

Implementation Of Linear Regression Method To Determine Predictions Of The Influence Of Religion On General Election Participation

Saniya Izza Fitria¹, Arif Senja Fitriani², Ade Eviyanti³

^{1,2,3}Informatika, Fakultas Sains dan Teknologi, Universitas Muhammadiyah Sidoarjo

^{1,2,3}Jl. Raya Gelam No. 250 Candi Sidoarjo, telp. 0813 3299 7037

e-mail: ¹saniaizza28@gmail.com, ²asfjim@umsida.ac.id, ³adeeviyanti@umsida.ac.id

Abstract

This research aims to predict election participation in places of worship through statistical data analysis methods and predictive algorithms. Election participation, as a complex phenomenon, is influenced by various factors, with religion often being a crucial element that motivates or inhibits voter turnout. This study uses variables from Central Statistics Agency (BPS) data and recapitulation of previous general elections, based on historical patterns. Using a statistical approach, the relationship between religious variables and the level of voter participation in places of worship is identified. The linear regression method is used to predict the influence of religion on election participation. In this research, a series of scenarios were carried out, and the research results showed different variations in R-squared (R-Square) and Mean Squared Error (MSE) results. The best scenario, namely the R-squared scenario with a value of around 0.00012 and an MSE of 0.09934, highlights the potential relationship between religion and voter participation. These findings suggest the need for further considerations in this context, as well as demonstrating the need for model adjustments to improve the accuracy of future election predictions.

Keywords: Election, Participation, Linear Regression, Prediction, Religion

Implementasi Metode Regresi Linier Untuk Menentukan Prediksi Pengaruh Agama Pada Partisipasi Pemilu

Abstrak

Penelitian ini bertujuan memprediksi partisipasi pemilu di tempat ibadah melalui metode analisis data statistik dan algoritma prediktif. Partisipasi pemilu, sebagai fenomena kompleks, dipengaruhi oleh berbagai faktor, di mana agama seringkali menjadi elemen krusial yang memotivasi atau menghambat jumlah pemilih. Studi ini menggunakan variabel dari data Badan Pusat Statistik (BPS) dan rekapitulasi pemilihan umum sebelumnya, berdasarkan pola historis. Dengan pendekatan statistik, hubungan antara variabel agama dan tingkat partisipasi pemilih

<https://doi.org/10.31849/digitalzone.v15i1.17027>

Digital Zone is licensed under a Creative Commons Attribution International (CC BY-SA 4.0)

di tempat ibadah diidentifikasi. Metode regresi linier digunakan untuk memprediksi pengaruh agama pada partisipasi pemilu. Dalam penelitian ini, dilakukan serangkaian skenario, dan hasil penelitian menunjukkan variasi hasil R-squared (*R-Square*) dan Mean Squared Error (*MSE*) yang berbeda. Skenario terbaik, yaitu Skenario R-squared dengan nilai sekitar 0.00012 dan MSE sebesar 0.09934, menyoroti potensi hubungan antara agama dan partisipasi pemilih. Temuan ini menunjukkan perlunya pertimbangan lebih lanjut dalam konteks ini, serta memperlihatkan kebutuhan untuk penyesuaian model guna meningkatkan akurasi prediksi pemilu di masa depan.

Kata kunci: Pemilu, Partisipasi, Regresi Linier, Prediksi, Agama

1. Pendahuluan

Pemilu adalah ekspresi kedaulatan rakyat, dan semua warga negara mempunyai hak untuk memilih dan memilih pemerintahannya. Konsep pemilu erat kaitannya dengan prinsip demokrasi. Indonesia adalah negara yang menganut prinsip demokrasi. Mengamalkan sila ke empat pancasila sebagai ideologi negara yang berbunyi “Kerakyatan yang dipimpin oleh hikmat kebijaksanaan dalam permusyawaratan/perwakilan”, yang berarti rakyat melaksanakan kedaulatannya secara benar-benar demokratis dan tanpa diskriminasi. Teori demokrasi menekankan pentingnya partisipasi aktif warga negara dalam proses politik, dan pemilu merupakan mekanisme utama yang memungkinkan warga negara berpartisipasi dalam pengambilan kebijakan[1].

Partisipasi dalam pemilihan umum (pemilu) merupakan salah satu pilar utama yang memberikan kesempatan kepada warga negara untuk berpartisipasi aktif dalam memimpin dan mengarahkan kebijakan pemerintah. Partisipasi pemilih dalam pemilu berdampak signifikan terhadap legitimasi, keterwakilan, dan stabilitas institusi demokrasi. Kesadaran masyarakat mencakup pemahaman masyarakat mengenai pentingnya hak untuk memilih, tanggung jawab untuk berpartisipasi dalam proses politik, dan dampak yang dapat ditimbulkan oleh partisipasi aktif[2].

Pada pemilu 2019 KPU melaporkan hasil rekapitulasi jumlah pemilih mencapai 81,93%, dibandingkan perkiraan awal sebesar 77,5%. Perihal ini menjadi tolak ukur meningkatnya kesadaran diri masyarakat akan hak dan tanggung jawabnya[3]. Kesuksesan pemilu tidak hanya tercermin dari hasil akhirnya, tetapi juga dari sejumlah faktor yang mendukung kelancaran dan integritas seluruh proses pemilihan. Banyak faktor berbeda yang dapat mempengaruhi tingkat keterlibatan masyarakat, karena partisipasi pemilih memiliki karakteristik dan dinamika tersendiri. Oleh karena itu, pemahaman menyeluruh tentang faktor-faktor yang mempengaruhi jumlah pemilih sangat penting[4].

