

Prediksi Tingkat Partisipasi Pemilu Menggunakan Algoritma Naive Bayes Berdasarkan Data Indeks Desa Membangun di Kabupaten Sidoarjo

Oleh:

Husnul Abidin

Arif Senja Fitrani, S.Kom., M.Kom

Progam Studi Informatika

Universitas Muhammadiyah Sidoarjo

2025



Pendahuluan

Pemilihan umum merupakan pilar utama demokrasi yang mencerminkan tingkat partisipasi politik masyarakat, dan dalam konteks pedesaan seperti Kabupaten Sidoarjo yang memiliki karakteristik sosial, ekonomi, dan infrastruktur yang beragam, penting untuk memahami faktor-faktor yang memengaruhi partisipasi tersebut. Penelitian ini menggunakan data Indeks Desa Membangun (IDM) sebagai variabel prediktor untuk mengukur sejauh mana aspek-aspek ketahanan sosial, ekonomi, dan ekologi desa dapat memengaruhi tingkat partisipasi pemilu, dengan mengimplementasikan algoritma Naïve Bayes berbasis klasifikasi probabilistik melalui platform Google Colab. Berbeda dari penelitian terdahulu yang hanya menggunakan variabel demografis dan wilayah umum, studi ini fokus pada desa-desa di Sidoarjo secara khusus, sehingga menghasilkan model prediksi yang lebih kontekstual dan akurat. Pendekatan ini diharapkan menjadi alternatif baru dalam mendukung strategi peningkatan partisipasi politik masyarakat perdesaan secara komprehensif dan berbasis teknologi terbuka..



Rumusan Masalah

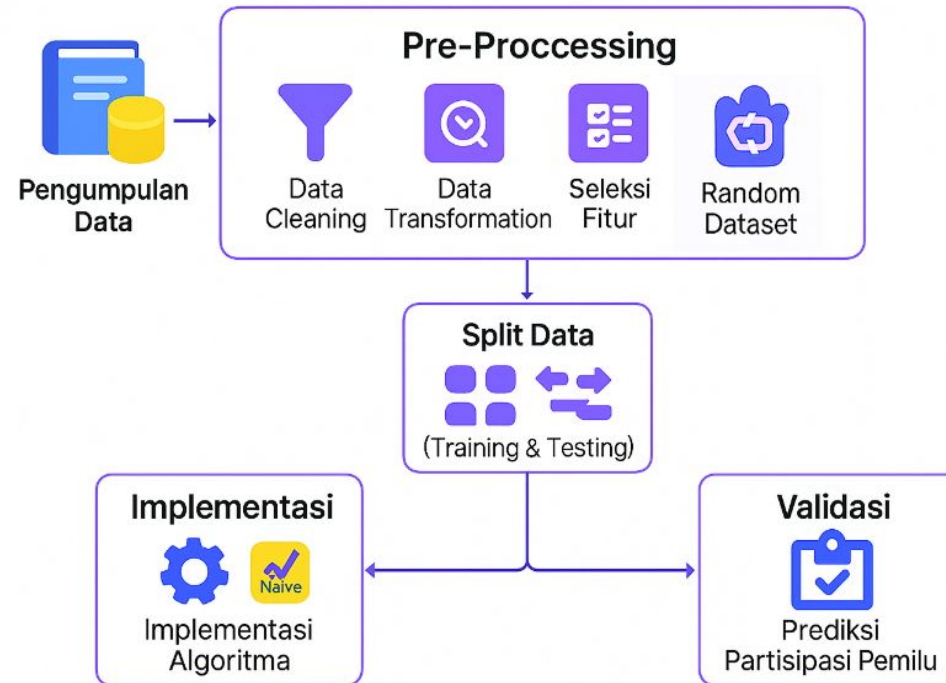
- Bagaimana tingkat partisipasi masyarakat Desa Membangun sebagai pemilih dalam Pemilihan Umum di Kabupaten Sidoarjo?
- Bagaimana implementasi algoritma Naïve Bayes menggunakan Google Colab untuk memprediksi tingkat partisipasi pemilu berdasarkan data Indeks Desa Membangun?

