

Yusuf
Raharja_201080200017_Karya
Tulis Ilmiah.docx
by 16 Perpustakaan UMSIDA

Submission date: 06-Feb-2024 01:46PM (UTC+0700)

Submission ID: 2287730334

File name: Yusuf Raharja_201080200017_Karya Tulis Ilmiah.docx (356.35K)

Word count: 3689

Character count: 22895

Classification of Election Participation Levels Based on Industrial Sectors Using Naïve Bayes Algorithm [Klasifikasi Tingkat Partisipasi Pemilu Berdasarkan Sektor Industri Menggunakan Algoritma Naïve Bayes]

Yusuf Raharja¹⁾, Arif Senja Fitriani^{*2)}

¹⁾Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

²⁾Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

*Email Penulis Korespondensi: asfjim@umsida.ac.id

Abstract. General elections are a form of the highest power that exists in the people and become a democratic means to produce representatives of the people and leaders. Public participation is an important factor in elections because it can be an indicator of the success of elections. Demographic aspects can be one of the factors that influence the level of public participation, especially in the industrial sector in the area. This research aims to classify the level of electoral participation in Pasuruan Regency based on demographic aspects, particularly in the industrial sector of Pasuruan Regency. The demographic dataset includes geography, population, industrial sector and other aspects at the village level sourced from the Pasuruan Regency Central Statistics Agency (BPS) with a total of 80 attributes designated as predictor attributes and 2019 election recapitulation data on the level of community participation at the voting place (TPS) specified as target attribute. The preprocessing steps include data cleaning, data transformation, data integration, attribute correlation, random dataset, forming two model datasets, and splitting the data with a 70:30 ratio. The method used in this research is the classification method with the Naïve Bayes algorithm. The results of the testing of model dataset 1 produced an accuracy of 61.8% and model dataset two produced an accuracy of 67.6%. The number of industries in Pasuruan Regency does not have a significant influence on the level of public participation in elections in Pasuruan Regency.

Keywords - Classification, Data Mining, Election Participation, Industrial Sector, Naïve Bayes

Abstrak. Pemilihan Umum merupakan bentuk dari kekuasaan tertinggi yang ada pada rakyat dan menjadi suatu sarana demokrasi untuk melahirkan wakil rakyat dan pemimpin. Partisipasi masyarakat merupakan faktor penting dalam pemilu karena dapat menjadi indikator suksesnya pemilu. Aspek demografi dapat menjadi salah satu faktor yang mempengaruhi tingkat partisipasi masyarakat, khususnya pada sektor industri di daerah tersebut. Penelitian ini bertujuan untuk mengklasifikasi tingkat partisipasi pemilu di Kabupaten Pasuruan berdasarkan aspek demografi, khususnya pada sektor industri di Kabupaten Pasuruan. Dataset demografi meliputi geografi, penduduk, sektor industri dan aspek-aspek lainnya pada tingkatan desa yang bersumber dari Badan Pusat Statistik (BPS) Kabupaten Pasuruan dengan total 80 atribut yang ditetapkan sebagai atribut prediktor dan data rekapitulasi pemilu 2019 pada tingkat partisipasi masyarakat di level Tempat Pemungutan Suara (TPS) yang ditetapkan sebagai atribut target. Tahapan preprocessing meliputi data cleaning, data transformation, data integration, korelasi atribut, random dataset, membentuk dua model dataset dan split data dengan rasio 70:30. Metode yang digunakan pada penelitian ini adalah metode klasifikasi dengan algoritma naïve bayes. Hasil pengujian model dataset 1 menghasilkan akurasi sebesar 61.8% dan model dataset dua menghasilkan akurasi sebesar 67.6%. Jumlah industri di Kabupaten Pasuruan tidak memiliki pengaruh yang signifikan terhadap tingkat partisipasi masyarakat dalam pemilu di Kabupaten Pasuruan.

Kata Kunci - Klasifikasi, Data Mining, Partisipasi Pemilu, Sektor Industri, Naïve Bayes

I. PENDAHULUAN

Indonesia merupakan negara yang menerapkan konsep demokrasi, di mana negara ini memiliki bentuk sistem pemerintahan yang mengedepankan kedaulatan rakyat [1]. Pemilihan umum adalah bentuk dari kekuasaan tertinggi yang ada pada rakyat dan menjadi suatu sarana demokrasi untuk melahirkan wakil rakyat dan pemimpin yang menjunjung tinggi aspirasi rakyat [2]. Pemilihan umum dilaksanakan berdasarkan prinsip langsung, umum, bebas, rahasia, jujur, dan adil [3]. Faktor penting dalam pemilihan umum yaitu partisipasi masyarakat [4].

