

Mengukur Tingkat Partisipasi Pemilu Menggunakan Algoritma Naïve Bayes (Studi Kasus : Kabupaten / Kota Matraman, Jawa Timur)

Oleh:

Nando Adinata Santoso,

Arif Senja Fitriani

Program Studi Teknik Informatika
Universitas Muhammadiyah Sidoarjo

Maret, 2025



Pendahuluan

Diantara penentu dalam keberhasilan demokrasi merupakan partisipasi pemilih dalam pemilihan umum. Pemilihan umum yang berkualitas tinggi akan menjamin pesta demokrasi yang sehat di mana setiap pemilih menerima perlakuan yang sama dan dihormati, dan setiap peserta pemilu memiliki kesempatan yang adil untuk bersaing. [1].

Menurut Ningsih [2], Partisipan dalam pemilihan umum menunjukkan keterlibatan dan peran mereka dalam kegiatan. Menurut Pasal 19 Ayat (1 dan 2) UU Nomor 10 Tahun 2008, pemilih adalah warga negara Indonesia yang terdaftar dalam daftar pemilih dan memiliki hak untuk memilih calon pasangan politik yang mereka pilih.

Jumlah konstituen yang tinggi menunjukkan betapa semangatnya masyarakat untuk menggunakan hak suara mereka untuk memilih orang yang akan memimpin pemerintah [3]. Namun, tingkat partisipasi pemilih di Indonesia terus berubah-ubah dan dipengaruhi oleh sejumlah variabel, termasuk mobilisasi politik, karakteristik pemilih, dan kondisi ekonomi dan sosial masyarakat [4].

Setiap tahun, Badan Pusat Statistik (BPS) Kabupaten/Kota Matraman (yang mencakup wilayah Ngawi, Madiun, Magetan, Nganjuk, Pacitan, dan Ponorogo) mengeluarkan publikasi data yang mencakup data tentang semua elemen yang ada di setiap desa yang termasuk dalam kecamatan Matraman. Publikasi tersebut membahas aspek demografi. Bidang keilmuan data mining dengan metode Naive Bayes dapat dimanfaatkan untuk menghubungkan data aspek demografi dari publikasi data dari Badan Pusat Statistik (BPS), serta data demografis dan partisipasi masyarakat dalam pemilu[5].

Pendahuluan

Teorema Bayes, yang menyatakan bahwa nilai antar variabel tidak dipengaruhi oleh nilai output, adalah dasar algoritma klasifikasi. Dalam keadaan seperti ini, dianggap bahwa kehadiran atau ketiadaan satu variabel tidak dikaitkan dengan kehadiran atau ketiadaan variabel lainnya [6].

Penelitian ini bertujuan untuk mengukur menggunakan algoritma Naive Bayes untuk menentukan tingkat partisipasi pemilih dalam pemilihan umum; algoritma ini dapat mengidentifikasi pola hubungan antara atribut demografi pemilih dan partisipasi mereka dalam pemilihan. Hasil penelitian sebelumnya menunjukkan bahwa Algoritma Naive Bayes sangat akurat dalam memprediksi partisipasi pemilih [7].

Selain itu, permasalahan dan fenomena yang terjadi partisipasi politik masyarakat dalam pemilihan kepala kampung juga menunjukkan bahwa faktor-faktor demografi, seperti tingkat pendidikan dan pendapatan, berpengaruh terhadap partisipasi masyarakat dalam pemilihan [8].

Dalam penelitian sebelumnya, Yusuf Raharja menggunakan algoritma naive bayes dan dua model dataset untuk mengklasifikasikan tingkat partisipasi pemilu berdasarkan industri. Peneliti menguji semua atribut prediktor terhadap atribut target, meskipun ada korelasi yang rendah antara atribut prediktor dan atribut target. Hasil penelitian menunjukkan akurasi sebesar 67% [9].