Pada penelitian sebelumnya yang dilakukan oleh Osbin Samosir dan Indah Novitasari membahas tentang isu pengalaman politik identitas salah satu faktor agama menjadi sasaran politik pada pemilihan nasional Indonesia tahun 2014 dan 2019. Indonesia terkenal karena adanya keragaman agama, seperti Islam, Kristen, Hindu, Buddha, dan kepercayaan-kepercayaan lainnya. Oleh karena itu, peran agama sangat penting dalam membentuk identitas masyarakat Indonesia. Kaitannya dengan kondisi politik terkini di Indonesia, agama memiliki peran sentral sebagai sumber nilai dan norma dalam masyarakat Indonesia. Politisi dan partai politik sering berupaya memanfaatkan nilai-nilai keagamaan guna membangun dukungan politik, terutama terkait dengan legislasi dan kebijakan yang bersinggungan dengan moral dan etika. Keterlibatan kelompok masyarakat tertentu, seperti ulama dan pemuka agama, memiliki dampak yang signifikan dalam ranah politik di Indonesia. Pengaruh mereka dapat membentuk pandangan politik serta mempengaruhi perumusan kebijakan yang diajukan oleh pemerintah atau partai politik[5].

Teori partisipasi politik menyatakan bahwa beragam faktor, termasuk agama dan tempat ibadah memiliki dampak signifikan pada tingkat partisipasi pemilih. Tempat ibadah, seperti

gereja, masjid, atau kuil, sering berperan sebagai pusat komunitas yang membentuk perilaku politik individu. Sebagai lingkungan sosial, tempat ibadah memiliki potensi untuk mempengaruhi partisipasi politik melalui berbagai mekanisme. Tempat ibadah dapat berfungsi sebagai wadah untuk menyampaikan informasi politik dan merangsang mobilitas pemilih. Pemimpin agama atau komunitas keagamaan juga memiliki peran penting dalam membentuk pandangan politik dan dapat mendorong anggota mereka untuk turut serta dalam proses politik[6].

Penelitian ini bertujuan menggali pemahaman dan menganalisis pengaruh agama terhadap partisipasi pemilih di tempat ibadah, memberikan wawasan bagi pengambil kebijakan dan penyelenggara pemilu. Wawasan ini diharapkan menjadi dasar strategi untuk meningkatkan partisipasi pemilih di tempat ibadah, terutama jika agama terbukti memiliki pengaruh yang signifikan. Selain itu, penelitian akan fokus pada evaluasi validitas dan generalisasi temuan, memastikan hasilnya dapat diterapkan secara luas pada pemilihan umum di tempat ibadah, termasuk di berbagai lokasi dan kelompok masyarakat.

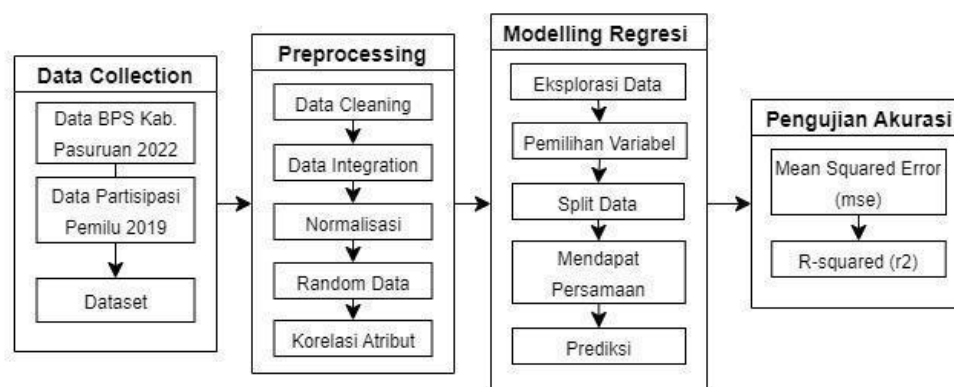
Badan Pusat Statistik (BPS) menyediakan data *ekstensif* tentang berbagai aspek masyarakat, seperti agama dan tempat ibadah di wilayah (kelurahan). Data ini dapat digunakan sebagai dasar untuk mengidentifikasi pola partisipasi di kalangan kelompok agama dan di tempat ibadah tertentu, memberikan dasar empiris yang kuat untuk analisis dan memungkinkan menggeneralisasi hasil ke tingkat yang lebih luas. Mengingat kondisi tersebut, penulis berupaya menghubungkan kondisi demografi agama ini dengan jumlah partisipasi pemilih (dengan cara yang bersifat prediktif)[7].

Menghadapi berbagai permasalahan tersebut, untuk dapat membantu memprediksi pengaruh tersebut dengan lebih baik, akurat dan cepat, sehingga memanfaatkan analisis data mining menggunakan algoritma regresi linier berganda dengan bahasa pemrograman python. Metode regresi linier berganda mudah diaplikasikan dan termasuk dalam model yang memiliki proses pelatihan dan pengujian yang cepat, cocok untuk dataset berjumlah besar. Algoritma ini memfasilitasi pemahaman arah hubungan antara variabel *independen* dan *dependen*, dengan koefisien regresi positif menunjukkan hubungan positif, sementara koefisien regresi negatif mencerminkan hubungan negatif[8].

