



Predicting Unemployment in West Nusa Tenggara Using Supervised Machine Learning-Based Multiple Linear Regression

Prediksi Pengangguran di Nusa Tenggara Barat Menggunakan Regresi Linear Berganda Berbasis *Supervised Learning*

Rukmiati^{1*}, Siti Mutmainah², Khairunnas³

^{1,2,3}Program Studi Ilmu Komputer, Universitas Muhammadiyah Bima, Indonesia

E-Mail: ¹rukmiatihz10@gmail.com, ²siti.mutmainah.id19@gmail.com ³assignment.khairunnas@gmail.com

Received May 30th 2026; Revised Jun 26th 2026; Accepted Jul 15th 2026; Available Online Jul 27th 2026

Corresponding Author: Rukmiati

Copyright © 2025 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

Abstract

The Open Unemployment Rate (OUR) is one of the key indicators for measuring labor market conditions and public welfare. This study aims to analyze the effects of the Human Development Index (HDI), the Poverty Rate, and Inflation on the OUR in West Nusa Tenggara Province and to develop a predictive model using Multiple Linear Regression. The data used are secondary data obtained from the Central Statistics Agency (BPS) of West Nusa Tenggara Province. The research stages include data collection, preprocessing, exploratory data analysis (EDA), testing of classical assumptions, modeling, and model evaluation. The results of the classical assumption tests indicate that the model satisfies the assumptions of normality, multicollinearity, heteroscedasticity, and autocorrelation. To strengthen the contribution of machine learning, a comparison was conducted between Multiple Linear Regression and Decision Tree Regression. The evaluation results show that Multiple Linear Regression achieved an MSE of 0.4357, an RMSE of 0.6601, and an R^2 of 0.3226, while Decision Tree Regression achieved an MSE of 0.5442, an RMSE of 0.7377, and an R^2 of 0.1540. Based on these results, Multiple Linear Regression produces lower error rates and better predictive ability than Decision Tree Regression. Thus, Multiple Linear Regression was selected as the best model for predicting the Open Unemployment Rate in West Nusa Tenggara Province.

Keywords: Inflation, Linear Regression, Machine Learning, Prediction, Unemployment

Abstrak

Tingkat Pengangguran Terbuka (TPT) merupakan salah satu indikator penting dalam mengukur kondisi ketenagakerjaan dan kesejahteraan masyarakat. Penelitian ini bertujuan untuk menganalisis pengaruh Indeks Pembangunan Manusia (IPM), Tingkat Kemiskinan, dan Inflasi terhadap TPT di Provinsi Nusa Tenggara Barat serta membangun model prediksi menggunakan Regresi Linear Berganda. Data yang digunakan merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Nusa Tenggara Barat. Tahapan penelitian meliputi pengumpulan data, preprocessing, analisis data eksploratif (EDA), pengujian asumsi klasik, pemodelan, dan evaluasi model. Hasil pengujian asumsi klasik menunjukkan bahwa model telah memenuhi asumsi normalitas, multikolinearitas, heteroskedastisitas, dan autokorelasi. Untuk memperkuat kontribusi machine learning, dilakukan perbandingan antara Regresi Linear Berganda dan *Decision Tree Regression*. Hasil evaluasi menunjukkan bahwa Regresi Linear Berganda memperoleh nilai MSE sebesar 0,4357, RMSE sebesar 0,6601, dan R^2 sebesar 0,3226, sedangkan *Decision Tree Regression* memperoleh nilai MSE sebesar 0,5442, RMSE sebesar 0,7377, dan R^2 sebesar 0,1540. Berdasarkan hasil tersebut, Regresi Linear Berganda menghasilkan tingkat kesalahan yang lebih rendah dan kemampuan prediksi yang lebih baik dibandingkan dengan *Decision Tree Regression*. Dengan demikian, Regresi Linear Berganda dipilih sebagai model terbaik untuk memprediksi Tingkat Pengangguran Terbuka di Provinsi Nusa Tenggara Barat.

Kata Kunci: Inflasi, Machine Learning, Pengangguran, Prediksi, Regresi Linear

1. PENDAHULUAN

Provinsi Nusa Tenggara Barat (NTB) adalah salah satu wilayah di Indonesia yang berkembang pesat, namun dalam beberapa tahun terakhir dilanda ketidakstabilan ekonomi, meskipun berbagai sektor seperti pariwisata, pertanian dan industri kreatif mengalami pertumbuhan, tingkat pengangguran di NTB masih



menjadi tantangan utama dalam pembangunan sosial dan ekonomi daerah. Data Badan Pusat Statistik NTB mencatat bahwa banyak angkatan kerja pada Agustus Tahun 2025 mencapai 3,20 juta orang, terjadi kenaikan sebanyak 12,10 ribu orang dibanding Tahun 2024, Ketidakseimbangan pertumbuhan ini menyebabkan peningkatan jumlah pengangguran menjadi 97,93 ribu orang, atau naik 10,92 ribu orang dari tahun sebelumnya [1]. Kondisi ketika angkatan kerja belum terserap pasar kerja meskipun aktif mencari pekerjaan, yang diukur melalui Tingkat Pengangguran Terbuka (TPT), menunjukkan bahwa ketersediaan lapangan kerja belum mampu mengimbangi pertumbuhan angkatan kerja. [2].

Pengangguran adalah kondisi ketika seseorang yang termasuk dalam angkatan kerja tidak memiliki pekerjaan, meskipun secara fisik dan mental masih mampu untuk bekerja serta memiliki keinginan untuk bekerja. Individu tersebut aktif mencari pekerjaan, bersedia menerima pekerjaan yang sesuai, atau sedang menunggu kesempatan kerja yang tersedia [3]. Permasalahan pengangguran menjadi semakin kompleks karena tidak hanya dipengaruhi oleh ketersediaan lapangan kerja, tetapi juga oleh kualitas tenaga kerja, pertumbuhan sektor informal serta dinamika upah minimum. Penelitian-penelitian sebelumnya menunjukkan bahwa faktor-faktor seperti kemiskinan, pertumbuhan ekonomi, dan pendidikan dapat memengaruhi tingkat kemiskinan. Namun, hasil empiris tersebut tidak selalu dapat diterapkan pada semua wilayah geografis [4].

Prediksi merupakan suatu proses analitis yang bertujuan untuk memperkirakan nilai atau kondisi suatu variabel target di masa yang akan datang dengan memanfaatkan pola hubungan yang diperoleh dari data historis. Proses ini dilakukan melalui pembentukan model matematis atau algoritma berbasis statistik dan *machine learning* yang mampu memetakan keterkaitan antara variabel independen dan variabel dependen secara sistematis. Dalam konteks data *mining*, prediksi tidak bersifat spekulatif, melainkan didasarkan pada proses komputasi yang mencakup tahapan pra-pemrosesan data, pembelajaran pola, serta evaluasi kinerja model menggunakan metrik kuantitatif tertentu. Dengan demikian, prediksi dapat dipahami sebagai estimasi terukur yang dihasilkan melalui pendekatan ilmiah untuk mendukung pengambilan keputusan berbasis data [5].

Pendekatan dengan *Regresi Linear Berganda* memungkinkan peneliti untuk memahami pola keterkaitan antarvariabel secara kuantitatif serta menghasilkan prediksi yang terukur. Penelitian-penelitian membuktikan bahwa model regresi linear dapat digunakan untuk memprediksi tingkat pengangguran maupun variabel lain yang berkaitan dengan pasar tenaga kerja, dengan tingkat kinerja yang memadai. Keberhasilan penerapan model tersebut sangat dipengaruhi oleh kualitas dan validitas data yang digunakan, serta ketepatan dalam melakukan evaluasi model melalui metrik pengukuran seperti *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan koefisien determinasi (R^2) [10].

