

# Analisis Model Pada Tingkat Partisipasi Pemilu 2019 Terhadap Demografi Di Kabupaten Pamekasan Menggunakan Metode Naive Bayes

Oleh:

Maulana Habib Firmansyah

Arif Senja Fitrani

Progam Studi Informatika

Universitas Muhammadiyah Sidoarjo

Juli, 2025



# Pendahuluan

- Pemilu merupakan kegiatan penting dalam demokrasi.
- Tingkat partisipasi dipengaruhi faktor demografi.
- Pamekasan memiliki partisipasi tinggi di Pemilu 2019.
- Data dari BPS digunakan untuk klasifikasi menggunakan metode Naive Bayes.

# Pertanyaan Penelitian (Rumusan Masalah)

- Apakah atribut demografi berpengaruh terhadap tingkat partisipasi pemilu?
- Seberapa akurat metode Naive Bayes dalam memprediksi partisipasi?
- Faktor demografi apa yang paling berpengaruh?

# Metode

- Data dari BPS Kabupaten Pamekasan.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	G1	G2	G3	G4	G5	G7	G9	G10	G11	G12	G15	G16	G22	G23	G24
2	Batumarmar	Pangereman	7,2	7,42	36	1.0	20	20	20	20	0.0	7	7.0	1.0	9.0
3	Batumarmar	Bujur Barat	14,59	15,03	33	1.0	20	20	20	20	1.0	10	10.0	1.0	11.0
4	Batumarmar	Bujur Tengah	14,51	14,95	30	1.0	17	17	17	17	3.0	8	8.0	1.0	11.0
5	Batumarmar	Bujur Timur	15,14	15,6	16	1.0	22	22	22	22	0.0	8	8.0	1.0	11.0
6	Batumarmar	Ponjanan Timur	7,4	7,62	27	1.0	15	15	15	17	0.0	7	7.0	1.0	7.0
7	Batumarmar	Ponjanan Barat	6,01	6,19	26	1.0	13	13	13	13	1.0	6	6.0	1.0	7.0
8	Batumarmar	Lesong Laok	4,28	4,41	18	1.0	20	20	20	20	0.0	4	4.0	1.0	7.0
9	Batumarmar	Bangsereh	6,05	6,23	25	1.0	10	10	10	10	0.0	7	7.0	1.0	7.0
10	Batumarmar	Blaban	3,99	4,11	22	1.0	3	3	45	3	0.0	12	12.0	1.0	7.0
11	Batumarmar	Tamberu	0,15	0,15	10	1.0	1	1	45	1	0.0	2	2.0	1.0	7.0
12	Batumarmar	Batubintang	10,05	10,36	17	1.0	2	2	45	2	0.0	14	14.0	1.0	11.0
13	Batumarmar	Lesong Daja	6,01	6,19	16	1.0	5	5	40	3	0.0	5	5.0	1.0	7.0
14	Batumarmar	Kapong	1,67	1,72	10	1.0	5	5	40	5	0.0	4	4.0	1.0	7.0
15	Galis	Konang	4,47	14,03	5	1.0	1	1,5	9	2,5	0.0	8	8.0	1.0	9.0
16	Galis	Pandan	8,37	26,27	1	1.0	5	5	15,5	3,5	1.0	5	5.0	1.0	5.0
17	Galis	Lembung	3,54	11,11	5	1.0	3	3	15	2	0.0	4	4.0	1.0	5.0
18	Galis	Galis	2,03	6,37	5	1.0	0,5	0,5	12	1	1.0	5	5.0	1.0	9.0
19	Galis	Bulav	2,2	6,91	16	1.0	1	1	13	1,5	0.0	4	4.0	1.0	6.0

# Metode

## Pre-processing

### 1. Data Cleaning

Proses menemukan dan menangani kesalahan seperti Duplikasi dan Data tidak lengkap. Tindakan yang dilakukan Menghapus data yang tidak sesuai, Memperbaiki data yang masih dapat digunakan dengan Tujuan Meningkatkan kualitas data sebelum proses analisis.

### 2. Korelasi atribut

Untuk mengevaluasi kekuatan hubungan antara atribut prediktor dengan target Tingkat Partisipasi Pemilu. Korelasi tinggi menandakan informasi yang serupa → atribut duplikatif. Atribut dengan kontribusi rendah dihilangkan agar proses klasifikasi efisien dan akurat.