Penelitian ini memiliki beberapa keterbatasan, seperti kesulitan dalam mengukur dampak langsung dari tempat ibadah terhadap partisipasi pemilu. Selain itu, sulitnya mengukur variabilitas individu dalam penafsiran ajaran agama dan interaksi dengan tempat ibadah juga merupakan tantangan tersendiri. Harapannya, temuan dari penelitian ini dapat memberikan kontribusi berharga sebagai dasar untuk merancang strategi kampanye yang lebih efektif dan sesuai dengan kondisi lokal. Penelitian ini diharapkan juga dapat memberikan informasi berharga untuk merencanakan kegiatan yang mendorong partisipasi pemilih yang lebih tinggi di wilayah tersebut dan memiliki implikasi yang signifikan untuk memperkuat proses kewarganegaraan, baik di tingkat lokal maupun nasional[9].

2. Metode Penelitian

Metode penelitian adalah metode utama yang dipilih peneliti untuk menemukan solusi terhadap permasalahan yang timbul atau diajukan, serta mencapai tujuan melalui proses tersebut. Dalam konteks ini, digunakan analisis data mining dengan menerapkan algoritma regresi linier berganda, menggunakan bantuan perangkat lunak Microsoft Excel dan bahasa pemrograman Python. Tujuannya adalah memprediksi apakah agama memiliki pengaruh terhadap partisipasi dalam pemilu. Adapun penjelasan pada Gambar 1



Gambar 1. Diagram Penelitian

2.1. Pengumpulan Data (*Data Collecting*)

Penelitian ini akan fokus pada Kabupaten Pasuruan sebagai studi kasus dengan tujuan menganalisis potensi pengaruh agama atau faktor lain terhadap jumlah partisipasi pemilih di Kabupaten Pasuruan dalam pemilu 2024. Pengumpulan data dilakukan dengan mengambil dari data resmi Badan Pusat Statistik (BPS) dan website pemilu <https://pemilu2019.kpu.go.id/>. Data demografi tersebut dengan total sebanyak 4377 instance dan 83 atribut kemudian mengambil atribut agama dan partisipasi untuk dijadikan variabel prediksi dan prediktor. Tabel 1 menampilkan atribut penelitian beserta keterangan dan skalanya[10].

Tabel 1. Atribut Penelitian

No	Atribut	Keterangan	Skala
1	Luas Total Area (km ²)	-	Ordinal
2-3	Jarak Ke Ibukota	-	Numerik
4-5	Kades	-	Nominal
6	Pasar Desa	-	Nominal
7	Kepemilikan Aset Desa Lainnya (Lapangan Olahraga, Pemandian Umum, dll)	-	Nominal
8-12	Penduduk	Jumlah, Kepadatan Penduduk, Rasio Jenis Kelamin, Jumlah Penduduk Laki, Jumlah Penduduk Perempuan	Ordinal
13-24	Sekolah	TK Bangunan, TK Guru, TK Murid, SD Bangunan, SD Guru, SD Murid, SMP Bangunan, SMP Guru, SMP Murid, SMA Bangunan, SMA Guru, SMA Murid	Numerik
25-38	Kemudahan Mencapai Sarana	SD, MI, SMP, MTS, SMA, MA, SMK, Akademi/Perguruan Tinggi, RS, RS Bersalin, Poliklinik, Puskesmas (Rawat Inap & Tidak), Apotik	Nominal
39	Penderita Gizi Buruk	Warga	Numerik
40	Pengguna Listrik	-	Numerik
41-46	Agama	Masjid, Mushola, Gereja Protestan, Gereja Katolik, Pura, Vihara	Ordinal
47-50	Bencana Alam	Gempa Bumi, Tanah Longsor, Banjir, Kekeringan	Nominal
51	Sistem Peringatan Dini Bencana Alam	-	Nominal
52-53	Lahan	Kosong, Bangunan & Pekarangan	Numerik

No	Atribut	Keterangan	Skala
54-58	Infrastruktur	Jenis Permukaan Jalan, Dapat Dilalui Roda 4 atau Lebih, Perusahaan/Agen Jasa Ekspedisi Swasta, Jumlah Menara Telepon Seluler, Kekuatan Sinyal Telepon Seluler	Nominal
59-65	Ekonomi	Bank Umum Pemerintah, Bank Umum Swasta, Bank Perkreditan, Koperasi Unit Desa (KUD), Koperasi Industri Kecil dan Kerajinan Rakyat (KOPINKRA), Koperasi Lainnya	Ordinal
66-71	Kelompok Pertokoan	Pasar dengan Bangunan Permanen, Pasar dengan Bangunan Semi Permanen, Pasar tanpa Bangunan, Mini Market / Swalayan / Supermarket, Restoran/Rumah Makan	Ordinal
72-80	Industri	Kulit, Furnitur Kayu, Barang Logam, Tekstil dan Pakaian Jadi, Barang Galian Bukan Logam, Anyaman Kayu, Makanan dan Minuman, Percetakan, Reparasi Mesin dan Peralatan	Ordinal
81-83	Pemilu	ID TPS, DPT (Daftar Pemilih Tetap), Partisipasi	Ordinal