Tujuan

- Menganalisis dan mengukur tingkat partisipasi masyarakat desa dalam Pemilihan Umum berdasarkan data partisipasi pemilu sebelumnya.
- Mengimplementasikan algoritma Naïve Bayes pada Google Colab untuk memprediksi tingkat partisipasi pemilu berdasarkan data Indeks Desa Membangun di Kabupaten Sidoarjo

Metode Penelitian

Pada penelitian ini, proses dimulai dari studi literatur, dilanjutkan dengan pengumpulan data, tahap pre-processing, pemrosesan data menggunakan algoritma Naïve Bayes, hingga tahap validasi model. Penelitian ini dilakukan dengan menggunakan IDM dan data partisipasi pemilu tahun 2024 di Kabupaten Sidoarjo. Dataset yang diperoleh telah diintegrasikan dalam satu file dan mencakup 48 atribut. Variabel target dalam penelitian ini adalah tingkat partisipasi pemilu, yang akan diklasifikasikan menggunakan algoritma Naïve Bayes



Metode Penelitian

Pre-Processing

Tahapan ini bertujuan menyiapkan data IDM dan partisipasi pemilu agar siap diklasifikasi menggunakan algoritma Naïve Bayes.

Data Cleaning

Membersihkan data dari nilai kosong, teks tidak konsisten, dan atribut tak relevan. Nilai kosong diganti dengan median/modus, teks seperti "Ada/Tidak Ada" diubah menjadi angka, dan kolom seragam dihapus.

Data Transformation

Mengubah data kategorikal menjadi numerik, misalnya "Ada" menjadi 1 dan "Tidak Ada" menjadi 0, agar sesuai dengan kebutuhan algoritma.

Seleksi Fitur

Dataset awal terdiri dari 48 atribut prediktor dari data IDM. Untuk menyederhanakan klasifikasi dan meningkatkan akurasi, dilakukan seleksi fitur dengan menghapus atribut yang kurang berpengaruh, seperti Utilitas Dasar, Produksi Desa, Pendukung Ekonomi, Pengelolaan Lingkungan, Penanggulangan Bencana, dan Tata Kelola Keuangan Desa. Hasilnya, dipilih 27 atribut dari 7 kelompok indikator yang relevan, serta 1 atribut target yaitu tingkat partisipasi pemilu.

Metode Penelitian

Random Dataset

Data diacak untuk memastikan distribusi merata dan mencegah model terpengaruh urutan data sebelum dibagi menjadi data latih dan data uji.

Split Data

Data dibagi 70% untuk pelatihan dan 30% untuk pengujian, guna melatih dan menguji performa model Naïve Bayes dalam memprediksi partisipasi pemilu.

Process

Proses dimulai dari penggabungan data IDM dan partisipasi pemilu, dilanjutkan pre-processing, lalu klasifikasi dengan Naïve Bayes. Model menghitung probabilitas tiap kelas berdasarkan fitur, dan hasil prediksi diuji untuk mengukur kinerja menggunakan metrik evaluasi.

Output

Model menghasilkan prediksi tingkat partisipasi pemilu desa dalam dua kategori: “tinggi” dan “rendah”, berdasarkan indikator IDM.

Analisis / Evaluasi

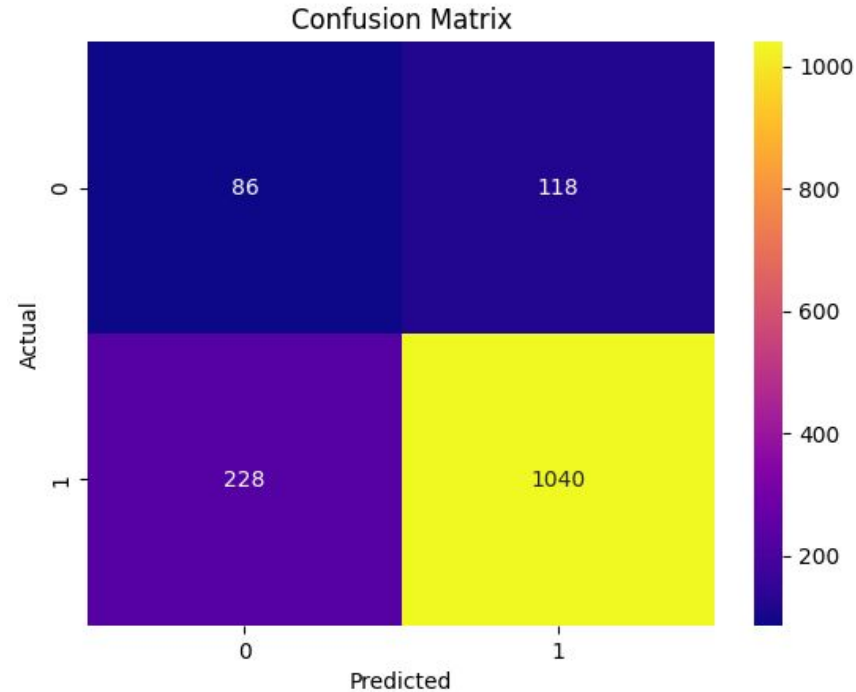
Evaluasi dilakukan dengan metrik akurasi, presisi, recall, dan F1-score untuk menilai seberapa baik model memprediksi partisipasi pemilu dan mengidentifikasi kelebihan serta kekurangannya.