Partisipasi masyarakat dalam pemilihan umum dapat menjadi indikator suksesnya pemilihan umum tersebut [5]. Keikutsertaan masyarakat dalam pemilu semakin tinggi menandakan bahwa masyarakat mengikuti, memahami dan berpartisipasi dalam kegiatan kenegaraan, sebaliknya jika keikutsertaan masyarakat dalam pemilu semakin rendah, hal ini menunjukkan bahwa masyarakat kurang menghargai atau berkepentingan terhadap masalah atau kegiatan negara [6]. Jumlah partisipasi masyarakat di seluruh Indonesia dalam pemilu tahun 2019 mencapai 81,93%, yang

awalnya diprediksi hanya 77,5% [7]. Tingkat partisipasi masyarakat dalam pemilu dapat dipengaruhi oleh beberapa faktor, salah satunya adalah pada aspek demografi di daerah tersebut.

Badan Pusat Statistik (BPS) Kabupaten Pasuruan setiap tahun rutin menerbitkan publikasi data, publikasi tersebut berisi data mengenai segala aspek yang terdapat di setiap desa di seluruh kecamatan di Kabupaten Pasuruan. Demografi menjadi salah satu aspek yang dibahas pada publikasi tersebut, sektor industri termasuk salah satu bagian dari demografi. Data aspek demografi yang terdapat pada publikasi data Badan Pusat Statistik (BPS) dan data aspek yang berkaitan dengan demografi dapat dihubungkan dengan data partisipasi masyarakat dalam pemilu menggunakan bidang keilmuan *data mining*.

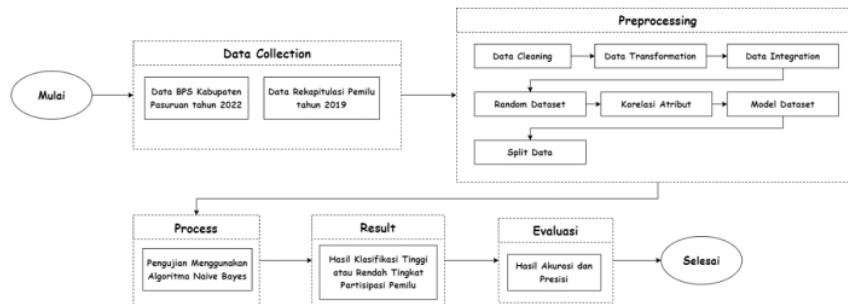
Data mining adalah proses penemuan informasi secara otomatis dengan mengenali pola serta hubungan yang tersembunyi dalam data [8]. Klasifikasi merupakan metode dalam data mining yang digunakan untuk mengelompokkan data berdasarkan keterikatan data terhadap data sampel, didalam metode klasifikasi terdapat berbagai algoritma, salah satunya yaitu algoritma naïve bayes [9]. Algoritma naïve bayes menggunakan teorema Bayes sebagai dasar dengan asumsi bahwa variabel-variabelnya tidak saling bergantung atau independen [10]. Algoritma naïve bayes memiliki kelebihan yaitu kecilnya data latih yang digunakan sehingga perhitungan dapat dilakukan secara efisien [11]. Kelemahan algoritma naïve bayes yaitu ketika data yang harus diklasifikasi memiliki banyak fitur dapat membuat algoritma naïve bayes menghasilkan akurasi yang rendah [12].

Penelitian terdahulu yang dilakukan oleh Dewi Eka Safitri untuk mengklasifikasi tingkat partisipasi pemilu berdasarkan demografi Kota Surabaya menggunakan algoritma SVM dengan kernel gaussian rbf dan menggunakan 2 model dataset, pada penelitian tersebut peneliti melakukan pengujian pada semua atribut prediktor terhadap atribut target meskipun atribut prediktor mempunyai tingkat korelasi yang rendah terhadap atribut target. Hasil pengujian pada penelitian tersebut mendapatkan hasil akurasi sebesar 67,91% pada rasio 80:20 dan 67,31% pada rasio 70:30 [7].

Penelitian ini bertujuan untuk mengklasifikasi tingkat partisipasi pemilu berdasarkan aspek demografi, lebih khusus pada sektor industri yang ada di Kabupaten Pasuruan menggunakan dataset demografi yang bersumber dari data BPS Kabupaten Pasuruan tahun 2022 yang ditetapkan sebagai atribut prediktor dan data rekapitulasi pemilu tahun 2019 yang ditetapkan sebagai atribut target. Tahap *preprocessing* pada penelitian ini meliputi meliputi data *cleaning*, data *transformation*, data *integration*, korelasi atribut, *random* dataset, membentuk dua model dataset dan yang terakhir melakukan *split* data dengan rasio 70:30. Tahap *process* atau pengujian dilakukan menggunakan algoritma naïve bayes dengan melakukan pengujian terhadap 2 model dataset, pengujian pada model dataset pertama yaitu melakukan pengujian pada semua atribut prediktor terhadap atribut target sedangkan pengujian pada model dataset kedua yaitu melakukan pengujian dengan mengurangi atribut prediktor berdasarkan tingkat korelasi antara atribut prediktor dengan atribut target. Evaluasi hasil dilakukan pada setiap kecamatan di Kabupaten Pasuruan untuk menjelaskan hubungan antara sektor industri yang di Kabupaten Pasuruan dengan tingkat partisipasi masyarakat dalam pemilu di Kabupaten Pasuruan.