2.2. Melakukan Pre-Processing

Preprocessing adalah proses mengolah kembali atau memodifikasi dokumen atau data yang telah diolah atau dianalisis. Tahapan ini terdiri dari pembersihan data (*data cleaning*), integrasi data (*data integration*), normalisasi (*data scaling*), dataset acak (*random dataset*) dan korelasi atribut[11].

proses awal *preprocessing* yaitu *data cleaning* berfungsi untuk menangani missing value, hasil dari pemeriksaan ulang *missing value* pada semua atribut ditunjukkan pada gambar 1.

```
Missing Values:
Masjid           0
Mushola          0
Gereja Protestan 0
Gereja Katholik  0
Pura             0
Vihara           0
Partisipasi      0
dtype: int64
```

Gambar 2. *Missing value* pada atribut

Proses kedua yaitu *data integration*, proses menggabungkan data dari berbagai sumber atau kumpulan data[12]. Pada penelitian ini menggabungkan 2 data dari sumber yang berbeda yakni BPS Kabupaten Pasuruan tahun 2022 dengan data partisipasi pemilu 2019 dengan menggunakan excel kemudian di export menjadi file csv lalu dimasukan menjadi dataset di google colaboration.



Gambar 3. Data BPS Kab. Pasuruan 2022 dan *website* pemilu 2019

<https://doi.org/10.31849/digitalzone.v15i1.17027>

Proses ketiga yaitu normalisasi melibatkan perubahan format, struktur, atau penyajian data untuk memenuhi kebutuhan analitis menggunakan fungsi *min-max normalization* dan mean data untuk memudahkan saat proses penelitian[13].

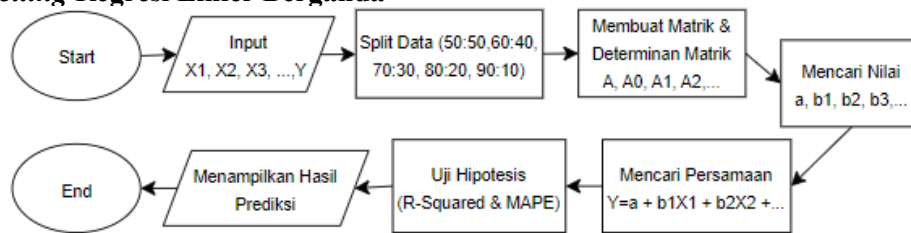
Proses keempat *random dataset* untuk menguji sejauh mana kinerja efektivitas, mengevaluasi dan membandingkan algoritma secara objektif untuk menentukan mana yang memberikan hasil yang lebih baik[14].

Proses terakhir yaitu *Attribute Correlation* mengacu pada hubungan statistik antara dua atau lebih atribut atau variabel dalam kumpulan data. Menggunakan *correlations heatmap* untuk melihat gambaran korelasi atribut pada dataset[15].

Tabel 2. Kategori korelasi

Kategori Korelasi	
0,00 - 0,199	Sangat Rendah
0,20 - 0,399	Rendah
0,40 - 0,599	Sedang
0,60 - 0,799	Kuat
0,80 - 1,00	Sangat Kuat

2.3. Modelling Regresi Linier Berganda



Gambar 4. Langkah-langkah Regresi

Pengujian ini menggunakan metode regresi linier berganda sebagai model untuk mendapatkan nilai prediksi. Sebelum menerapkan metode regresi linier berganda, menentukan terlebih dahulu variabel-variabelnya. Dalam regresi, terdapat dua jenis variabel utama yang mempunyai peran berbeda yaitu variabel terikat (*dependen*) dan variabel bebas (*independen*). Di penelitian ini partisipasi pemilu menjadi variabel dependen karena menjadi variabel target yang ingin diprediksi disimbolkan dengan Y dan atribut yang lain sebagai variabel independen atau variabel prediktor dengan simbol X[16].

Tabel.3 Datasheet Agama

No	Tempat Ibadah						Partisipasi
	Masjid X1	Mushola X2	Gereja Protestan X3	Gereja Katolik X4	Pura X5	Vihara X6	
1	2	3	1	1	1	1	3
2	3	3	1	1	1	1	1
3	1	3	1	1	1	1	2
...
...
4376	1	2	1	1	1	1	2
4377	1	2	1	1	1	1	3

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4382 entries, 0 to 4381
Data columns (total 1 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Partisipasi 4377 non-null   float64
dtypes: float64(1)
memory usage: 34.4 KB
```

Gambar 5. Atribut yang menjadi variabel target

Kemudian dataset akan di split menjadi 2 bagian yaitu data *training* dan data *testing*, *Training set* digunakan untuk melatih model dan mempelajari pola hubungan menggunakan metode regresi. *Testing set* digunakan untuk mengevaluasi seberapa baik suatu model dapat memprediksi data baru. Pada penelitian ini terdapat 2 model cara pengerjaan[17]. Model 1 dengan perbandingan 70:30, sedangkan model 2 dengan perbandingan 50:50, 60:40, 70:30, 80:20 dan 90:10.