Regresi Linear berganda memiliki kelebihan dalam menganalisis hubungan multivariat pada variabel-variabel ekonomi, karena dapat mengukur besarnya kontribusi masing-masing variabel independen terhadap variabel dependen model regresi mampu memberikan gambaran yang lebih menyeluruh mengenai pola dan tren kemiskinan di suatu daerah, sehingga dapat mendukung proses evaluasi kebijakan ekonomi [6]. Penelitian menunjukkan bahwa tingkat kemiskinan berkaitan erat dengan indikator pembangunan ekonomi daerah, seperti Produk Domestik Regional Bruto (PDRB), Indeks Pembangunan Manusia (IPM), dan kondisi ekonomi makro lainnya. Hal ini menegaskan bahwa kemiskinan tidak hanya menjadi persoalan ekonomi, tetapi juga memengaruhi tingkat kesejahteraan masyarakat secara keseluruhan [6].

Dalam analisis data ekonomi, metode statistik dan *machine learning* semakin banyak digunakan untuk memodelkan fenomena sosial ekonomi, termasuk kemiskinan. Salah satu metode yang umum diterapkan adalah regresi Linear berganda karena mampu menganalisis hubungan antara variabel terikat dan beberapa variabel bebas sekaligus guna mengidentifikasi faktor yang berpengaruh serta memprediksi perubahan di masa mendatang [7]. Sebagian besar penelitian tentang kemiskinan di Indonesia masih berfokus pada hubungan antarvariabel dan belum mengembangkan pendekatan prediktif secara komprehensif. Padahal, prediksi tingkat kemiskinan penting untuk membantu pemerintah merancang kebijakan yang lebih tepat sasaran melalui model regresi berbasis indikator ekonomi sebagai dasar perencanaan pembangunan daerah [8].

Penelitian sebelumnya hanya menggunakan metode statistik konvensional tanpa menguji aspek performa prediksi dan hasil penelitian berbasis model prediktif. Penggunaan pendekatan *Machine Learning* dalam prediksi pengangguran di tingkat daerah masih relatif terbatas, termasuk di NTB [9]. Penelitian oleh Halif [7] menunjukkan bahwa Regresi Linear Berganda efektif dalam memprediksi tingkat pengangguran menggunakan variabel inflasi, PDRB, IPM, dan jumlah pengangguran, dengan hasil evaluasi menggunakan RMSE dan *R-squared*, dengan hasil RMSE 0,0148 dan R^2 0,5716. Penelitian oleh Arianto [8] menegaskan bahwa Regresi Linear Berganda dapat digunakan memprediksi jumlah pengangguran berdasarkan jumlah penduduk usia kerja.

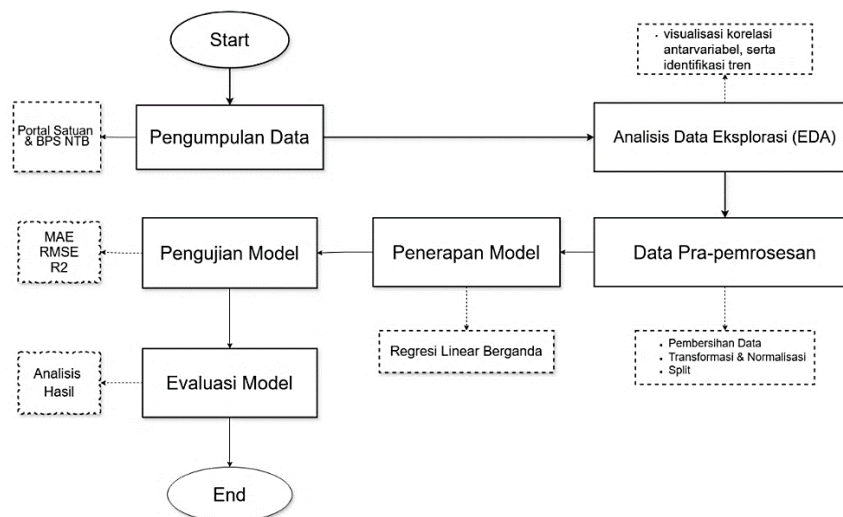
Penelitian mengenai prediksi tingkat pengangguran dengan pendekatan *machine learning* telah berkembang di berbagai negara. Annastasya dkk. (2025) menggunakan beberapa algoritma *machine learning*, seperti Linear Regression, Random Forest, Support Vector Regression, dan Gradient Boosting untuk memprediksi tingkat pengangguran di Indonesia dengan memanfaatkan indikator makroekonomi. Hasil penelitian menunjukkan bahwa model Gradient Boosting memberikan performa prediksi terbaik. Selain itu, berbagai penelitian internasional juga telah menerapkan algoritma *machine learning* seperti *Random Forest*,

Support Vector Machine (SVM), *Neural Network*, serta *Long Short-Term Memory (LSTM)* dalam memprediksi tingkat pengangguran dan memperoleh hasil yang lebih akurat dibandingkan dengan metode statistik konvensional. Temuan tersebut menunjukkan bahwa *machine learning* memiliki kemampuan yang baik dalam menangkap hubungan nonLinear antar variabel ekonomi sehingga semakin banyak digunakan dalam penelitian peramalan pasar tenaga kerja dan tingkat pengangguran [10].

Berdasarkan kompleksitas dinamika ketenagakerjaan di Provinsi Nusa Tenggara Barat, diperlukan pendekatan kuantitatif yang mampu menghasilkan prediksi tingkat pengangguran secara terukur melalui pemanfaatan indikator ekonomi. Penelitian ini berfokus pada pengembangan model prediksi tingkat pengangguran dengan menggunakan metode Regresi Linear Berganda berbasis *supervised learning* dengan rumusan masalah, yaitu bagaimana membangun model prediksi pengangguran serta bagaimana mengevaluasi kinerjanya menggunakan indikator *Mean Absolute Error (MAE)*, *Root Mean Squared Error (RMSE)*, dan *Coefficient of Determination (R²)*. Oleh karena itu, penelitian ini diarahkan untuk menghasilkan model prediktif yang reliabel, terukur, dan tervalidasi sebagai instrumen berbasis data guna mendukung perencanaan tenaga kerja dan mitigasi risiko peningkatan pengangguran di Provinsi Nusa Tenggara Barat.

2. BAHAN DAN METODE

Rancangan penelitian disusun melalui serangkaian tahapan yang dilakukan secara sistematis dan berurutan, dimulai dari proses pengumpulan data hingga analisis hasil penelitian. Penelitian ini menggunakan platform Google Colab dengan bahasa pemrograman Python. Hasil yang diperoleh bersifat berskala uji dan dimanfaatkan untuk kepentingan analisis akademik. Adapun tahapan penelitian dijelaskan melalui alur penelitian berikut.



Gambar 1. Alur Penelitian

2.1. Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan melalui penerapan metode pengumpulan data sekunder yang bersumber dari instansi resmi, yaitu Badan Pusat Statistik (BPS) Provinsi Nusa Tenggara Barat serta Portal Satuan Data Daerah (Data Portal) Nusa Tenggara Barat. Data yang digunakan dalam penelitian ini mencakup periode 2016–2025 dan merupakan data panel yang meliputi semua Kabupaten/Kota di Provinsi Nusa Tenggara Barat.