II. METODE

Penelitian ini berjenis eksperimental, yakni eksperimen model klasifikasi terhadap tingkat partisipasi masyarakat dalam pemilihan umum dengan menerapkan algoritma naïve bayes.



Gambar 1. Tahapan Penelitian

Gambar 1 menunjukkan tahapan dalam penelitian ini yang diawali dengan data collection atau pengumpulan data, kemudian melakukan preprocessing terhadap data yang sudah dikumpulkan, selanjutnya melakukan process atau pengujian menggunakan algoritma naïve bayes, setelah melakukan pengujian akan muncul hasil dari pengujian dan hasil tersebut akan dievaluasi untuk mendapatkan nilai akurasi dan presisi.

A. Data Collection

Data collection adalah tahap awal dalam penelitian ini, pada tahap ini dilakukan pengumpulan data. Data yang digunakan pada penelitian ini adalah data yang berasal dari publikasi data oleh Badan Pusat Statistik (BPS) Kabupaten Pasuruan tahun 2022 yang diperoleh dari situs resmi Badan Pusat Statistik (BPS) Kabupaten Pasuruan serta data rekapitulasi pemilu di Kabupaten Pasuruan tahun 2019 yang diperoleh dari situs resmi Komisi Pemilihan Umum (KPU). Tabel 1 menunjukkan atribut dataset yang terdiri dari 80 atribut prediktor dan 1 atribut target.

Tabel 1. Atribut Dataset

No Atribut	Kelompok Atribut	Atribut
X1 - X3	Geografi	Luas Total Area {X1}, Jarak Ke Ibukota Kecamatan {X2}, Jarak Ke Ibukota Kabupaten {X3}
X4 - X7	Pemerintah	Jenis Kelamin Kepala desa {X4}, Pendidikan Tertinggi Kepala Desa {X5}, Pasar Desa {X6}, Kepemilikan Aset Desa Lainnya (Lapangan Olahraga, Pemandian Umum, dll) {X7}
X8 - X12	Penduduk	Jumlah Penduduk {X8}, Kepadatan Penduduk per km ² {X9}, Rasio Jenis Kelamin {X10}, Jumlah Penduduk Laki-Laki {X11}, Jumlah Penduduk Perempuan {X12}
X13 - X24	Pendidikan	Jumlah Bangunan TK / Sederajat {X13}, Jumlah Guru TK / Sederajat {X14}, Jumlah Murid TK / Sederajat {X15}, Jumlah Bangunan SD / Sederajat {X16}, Jumlah Guru SD / Sederajat {X17}, Jumlah Murid SD / Sederajat {X18}, Jumlah Bangunan SMP / Sederajat {X19}, Jumlah Guru SMP / Sederajat {X20}, Jumlah Murid SMP / Sederajat {X21}, Jumlah Bangunan SMA / Sederajat {X22}, Jumlah Guru SMA / Sederajat {X23}, Jumlah Murid SMA / Sederajat {X24}
X25 - X38	Kemudahan Mencapai Sarana Pendidikan dan Kesehatan Terdekat	SD {X25}, MI {X26}, SMP {X27}, MTs {X28}, SMA {X29}, MA {X30}, SMK {X31}, Akademi / Perguruan Tinggi {X32}, Rumah Sakit {X33}, Rumah Sakit Bersalin {X34}, Poliklinik/Balai Pengobatan {X35}, Puskesmas (Rawat Inap) {X36}, Puskesmas (Tanpa Rawat Inap) {X37}, Apotik {X38}
X39	Kesehatan	Jumlah Warga Penderita Gizi Buruk {X39}
X40	Lingkungan	Jumlah Pengguna Listrik {X40}
X41 - X46	Agama	Jumlah Masjid {X41}, Jumlah Mushola {X42}, Jumlah Gereja Protestan {X43}, Jumlah Gereja Katolik {X44}, Jumlah Pura {X45}, Jumlah Vihara {X46}
X47 - X51	Bencana Alam	Jumlah Gempa Bumi {X47}, Jumlah Tanah Longsor {X48}, Jumlah Banjir {X49}, Jumlah Kekeringan {X50}, Jumlah Sistem Peringatan Dini Bencana Alam {X51}
X52 - X53	Lahan	Luas Lahan Kosong {X52}, Luas Lahan Untuk Bangunan Dan Pekarangan {X53}
X54 - X58	Infrastruktur	Jenis Permukaan Jalan {X54}, Dapat Dilalui Roda 4 atau Lebih {X55}, Perusahaan/Agen Jasa Ekspedisi Swasta {X56}, Jumlah Menara Telepon Seluler {X57}, Kekuatan Sinyal Telepon Seluler {X58}
X59 - X65	Lembaga Keuangan	Jumlah Bank Umum Pemerintah {X59}, Jumlah Bank Umum Swasta {X60}, Jumlah Bank Perkreditan {X61}, Jumlah Koperasi Unit Desa {X62}, Jumlah Koperasi Industri Kecil dan Kerajinan Rakyat {X63}, Jumlah Koperasi Simpan Pinjam {X64}, Jumlah Koperasi Lainnya {X65}