```
Training Set Length: 2191 Training Set Length: 2629 Training Set Length: 3067
Testing Set Length : 2191 Testing Set Length : 1753 Testing Set Length : 1315
Training Set Length: 3505 Training Set Length: 2629
Testing Set Length : 877 Testing Set Length : 1753
```

Gambar 6. Split data 50:50, 60:40, 70:30, 80:20 dan 90:10.

Persamaan regresi linier memiliki bentuk umum:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_px_p + \varepsilon \tag{1}$$

Dimana b_0 adalah nilai *intercept* dan $[\beta_1, \beta_2, \dots, \beta_p]$ adalah koefisien yang dihitung dari data *training*. Perhitungan nilai *intercept* dan *coefficients* dapat menggunakan Metode Kuadrat Terkecil (*Least Square Method*) yang dapat disederhanakan menggunakan persamaan matriks.

$$\begin{bmatrix} n & \sum X_1 & \sum X_2 \\ \sum X_1 & \sum X_1^2 & \sum X_1X_2 \\ \sum X_2 & \sum X_1X_2 & \sum X_2^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum Y \\ \sum X_1Y \\ \sum X_2Y \end{bmatrix} \tag{2}$$

2.3. Pengujian Akurasi

Perkiraan kesalahan dalam perumusan timbul bukan hanya dari komponennya saja tetapi juga dari kegagalan model prediksi dalam mendeteksi faktor-faktor lain dalam rangkaian data yang mempengaruhi dalam prediksi[18].

Tahap pengujian akurasi peneliti menggunakan Mean Absolute Percentage Error (MAPE) dan R-squared (R2). MAPE memberikan informasi tentang tingkat kesalahan relatif sementara R2 memberikan gambaran tentang seberapa baik model dapat menjelaskan variasi dalam data.

$$MAPE = \sum_{t=1}^n \left| \frac{y_i - \hat{y}_i}{\hat{y}_i} \right| \times 100\% \tag{3}$$

dimana n adalah jumlah observasi, Y_i adalah nilai sebenarnya, dan \hat{Y}_i adalah nilai prediksi. Semakin rendah nilai MAPE, semakin baik model dalam memprediksi nilai dengan kesalahan yang lebih rendah secara relatif.

$$R^2 = 1 - \frac{SS\ Error}{SS\ Total} = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \tag{4}$$

R-squared berkisar antara 0 dan 1. Nilai yang lebih tinggi menandakan bahwa model memiliki kinerja lebih baik dalam menggambarkan variasi dalam data.

<https://doi.org/10.31849/digitalzone.v15i1.17027>

3. Hasil dan Pembahasan

Penerapan algoritma regresi linier dilakukan pada tahapan ini, Untuk menghasilkan nilai prediksi membutuhkan data *training* dan *Testing* yang digunakan sebagai justifikasi sistem untuk memperoleh persamaan regresi linier. Untuk melihat perbandingan penerapan algoritma regresi linier pada penelitian ini menggunakan dua cara perhitungan dengan 6 skenario pengujian. pertama menggunakan perhitungan manual perbandingan 70:30, kedua menggunakan perhitungan menggunakan bahasa pemrograman python, dengan perbandingan 50:50, 60:40, 70:30, 80:20 dan 90:10.

3.1 Perhitungan Manual

Regresi dapat dilakukan dengan menggunakan perhitungan manual, menggunakan persamaan (2) dilakukan untuk mencari nilai *intercept* dan *coefficients* untuk menghasilkan nilai matriks yang diketahui (A) dan nilai vektor kolom yang diketahui (H).

Tabel 4. Nilai matriks metode kuadrat terkecil

A =	3067	4986	7057	3067	3067	3067	3067	H =	5738
	4986	9242	11858	4986	4986	4986	4986		9319
	7057	11858	17815	7057	7057	7057	7057		13162
	3067	4986	7057	3067	3067	3067	3067		5738
	3067	4986	7057	3067	3067	3067	3067		5738
	3067	4986	7057	3067	3067	3067	3067		5738
	3067	4986	7057	3067	3067	3067	3067		5738

Nilai a dan b diambil dari nilai perhitungan matrix X, Y, XY dan XX. Model persamaan regresi dibentuk dari koefisien dari nilai a dan b yang selanjutnya digunakan untuk mendapatkan persamaan regresi linier berganda

A1	5738	4986	7057	3067	3067	3067	3067	A2	3067	5738	7057	3067	3067	3067	3067	
	9319	9242	11858	4986	4986	4986	4986		4986	9319	11858	17815	4986	4986	4986	4986
	13162	11858	17815	7057	7057	7057	7057		7057	7057	13162	17815	7057	7057	7057	7057
	5738	4986	7057	3067	3067	3067	3067		3067	3067	5738	7057	3067	3067	3067	3067
	5738	4986	7057	3067	3067	3067	3067		3067	3067	5738	7057	3067	3067	3067	3067