2.2. Analisis Data Eksplorasi / *Exploratory Data Analysis (EDA)*

Exploratory Data Analysis (EDA) merupakan tahap awal dalam analisis data yang dilakukan untuk memahami karakteristik dataset, mengidentifikasi pola serta hubungan antarvariabel, dan mendeteksi anomali sebelum pemodelan lebih lanjut [11]. Melalui tahap ini, peneliti dapat memperoleh gambaran awal mengenai struktur data dan potensi variabel yang berpengaruh terhadap tingkat pengangguran di Provinsi Nusa Tenggara Barat. Proses EDA dalam penelitian ini meliputi analisis statistik deskriptif, visualisasi korelasi antarvariabel, serta identifikasi tren spasial dan temporal. Tahap EDA ini bertujuan memastikan bahwa data yang digunakan dalam pemodelan telah dipahami secara menyeluruh, serta untuk mengidentifikasi potensi permasalahan data sebelum memasuki tahap preprocessing dan pembangunan model [11].

2.3. *Preprocessing Data*

Pre-processing merupakan tahap untuk mengubah data mentah menjadi format yang siap digunakan algoritma serta memastikan data bebas dari *noise*, *missing values*, redundansi, dan inkonsistensi agar

meningkatkan akurasi serta stabilitas model [12]. Tahap preprocessing dilakukan untuk memastikan kualitas dan kesiapan data sebelum pemodelan regresi Linear berganda untuk mengurangi bias, meminimalkan kesalahan, dan meningkatkan stabilitas model. Tahapan yang dilakukan meliputi:

1. *Data Cleaning*: Tahap ini meliputi identifikasi dan penanganan *missing values*, penghapusan data duplikat, pemeriksaan konsistensi tipe data, serta deteksi outlier untuk memastikan kualitas dan kestabilan data sebelum analisis.
2. *Transformasi dan Normalisasi Data*: Proses ini bertujuan menyamakan skala antarvariabel untuk mencegah dominasi nilai besar serta meningkatkan stabilitas model selama pelatihan.
3. *Pembagian Data*: Dataset dibagi secara proporsional. Data latih 80% dan data uji 20% dari seluruh dataset untuk melatih serta mengevaluasi model sekaligus menghindari *overfitting*.

2.4. Penerapan Model

Tahap pemodelan merupakan inti penelitian yang berfokus pada pembangunan model Regresi Linear Berganda untuk menganalisis hubungan variabel independen terhadap Tingkat Pengangguran Terbuka (TPT) di Provinsi Nusa Tenggara Barat melalui proses pemodelan dan evaluasi performa prediksi. Model yang diterapkan dalam penelitian ini merupakan Regresi Linear Berganda, sebab metode ini bisa menarangkan ikatan antara satu variabel dependen dengan lebih dari satu variabel leluasa terhadap variabel terikat secara serentak. Regresi Linear Berganda juga memberikan interpretasi koefisien yang jelas untuk melihat arah dan kekuatan hubungan antar variabel bebas terhadap tingkat pengangguran [13]. Regresi linear digunakan untuk membentuk persamaan prediktif dalam memprediksi pengangguran berdasarkan hubungan antara variabel independen yang memengaruhinya [14]. Adapun rumus yang digunakan dijelaskan pada Persamaan 1.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon \quad (1)$$

Dimana :

- Y : TPT (Tingkat Pengangguran Terbuka)
 $X_1, X_2, \dots, X_n$: Variabel independen
 β_0 : Intersep model (nilai rata-rata Y ketika semua $X_i = 0$)
 $\beta_1, \beta_2, \dots, \beta_n$: Koefisien regresi untuk setiap variabel independen, (pengaruh masing-masing variabel terhadap Y).
 ε : Error term (perbedaan antara nilai aktual dan prediksi) [15]

2.5. Pengujian Model

Model kemudian diuji menggunakan data uji untuk mengetahui kemampuan prediksi di luar data pelatihan. Evaluasi menggunakan *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), dan *Root Mean Squared Error* (RMSE). Penggunaan ketiga metrik ini memberikan gambaran yang komprehensif mengenai akurasi prediksi serta ukuran deviasi kesalahan model terhadap nilai actual [16].

Dalam evaluasi performa prediktif, nilai hasil prediksi model dibandingkan dengan nilai aktual menggunakan beberapa ukuran kesalahan (*error metrics*), yaitu :

1. *Mean Absolute Error* (MAE) merupakan salah satu evaluasi yang diterapkan untuk mengukur rata-rata nilai absolut dari perbedaan antara nilai aktual dan nilai prediksi yang dihasilkan oleh model [17]. Secara matematis, nilai MAE dinyatakan dalam bentuk Persamaan 2.

$$MAE = \frac{\sum (y_n - \hat{y}_n)^2}{n} \quad (2)$$

2. *Root Mean Squared Error* (RMSE) diterapkan untuk mengukur rata-rata besaran kesalahan antara nilai aktual dan nilai yang dihasilkan oleh model prediktif melalui akar kuadrat dari rata-rata selisih kuadrat keduanya [18]. Secara matematis, RMSE dinyatakan pada Persamaan 3.

$$RMSE = \sqrt{\frac{\sum (y_n - \hat{y}_n)^2}{n}} \quad (3)$$

3. *Koefisien determinasi* (R^2) dimanfaatkan untuk mengukur tingkat kemampuan variabel independen dalam menjelaskan variasi atau fluktuasi yang terjadi pada variabel dependen. Hal tersebut didasarkan pada prinsip bahwa semakin dekat nilai R^2 mendekati 1 (satu), maka akan mengindikasikan bahwa performansi model dalam menjelaskan variabilitas data secara bersamaan juga semakin tinggi [17]. Nilai (R^2) dihitung menggunakan Persamaan 4.

$$R^2 = 1 - \frac{SS_{error}}{SS_{total}} \quad (4)$$

2.6. Evaluasi

Hasil pengujian model dianalisis menggunakan nilai RME, RMSE dan *R-squared* untuk mengukur tingkat keakuratan prediksi terhadap data aktual. Apabila performa model belum optimal, perbaikan dapat dilakukan melalui *preprocessing* tambahan, penambahan variabel, atau penerapan metode lain guna meningkatkan kualitas prediksi [16].

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Pengumpulan data dalam penelitian diperoleh dari sumber resmi yaitu Badan Pusat Statistik (BPS) Provinsi Nusa Tenggara Barat <https://ntb.bps.go.id/id>. dan Portal Satuan Data NTB <https://data.ntbprov.go.id/>. Data yang digunakan dalam penelitian ini berupa data panel yang mencakup seluruh kabupaten/kota di Provinsi Nusa Tenggara Barat selama periode 2016–2025. Meskipun data memiliki struktur panel, penelitian ini menggunakan metode Regresi Linear Berganda (*Ordinary Least Squares/OLS*) dengan pendekatan data gabungan. Pemilihan metode ini didasarkan pada tujuan penelitian yang berfokus pada identifikasi hubungan dan pengaruh variabel independen terhadap tingkat pengangguran serta pembangunan model prediksi yang sederhana dan mudah diinterpretasikan.

```
data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 12 columns):
 #   Column                                     Non-Null Count  Dtype
---  -
 0   id_tpt                                   100 non-null    int64
 1   kode_kabupaten_kota                     100 non-null    int64
 2   nama_kabupaten_kota                     100 non-null    object
 3   tahun                                   100 non-null    int64
 4   Tingkat Pengangguran Terbuka (TPT) (%) 100 non-null    float64
 5   Jumlah Penduduk (jiwa)                  100 non-null    int64
 6   PDRB per Kapita (juta rupiah)           100 non-null    object
 7   Indeks Pembangunan Manusia (IPM)        100 non-null    float64
 8   Tingkat Kemiskinan (%)                  100 non-null    float64
 9   Tingkat Partisipasi Angkatan Kerja (TPAK) (%) 100 non-null    float64
10   Upah Minumun Kota/Kab (Rupiah)         100 non-null    object
11   Inflasi (%)                             100 non-null    float64
dtypes: float64(5), int64(4), object(3)
memory usage: 9.5+ KB
```

Gambar 2. Data Pengangguran

Gambar 2 menampilkan dataset yang digunakan dalam penelitian, yang memuat informasi tingkat pengangguran, jumlah penduduk, PDRB, IPM, Tingkat Kemiskinan, Inflasi serta indikator ekonomi lainnya pada kabupaten/kota di Nusa Tenggara Barat periode 2016–2025, dengan total 100 baris data dan 11 variabel.