X66 - X71	Sarana Perdagangan	Jumlah Kelompok Pertokoan {X66}, Jumlah Pasar dengan Bangunan Permanen {X67}, Pasar dengan Bangunan Semi Permanen {X68}, Pasar tanpa Bangunan {X69}, Jumlah Minimarket / Swalayan / Supermarket {X70}, Restoran/Rumah Makan {X71}
X72 - X80	Industri	Jumlah Industri dari Kulit {X72}, Jumlah Industri Furniture dari Kayu {X73}, Jumlah Industri Barang Logam {X74}, Jumlah Industri Tekstil dan Pakaian Jadi {X75}, Jumlah Industri Barang Galian Bukan Logam {X76}, Jumlah Industri Anyaman Kayu {X77}, Jumlah Industri dari Makanan dan Minuman {X78}, Jumlah Industri Percetakan {X79}, Jumlah Reparasi Mesin dan Peralatan {X80}
Y	Pemilu	Tingkat Partisipasi Masyarakat

B. Preprocessing

Preprocessing merupakan tahapan yang dilakukan setelah tahap *input*, pada tahap ini data yang telah terkumpul akan diolah menjadi data yang siap untuk diproses menggunakan algoritma naïve bayes [13]. Tahapan *preprocessing* terdiri dari data *cleaning*, data *integration*, data *transformation*, *random* dataset, korelasi atribut, model dataset dan *split* data.

1. Data cleaning

Data cleaning adalah langkah untuk mengidentifikasi kesalahan yang ada pada data seperti duplikasi data, inkonsistensi data, data tidak lengkap, yang kemudian akan diambil keputusan terhadap data tersebut seperti penghapusan data yang tidak sesuai atau perbaikan terhadap data [14]. Median atau mean *cleaning* akan dilakukan pada atribut yang memiliki *missing values*.

2. Data transformation

Data transformation merupakan langkah untuk melakukan perubahan format dan bentuk data yang bertujuan untuk mempermudah dalam proses *mining* [15]. Transformasi data yang dilakukan pada penelitian ini mengubah tipe data pada dataset menjadi numerik, salah satu contohnya seperti mengubah baris data yang memiliki nilai “tinggi” menjadi 3, “sedang” diubah menjadi 2, “rendah” diubah menjadi 1.

3. Data integration

Data integration adalah penggabungan data yang berasal dari sumber yang berbeda [16]. Integrasi data dalam penelitian ini yaitu penggabungan data BPS Kabupaten Pasuruan tahun 2022 dengan data rekapitulasi pemilu di Kabupaten Pasuruan tahun 2019. Integrasi data dapat dilakukan jika kondisi dua data yang mempunyai kesamaan pada hirarki konsep, dalam hal ini data Badan Pusat Statistik (BPS) dan rekapitulasi pemilu memiliki kesamaan hirarki konsep yaitu pada tingkatan desa.

4. Random dataset

Random dataset dilakukan untuk memaksimalkan keterwakilan bobot setiap baris data pada semua atribut [7].

5. Korelasi atribut

Korelasi atribut dilakukan untuk melihat seberapa besar tingkat korelasi antara atribut prediktor terhadap atribut target.

6. Model dataset

Dua model dataset yang digunakan pada pengujian, model dataset 1 terdiri dari seluruh atribut dan model dataset 2 dilakukan pengurangan atribut dengan melihat tingkat korelasi atribut prediktor terhadap atribut target. Model dataset 1 terdiri dari 81 atribut dan model dataset 2 terdiri dari 61 atribut, pengurangan dilakukan pada atribut X2, X10, X15, X20, X21, X26, X37, X39, X40, X41, X42, X49, X50, X51, X55, X57, X61, X62, X64, X67.

7. Split data

Split data merupakan langkah terakhir dalam tahap *preprocessing*, dataset akan dibagi menjadi 2 bagian yaitu data *training* dan data *testing* [17]. Data *training* digunakan oleh algoritma klasifikasi untuk membentuk model klasifikasi [18]. Data *testing* berfungsi untuk mengukur keberhasilan *classifier* dalam melakukan klasifikasi dengan akurat [19]. Rasio yang digunakan pada penelitian ini untuk membagi antara data *training* dan data *testing* yaitu 70:30.