	5738	4986	7057	3067	3067	3067	3067		3067	3067	5738	7057	3067	3067	3067	3067
A3	3067	4986	7057	3067	3067	3067	3067	A4	3067	4986	7057	3067	3067	3067	3067	
	4986	9242	11858	4986	4986	4986	4986		4986	9242	11858	17815	4986	4986	4986	4986
	7057	11858	17815	7057	7057	7057	7057		7057	7057	11858	17815	13162	7057	7057	7057
	3067	4986	7057	3067	3067	3067	3067		3067	3067	4986	7057	3067	3067	3067	3067
	3067	4986	7057	3067	3067	3067	3067		3067	3067	4986	7057	3067	3067	3067	3067
	3067	4986	7057	3067	3067	3067	3067		3067	3067	4986	7057	3067	3067	3067	3067
A5	3067	4986	7057	3067	3067	3067	3067	A6	3067	4986	7057	3067	3067	3067	3067	
	4986	9242	11858	4986	4986	4986	4986		4986	9242	11858	17815	4986	4986	4986	4986
	7057	11858	17815	7057	7057	7057	7057		7057	7057	11858	17815	7057	7057	7057	7057
	3067	4986	7057	3067	3067	3067	3067		3067	3067	4986	7057	3067	3067	3067	3067
	3067	4986	7057	3067	3067	3067	3067		3067	3067	4986	7057	3067	3067	3067	3067
	3067	4986	7057	3067	3067	3067	3067		3067	3067	4986	7057	3067	3067	3067	3067
A7	3067	4986	7057	3067	3067	3067	3067	A8	3067	4986	7057	3067	3067	3067	3067	
	4986	9242	11858	4986	4986	4986	4986		4986	9242	11858	17815	4986	4986	4986	4986
	7057	11858	17815	7057	7057	7057	7057		7057	7057	11858	17815	7057	7057	7057	7057
	3067	4986	7057	3067	3067	3067	3067		3067	3067	4986	7057	3067	3067	3067	3067
	3067	4986	7057	3067	3067	3067	3067		3067	3067	4986	7057	3067	3067	3067	3067
	3067	4986	7057	3067	3067	3067	3067		3067	3067	4986	7057	3067	3067	3067	3067

Hasil matriks A menghasilkan nilai

$$B = \begin{vmatrix} 1,930 & 0,010 & -0,026 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 \end{vmatrix} \text{ Meng}$$

gunakan persamaan (1) sehingga dapat menghasilkan nilai sebagai berikut:

$$y = 1,930 + 0,010(x_1) + (-0,026)(x_2) + 0,000(x_3) + 0,000(x_4) + 0,000(x_5) + 0,000(x_6)$$

- Pengujian Akurasi

Tabel 5. summary nilai pengujian

R	R Square	Std. Error of the Estimate
0,055	0.03	0.33498

3.2 Program

Python menjadi bahasa populer untuk analisis data dan ilmu data dengan dukungan pustaka seperti NumPy, Pandas, dan Matplotlib. Alat-alat ini memungkinkan pengolahan, analisis, dan visualisasi data dengan mudah. Pada Gambar 7 serta Gambar 8 menunjukkan hasil korelasi pada atribut agama



Gambar 7. Correlations Heatmap

Berdasarkan Gambar 7 dan 8 dapat dilihat ternyata korelasi atribut selain masjid dan mushola diklasifikasikan tidak memiliki adanya korelasi dengan atribut partisipasi.

	Masjid	Mushola	Gereja Protestan	Gereja Katholik	Pura	Vihara	Partisipasi
Masjid	1.0	0.4	nan	nan	nan	nan	0.0
Mushola	0.4	1.0	nan	nan	nan	nan	-0.0
Gereja Protestan	nan	nan	nan	nan	nan	nan	nan
Gereja Katholik	nan	nan	nan	nan	nan	nan	nan
Pura	nan	nan	nan	nan	nan	nan	nan
Vihara	nan	nan	nan	nan	nan	nan	nan
Partisipasi	0.0	-0.0	nan	nan	nan	nan	1.0

Gambar 8. Korelasi atribut agama

Maka untuk membuktikan dan mengevaluasi model dengan lebih akurat, meminimalkan risiko *overfitting*, dan menghasilkan model yang lebih optimal dibutuhkan perhitungan.

1) Rasio 70:30

Intercept :

[1.89908595]

Coefficients:

[[0.004743 -0.01171564 0. 0. 0. 0.]]

Gambar 9. summary nilai coefficients dan intercept

Menggunakan persamaan (1) sehingga dapat menghasilkan nilai sebagai berikut:

$$y = 1.89908595 + 0.004743(x_1) + (-0.01171564)(x_2) + (0)(x_3) + (0)(x_4) + (0)(x_5) + (0)(x_6)$$

- Pengujian Akurasi

Mean Squared Error: 0.10799287911862852

R-squared : 0.0008783805153274038

Gambar 10. Hasil pengujian 70:30

2) Rasio 80:20

Intercept :
 [1.89729109]
 Coefficients:
 [[0.00469879 -0.01045901 0. 0. 0. 0.]]

Gambar 11. summary nilai coefficients dan intercept

Menggunakan persamaan (1) sehingga dapat menghasilkan nilai sebagai berikut:

$$y = 1.89729109 + 0.00469879(x_1) + (-0.01045901)(x_2) + (0)(x_3) + (0)(x_4) + (0)(x_5) + (0)(x_6)$$

- Pengujian Akurasi

Mean Squared Error: 0.11228823680471636
 R-squared : 0.0005996589353158832

Gambar 12. Hasil pengujian 80:20

3.3 Pembahasan