3.2. Analisis Data Eksplorasi

Tahap analisis data meliputi statistik, analisis korelasi dan tren yang bertujuan untuk mengidentifikasi dan memahami karakteristik, hubungan antarvariabel, pola sebaran, serta mengamati dinamika perubahan data secara temporal.

1. Pengujian Asumsi Klasik

Pengujian asumsi klasik dilakukan untuk memastikan model memenuhi asumsi statistik yang diperlukan. Pengujian yang dilakukan meliputi uji normalitas, multikolinearitas, heteroskedastisitas, dan autokorelasi. Hasil pengujian digunakan untuk memastikan bahwa model regresi yang dibangun menghasilkan prediksi yang valid dan dapat diinterpretasikan dengan baik.

a. Uji Multikolinearitas

Hasil uji *Shapiro-Wilk* menunjukkan nilai signifikansi sebesar 0,092 ($> 0,05$), sehingga residual model berdistribusi normal dan memenuhi asumsi normalitas.

b. Uji Normalitas

Uji multikolinearitas dilakukan untuk mengetahui ada tidaknya korelasi yang tinggi antarvariabel independen. Hasil pengujian menunjukkan bahwa variabel Indeks Pembangunan Manusia (IPM), Tingkat Kemiskinan, dan Inflasi memiliki nilai VIF masing-masing sebesar 8,67; 2,70; dan 7,41. Seluruh nilai VIF berada di bawah batas 10, sehingga dapat disimpulkan bahwa tidak terdapat gejala multikolinearitas yang signifikan pada model regresi.

c. Uji Heteroskedastisitas

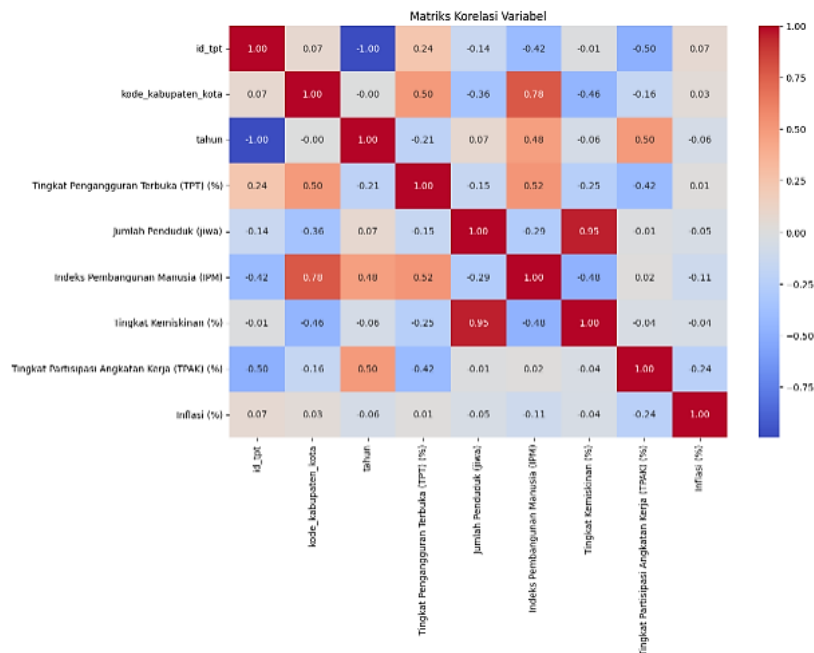
Hasil uji heteroskedastisitas menggunakan metode Breusch-Pagan menghasilkan nilai signifikansi (p-value) sebesar 0,584. Nilai tersebut lebih besar dari taraf signifikansi 0,05, sehingga dapat disimpulkan bahwa model regresi tidak mengalami gejala heteroskedastisitas.

d. Uji Autokorelasi

Nilai Durbin-Watson sebesar 1,81 menunjukkan tidak adanya autokorelasi karena nilainya mendekati 2.

2. Analisis Korelasi

Matriks korelasi merupakan tabel yang digunakan untuk menggambarkan hubungan Linear antarvariabel melalui koefisien korelasi Pearson dengan rentang nilai -1 hingga +1. Koefisien korelasi dengan nilai positif mengindikasikan adanya koneksi antara variabel-variabel yang berkorelasi. Sebaliknya, koefisien korelasi dengan nilai negatif menandakan adanya hubungan berlawanan arah (negatif) di antara variabel-variabel tersebut. Sementara itu, koefisien korelasi yang bernilai nol (0) menunjukkan bahwa tidak terdapat hubungan linear yang signifikan antarvariabel [19].



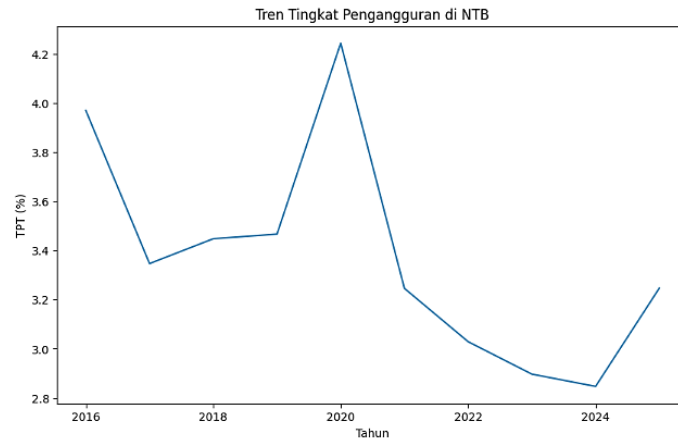
Gambar 3. Korelasi Matriks

Hasil analisis korelasi menunjukkan hubungan positif terkuat antara jumlah penduduk dan TPAK (0,95), serta hubungan negatif terkuat antara IPM dan tingkat kemiskinan (-0,90). TPT memiliki korelasi sedang dengan IPM, kemiskinan, dan TPAK, sedangkan inflasi menunjukkan hubungan yang sangat lemah terhadap sebagian besar variabel. Secara umum, IPM, kemiskinan, jumlah penduduk, dan TPAK menjadi variabel dengan hubungan paling kuat. Berdasarkan hasil uji VIF, seluruh variabel memiliki nilai VIF di bawah 10 sehingga tidak terdapat multikolinearitas yang serius dalam model. Meskipun korelasi antara jumlah penduduk dan TPAK tergolong sangat kuat, hasil VIF menunjukkan bahwa kedua variabel masih berada dalam batas yang dapat diterima untuk digunakan dalam pemodelan regresi.