Pertanyaan Penelitian (Rumusan Masalah)

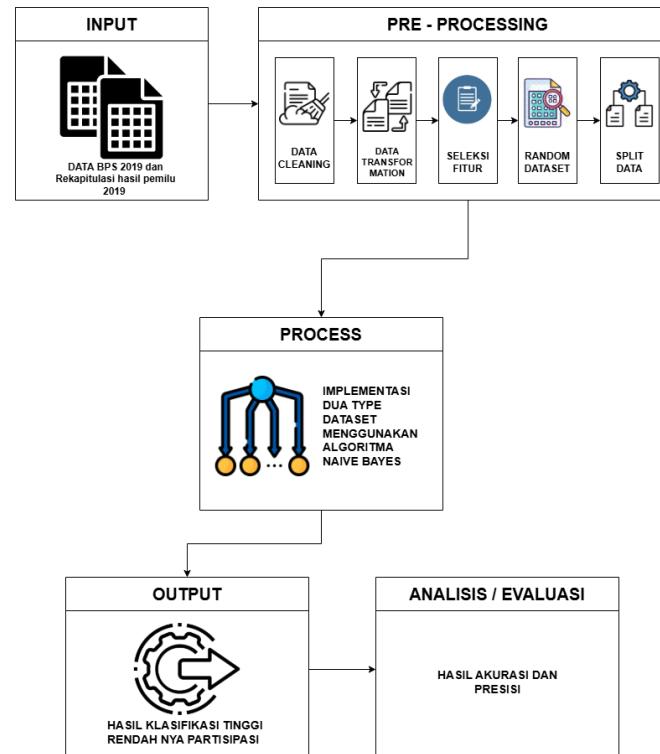
1. Bagaimana tingkat partisipasi pemilih dalam pemilihan di Kabupaten / Kota Matraman, Jawa Timur?
2. Bagaimana kinerja algoritma Naïve Bayes dalam memprediksi partisipasi pemilih berdasarkan faktor-faktor yang diidentifikasi?



Metode

Studi ini melakukan penelitian eksperimental. Penelitian ini bertujuan untuk memprediksi partisipasi pemilihan dengan menggunakan data demografi dari Badan Pusat Statistik (BPS) dari kota dan kabupaten Matraman seperti Ngawi, Madiun, Magetan, Nganjuk, Pacitan, dan Ponorogo. Untuk mengetahui tingkat partisipasi masyarakat dalam pemilu, Data akan diproses dengan menggunakan teknik klasifikasi Naive Bayes. Data awal mencakup lebih dari 1000 instance dan mencakup 85 fitur yang dirangkum dari data BPS[10].

Gambar Alur Penelitian



Metode

2.1 Input

Data dikumpulkan pada tahap awal penelitian ini. Penelitian ini menggunakan data publikasi data tahun 2019 dari beberapa Kota dan Kabupaten di wilayah Mataraman dan data rekapitulasi hasil pemilu tahun 2019. Publikasi data ini dapat diperoleh secara langsung dari situs resmi Badan Pusat Statistik (BPS). Data terdiri dari 85 atribut prediktor serta satu atribut target

2.2 Pre-Processing

Pada tahap ini melakukan *pre- procesing* data, berikut penjelasannya.

Data Cleaning

Data bersih adalah proses menemukan kesalahan seperti inkonsistensi, duplikat, dan data tidak lengkap. Kemudian, keputusan dibuat tentang data, seperti menghapus data yang tidak sesuai atau memperbaiknya[11].

Data Transformation

Data *trasformation* di fase ini, data akan disesuaikan menjadi format yang relevan untuk data mining. Data bakal dikelola menjadi format numerik[12].

Seleksi Fitur

Di dalam data BPS terdapat 11 kelompok atribut prediktor yang terdiri dari 85 atribut. Sehingga di dalam tahap seleksi fitur ini akan dilakukan drop beberapa atribut agar dapat menentukan fitur fitur yang relevan untuk model klasifikasi. Oleh karena itu akan dilakukan drop pada 5 kelompok atribut (Kemudahan Mencapai Sarana Pendidikan dan Kesehatan Terdekat, Bencana Alam, Infrastruktur, Lembaga Keuangan, dan Sarana Perdagangan). Jadi akan tersisa 7 kelompok atribut prediktor dan satu atribut target dalam 39 atribut

Metode

Random dataset

Untuk mengacak dan dataset random digunakan untuk memaksimalkan keterwakilan bobot pada setiap atribut dari setiap baris data.

Split data

Data yang telah di proses selanjutnya dibagi 75 % sebagai data *training* dan 25% sebagai data *testing*.

2.3 Process

Setelah tahap preprocessing, process dilakukan. Pada tahap ini, dataset akan diuji menggunakan metode klasifikasi menggunakan algoritma naïve Bayes[13]. Data dibagi dengan rasio 75:25. Ada 1028 baris data dalam data pelatihan dan 343 baris data dalam data pengujian.

2.4 Output

Setelah tahap proses selesai, tahap output menampilkan hasil pengujian menggunakan algoritma Naive Bayes yang menunjukkan tingkat partisipasi masyarakat dalam pemilu yang tinggi atau rendah[14].