C. Process

Process merupakan tahapan yang dilakukan setelah tahap *preprocessing*, pada tahap ini akan dilakukan proses pengujian terhadap dataset menggunakan metode klasifikasi dengan algoritma naïve bayes. Pengujian dilakukan menggunakan 2 tipe model dataset. Rasio yang digunakan untuk membagi data yaitu 70:30. Data *training* memiliki jumlah baris data sebanyak 3063 dan data *testing* memiliki jumlah baris data sebanyak 1314.

D. Result

Result adalah tahapan yang dilakukan setelah tahap *process* selesai dilakukan, pada tahap ini akan menampilkan hasil dari proses pengujian menggunakan algoritma naïve bayes yang berupa hasil klasifikasi tinggi atau rendah tingkat partisipasi masyarakat dalam pemilu.

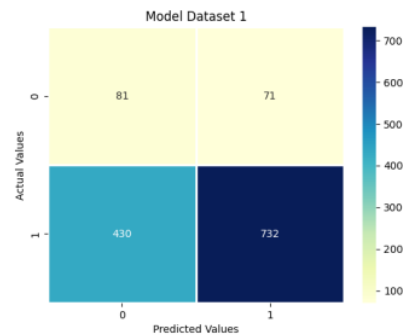
E. Evaluasi

Evaluasi adalah tahap terakhir dalam penelitian ini. Evaluasi kinerja algoritma naïve bayes dalam penelitian ini menggunakan *confusion matrix* untuk memperoleh nilai akurasi dan presisi [20].

III. HASIL DAN PEMBAHASAN

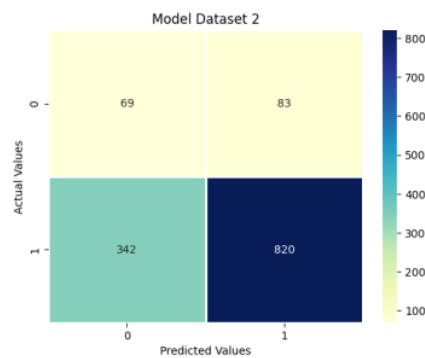
Penelitian ini menggunakan dua tipe model dataset. Model dataset 1 terdiri dari 81 atribut dan 4377 baris data, model dataset 2 terdiri dari 61 atribut dan 4377 baris data. Implementasi algoritma naïve bayes menggunakan bahasa pemrograman python.

A. Akurasi dan Presisi



Gambar 2. *Confusion matrix* Model Dataset 1

Gambar 2 menunjukkan hasil *confusion matrix* pada model dataset 1. 430 data diklasifikasikan dengan salah sebagai rendah, 732 data diklasifikasikan dengan benar sebagai tinggi, 81 data diklasifikasikan dengan benar sebagai rendah dan 71 data diklasifikasikan dengan salah sebagai tinggi.



Gambar 3. *Confusion Matrix* Model Dataset 2

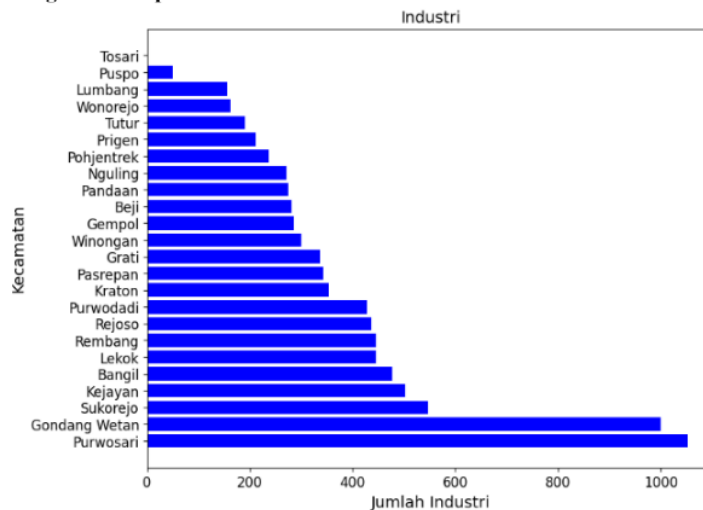
Gambar 3 menunjukkan hasil *confusion matrix* pada model dataset 2. 342 data diklasifikasikan dengan salah sebagai rendah, 820 data diklasifikasikan dengan benar sebagai tinggi, 69 data diklasifikasikan dengan benar sebagai rendah, dan 83 data diklasifikasikan dengan salah sebagai tinggi.