3. Tren Data TPT

Analisis tren tingkat pengangguran dilakukan dengan tujuan untuk mengidentifikasi tren dan pola yang menjadi ciri perkembangan tingkat kemiskinan tersebut seiring berjalannya waktu. Tren Tingkat Pengangguran Terbuka ditunjukkan pada Gambar 4.

Berdasarkan grafik tren Tingkat Pengangguran Terbuka (TPT) di Provinsi Nusa Tenggara Barat (NTB) tahun 2016–2025, terlihat bahwa tingkat pengangguran mengalami fluktuasi dengan kecenderungan menurun. TPT mengalami penurunan dari 3,97% pada tahun 2016 menjadi 3,35% pada tahun 2017, kemudian relatif stabil hingga tahun 2019. Pada tahun 2020 terjadi peningkatan signifikan hingga mencapai 4,24%, yang diperkirakan dipengaruhi oleh implikasi pandemi *Coronavirus Disease 2019* (COVID-19) terhadap dinamika sektor ketenagakerjaan. Setelah itu, TPT terus menurun dari 3,24% pada tahun 2021 menjadi 2,85% pada tahun 2024. Namun, pada tahun 2025 terjadi kenaikan kembali menjadi 3,25%. Secara umum, tren pengangguran di NTB menunjukkan perbaikan kondisi ketenagakerjaan pascapandemi meskipun masih terdapat fluktuasi pada akhir periode pengamatan.



Gambar 4. Tren Tingkat Pengangguran Terbuka

3.3. Preprocessing Data

Tahap *pre-processing* meliputi penanganan nilai yang hilang (*missing value*), penghapusan atribut-atribut yang tidak relevan, transformasi atau konversi data, serta pembagian dataset ke dalam kelompok yang telah ditentukan.

1. Data Cleaning

Tahap data *cleaning* dilakukan untuk meminimalkan kesalahan pada data melalui identifikasi *error*, perbaikan data yang tidak lengkap, serta pemilihan data yang layak digunakan agar kualitas data lebih akurat dan konsisten [20].

```
data.isnull().sum()
```

	0
id_tpt	0
kode_kabupaten_kota	0
nama_kabupaten_kota	0
tahun	0
Tingkat Pengangguran Terbuka (TPT) (%)	0
Jumlah Penduduk (jiwa)	0
PDRB per Kapita (juta rupiah)	0
Indeks Pembangunan Manusia (IPM)	0
Tingkat Kemiskinan (%)	0
Tingkat Partisipasi Angkatan Kerja (TPAK) (%)	0
Upah Minimum Kota/Kab (Rupiah)	0
Inflasi (%)	0

Gambar 5. Hasil Data *cleaning*

Tahap data *cleaning* menunjukkan bahwa yang digunakan bebas dari nilai yang hilang (*missing values*) maupun data yang terduplikasi. Seluruh variabel telah memiliki tipe data yang sesuai dan dapat digunakan langsung dalam proses analisis. Beberapa nilai ekstrem teridentifikasi pada variabel tertentu, namun tetap dipertahankan karena merepresentasikan kondisi aktual wilayah penelitian.

2. Transformasi dan Normalisasi Data

Transformasi data dilakukan menggunakan metode *Min-Max Scaling* untuk mengubah seluruh variabel ke dalam rentang nilai 0–1. Proses ini bertujuan menyamakan skala antarvariabel sehingga tidak terjadi dominasi variabel yang memiliki nilai lebih besar serta meningkatkan stabilitas model selama proses pelatihan. Metode *Min-Max Scaling* dipilih karena mampu mempertahankan distribusi dan hubungan antardata asli tanpa mengubah pola data secara signifikan. Selain itu, metode ini cocok digunakan pada algoritma *machine learning* yang sensitif terhadap perbedaan skala data. Secara matematis, normalisasi *Min-Max Scaling* dihitung menggunakan Persamaan 5.

$$X_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (5)$$

Keterangan:

X_{norm} : nilai hasil normalisasi

x : nilai data asli
 x_{min} : nilai minimum pada variabel
 x_{max} : nilai maksimum pada variabel

Dengan persamaan tersebut, seluruh nilai data akan ditransformasikan ke rentang 0 hingga 1 sehingga setiap variabel memiliki kontribusi yang lebih seimbang dalam proses pemodelan.

3. Pembagian Data

Dataset dibagi menjadi data latih sebesar 80% dan data uji sebesar 20%. Data latih digunakan untuk membangun model prediksi pelatihan, sedangkan data uji digunakan untuk menilai performa model terhadap data yang belum pernah terekspos selama proses pelatihan. Pembagian ini bertujuan untuk meningkatkan kemampuan generalisasi model serta mengurangi risiko overfitting.

Tabel 1. Split Data

Data Test	Data Uji
80	20

3.4. Penerapan Model

3.4.1. Regresi Linear Berganda

Proses evaluasi model dilakukan untuk menilai kapabilitas model Regresi Linear Berganda dalam memprediksi tingkat pengangguran terhadap data yang belum pernah terekspos selama proses pelatihan (*training process*). Pengujian ini memanfaatkan data testing yang telah dipisahkan pada tahap pembagian dataset, sehingga model dapat diuji terhadap data baru untuk menilai tingkat akurasi, kemampuan generalisasi, serta konsistensi hasil prediksi yang dihasilkan. Melalui tahap ini, performa model dapat dianalisis untuk mengetahui sejauh mana model mampu merepresentasikan pola data dengan tepat.

Tabel 2. Nilai Koefisien dan intercept

Variabel	Koefisien	p-value	Keterangan
Jumlah Penduduk	1.06E-06	0,259	Tidak Signifikan
PDRB per Kapita	6.55E-06	0,001	Signifikan
IPM	0,1690	0,000	Signifikan
Tingkat Kemiskinan	-0,0047	0,506	Tidak Signifikan
TPAK	-0,0665	0,006	Signifikan
Upah Minimum	-1.48E-03	0,000	Signifikan
Inflasi	0,0141	0,817	Tidak Signifikan

Berdasarkan hasil uji t, tidak seluruh variabel independen berpengaruh signifikan terhadap Tingkat Pengangguran Terbuka (TPT). Variabel yang terbukti signifikan pada taraf signifikansi 5% adalah PDRB per Kapita, Indeks Pembangunan Manusia (IPM), Tingkat Partisipasi Angkatan Kerja (TPAK), dan Upah Minimum Kabupaten/Kota. Sementara itu, variabel Jumlah Penduduk, Tingkat Kemiskinan, dan Inflasi tidak menunjukkan pengaruh yang signifikan terhadap TPT. Hasil ini menunjukkan bahwa interpretasi koefisien regresi tidak hanya didasarkan pada besarnya dan arah koefisien, tetapi juga mempertimbangkan signifikansi statistik masing-masing variabel.

3.4.2. Decision Tree Regression

Selain Regresi Linear Berganda, penelitian ini juga menerapkan algoritma *Decision Tree Regression* sebagai model pembanding. *Decision Tree Regression* merupakan metode *machine learning* yang membentuk struktur pohon keputusan melalui proses pembagian data secara bertahap (*recursive splitting*) untuk menghasilkan prediksi nilai target. Model ini mampu menangkap hubungan nonLinear antara variabel sehingga berpotensi memberikan hasil prediksi yang lebih baik pada data dengan pola yang kompleks.

