (misalnya  $K = 1$  hingga 40). Pada setiap perulangan, model akan menghitung error rate (tingkat kesalahan). Nilai  $K$  yang menghasilkan error rate paling rendah (atau akurasi tertinggi) pada plot visualisasi matriks dipilih sebagai parameter optimal model untuk mahasiswa ISB Atma Luhur.

### 3.1. Landasan KNN

Walaupun perhitungan dilakukan secara otomatis oleh pustaka Sklearn, secara matematis, modul *K-Neighbors Classifier* dengan parameter default mengukur kedekatan menggunakan jarak *Euclidean*. Formula dari jarak *Euclidean* adalah:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Keterangan

$d(x, y)$  : Jarak antara data mahasiswa yang diuji dengan data mahasiswa pada *training set*.

$n$  : Jumlah atribut/fitur yang digunakan.

$x_i$  : Nilai atribut ke-  $i$  pada data uji.

$y_i$  : Nilai atribut ke-  $i$  pada data latih.

Selain itu, perhitungan standarisasi fitur dari kelas *StandardScaler* (yang Anda gunakan pada langkah *preprocessing*) mengikuti rumus Z-Score:

$$z = \frac{x - \mu}{\sigma} \quad (2)$$

Keterangan:  $x$  adalah nilai asli,  $\mu$  adalah rata-rata fitur, dan  $\sigma$  adalah standar deviasi.

### 3.2. Evaluasi Kinerja Model

Evaluasi dilakukan dengan membandingkan hasil prediksi model terhadap data aktual pada data uji. Menggunakan modul `confusion_matrix` dan `classification_report`, evaluasi ini menghasilkan tiga metrik utama:

- Akurasi (*Accuracy*): Tingkat ketepatan model dalam memprediksi seluruh kelas secara benar.

- Presisi (*Precision*): Tingkat ketepatan model dalam memprediksi kelas target spesifik (misal: "Tepat Waktu").
- Recall: Tingkat keberhasilan model dalam menemukan semua sampel mahasiswa yang sebenarnya lulus tepat waktu.

## 4. HASIL DAN PEMBAHASAN

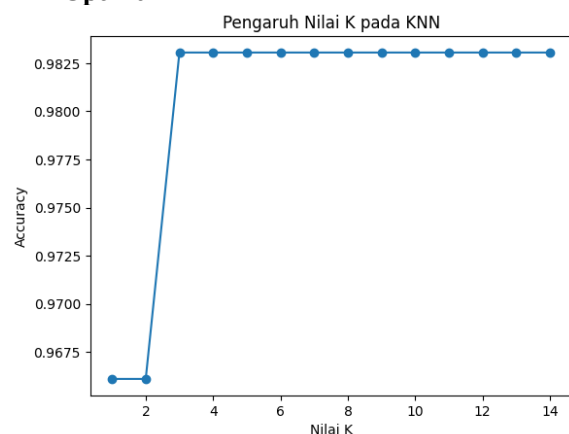
### 4.1. Hasil Penelitian

Penelitian ini menghasilkan sebuah model klasifikasi menggunakan algoritma K-Nearest Neighbor (KNN) untuk memprediksi kelulusan mahasiswa tepat waktu di ISB Atma Luhur. Dataset yang digunakan merupakan data akademik mahasiswa yang telah melalui tahapan preprocessing, meliputi pembersihan data, seleksi fitur, dan standarisasi menggunakan metode Z-Score.

Setelah proses preprocessing, dataset dibagi menjadi data latih (training data) sebesar 80% dan data uji (testing data) sebesar 20%. Model KNN kemudian dibangun menggunakan data latih dan diuji menggunakan data uji untuk mengetahui performa model dalam melakukan prediksi.

Selain itu, dilakukan pengujian terhadap beberapa nilai parameter  $K$  untuk menentukan nilai optimal yang memberikan performa terbaik. Hasil pengujian divisualisasikan dalam bentuk grafik error rate terhadap nilai  $K$ , yang menunjukkan adanya variasi tingkat kesalahan pada setiap nilai  $K$ .