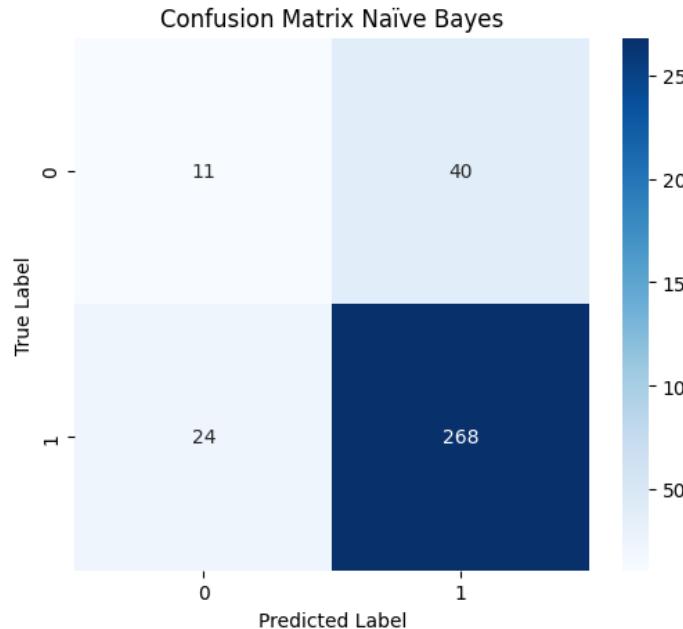
2.5 Analisis / Evaluasi

Evaluasi adalah tahap terakhir dalam penelitian ini. Evaluasi model dilakukan untuk mengukur performa model menggunakan berbagai *matrix* evaluasi seperti akurasi, presisi, recal dan F1-score[15].

Hasil dan Pembahasan

Penelitian ini memanfaatkan 1371 instance data dengan 39 fitur atau atribut, di mana atribut target diklasifikasikan ke dalam dua kategori: kelas 1 menunjukkan partisipasi tinggi, sedangkan kelas 0 merepresentasikan partisipasi rendah. Implementasi dilakukan menggunakan algoritma Naïve Bayes dengan Bahasa pemrograman python.

3.1 Akurasi dan Presisi



menampilkan hasil confusion matrix dari model Naive Bayes. Dari data yang dianalisis, sebanyak 24 data salah diklasifikasikan sebagai partisipasi rendah, sementara 268 data berhasil diklasifikasikan dengan benar sebagai partisipasi tinggi. Selain itu, terdapat 11 data yang diklasifikasikan dengan benar sebagai partisipasi rendah, sedangkan 40 data lainnya salah diklasifikasikan sebagai partisipasi tinggi.

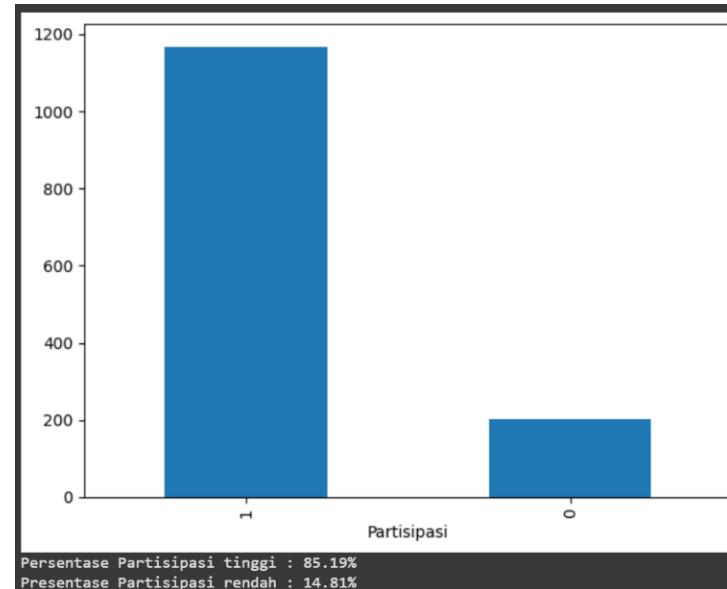
Hasil dan Pembahasan

	precision	recall	f1 - score	support
0	0.31	0.22	0.26	51
1	0.87	0.92	0.89	292
accuracy			0.81	343
macro avg	0.59	0.57	0.57	343
weighted avg	0.79	0.81	0.80	343