Proses pelatihan dan pengujian model *Decision Tree* menggunakan pembagian data yang sama dengan model Regresi Linear Berganda, yaitu data training dan data testing. Penggunaan dataset yang sama bertujuan agar hasil evaluasi kedua model dapat dibandingkan secara objektif. Selanjutnya, performa model dievaluasi menggunakan metrik *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), dan koefisien determinasi (R^2) untuk mengukur tingkat kesalahan prediksi serta kemampuan model dalam menjelaskan variasi data Tingkat Pengangguran Terbuka (TPT).

3.5. Pengujian Model

Proses pengujian model dilakukan untuk mengevaluasi kemampuan regresi Linear berganda dan juga menggunakan *Decision Tree Regression* sebagai perbandingan dalam memprediksi rasio penurunan pada data baru dengan menggunakan pengujian data yang telah dipisahkan pada tahap pembagian dataset.

Tabel 3. Hasil Pengujian Model Regresi Linear Berganda

Kabupaten Kota	Actual	Prediksi
Kabupaten Lombok Barat	2,82	3.713464
Kabupaten Lombok Tengah	3,36	2.908950
Kabupaten Lombok Timur	2,82	3.663963
Kabupaten Sumbawa	2,6	3.357933
Kabupaten Dompu	2,84	3.198514
Kabupaten Bima	2,13	3.003346
Kabupaten Sumbawa Barat	4,13	4.861946
Kabupaten Lombok Utara	2,68	2.228737
Kota Mataram	4,8	5.982075
Kota Bima	4,28	4.428572

Hasil pengujian model prediksi menunjukkan bahwa model mampu menggambarkan pola tingkat pengangguran antarkabupaten dan kota di Provinsi Nusa Tenggara Barat, meskipun masih terdapat selisih antara nilai aktual dan nilai prediksi. Sebagian besar hasil prediksi cenderung lebih tinggi dibandingkan dengan nilai aktual (*overestimate*), yang terlihat pada 8 dari 10 wilayah pengamatan. Prediksi paling akurat terdapat pada Kota Bima dengan selisih sebesar 0,15%, sedangkan kesalahan prediksi terbesar terjadi pada Kota Mataram dengan selisih sebesar 1,18%. Secara keseluruhan, model telah mampu mengidentifikasi wilayah dengan tingkat pengangguran tinggi dan rendah, namun masih memerlukan penyempurnaan untuk meningkatkan akurasi prediksi dan mengurangi bias *overestimasi*.

Tabel 4. Hasil Pengujian Model *Decision Tree Regression*

Kabupaten Kota	Actual	Prediksi
Kabupaten Lombok Barat	2,82	3,29
Kabupaten Lombok Tengah	3,36	2,72
Kabupaten Lombok Timur	2,82	2,72
Kabupaten Sumbawa	2,6	3,29
Kabupaten Dompu	2,84	3,30
Kabupaten Bima	2,13	2,72
Kabupaten Sumbawa Barat	4,13	5,36
Kabupaten Lombok Utara	2,68	1,97
Kota Mataram	4,8	5,22
Kota Bima	4,28	3,29

Berdasarkan Tabel 4, model *Decision Tree Regression* menghasilkan nilai prediksi yang relatif mendekati nilai aktual pada beberapa kabupaten/kota, namun masih terdapat perbedaan yang cukup besar pada beberapa wilayah. Sebagai contoh, pada Kabupaten Lombok Barat nilai aktual Tingkat Pengangguran Terbuka (TPT) sebesar 2,82%, sedangkan model memprediksi sebesar 3,29%. Di Kabupaten Lombok Tengah, nilai aktual sebesar 3,36% dan diprediksi sebesar 2,72%. Sementara itu, di Kota Mataram nilai aktual sebesar 4,80% dan diprediksi sebesar 5,22%.

Secara umum, model *Decision Tree* mampu mengikuti pola sebaran data TPT pada masing-masing kabupaten/kota, namun beberapa hasil prediksi menunjukkan deviasi yang cukup besar dari nilai aktual. Hal ini mengindikasikan bahwa model belum mampu menangkap seluruh karakteristik data secara optimal. Selain itu, terlihat bahwa beberapa wilayah memiliki nilai prediksi yang sama, seperti 2,72% dan 3,29%, yang merupakan karakteristik algoritma *Decision Tree* karena prediksi dihasilkan berdasarkan pembagian data ke dalam kelompok-kelompok tertentu (leaf node).

Hasil tersebut menunjukkan bahwa meskipun *Decision Tree* dapat digunakan untuk memprediksi Tingkat Pengangguran Terbuka (TPT), tingkat akurasinya masih terbatas sehingga menghasilkan kesalahan prediksi yang relatif lebih besar dibandingkan dengan nilai aktual pada beberapa wilayah pengamatan.

Tabel 5. Akurasi Prediksi

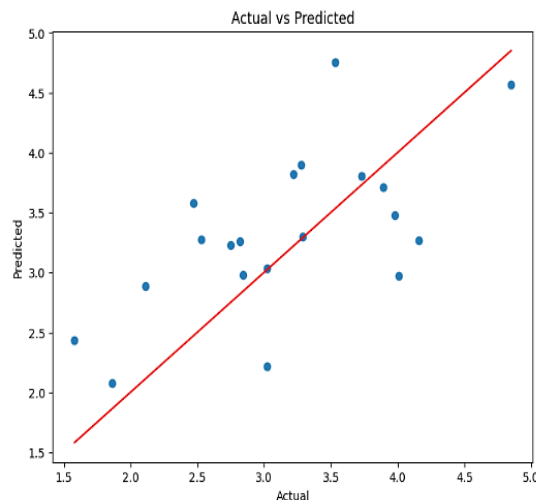
Model	MSE	RMSE	R ²
Regresi Linear Berganda	0.43573	0.66010	0.32256
<i>Decision Tree Regression</i>	0.54416	0.73767	0.15400

Hasil evaluasi menunjukkan bahwa model Regresi Linear Berganda menghasilkan nilai MSE sebesar 0,4357, lebih rendah dibandingkan dengan *Decision Tree Regression* yang menghasilkan nilai 0,5442. Nilai MSE yang lebih kecil menunjukkan bahwa rata-rata kesalahan prediksi pada model Regresi Linear Berganda lebih rendah dibandingkan dengan model *Decision Tree Regression*.

Berdasarkan nilai RMSE, Regresi Linear Berganda memperoleh nilai 0,6601, sedangkan *Decision Tree Regression* memperoleh nilai 0,7377. Hasil ini menunjukkan bahwa tingkat penyimpangan prediksi terhadap

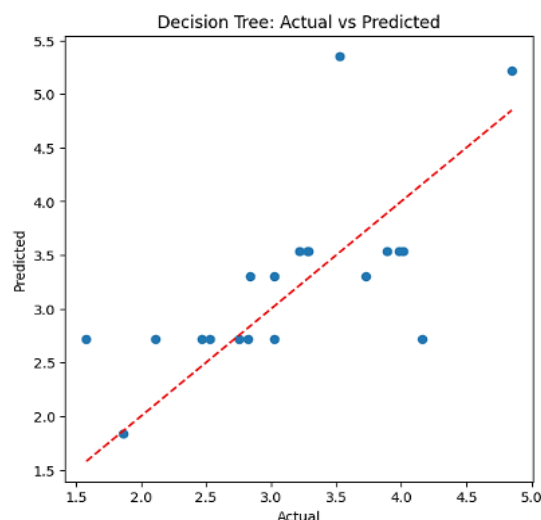
nilai aktual pada Regresi Linear Berganda lebih kecil, sehingga model mampu menghasilkan prediksi yang lebih akurat. Selain itu, nilai koefisien determinasi (R^2) pada Regresi Linear Berganda sebesar 0,3226 atau 32,26%, sedangkan *Decision Tree Regression* hanya sebesar 0,1540 atau 15,40%. Hal ini menunjukkan bahwa Regresi Linear Berganda memiliki kemampuan yang lebih baik dalam menjelaskan variasi data Tingkat Pengangguran Terbuka (TPT) dibandingkan dengan *Decision Tree Regression*.