### 4.2. Pengujian Model dan Penentuan Nilai K Optimal



Gambar 2. Pengaruh Nilai K terhadap Akurasi Model KNN

Gambar 2 menunjukkan pengaruh variasi nilai  $K$  terhadap akurasi model KNN. Pada  $K = 1$  dan  $K = 2$ , akurasi model masih relatif rendah yaitu sekitar 96,61%. Mulai dari  $K = 3$  hingga  $K = 14$ , akurasi meningkat dan stabil pada nilai 98,31%. Hal ini menunjukkan bahwa nilai  $K$  yang terlalu kecil menyebabkan model terlalu sensitif terhadap data latih (*overfitting*), sedangkan nilai  $K$  yang lebih besar memberikan keseimbangan yang lebih baik antara bias dan varians.

Berdasarkan hasil tersebut, nilai K optimal yang digunakan dalam penelitian ini adalah  $K = 3$ , karena telah memberikan akurasi maksimum dan stabil pada rentang nilai K selanjutnya.

#### 4.3. Hasil Evaluasi Model

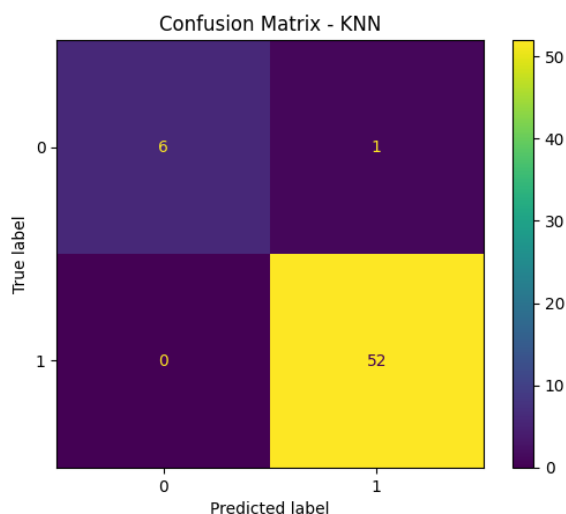
Model KNN dengan nilai  $K = 3$  kemudian dievaluasi menggunakan *confusion matrix* dan *classification report*.

Tabel 1. Classification Report Model KNN ( $K = 3$ )

| Kelas        | Precision | Recall | F1-score | Support |
|--------------|-----------|--------|----------|---------|
| 0            | 1.00      | 0.86   | 0.92     | 7       |
| 1            | 0.98      | 1.00   | 0.99     | 52      |
| Accuracy     |           |        | 0.98     | 59      |
| Macro Avg    | 0.99      | 0.93   | 0.96     | 59      |
| Weighted Avg | 0.98      | 0.98   | 0.98     | 59      |

Berdasarkan Tabel 2, model KNN dengan  $K = 3$  menghasilkan akurasi sebesar 98,31% pada data uji. Nilai *precision* untuk kelas 1 (tepat waktu) sebesar 0,98 dan *recall* sebesar 1,00, yang berarti model mampu mengidentifikasi seluruh mahasiswa yang lulus tepat waktu tanpa ada yang terlewat (*false negative* = 0). Nilai *F1-score* sebesar 0,99 menunjukkan keseimbangan yang sangat baik antara *precision* dan *recall*.

#### 4.4. Analisis Confusion Matrix



Gambar 3. Confusion Matrix Model KNN ( $K = 3$ )

Berdasarkan Gambar 3, diperoleh nilai:

- True Negative (TN) = 6
- False Positive (FP) = 1
- False Negative (FN) = 0
- True Positive (TP) = 52

Hasil ini menunjukkan bahwa hanya terdapat 1 kesalahan prediksi, yaitu mahasiswa yang diprediksi lulus tepat waktu namun sebenarnya tidak. Sementara itu, seluruh mahasiswa yang benar-benar lulus tepat waktu berhasil diprediksi dengan benar.