Hasil analisis data dari BPS Kabupaten Pasuruan dan rekapitulasi pemilihan umum tahun 2019 diperoleh melalui penggunaan metode Regresi Linear Berganda dengan jumlah total 4377 dataset, terdiri dari 7 atribut. Dalam proses analisis ini, dilakukan 2 cara perhitungan dan diuji melalui 6 skenario berbeda. Temuan hasil diperoleh dan disajikan berdasarkan informasi yang tertera pada Tabel 8.

Tabel 6. Hasil pengujian metode regresi linier berganda

Perhitungan	Skenario	Rasio	R-squared	Error of the Estimate
Manual	1	70:30	0.03	0.33498
	2	50:50	-0.0007	0.10334
	3	60:40	0.00043	0.10332
Program	4	70:30	0.0009	0.10799
	5	80:20	0.0006	0.11228
	6	90:10	0.00012	0.09934

Penelitian yang dilaksanakan oleh Osbin Samosir dan Indah Novitasari hanya mengeksplorasi isu seputar pengalaman politik identitas, dengan fokus pada bagaimana faktor agama menjadi target dalam konteks pemilihan nasional Indonesia pada tahun 2014 dan 2019. Dengan demikian, temuan ini mengkonfirmasi bahwa agama dalam konteks tempat ibadah juga termasuk salah satu faktor yang mempengaruhi partisipasi pemilu, dengan bukti adanya perhitungan dari data yang sudah diperoleh.

Dari hasil pengujian, Skenario keenam (R-squared: 0.00012, MSE: 0.09934) memiliki nilai R-squared yang relatif lebih tinggi dan MSE yang lebih rendah, yang menunjukkan adanya peningkatan dalam akurasi model tersebut dibandingkan dengan skenario lainnya. Namun, perlu diingat bahwa nilai-nilai R-squared yang sangat rendah seperti ini menunjukkan bahwa model belum mampu menjelaskan variasi yang signifikan dalam data. Oleh karena itu, meskipun satu skenario mungkin memiliki kinerja yang sedikit lebih baik, hasil prediksi dari model-model ini

mungkin memiliki tingkat ketidakpastian yang tinggi dalam konteks prediksi partisipasi pemilu berdasarkan agama.

4. Kesimpulan

Dalam konteks penelitian pengaruh agama pada tingkat partisipasi pemilu di Kabupaten Pasuruan menggunakan metode regresi linier, hasil evaluasi model menunjukkan bahwa skenario keenam memiliki kinerja tertinggi dengan nilai E-squared yang lebih tinggi dan Mean Squared Error (MSE) yang lebih rendah. Meskipun demikian, penerapan metode regresi linier berganda belum sepenuhnya memberikan hasil optimal dalam memprediksi dampak agama terhadap tingkat partisipasi pemilu. Temuan ini menyoroti bahwa agama dalam konteks tempat ibadah tidak menjadi faktor yang signifikan dalam menentukan jumlah partisipasi pemilu, dan penelitian ini juga mencatat bahwa faktor lain yang belum diteliti. Oleh karena itu, hipotesis yang menyatakan "Agama Berpengaruh dalam Partisipasi Pemilu di Kabupaten Pasuruan" secara parsial tidak dapat diterima. Kesimpulan ini menunjukkan kompleksitas dinamika yang mempengaruhi partisipasi pemilu di wilayah tersebut, dengan potensi faktor lain yang perlu diperhatikan dalam analisis mendalam.

Temuan ini menunjukkan perlunya penelitian lebih lanjut dan penyesuaian model untuk meningkatkan kapasitas variabilitas dalam partisipasi pemilu. Mengenai saran untuk penelitian selanjutnya yakni meneliti lebih lanjut dengan menggunakan variabel lain dan pengembangan metode atau menggunakan metode yang lain untuk mengetahui hasil persentase yang lebih baik.

Daftar Pustaka

- [1] J. Penelitian, P. Pancasila, D. Kewarganegaraan, and D. N. Ramadhanti, "De Cive: Demokrasi dan Pemilu Indonesia dalam Tinjauan Sosiologi Politik," vol. 2, no. 9, pp. 361–366, 2022, [Online]. Available: <https://doi.org/10.56393/decive.v2i9.1677>.
- [2] I. P. A. P. Wibawa, I. K. A. Purnawan, D. P. S. Putri, and N. K. D. Rusjyanthi, "Prediksi Partisipasi Pemilih dalam Pemilu Presiden 2014 dengan Metode Support Vector Machine," *J. Ilm. Merpati (Menara Penelit. Akad. Teknol. Informasi)*, vol. 7, no. 3, p. 182, 2019, [doi: 10.24843/jim.2019.v07.i03.p02](https://doi.org/10.24843/jim.2019.v07.i03.p02).
- [3] A. Intani, "Pemilu Serentak Tahun 2024: Prediksi Partisipasi Politik Masyarakat Mengacu Pada Perbandingan Anggaran Kampanye Pemilu Serentak Tahun 2019," *Madani J. Ilm. Multidisiplin*, vol. 1, no. 8, pp. 1–13, 2023, [Online].