Hasil & Pembahasan

| | Precisi on | Recall | F1-Score | Support |
|---------------------|---------------|--------|-------------|-------------|
| 0 | 0.27 | 0.42 | 0.33 | 204 |
| 1 | 0.90 | 0.82 | 0.86 | 1268 |
| Accuracy | | | 0.76 | 1472 |
| Macro Avg | 0.59 | 0.62 | 0.59 | 1472 |
| Weighted Avg | 0.81 | 0.76 | 0.78 | 1472 |

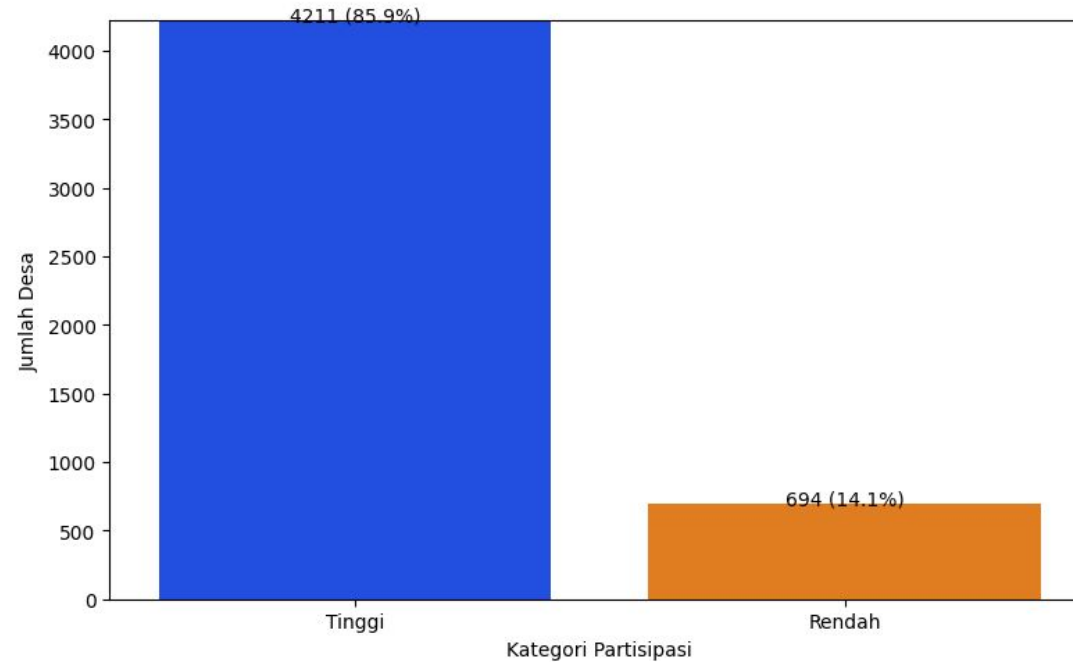
Pada Tabel 3 merupakan **Classification Report** yang menunjukkan bahwa performa model sangat baik dalam mengklasifikasikan desa dengan partisipasi tinggi, dengan nilai precision, recall, dan F1-score masing-masing 0.81, 0.82, dan 0.81. Namun, model mengalami kesulitan besar saat mengenali partisipasi rendah, terbukti dari nilai precision dan recall yang hanya sekitar 0.30. Hal ini mengindikasikan bahwa model sering salah mengklasifikasikan desa yang sebenarnya memiliki partisipasi rendah sebagai tinggi [12]. Performa buruk pada kelas "Rendah" kemungkinan besar disebabkan oleh ketimpangan jumlah data antara kelas tinggi (2228 data) dan rendah (614 data), yang membuat model lebih condong mempelajari pola dari kelas mayoritas.

Hasil & Pembahasan



Confusion Matrix Pada Gambar 2 menampilkan bahwa Model Naïve Bayes mengklasifikasikan 1.040 data partisipasi tinggi dan 86 data partisipasi rendah dengan benar. Namun, terdapat kesalahan pada 118 data rendah yang diprediksi tinggi dan 228 data tinggi yang diprediksi rendah. Hasil ini menunjukkan model cukup mengenali kelas mayoritas, namun masih perlu perbaikan dalam membedakan kelas partisipasi rendah.