Perbedaan performa tersebut mengindikasikan bahwa hubungan antara variabel independen yang digunakan, yaitu Indeks Pembangunan Manusia (IPM), Tingkat Kemiskinan, dan Inflasi, terhadap Tingkat Pengangguran Terbuka cenderung mengikuti pola Linear. Oleh karena itu, model Regresi Linear Berganda lebih sesuai untuk merepresentasikan hubungan antarvariabel pada dataset penelitian dibandingkan dengan model *Decision Tree Regression* yang umumnya lebih efektif untuk data dengan pola nonLinear yang kompleks.



Gambar 6. Scatter plot Prediksi vs Realita Regresi Linear

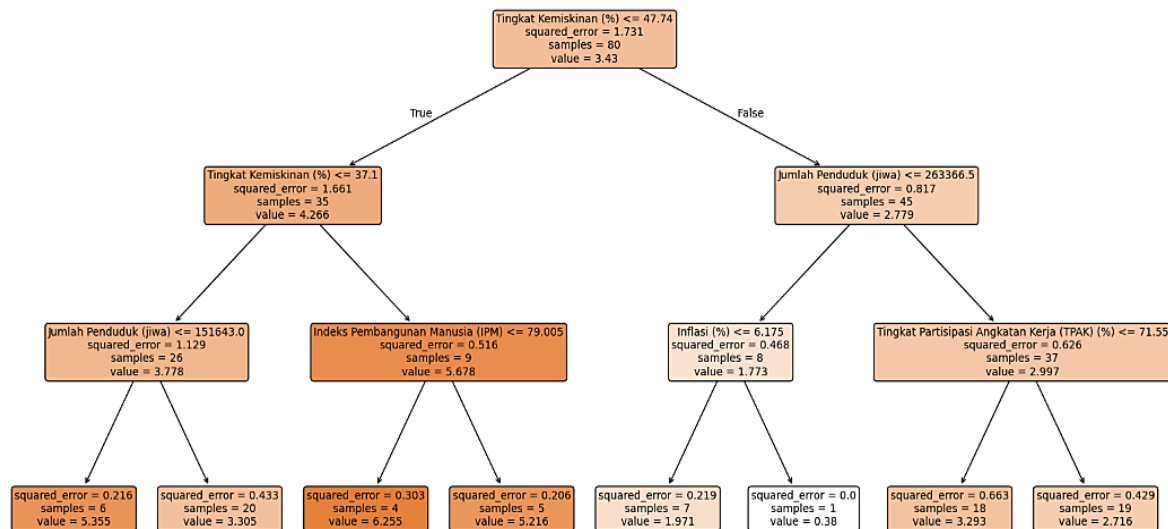
Hasil *scatter plot Actual vs Predicted* pada Gambar 6 menunjukkan bahwa sebagian besar nilai prediksi berada di sekitar garis diagonal, yang menandakan bahwa hasil prediksi model relatif mendekati nilai aktual. Meskipun terdapat beberapa penyimpangan berupa overestimasi dan underestimasi, pola sebaran titik yang cukup mengikuti garis menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam memprediksi Tingkat Pengangguran Terbuka (TPT).



Gambar 7. Scatter plot Prediksi vs Realita *Decision Tree Regression*

Gambar 7 menunjukkan perbandingan antara nilai aktual dan nilai prediksi yang dihasilkan oleh model *Decision Tree Regression*. Garis merah putus-putus merupakan garis ideal ($y = x$), yang menunjukkan kondisi ketika nilai prediksi sama dengan nilai aktual. Sebagian besar titik data berada cukup dekat dengan garis tersebut, yang menandakan bahwa model mampu memprediksi data dengan cukup baik.

Namun, masih terdapat beberapa titik yang menyimpang dari garis ideal, yang menunjukkan adanya kesalahan prediksi (*error*). Hal ini mengindikasikan bahwa model belum sepenuhnya mampu menangkap seluruh pola dalam data. Secara keseluruhan, *Decision Tree Regression* menunjukkan performa yang cukup baik dalam memprediksi tingkat pengangguran, meskipun masih terdapat beberapa perbedaan antara nilai aktual dan nilai prediksi.



Gambar 8. Visualisasi *Decision Tree Regression*

Berdasarkan hasil pemodelan *Decision Tree Regression* pada Gambar 8, variabel yang paling berpengaruh dalam memprediksi Tingkat Pengangguran Terbuka (TPT) adalah Tingkat Kemiskinan (%), yang ditunjukkan sebagai akar pohon keputusan dengan ambang nilai 47,74%. Selanjutnya, model melakukan pemisahan data berdasarkan variabel Jumlah Penduduk, Indeks Pembangunan Manusia (IPM), Inflasi, dan Tingkat Partisipasi Angkatan Kerja (TPAK). Nilai prediksi TPT tertinggi sebesar 6,255 ditemukan pada kelompok wilayah dengan tingkat kemiskinan di atas 37,1% dan $IPM \leq 79,005$, sedangkan nilai prediksi terendah sebesar 0,38 terdapat pada kelompok wilayah dengan tingkat kemiskinan di atas 47,74%, jumlah penduduk $\leq 263.366,5$ jiwa, dan inflasi $> 6,175\%$. Hasil ini menunjukkan bahwa tingkat kemiskinan merupakan faktor dominan yang memengaruhi TPT, diikuti oleh kondisi demografis dan indikator pembangunan daerah lainnya.

3.6 Evaluasi

Hasil evaluasi model Regresi Linear Berganda menunjukkan performa yang lebih baik dibandingkan dengan *Decision Tree Regression* dalam memprediksi Tingkat Pengangguran Terbuka (TPT) di Provinsi Nusa Tenggara Barat. Regresi Linear Berganda menghasilkan nilai MSE sebesar 0,4357, RMSE sebesar 0,6601, dan R^2 sebesar 0,3226, sedangkan *Decision Tree Regression* memperoleh nilai MSE sebesar 0,5442, RMSE sebesar 0,7377, dan R^2 sebesar 0,1540. Nilai MSE dan RMSE yang lebih rendah pada Regresi Linear Berganda menunjukkan bahwa tingkat kesalahan prediksi model lebih kecil dibandingkan dengan *Decision Tree Regression*. Selain itu, nilai R^2 yang lebih tinggi menunjukkan bahwa regresi Linear berganda mampu menjelaskan variasi data tingkat pengangguran terbuka dengan lebih baik.

Hasil perbandingan ini mengindikasikan bahwa hubungan antara variabel independen yang digunakan, yaitu Indeks Pembangunan Manusia (IPM), Tingkat Kemiskinan, dan Inflasi, terhadap Tingkat Pengangguran Terbuka cenderung mengikuti pola Linear. Oleh karena itu, Regresi Linear Berganda lebih sesuai digunakan untuk memodelkan data penelitian dibandingkan dengan *Decision Tree Regression*. Berdasarkan keseluruhan hasil evaluasi, Regresi Linear Berganda terpilih sebagai model terbaik karena memiliki tingkat kesalahan prediksi yang lebih rendah serta kemampuan prediksi yang lebih baik dibandingkan *Decision Tree Regression* dalam memprediksi Tingkat Pengangguran Terbuka di Provinsi Nusa Tenggara Barat.