Tabel Classification Report

Tabel diatas menunjukkan kinerja yang sangat baik untuk kelas 1, dengan precision 0.87, recall 0.92, dan f1-score 0.89. Artinya, model cukup baik dalam mengklasifikasikan kelas 1 dan jarang membuat kesalahan. Sedangkan model sangat buruk dalam mengklasifikasikan kelas 0, dengan precision 0.31, recall 0.22, dan f1-score 0.26. Ini menunjukkan bahwa model gagal mengenali kelas 0 secara efektif dan sering salah mengklasifikasikan kelas lain sebagai kelas 0. Perbedaan besar antara jumlah sampel kelas 1 (292) dan kelas 0 (51) menunjukkan adanya ***ketidakseimbangan kelas***. Model lebih banyak berfokus pada kelas 1 karena kelas tersebut jauh lebih sering muncul dalam dataset. Ketidakseimbangan ini bisa menjadi alasan mengapa kinerja model pada kelas 0 sangat buruk. Secara keseluruhan, model memiliki akurasi yang cukup baik (81%)

3.2 Presentase Tingkat Partisipasi



Gambar Presentase partisipasi

Gambar diatas menunjukkan dari 1371 baris data yang telah dikumpulkan dari tujuh Kabupaten / Kota di wilayah Matraman memberikan kesimpulan, bahwasannya data menunjukkan angka presentase 85,19% atau 1168 data adalah partisipasi tinggi, sedangkan presentase 14,81% atau 203 data adalah partisipasi rendah. Hasil ini mengindikasikan bahwa mayoritas wilayah Kabupaten / Kota Matraman memiliki tingkat partisipasi yang tinggi. Hal ini memberikan gambaran umum mengenai pola partisipasi di wilayah Kabupaten / Kota Matraman dan dapat menjadi dasar untuk analisis lebih lanjut

Hasil dan Pembahasan

3.3 korelasi atribut

index	Attribute 1	Attribute 2	Correlation
0	Perempuan	JumlahPenduduk	0.988351469122888
1	JumlahPenduduk	Perempuan	0.988351469122888
2	Islam	JumlahPenduduk	0.9053623068286142
3	JumlahPenduduk	Islam	0.9053623068286142
4	Islam	Perempuan	0.8998681809088424
5	Perempuan	Islam	0.8998681809088424
6	JumlahPenduduk	LakiLaki	0.8881724698351245
7	LakiLaki	JumlahPenduduk	0.8881724698351245

Gambar Korelasi Atribut

Gambar diatas menunjukkan hasil dari tujuh korelasi tertinggi dari 39 atribut ini menunjukkan bahwa jumlah total penduduk memiliki keterkaitan yang sangat kuat dengan jumlah laki – laki dan perempuan dalam populasi, serta dengan jumlah pemeluk agama islam. Korelasi yang tinggi menunjukkan bahwa atribut – atribut ini saling berhubungan dalam membentuk komposisi demografi. Selain itu, hubungan yang kuat diantara variabel – variabel ini menunjukkan bahwa mereka dapat berfungsi sebagai variabel prediktor yang signifikan dalam mempengaruhi tingkat partisipasi pemilu. Analisis ini memberikan wawasan penting dalam memahami faktor – faktor yang mempengaruhi keterlibatan masyarakat dalam pemilu.



Temuan Penting Penelitian

Metode ini menghasilkan Tingkat akurasi yang cukup tinggi, mencapai 81,34%. Berdasarkan evaluasi yang dilakukan, sistem menunjukkan performa yang baik dalam mengenali dan mengklasifikasikan kategori partisipasi tinggi dengan akurasi yang memadai, dan juga dari korelasi atribut menjelaskan hubungan yang kuat diantara variabel – variabel ini menunjukkan bahwa mereka dapat berfungsi sebagai variabel prediktor yang signifikan dalam mempengaruhi tingkat partisipasi pemilu. Dengan demikian, Tingkat partisipasi pemilu di wilayah Kabupaten / Kota dapat dianggap cukup baik jika diukur berdasarkan efektivitas sistem dalam mengelompokkan data. Hasil ini membuktikan bahwa aspek demografi juga dapat mempengaruhi Tingkat partisipasi pemilu



Manfaat Penelitian

Penelitian ini memberikan manfaat dalam menganalisis dan mengukur tingkat partisipasi pemilih dalam pemilu di Kabupaten/Kota Matraman, Jawa Timur. Dengan menggunakan data demografi, penelitian ini memanfaatkan algoritma Naïve Bayes untuk memprediksi tingkat partisipasi pemilih secara lebih akurat. Selain itu, penelitian ini juga bertujuan untuk mengevaluasi akurasi dan efektivitas algoritma tersebut dalam konteks pemilu, yang dapat memberikan wawasan penting bagi perbaikan sistem pemilihan di masa mendatang dan meningkatkan kesadaran masyarakat akan pentingnya berpartisipasi.