Hasil & Pembahasan



Presentase Tingkat Partisipasi pada Gambar diatas menunjukkan bahwa 85,9% desa tergolong partisipasi tinggi (4211 desa), dan 14,1% partisipasi rendah (694 desa). Ketidakseimbangan ini umum dalam data sosial dan dapat memengaruhi hasil klasifikasi. Penyeimbangan data seperti resampling atau oversampling dapat dipertimbangkan untuk penelitian lanjutan agar model dapat belajar lebih optimal dari kedua kelas.

Hasil & Pembahasan

| No | Desa | Kecamatan | Partisipasi (%) |
|-----|----------------|-------------|-----------------|
| 1 | Krembung | Krembung | 95.97 |
| 2 | Mojorangagung | Wonoayu | 94.96 |
| 3 | Gempolklutuk | Tarik | 93.64 |
| 4 | Sumokembangsri | Balongbendo | 93.61 |
| 5 | Plumbon | Porong | 92.72 |
| ... | ... | ... | ... |
| 6 | Sawotratap | Gedangan | 72.67 |
| 7 | Gelang | Tulangan | 72.31 |
| 8 | Betro | Sedati | 71.89 |
| 9 | Wadungasri | Waru | 68.03 |
| 10 | Tambakoso | Waru | 53.36 |

Tabel 4 menampilkan persentase partisipasi pemilih di berbagai desa di Kabupaten Sidoarjo. Desa Krembung di Kecamatan Krembung mencatat partisipasi tertinggi sebesar 95,97%, diikuti Mojorangagung di Kecamatan Wonoayu (94,96%) dan Gempolklutuk di Kecamatan Tarik (93,64%). Ketiganya menunjukkan keterlibatan masyarakat yang tinggi dalam pemilu. Sebaliknya, partisipasi terendah terdapat di Desa Tambakoso di Kecamatan Waru dengan 53,36%, diikuti Wadungasri (68,03%) dan Betro di Kecamatan Sedati (71,89%).

Kesimpulan

Berdasarkan hasil analisis, mayoritas desa di Kabupaten Sidoarjo termasuk dalam kategori partisipasi tinggi, yakni sebesar 85,9%, sementara hanya 14,1% desa yang tergolong dalam partisipasi rendah. Ketimpangan distribusi ini menyebabkan model cenderung lebih akurat dalam mengklasifikasikan desa dengan partisipasi tinggi, sebagaimana terlihat pada nilai precision dan recall masing-masing 0.90 dan 0.82, dibandingkan dengan desa berpartisipasi rendah yang hanya mencapai precision 0.27 dan recall 0.42. Akurasi model secara keseluruhan tercatat sebesar 76%. Hasil confusion matrix juga menunjukkan bahwa model lebih sering salah dalam memprediksi partisipasi rendah sebagai tinggi, yang merupakan konsekuensi dari dominasi kelas mayoritas dalam data latih. Selain itu, analisis pada perbedaan indikator antara desa dengan partisipasi tertinggi dan terendah menunjukkan bahwa perbedaan signifikan hanya terdapat pada beberapa aspek, seperti akses pendidikan menengah dan layanan kesehatan dasar, sementara sebagian besar indikator lainnya memiliki selisih nol, bahkan beberapa bernilai negatif. Hal ini mengindikasikan bahwa tingkat partisipasi pemilu tidak hanya ditentukan oleh faktor layanan publik, tetapi juga oleh aspek sosial, budaya, serta kesadaran politik masyarakat. Oleh karena itu, pendekatan penyeimbangan data serta analisis fitur yang lebih mendalam menjadi penting untuk diterapkan pada penelitian selanjutnya guna meningkatkan performa model terhadap kelas minoritas.