#### 4.5. Pembahasan

Berdasarkan hasil pengujian dan evaluasi, algoritma KNN terbukti memiliki performa yang sangat baik dalam memprediksi kelulusan mahasiswa tepat waktu. Tingginya nilai akurasi sebesar 98,31% menunjukkan bahwa model mampu melakukan klasifikasi dengan tingkat kesalahan yang sangat rendah.

Nilai *recall* sebesar 1,00 pada kelas “tepat waktu” menunjukkan bahwa model mampu mendeteksi seluruh mahasiswa yang benar-benar lulus tepat waktu tanpa kesalahan. Hal ini sangat penting dalam implementasi sistem prediksi karena tidak ada data penting yang terlewat. Selain itu, nilai *precision* sebesar 0,98 menunjukkan bahwa prediksi model juga memiliki tingkat ketepatan yang tinggi. Nilai *F1-score* yang mendekati 1 menunjukkan keseimbangan yang baik antara *precision* dan *recall*.

Hasil *confusion matrix* juga memperkuat bahwa model memiliki performa tinggi dengan hanya satu kesalahan prediksi (FP = 1). Tidak adanya *false negative* menunjukkan bahwa model sangat andal dalam mendeteksi mahasiswa yang lulus tepat waktu. Keberhasilan model ini dipengaruhi oleh beberapa faktor:

- Pemilihan nilai K yang optimal
- Proses prapemrosesan data yang baik
- Standardisasi data yang meningkatkan akurasi perhitungan jarak
- Kesesuaian algoritma KNN dengan karakteristik data akademik

Hasil penelitian ini membuktikan bahwa algoritma KNN sangat efektif untuk prediksi kelulusan mahasiswa tepat waktu pada data ISB Atma Luhur. Model ini berpotensi diimplementasikan sebagai sistem peringatan dini (*early warning system*) untuk membantu pihak program studi dalam:

- Mengidentifikasi mahasiswa berisiko terlambat lulus.
- Memberikan intervensi akademik lebih awal
- Meningkatkan persentase kelulusan tepat waktu

#### 5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian, algoritma K-Nearest Neighbor (KNN) terbukti efektif dalam memprediksi kelulusan mahasiswa tepat waktu di ISB Atma Luhur dengan nilai parameter optimal pada  $K = 3$  yang menghasilkan akurasi sebesar 98,31%, *precision* 0,98, *recall* 1,00, dan *F1-score* 0,99. Hasil ini menunjukkan bahwa model memiliki kemampuan klasifikasi yang sangat baik dengan tingkat kesalahan yang sangat rendah, ditunjukkan oleh *confusion matrix* dengan nilai TP = 52, TN = 6, FP = 1, dan FN = 0. Dengan demikian, model KNN berpotensi diimplementasikan sebagai sistem peringatan dini (*early warning system*) untuk membantu program studi dalam mengidentifikasi mahasiswa yang berisiko tidak lulus tepat waktu. Adapun saran untuk penelitian selanjutnya adalah menggunakan dataset yang lebih besar dan beragam, melakukan

perbandingan dengan algoritma lain seperti Decision Tree atau Naive Bayes, menerapkan teknik penanganan data tidak seimbang seperti SMOTE, serta mengembangkan sistem berbasis aplikasi agar hasil prediksi dapat dimanfaatkan secara real-time dalam pengambilan keputusan akademik.