# Metode

## 3. Data Transformation

Mengubah seluruh atribut menjadi format numerik agar kompatibel dengan algoritma data mining.

```

G1      object
G2      object
G3      object
G4      object
G5      object
...
G265    int64
G266    int64
G267    int64
G268    int64
P       object
Length: 208, dtype: object

```

Hasil setelah Transformasi Data:

	G1	G2	G3	G4	G5	G7	G9	G10	G11	G12	...	G260	G261	G262	G263	\
0	0	118	150	154	72	0.0	24	24	37	29	...	1.0	1.0	0.0	0.0	
1	0	32	63	43	67	0.0	24	24	37	29	...	1.0	1.0	0.0	0.0	
2	0	33	62	41	58	0.0	19	19	28	19	...	1.0	1.0	0.0	0.0	
3	0	34	65	46	23	0.0	25	25	41	30	...	1.0	1.0	0.0	0.0	
4	0	131	153	156	49	0.0	17	18	22	19	...	1.0	1.0	0.0	0.0	

	G264	G265	G266	G267	G268	P
0	0.0	0.0	0.0	0.0	0.0	71
1	0.0	0.0	0.0	0.0	0.0	47
2	0.0	0.0	0.0	0.2	0.0	172
3	0.0	0.0	0.0	0.0	0.0	142
4	0.0	0.0	0.0	0.0	0.0	177

## 4. Seleksi Fitur

Pemilihan fitur yang paling relevan untuk pemodelan klasifikasi.

# Metode

- 5. Random Dataset Tujuan: Mengacak baris data untuk Memastikan distribusi yang merata Meningkatkan representasi bobot atribut Menghindari bias urutan
- 5. Split Data Tujuan Membagi data untuk proses pelatihan dan pengujian model Proporsi 80% Training 20% Testing Manfaat Melatih model secara optimal Mengukur performa model secara objektif

# Hasil

## Akurasi Model Naive Bayes: 65.79%

Model Naive Bayes menghasilkan akurasi sebesar 65,79%, yang berarti model berhasil mengklasifikasikan 66% data uji dengan benar. Nilai ini menunjukkan performa yang cukup baik secara keseluruhan, namun tetap perlu didukung oleh metrik lain seperti precision, recall, dan f1-score untuk evaluasi yang lebih menyeluruh.



# Hasil

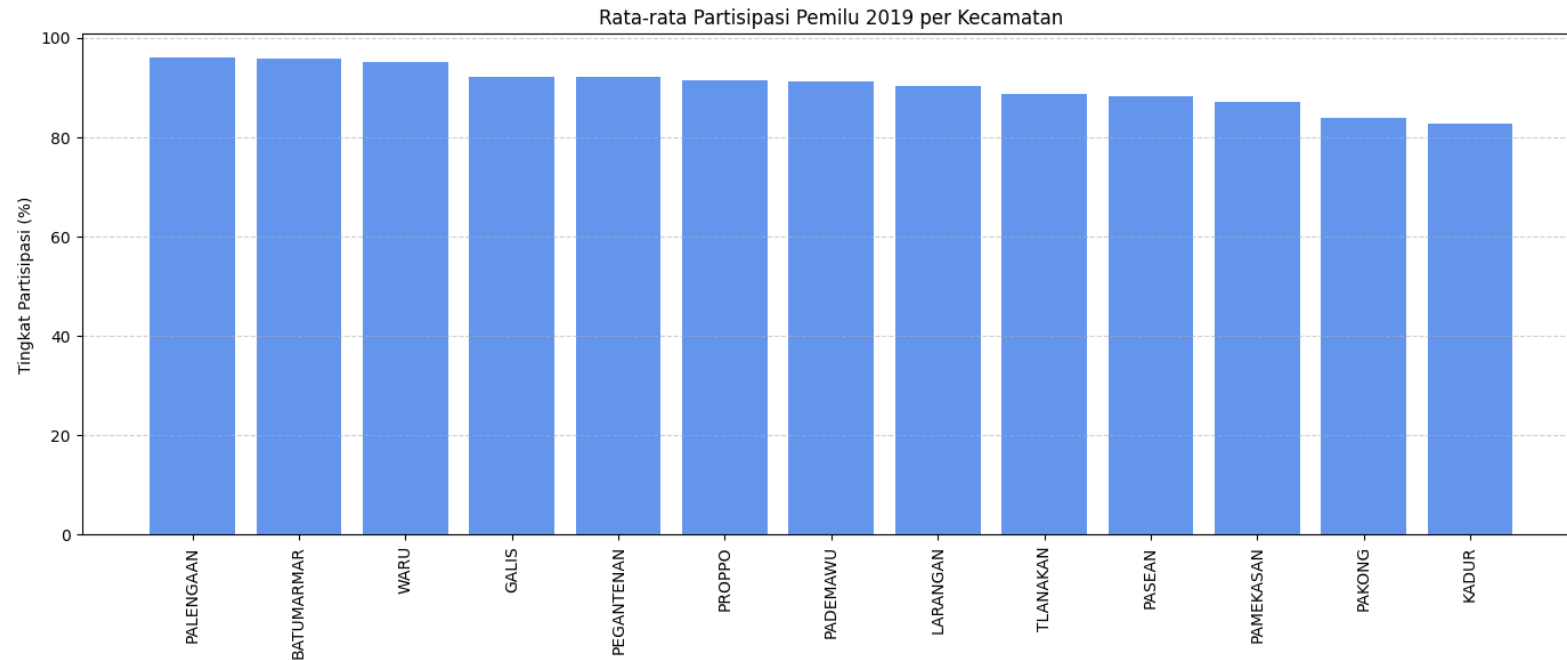
## Classification Report:

	precision	recall	f1-score	support
Rendah	0.82	0.45	0.58	20
Tinggi	0.59	0.89	0.71	18
accuracy			0.66	38
macro avg	0.71	0.67	0.65	38
weighted avg	0.71	0.66	0.64	38

Hasil classification report menunjukkan perbedaan kinerja model Naive Bayes dalam mengenali dua kelas Kelas 1 (Partisipasi Tinggi) Precision 0.59, Recall 0.89, F1-score 0.71 Model sangat baik dalam mendeteksi partisipasi tinggi meskipun presisinya sedang.