Tabel 2. Hasil Pengujian Algoritma Naïve Bayes

Model Dataset	Akurasi	Presisi
Model Dataset 1	0.618	0.911
Model Dataset 2	0.676	0.908

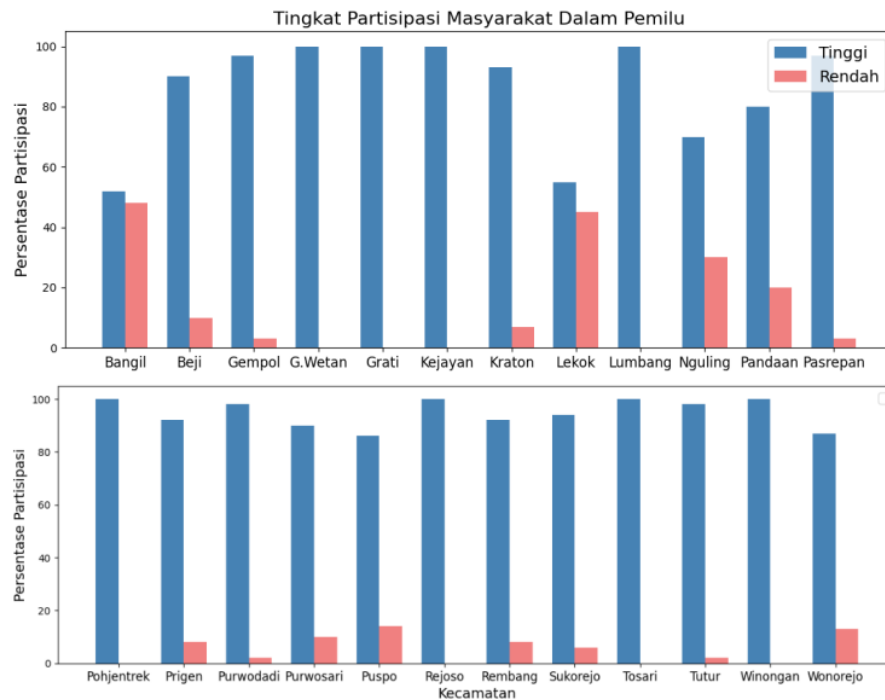
Tabel 2 menunjukkan bahwa hasil akurasi terbaik dalam pengujian menggunakan algoritma naïve bayes adalah pada model dataset 2 dengan nilai akurasi sebesar 0.676 atau 67.6%. Nilai presisi terbaik adalah pada model dataset 1 dengan nilai presisi sebesar 0.911 atau 91.1%. Hasil akurasi model dataset 1 dan 2 menunjukkan bahwa tingkat korelasi antara atribut prediktor dan atribut target memiliki pengaruh yang signifikan terhadap nilai akurasi model. Penelitian [7] mendapatkan hasil akurasi sebesar 67.38% pada model dataset 1 dengan menggunakan rasio 70:30 dan hasil akurasi sebesar 67.91% pada model dataset 2 dengan menggunakan rasio 80:20, hasil pada penelitian tersebut tidak meningkat secara signifikan karena peneliti tetap melakukan pengujian pada semua atribut prediktor terhadap atribut target meskipun atribut prediktor memiliki tingkat korelasi yang rendah terhadap atribut target.

B. Industri dan Tingkat Partisipasi Pemilu



Gambar 4. Jumlah Industri

Gambar 4 menunjukkan jumlah industri pada setiap kecamatan di Kabupaten Pasuruan. Kecamatan Purwosari dan Kecamatan Gondang Wetan merupakan dua kecamatan yang memiliki jumlah industri terbanyak di Kabupaten Pasuruan, dengan jumlah industri masing-masing sebanyak 1052 dan 1000. Kecamatan Puspo dan Kecamatan Tosari merupakan dua kecamatan yang memiliki jumlah industri paling sedikit di Kabupaten Pasuruan, dengan jumlah industri masing-masing sebanyak 50 dan 2.



Gambar 5. Tingkat Partisipasi Pemilu

Gambar 5 menunjukkan tingkat partisipasi masyarakat dalam pemilu pada setiap kecamatan di Kabupaten Pasuruan, dengan kategori rendah dan tinggi di level Tempat Pemungutan Suara (TPS). Tingkat partisipasi pemilu pada setiap kecamatan digambarkan dalam bentuk persentase untuk memberikan gambaran yang lebih jelas tentang tingkat partisipasi pemilu.

Tabel 3. Jumlah Industri dan Tingkat Partisipasi Pemilu

Kecamatan	Jumlah Industri	Tingkat Partisipasi Masyarakat Dalam Pemilu	
		Rendah	Tinggi
Bangil	477	48%	52%
Beji	281	10%	90%
Gempol	286	3%	97%
Gondang Wetan	1000	0%	100%
Grati	337	0%	100%
Kejayan	502	0%	100%
Kraton	354	7%	93%
Lekok	445	45%	55%
Lumbang	156	0%	100%
Nguling	272	30%	70%
Pandaan	274	20%	80%
Pasrepan	343	3%	97%

Pohjentrek	237	0%	100%
Prigen	211	8%	92%
Purwodadi	429	2%	98%
Purwosari	1052	10%	90%
Puspo	50	14%	86%
Rejoso	436	0%	100%
Rembang	445	8%	92%
Sukorejo	547	6%	94%
Tosari	2	0%	100%
Tutur	191	2%	98%
Winongan	300	0%	100%
Wonorejo	162	13%	87%