Referensi

- [1] A. Sentimen, A. Pemilu, and A. N. Bayes, "SWADHARMA (JEIS)," vol. 05, pp. 131–139, 2025.
- [2] Y. Raharja, A. S. Fitriani, and R. Dijaya, "Klasifikasi Tingkat Partisipasi Pemilu Berdasarkan Sektor Industri Menggunakan Algoritma Naïve Bayes," J. Tekinkom (Teknik Inf. dan Komputer), vol. 7, no. 1, pp. 135–143, 2024, doi: 10.37600/tekinkom.v7i1.1204.
- [3] D. E. Safitri and A. S. Fitriani, "Implementasi Metode Klasifikasi Dengan Algoritma Support Vector Machine Kernel Gaussian Rbf Untuk Prediksi Partisipasi Pemilu Terhadap Demografi Kota Surabaya," Indones. J. Bus. Intell., vol. 5, no. 1, p. 36, 2022, doi: 10.21927/ijubi.v5i1.2259.
- [4] B. A. B. Iii, "digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id," pp. 66–90.
- [5] T. dkk Madjid, "Peringkat Indeks Desa Membangun," p. 871, 2020.
- [6] H. Rusli, "PREDIKSI HASIL PEMILIHAN UMUM BERDASARKAN DATA MEDIA SOSIAL MENGGUNAKAN TEKNIK DATA MINING NAIVE BAYES," 2025. [Online]. Available: <http://jurnal.goretanpena.com/index.php/JSSR>
- [7] A. S. Fitriani, "JTAM (Jurnal Teori dan Aplikasi Matematika) Penerapan Data Mining Menggunakan Metode Klasifikasi Naïve Bayes untuk Memprediksi Partisipasi Pemilihan Gubernur," vol. 3, no. 2, pp. 98–104, 2019, doi: 10.31764/jtam.v3i2.995.

Referensi

- [8] M. N. Zarti, E. Sahputra,) ; Anisya Sonita, and Y. Apridiansyah, "Application Of Data Mining Using The Naïve Bayes Classification Method To Predict Public Interest Participation In The 2024 Elections Penerapan Data Mining Menggunakan Metode Klasifikasi Naïve Bayes Untuk Memprediksi Partisipasi Minat Masyarakat Pada Pemilu 2024," J. Kom., vol. 3, no. 1, pp. 105–114, doi: 10.53697/jkomitek.v3i1.
- [9] S. Pokhrel, "No TitleEΛENH," Ayan, vol. 15, no. 1, pp. 37–48, 2024.
- [10] D. Syahputra, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Menentukan Produk Laptop Terlaris," 2022.
- [11] A. W. Anggraeni, A. S. Fitrani, and A. Eviyanti, "Penerapan Algoritma Support Vector Machine untuk Memprediksi Tingkat Partisipasi Pemilu terhadap Kualitas Pendidikan," Edumatic J. Pendidik. Inform., vol. 8, no. 1, pp. 21–27, 2024, doi: 10.29408/edumatic.v8i1.24838.
- [13] D. M. Chulloh, A. S. Fitrani, I. R. Indra Astutik, and A. Eviyanti, "Uji Akurasi K-Means dalam Prediksi Partisipasi Pemilu pada Demografi Wilayah Kabupaten Pasuruan," Jutisi : Jurnal Ilmiah Teknik Informatika dan Sistem Informasi, vol. 13, no. 1, p. 201, 2024, doi: 10.35889/jutisi.v13i1.1753.
- [14] D. Pilendia, S. Muhammadiyah, and S. Penuh, "PEMANFAATAN ADOBE FLASH SEBAGAI DASAR PENGEMBANGAN BAHAN AJAR FISIKA : STUDI LITERATUR," 2020. [Online]. Available: <http://ejournal.stkip-mmb.ac.id/index.php/pgsd/login>
- [15] W. Kurniawati, A. S. Fitrani, S. Aji, and S. Suprianto, "Analisis Kehadiran Pemilih di Tempat Pemungutan Suara Pada Pemilihan Presiden dengan Menggunakan Algoritma C4.5 (Studi Kasus di Desa Wonokasian Sidoarjo)," Journal of Electrical Engineering, vol. 1, no. 1, p. 18, 2024, doi: 10.47134/jte.v1i1.2475.
- [16] J. Riset and S. Informasi, "ANALISIS POLA KEHADIRAN MAHASISWA MENGGUNAKAN," vol. 2, no. 1, pp. 60–66, 2025.

Referensi

- [15] I. M. Widiarta, Y. Mulyanto, and A. Sutrianto, "Rancang Bangun Sistem Informasi Inventory Menggunakan Metode Agile Software Development (Studi Kasus Toko Nada)," Digital Transformation Technology (Digitech), vol. 3, no. Maret, p. 20, 2023.
- [16] M. A. Afandi and N. L. Azizah, "Digital Based Posyandu Service Design in Ganggang Panjang Village," Jul. 05, 2024. doi: 10.21070/ups.4924.
- [17] P. A. T. Subur et al., "Implementation of Agile Software Development in the Design of Management Information Systems Inventory and Finance at," 2025.