www.umsida.ac.id



[umsida1912](https://www.instagram.com/umsida1912/)



[@umsida1912](https://twitter.com/umsida1912)



universitas
muhammadiyah
sidoarjo



[umsida1912](https://www.youtube.com/umsida1912)



Referensi

- [1] S. Puad and A. Susilo Yuda Irawan, “ANALISIS SENTIMEN MASYARAKAT PADA TWITTER TERHADAP PEMILIHAN UMUM 2024 MENGGUNAKAN ALGORITMA NAÏVE BAYES,” 2023.
- [2] S. Listyaningrum, “Penerapan Data Mining Untuk Analisis Karakteristik DPT Non-Participate sebagai Prediksi Partisipan Pemilu dengan Menggunakan Metode Naive Bayes Classifier.”
- [3] I. Putu *et al.*, “Prediksi Partisipasi Pemilih dalam Pemilu Presiden 2014 dengan Metode Support Vector Machine.”
- [4] S. Ode, B. Dalupe, S. Y. Regif, and D. Al Jannah, “Pemilu dan COVID-19 di Indonesia: Partisipasi Pemilih dalam Pilkada Serentak 2020,” *Journal of Political Issues*, vol. 4, no. 1, pp. 29–41, 2022, doi: 10.33019/jpi.v4i.
- [5] 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.
- [6] I. Widhi Saputro and B. Wulan Sari, “Uji Performa Algoritma Naïve Bayes untuk Prediksi Masa Studi Mahasiswa Naïve Bayes Algorithm Performance Test for Student Study Prediction,” *Citec Journal*, vol. 6, no. 1, 2019.
- [7] A. W. Anggraeni, A. S. Fitriani, and A. Eviyanti, “Penerapan Algoritma Support Vector Machine untuk Memprediksi Tingkat Partisipasi Pemilu terhadap Kualitas Pendidikan,” *Edumatic: Jurnal Pendidikan Informatika*, vol. 8, no. 1, pp. 21–27, Jun. 2024, doi: 10.29408/edumatic.v8i1.24838.

Referensi

- [8] L. Endang Nurrokhmah and I. Jaquillen Loppies, “Partisipasi Politik Masyarakat dalam Pemilihan Kepala Kampung (Studi Kasus Pilkam Mandouw Distrik Samofa Kabupaten Biak Numfor),” *Jurnal Syntax Transformation*, vol. 2, no. 09, pp. 1300–1307, Sep. 2021, doi: 10.46799/jst.v2i9.414.
- [9] Y. Raharja, A. Senja Fitriani, R. Dijaya, and F. Sains dan Teknologi, “KLASIFIKASI TINGKAT PARTISIPASI PEMILU BERDASARKAN SEKTOR INDUSTRI MENGGUNAKAN ALGORITMA NAÏVE BAYES,” *Jurnal TEKINKOM*, vol. 7, no. 1, 2024, doi: 10.37600/tekinkom.v7i1.1204.
- [10] D. Mizta Chulloh, A. Senja Fitriani, I. Ratna Indra Astutik, and A. Eviyanti, “Uji Akurasi K-Means dalam Prediksi Partisipasi Pemilu pada Demografi Wilayah Kabupaten Pasuruan”.
- [11] 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.
- [12] A. Hakim and D. Suherman, “Prediksi Kehadiran Masyarakat Dalam Pemilihan Umum Dengan Menggunakan Metode Naïve Bayes Classification,” *Jakarta STI&K Jl. BRI*, vol. 3, no. 1, p. 42116, 2019.
- [13] S. Adi Mulyanto, “58) Prediksi Minat Masyarakat dalam Pemilihan Umum Kabupaten Banyumas Menggunakan Bayesian Classification (Public Interest Prediction Research in Banyumas District Election Using Bayesian Classification),” 2017.
- [14] 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,” *Indonesian Journal of Business Intelligence (IJUBI)*, vol. 5, no. 1, p. 36, Jun. 2022, doi: 10.21927/ijubi.v5i1.2259.
- [15] 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 Mayarakat Pada Pemilu 2024,” *JURNAL KOMITEK*, vol. 3, no. 1, pp. 105–114, doi: 10.53697/jkomitek.v3i1.