Tabel 3 menunjukkan jumlah industri pada setiap kecamatan di Kabupaten Pasuruan, serta persentase tingkat partisipasi masyarakat dalam pemilu di setiap kecamatan di Kabupaten Pasuruan dengan kategori rendah dan tinggi di level Tempat Pemungutan Suara (TPS). Kecamatan Bangil dan Kecamatan Lekom memiliki tingkat partisipasi pemilu dalam kategori rendah yang persentasenya terbesar diantara kecamatan lainnya dengan masing-masing kecamatan memiliki persentase sebesar 48% dan 45%, namun jumlah industri di Kecamatan Bangil dan Kecamatan Lekom tergolong tinggi atau diatas rata-rata dengan jumlah industri masing-masing sebanyak 477 dan 445. Kecamatan Gondang Wetan dan Kecamatan Lumbang merupakan dua kecamatan yang memiliki tingkat partisipasi pemilu dalam kategori tinggi yang persentasenya terbesar diantara kecamatan lainnya dengan masing-masing kecamatan memiliki persentase yang sama yaitu 100%, jumlah industri pada Kecamatan Gondang Wetan tergolong tinggi atau diatas rata-rata dengan jumlah industri sebanyak 1000 sedangkan Kecamatan Lumbang memiliki jumlah industri yang tergolong rendah atau dibawah rata-rata dengan jumlah industri sebanyak 156. Evaluasi hasil menunjukkan bahwa jumlah industri di Kabupaten Pasuruan tidak memiliki pengaruh yang signifikan terhadap tingkat partisipasi masyarakat dalam pemilu di Kabupaten Pasuruan.

IV. SIMPULAN

Kesimpulan dari penelitian dengan algoritma naïve bayes yaitu dapat mengklasifikasi tingkat partisipasi pemilu berdasarkan aspek demografi dan menciptakan model dataset dengan korelasi atribut dalam mempertimbangkan kualitas di masing-masing atribut. Khusus pada sektor industri di Kabupaten Pasuruan, akurasi tertinggi yakni pada model dataset 2 dengan nilai akurasi sebesar 67.6%. Evaluasi hasil pada penelitian ini menunjukkan bahwa jumlah sektor industri di Kabupaten Pasuruan tidak memiliki pengaruh yang signifikan terhadap tingkat partisipasi masyarakat dalam pemilu di Kabupaten Pasuruan

UCAPAN TERIMA KASIH

Terima kasih kepada Badan Pusat Statistik (BPS) Kabupaten Pasuruan yang telah menyediakan data demografi Kabupaten Pasuruan dan Komisi Pemilihan Umum (KPU) yang telah menyediakan data partisipasi masyarakat dalam pemilu tahun 2019 serta Universitas Muhammadiyah Sidoarjo yang telah menyediakan fasilitas laboratorium sehingga penelitian ini dapat diselesaikan dengan baik.

REFERENSI

- [1] U. Nufikha, H. Yayuk, and T. Meiwatizal, "Urgensi Etika Demokrasi Di Era Global: Membangun Etika Dalam Mengemukakan Pendapat Bagi Masyarakat Akademis Melalui Pendidikan Kewarganegaraan," *J. Kewarganegaraan*, vol. 5, no. 2, p. 18, 2021, doi: 10.31316/jk.v5i2.1576.
- [2] A. Kristian Sumual, M. Gerson Lontaan, and Y. Supit, "Pelaksanaan Pemilu Di Indonesia Berdasarkan Perspektif Undang Undang Dasar 1945," *J. Law Nation*, vol. 2, no. Mei, pp. 103–112, 2023.
- [3] A. Dedi, "Analisis Sistem Pemilihan Umum Serentak 2019," *Moderat J. Ilm. Ilmu Pemerintah.*, vol. 5, no. 3, pp. 213–226, 2019, doi: 10.25147/moderat.v5i3.2676.
- [4] W. Kusuma, B. Permatasari, and R. A. Suntara, "Peningkatan Pengawasan Partisipatif Masyarakat Desa dalam Pelaksanaan Pemilihan Umum Serentak Tahun 2024 melalui Penyuluhan Hukum," *DAS SEIN J. Pengabd.*