#### DAFTAR PUSTAKA

- [1] A. Rahmayanti *et al.*, "Perbandingan Metode Algoritma C4.5 dan Naïve Bayes Untuk Memprediksi Kelulusan Mahasiswa," *Walisongo Journal of IT*, vol. 4, no. 1, 2022.
- [2] T. H. Hasibuan and D. Mahdiana, "Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Algoritma C4.5 Pada UIN Syarif Hidayatullah Jakarta," *SKANIKA*, vol. 6, no. 1, 2023.
- [3] N. Khasanah *et al.*, "Prediksi Kelulusan Mahasiswa Menggunakan Algoritma C4.5 dengan RapidMiner," *Indonesian Journal Computer Science*, vol. 4, no. 2, 2025.
- [4] H. Husain *et al.*, "Analisis Komparatif Algoritma Klasifikasi untuk Prediksi Kelulusan Tepat Waktu Mahasiswa," *Bulletin of Information Technology (BIT)*, vol. 7, no. 1, 2026.
- [5] M. R. Qisthiano *et al.*, "Pengembangan Model Untuk Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu dengan Metode Naïve Bayes," *Jurnal Media Informatika Budidarma*, vol. 5, no. 3, 2021.
- [6] D. A. R. A. Adha *et al.*, "Performansi Algoritma C4.5 untuk Prediksi Kelulusan Mahasiswa," *Jurnal Cakrawala Informasi*, vol. 3, no. 2, 2023.
- [7] C. Mulyadi and L. Sugiarto, "Penggunaan Algoritma Naïve Bayes untuk Prediksi Ketepatan Waktu Lulus Mahasiswa Diploma 3," *Teknomatika*, vol. 11, no. 1, 2021.
- [8] I. R. M. Tomu *et al.*, "Klasifikasi Ketepatan Waktu Lulus Mahasiswa Jurusan Matematika Universitas Pattimura Menggunakan Metode Naïve Bayes Classifier," *Proximal*, 2024.
- [9] R. Mubarak *et al.*, "Komparasi Performa Naive Bayes Gaussian dan K-NN Untuk Prediksi Kelulusan Mahasiswa dengan CRISP-DM," *KLIK*, vol. 4, no. 6, 2024.
- [10] N. Sofiyanto *et al.*, "Prediksi Kelulusan Mahasiswa Menggunakan Algoritma K-Nearest Neighbor (KNN) dengan Klasifikasi Biner," *SENATIB*, 2025.
- [11] N. A. W. Pertiwi and I. S. Wibisono, "Comparison of K-Nearest Neighbor, Naïve Bayes, and C4.5 Algorithms for Predicting Academic Stress Risk," *JAIC*, vol. 10, no. 1, 2026.
- [12] M. S. Fiqri and H. D. Bhakti, "Klasifikasi Potensi Penyakit Diabetes Mellitus Tipe II Pada Pasien Menggunakan Algoritme KNN," *JATI*, vol. 8, no. 4, 2024.
- [13] S. Yulihartati and A. Ramadhanu, "Implementasi Algoritma KNN, PCA Pada Klasifikasi Buah Pir Madu, Pir Hijau dan Apel Merah," *JATI*, vol. 9, no. 4, 2025.
- [14] M. B. Alfayyadh *et al.*, "Perbandingan Algoritma C4.5 Dan Naïve Bayes Dalam Machine Learning Untuk Klasifikasi Performa Pelajar," *JMS*, vol. 5, no. 2, 2025.
- [15] A. Zamsuri *et al.*, "Perbandingan Akurasi Algoritma C4.5 Dan Naïve Bayes Terhadap Klasifikasi Kelulusan Berdasarkan D
- [16] M. Riyyan and H. Firdaus, "Perbandingan Algoritme Naïve Bayes dan KNN Terhadap Data Penerimaan Beasiswa," *JIRE*, vol. 5, no. 1, 2022.
- [17] Fathoni *et al.*, "Komparasi Kinerja KNN Dan Naïve Bayes Untuk Klasifikasi Sentimen Pengguna Aplikasi DeepSeek AI," *JATI*, vol. 9, no. 4, 2025.
- [18] A. S. Firmansyah *et al.*, "Optimasi K-Nearest Neighbor Menggunakan Algoritma SMOTE Untuk Mengatasi Imbalance Class," *JATI*, vol. 7, no. 6, 2023.
- [19] A. Fakhri *et al.*, "Perbandingan Akurasi Algoritma C4.5 dan K-NN Untuk Prediksi Kelulusan Mahasiswa Penerima Beasiswa," *Jurnal Komputer Antartika*, vol. 3, no. 1, 2025.
- [20] M. S. Fiqri and H. D. Bhakti, "Klasifikasi Potensi Penyakit Diabetes Mellitus Tipe II Pada Pasien Menggunakan Algoritme KNN," *JATI*, vol. 8, no. 4, 2024.