- [4] A. S. Fitriani, "Penerapan Data Mining Menggunakan Metode Klasifikasi Naïve Bayes untuk Memprediksi Partisipasi Pemilihan Gubernur," *JTAM (Jurnal Teor. dan Apl. Mat.)*, vol. 3, no. 2, pp. 98–104, 2019. <https://doi.org/10.31764/jtam.v3i2.995>
- [5] O. Samosir and I. Novitasari, "Hak Politik Warga Negara Dalam Cengkeraman Politik Identitas: Refleksi Menuju Pemilu Serentak Nasional Tahun 2024," *J. Ilmu Hukum, Hum. dan Polit.*, vol. 2, no. 3, pp. 332–346, 2022., <https://doi.org/10.38035/jihhp.v2i3.1052>
- [6] E. Trianggorowati and R. Al-Hamdi, "Jurnal Tapis: Jurnal Teropong Aspirasi Politik Islam 16 (1) (2020) 65-82," *Strateg. Kampanye Partai Islam Kasus Partai Keadilan Sejaht. pada Pemilu 2019*, vol. 16 (1), no. 1, pp. 65–82, 2020, [Online]. Available: <https://http://ejournal.radenintan.ac.id/index.php/TAPIS/index>. <http://dx.doi.org/10.24042/tps.v16i1.6814>
- [7] P. Purwadi, P. S. Ramadhan, and N. Safitri, "Penerapan Data Mining Untuk Mengestimasi Laju Pertumbuhan Penduduk Menggunakan Metode Regresi Linier Berganda Pada BPS Deli Serdang," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 18, no. 1, p. 55, 2019, [doi: 10.53513/jis.v18i1.104](https://doi.org/10.53513/jis.v18i1.104). <https://doi.org/10.53513/jis.v18i1.104>
- [8] Y. Goktua Siadari and D. Saripuna, "Data Mining Untuk Mengestimasi Jumlah Penumpang Pada Pt. Pinem Lau Guna Medan Dengan Menggunakan Metodere Gresi Linear Berganda," *J. CyberTech*, vol. x, No.x, no. x, 2020. <https://doi.org/10.31849/digitalzone.v15i1.17027>

- <https://doi.org/10.53513/jct.v4i2.1998>
- [9] M. C. Pulungan, M. Rahmatunnisa, and A. G. Herdiansyah, "Strategi KPU Kota Bekasi dalam meningkatkan Partisipatif Pemilih Disabilitas," *J. Polit. Islam*, vol. 3, no. 2, pp. 251–272, 2020. <https://doi.org/10.20414/politea.v3i2.2439>
- [10] A. N. Latifah, M. Sulistiyono, A. Sidauruk, B. Satria, and M. T. Nurcholis, "Prediksi Curah Hujan Menggunakan Algoritma Regresi Linear Berganda," vol. 23, pp. 39–44, 2023.
- [11] M. Sholeh, R. Y. Rachmawati, and E. N. Cahyo, "Penerapan Regresi Linear Ganda Untuk Memprediksi Hasil Nilai Kuesioner Mahasiswa Dengan Menggunakan Python," *J. Din. Inform.*, vol. 11, no. 1, pp. 13–24, 2022, [Online]. Available: <https://jdi.upy.ac.id/index.php/jdi/article/view/124>.
- [12] Achmad Saiful Rizal and Moch. Lutfi, "Prediksi Hasil Pemilu Legislatif Menggunakan Algoritma K-Nearest Neighbor Berbasis Backward Elimination," *J. Resist. (Rekayasa Sist. Komputer)*, vol. 3, no. 1, pp. 27–41, 2020, doi: 10.31598/jurnalresistor.v3i1.517. <https://doi.org/10.31598/jurnalresistor.v3i1.517>
- [13] M. P. Handayani, A. Azhar, and V. C. G. Hura, "Perhitungan Nilai Ekuivalensi Mobil Penumpang Dengan Metode Regresi Linier Berganda Pada Ruas Jalan Thamrin Kota Padang," *Ekasakti J. Penelit. Pengabd.*, vol. 2, no. 2, pp. 106–121, 2022. <https://doi.org/10.31933/ejpp.v2i2.545>
- [14] D. Selent, T. Patikorn, and N. Heffernan, "ASSIStments dataset from multiple randomized controlled experiments," *L@S 2016 - Proc. 3rd 2016 ACM Conf. Learn. Scale*, pp. 181–184, 2016, doi: 10.1145/2876034.2893409.
- [15] S. V., S. Chandurkar, and S. Bansode, "Attribute Selection to Improve Accuracy of Classification," *Int. J. Comput. Appl.*, vol. 173, no. 5, pp. 18–22, 2017, doi: 10.5120/ijca2017915117.
- [16] Y. S. Bhakti, A. B. Kusdinar, and A. A. Sunarto, "Model Peramalan Penerimaan Calon Mahasiswa Menggunakan Metode Regresi," *Progresif J. Ilm. Komput.*, vol. 16, no. 2, p. 113, 2020, doi: 10.35889/progresif.v16i2.515.
- [17] A. A. Khan, "Balanced Split: A new train-test data splitting strategy for imbalanced datasets," 2022, [Online]. Available: <http://arxiv.org/abs/2212.11116>. <https://doi.org/10.48550/arXiv.2212.11116>
- [18] G. N. Ayuni and D. Fitriana, "Penerapan metode Regresi Linear untuk prediksi penjualan properti pada PT XYZ," *J. Telemat.*, vol. 14, no. 2, pp. 79–86, 2019, [Online]. Available: <https://doi.org/10.61769/telematika.v14i2.321>
-