4. KESIMPULAN

Penelitian ini bertujuan untuk menganalisis dan memprediksi Tingkat Pengangguran Terbuka (TPT) di Provinsi Nusa Tenggara Barat menggunakan metode Regresi Linear Berganda dengan variabel independen berupa Indeks Pembangunan Manusia (IPM), Tingkat Kemiskinan, dan Inflasi. Berdasarkan hasil analisis korelasi dan pengujian asumsi klasik, model yang dibangun telah memenuhi asumsi regresi yang meliputi uji multikolinearitas, normalitas, heteroskedastisitas, dan autokorelasi, sehingga layak digunakan untuk proses pemodelan. Hasil regresi menunjukkan bahwa variabel IPM, Tingkat Kemiskinan, dan Inflasi memiliki

hubungan dengan Tingkat Pengangguran Terbuka, dengan kemampuan model dalam menjelaskan variasi data sebesar 32,26%.

Untuk memperkuat kontribusi machine learning, penelitian ini juga melakukan komparasi dengan algoritma *Decision Tree Regression*. Hasil evaluasi menunjukkan bahwa Regresi Linear Berganda memperoleh nilai MSE sebesar 0,4357, RMSE sebesar 0,6601, dan R^2 sebesar 0,3226, sedangkan *Decision Tree Regression* memperoleh nilai MSE sebesar 0,5442, RMSE sebesar 0,7377, dan R^2 sebesar 0,1540. Berdasarkan perbandingan tersebut, Regresi Linear Berganda menghasilkan tingkat kesalahan prediksi yang lebih rendah dan kemampuan prediksi yang lebih baik dibandingkan dengan *Decision Tree Regression*. Dengan demikian, dapat disimpulkan bahwa Regresi Linear Berganda merupakan model yang paling sesuai dan memberikan performa terbaik dalam memprediksi Tingkat Pengangguran Terbuka (TPT) di Provinsi Nusa Tenggara Barat pada dataset yang digunakan dalam penelitian ini.

REFERENCES

- [1] BPS, "Keadaan Ketenagakerjaan Nusa Tenggara Barat Agustus 2025," *Statistik*, 2025.
- [2] A. S. Darmawan and M. N. Mifrahi, "Jurnal Kebijakan Ekonomi dan Keuangan Analisis tingkat pengangguran terbuka di Indonesia periode sebelum dan saat pandemi covid-19," vol. 1, no. 1, pp. 111–118, 2022.
- [3] S. Frisnoiry, H. M. W. Sihotang, N. Indri, and T. Munthe, "Analisis Permasalahan Pengangguran Di Indonesia," vol. 17, no. 1, 2024.
- [4] S. D. Sari and E. Setyowati, "Procedia of Social Sciences and Humanities Analisis Pengangguran , Pendapatan Perkapita , Dan IPM Terhadap Pertumbuhan Ekonomi di Indonesia Tahun 2017- Analysis of Unemployment , Capita Income , and HDI on Economic Growth in Indonesia , 2017-2020 . Procedia of Social Sciences and Humanities," vol. 0672, no. c, pp. 8–18, 2022.
- [5] R. Y. Endra and U. Rizal, "Teknik Prediksi Data Mining pada Perguruan Tinggi sebagai Kajian Literatur," no. 2.
- [6] D. K. Malum Tahan Manaon Siregar, "Analisis Faktor-Faktor Yang Mempengaruhi Tingkat Pengangguran Di Kabupaten Dairi Menggunakan Regresi Linear Berganda," vol. 12, no. 03, pp. 697–703, 2024.
- [7] J. Halif, D. Wahiddin, I. Sanjaya, and S. Faisal, "Model Regresi Linear Berganda untuk Prediksi Tingkat Pengangguran di Provinsi Jawa Barat," pp. 324–335, 2025, doi: 10.33364/algoritma/v.22-1.2312.
- [8] A. S. Singarimbun and D. B. Arianto, "Prediksi Jumlah Pengangguran Di Indonesia Berdasarkan Data," vol. 4, no. 2, pp. 58–62, 2024.
- [9] M. W. Musthofa *et al.*, "Efek peningkatan upah minimum terhadap tingkat pengangguran pada kabupaten kota di nusa tenggara barat (ntb) analisi data panel pada tahun 2014-2023," vol. 8, pp. 260–275, 2025.
- [10] T. Annastasya, R. Passarella, and Z. Yamani, "Unemployment rate forecasting in Indonesia using macroeconomic indicators with a machine learning approach," 2025.
- [11] M. Ogryzek, "Applying Methods of Exploratory Data Analysis and Methods of Modeling the Unemployment Rate in Spatial Terms in Poland," pp. 1–18, 2025.
- [12] R. Rachman *et al.*, "Penerapan Regresi Linear Berganda Dalam Prediksi Dan Optimalisasi Persediaan Barang Toko Mungil," vol. 8, no. 5, pp. 10499–10506, 2024.
- [13] A. N. Maharadja, I. Maulana, and B. A. Dermawan, "Penerapan Metode Regresi Linear Berganda untuk Prediksi Kerugian Negara Berdasarkan Kasus Tindak Pidana Korupsi," vol. 5, no. 1, pp. 95–102, 2021.
- [14] J. Riset, T. Informasi, J. Penelitian, M. Jppm, A. Juli, and P. Handayani, "Analisis Regresi Linear Berganda Pada Hubungan Antara Tinggi Badan , Berat Badan Dan Usia Terhadap Perkembangan Potensi Akademik Anak Usia Dini (Studi Kasus : Ra - Adduriyat)," vol. 1, no. 3, pp. 88–94, 2024.
- [15] R. N. T. Siregar, "Analisis Prediksi Harga Rumah di Bandung Menggunakan Regresi Linear Berganda," vol. 1, no. 6, 2023.
- [16] D. Chicco, M. J. Warrens, and G. Jurman, "The coef fi cient of determination R-squared is more informative than SMAPE , MAE , MAPE , MSE and RMSE in regression analysis evaluation," pp. 1–24, 2021, doi: 10.7717/peerj-cs.623.
- [17] B. S. Qolbi, N. Puteh, and C. Rozikin, "Prediksi Harga Rumah Di Jakarta Pusat Menggunakan Algoritma Machine Learning," vol. 16, no. 1, pp. 16–24, 2025.
- [18] H. Rezkian, A. Dama, A. A. Supianto, and N. Y. Setiawan, "Analisis Penggunaan Model Regresi untuk Prediksi Penjualan Spare Part pada AHASS Nur Andhita Grogol," vol. 5, no. 12, pp. 5591–5603, 2021.
- [19] C. Haryanto, N. Rahaningsih, F. M. Basysyar, K. Cirebon, R. F. Regression, and H. Rumah, "KOMPARASI ALGORITMA MACHINE LEARNING DALAM MEMPREDIKSI HARGA RUMAH," vol. 1, no. 1, pp. 533–539, 2023.
- [20] S. A. Zega, A. Saropi, T. Ronaldy, Y. Ilhami, and M. Farabi, "Mengatasi Data Error Pada Proses Data Cleaning Motion Capture Motive Optitrack," vol. 6, no. 1, pp. 89–95, 2022.