- Huk. dan Hum.*, vol. 2, no. 2, pp. 93–104, 2022, doi: 10.33756/jds.v2i2.15256.
- [5] M. Gultom, M. Silalahi, G. Hutagalung, and J. A. Marbun, “Pengaruh Sosialisasi Terhadap Tingkat Partisipasi Pemilih Dalam Menggunakan Hak Pilihnya Pada Pemilu Di Kabupaten Toba Samosir,” *J. Reg. Plan.*, vol. 2, no. 1, pp. 26–41, 2020, doi: 10.36985/jrp.v2i1.591.
- [6] Z. Abidin AS and D. Kurnia, “Implementasi Nilai-Nilai Demokrasi Dalam Pemilihan Kepala Desa Di Kabupaten Bandung Barat,” *J. Acad. Praja*, vol. 5, no. 1, pp. 1–12, 2022, doi: 10.36859/jap.v5i1.700.
- [7] 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.
- [8] O. Manurung and P. S. Hasugian, “Analisa Algoritma Apriori Untuk Peminjaman Buku Pada Perpustakaan SMA 1 Silima Pungga-Pungga Parongil,” *Remik*, vol. 4, no. 1, pp. 154–160, 2019, doi: 10.33395/remik.v4i1.10445.
- [9] M. Wibowo and R. Ramadhani, “Perbandingan Metode Klasifikasi Data Mining Untuk Rekomendasi Tanaman Pangan,” vol. 5, pp. 913–921, 2021, doi: 10.30865/mib.v5i3.3086.
- [10] M. M. Arif, H. Setiawan, A. S. Fitriani, F. Sains, and U. Muhammadiyah, “Penggunaan Datamining Untuk Memprediksi Masa Studi Mahasiswa di Universitas Muhammadiyah Sidoarjo Dengan Algoritma Naive Bayes,” vol. 4, no. 3, pp. 622–629, 2023, doi: 10.30645/kesatria.v4i3.210.
- [11] B. Laurensz and Eko Sedyono, “Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19,” *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 10, no. 2, pp. 118–123, 2021, doi: 10.22146/jnteti.v10i2.1421.
- [12] A. N. Sihananto and H. Maulana, “STUDI LITERATUR TENTANG PERFORMA NAÏVE BAYES DALAM KLASIFIKASI DATA,” *Pros. Semin. Nas. Inform. Bela Negara*, vol. 2, pp. 132–135, 2021, doi: 10.33005/santika.v2i0.134.
- [13] M. Saputra *et al.*, “Analisis metode algoritma k- nearest neighbor (knn) dan naive bayes untuk klasifikasi diabetes mellitus,” vol. 6, pp. 723–729, 2023, doi: 10.37600/tekinkom.v6i2.942.
- [14] R. I. Borman and M. Wati, “Penerapan Data Mining Dalam Klasifikasi Data Anggota Kopdit Sejahtera Bandarlampung Dengan Algoritma Naive Bayes,” *J. Ilm. Fak. Ilmu Komput.*, vol. 09, no. 01, pp. 25–34, 2020.
- [15] F. A. Pratama, R. Narasati, and D. R. Amalia, “Pengaruh Kata Cashback Terhadap Peningkatan Penjualan Menggunakan Data Mining,” *J. Ilm. Manaj. Inform. dan Komput.*, vol. 3, no. 2, pp. 1–5, 2019, doi: 10.32485/kopertip.v3i2.66.
- [16] S. Anastassia Amellia Kharis and A. Haqqi Anna Zili, “Learning Analytics dan Educational Data Mining pada Data Pendidikan,” *J. Ris. Pembelajaran Mat. Sekol.*, vol. 6, pp. 12–20, 2022, doi: 10.21009/jrpms.06i1.02.
- [17] M. Thoriq, “Peramalan Jumlah Permintaan Produksi Menggunakan Jaringan Saraf Tiruan Algoritma Backpropagation,” *J. Inf. dan Teknol.*, vol. 4, pp. 27–32, 2022, doi: 10.37034/jidt.v4i1.178.
- [18] A. S. Fitriani, “Penerapan Data Mining Menggunakan Metode Klasifikasi Naive Bayes untuk Memprediksi Partisipasi Pemilihan Gubernur,” *JTAM (Jurnal Teor. dan Apl. Mat.)*, vol. 3, no. 2, pp. 98–104, 2019, doi: 10.31764/jtam.v3i2.995.
- [19] S. Melangi, “Klasifikasi Usia Berdasarkan Citra Wajah Menggunakan Algoritma Artificial Neural Network dan Gabor Filter,” *Jambura J. Electr. Electron. Eng.*, vol. 2, no. 2, pp. 60–67, 2020, doi: 10.37905/jjee.v2i2.6956.
- [20] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, “Implementasi Data Mining dengan Algoritma Naive Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 8, no. 6, p. 219, 2021, doi: 10.30865/jurikom.v8i6.3655.

Conflict of Interest Statement:

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORIGINALITY REPORT

17%

SIMILARITY INDEX

17%

INTERNET SOURCES

15%

PUBLICATIONS

18%

STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Universitas Muhammadiyah Sidoarjo Student Paper	13%
2	fai.umsida.ac.id Internet Source	2%
3	www.researchgate.net Internet Source	1%
4	repository.its.ac.id Internet Source	1%
5	media.neliti.com Internet Source	1%

Exclude quotes On

Exclude matches < 1%

Exclude bibliography On