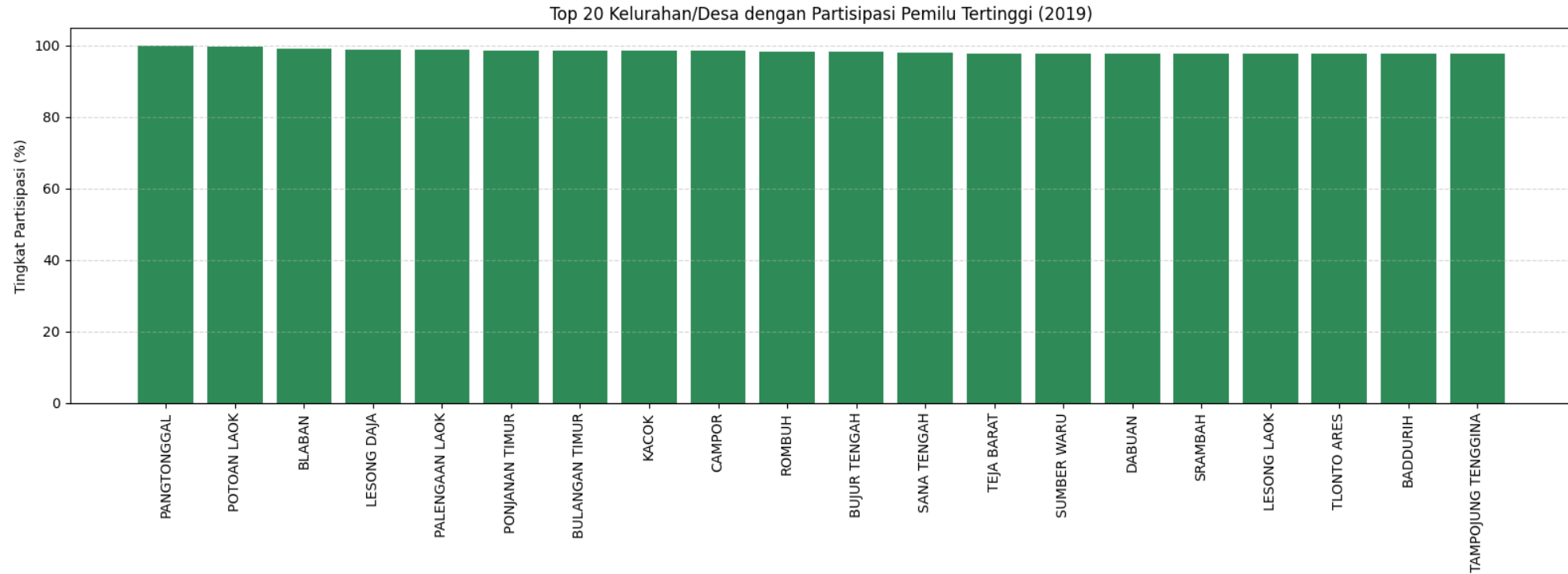
Kelas 0 (Partisipasi Rendah) Precision 0.82, Recall 0.45, F1-score 0.58 Model cukup tepat dalam prediksi, namun banyak data kelas ini yang tidak terdeteksi. Secara keseluruhan, model memperoleh akurasi 65,79% dan macro average F1-score 0.65, yang menunjukkan performa cukup seimbang namun belum optimal di kedua kelas.

# Hasil



Rata-rata partisipasi pemilu di seluruh kecamatan di Kabupaten Pamekasan tergolong tinggi, berkisar antara 82,87% hingga 96,03%. Hal ini menunjukkan bahwa partisipasi masyarakat dalam pemilu di wilayah ini secara umum sangat baik.

# Hasil



Terdapat 20 kelurahan/desa di Kabupaten Pamekasan dengan tingkat partisipasi pemilu tertinggi, seluruhnya di atas **97%**. Ini mencerminkan antusiasme tinggi masyarakat serta keberhasilan dalam mendorong kesadaran politik dan partisipasi demokratis di tingkat lokal.

# Manfaat Penelitian

- Memberikan wawasan tentang faktor demografi dan partisipasi pemilu.
- Menjadi dasar strategi peningkatan partisipasi oleh pemerintah.
- Dapat digunakan dan dikembangkan dengan algoritma klasifikasi lain.

