



## Analisis Perbandingan Prediksi Tingkat Kemiskinan Menggunakan Metode XGBoost dan Random Forest Regression

Isnan Wisnu Prastiyo, Arafat Febriandirza\*

Fakultas Teknologi Industri dan Informatika, Teknik Informatika, Universitas Muhammadiyah Prof. Dr. Hamka, Jakarta, Indonesia

Email: <sup>1</sup><mailto:isnanpras@gmail.com>, <sup>2\*</sup><mailto:arafat@uhamka.ac.id>

Email Penulis Korespondensi: <mailto:arafat@uhamka.ac.id>

**Abstrak**—Penelitian ini bertujuan untuk membandingkan kinerja dua algoritma prediksi, XGBoost Regression dan Random Forest Regression, dalam memprediksi tingkat kemiskinan di wilayah DKI Jakarta. Untuk penelitian ini peneliti memperoleh data dari Badan Pusat Statistik (BPS) DKI Jakarta yang mencakup periode tahun 2010 hingga 2023. Metode pengujian yang digunakan melibatkan pengukuran Mean Squared Error (MSE) dan Mean Absolute Percentage Error (MAPE) untuk menilai keakuratan prediksi dari kedua algoritma. Temuan menunjukkan bahwa algoritma Random Forest Regression umumnya menghasilkan prediksi yang lebih akurat dibandingkan dengan algoritma XGBoost Regression terlihat dari hasil pengujian pada (MSE) dan (MAPE) pada untuk sebagian besar wilayah yang dianalisis. Seperti pada MAPE ini untuk wilayah Jakarta Barat hasil pengujiannya pada XGBoost Regression ini sebesar 1,43 sementara Random Forest Regression menghasilkan 1,42 maka lebih baik Random Forest Regression dari pada XGBoost Regression. Akan tetapi pada Kepulauan Seribu ini MAPE untuk XGBoost lebih baik dengan nilai sebesar 4,49 dari pada Random Forest Regression yang nilai sebesar 4,56. Lalu pada MSE Random Forest lebih baik dari pada XGBoost ini dalam pengujian prediksi ini. Seperti pada wilayah Jakarta Pusat dengan nilai sebesar 0,02 untuk XGBoost Regression ini, sementara Random Forest Regression mempunyai hasil pengujian lebih kecil dengan nilai 0,01.

**Kata Kunci:** Prediksi; XGBoost Regression; Random Forest Regression; Tingkat Kemiskinan; DKI Jakarta; MSE; MAPE

**Abstract**—This research aims to compare the performance of two prediction algorithms, XGBoost Regression and Random Forest Regression, in predicting poverty levels in the DKI Jakarta area. For this research, researchers obtained data from the DKI Jakarta Central Statistics Agency (BPS) covering the period 2010 to 2023. The testing method used involved measuring Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE) to assess the accuracy of predictions from the two algorithms. The findings show that the Random Forest Regression algorithm generally produces more accurate predictions compared to the XGBoost Regression algorithm as seen from the test results on (MSE) and (MAPE) for most of the areas analyzed. As with MAPE for the West Jakarta area, the test results for XGBoost Regression were 1.43, while Random Forest Regression produced 1.42, so Random Forest Regression is better than XGBoost Regression. However, in the Seribu Islands, the MAPE for XGBoost is better with a value of 4.49 than for Random Forest Regression which has a value of 4.56. Then MSE Random Forest is better than XGBoost in this prediction test. For example, in the Central Jakarta area with a value of 0.02 for XGBoost Regression, while Random Forest Regression has a smaller test result with a value of 0.01.

**Keywords:** Prediction; XGBoost Regression; Random Forest Regression; Poverty Level; DKI Jakarta; MSE; MAPE

### 1. PENDAHULUAN

Kemiskinan merupakan definisi kelompok atau individu dalam masyarakat menjadi kurang sejahtera dan berjuang untuk memenuhi seluruh kebutuhan dasarnya [1]. Kemiskinan sebagai penyakit sosial ekonomi tidak hanya terjadi di negara berkembang, bahkan negara maju pun masih menghadapi permasalahan kemiskinan [2]. Kemiskinan di negara-negara berkembang ialah masalah yang rumit, akan tetapi beberapa di antaranya sudah membuat kemajuan dalam hal produksi dan pendapatan nasional. Tingkat kemiskinan sebuah negara atau wilayah mencerminkan kesejahteraan masyarakat yang tinggal di sana [3]. Kemiskinan telah lama menjadi masalah serius dalam kehidupan manusia. Hal ini tidak hanya mempengaruhi individu, namun pula pada kelompok masyarakat, daerah, negara, dan bahkan populasi dunia secara keseluruhan [4]. Setiap provinsi di Indonesia memiliki jumlah penduduk miskin, namun tingkat kemiskinan bervariasi antar provinsi. Sebagai contoh, DKI Jakarta memiliki garis kemiskinan yang merupakan yang tertinggi ketiga di Indonesia, menunjukkan bahwa biaya di Jakarta untuk mencukupi kebutuhan dasar hidup layak lebih tinggi yang berbeda dari beberapa provinsi di Indonesia [5].

Salah satu kota terbesar di Indonesia ialah Jakarta. Yang merupakan sebagai pusat politik, ekonomi, dan budaya negara. Jakarta memiliki jumlah penduduk yang sangat besar dan beragam. Kemiskinan di Jakarta merupakan tantangan sosial yang cukup serius dengan tingkat kemiskinan di Jakarta yang dapat dilihat melalui berbagai macam indikator seperti tingkat pengangguran, tingkat pendapatan, dan tingkat kesejahteraan masyarakat. Hingga saat ini, DKI Jakarta masih menghadapi tantangan kemiskinan. Pertumbuhan perkotaan yang pesat di Jakarta telah membuatnya menjadi salah satu kota besar dengan kepadatan penduduk tinggi. Fenomena ini dipengaruhi oleh aktivitas sosial, ekonomi, dan bisnis yang beroperasi dalam skala regional, nasional, bahkan global [6]. Beberapa provinsi di Pulau Jawa mengalami tingkat kemiskinan yang signifikan, dan data kemiskinan dari tahun 2018 hingga 2022 menjadi perhatian serius. DKI Jakarta, misalnya, memiliki tingkat kemiskinan sebesar 4,19% [7]. Dengan jumlah penduduk Jakarta yang terus meningkat karena lebih banyak orang yang pindah ke Jakarta daripada yang pindah keluar, hal ini menjadi faktor utama masalah di Jakarta yang berkontribusi pada tingginya tingkat kemiskinan. Banyak dari mereka datang ke Jakarta dengan harapan memiliki pekerjaan dan kehidupan yang lebih baik, tetapi tidak semua dari mereka berhasil mencapainya [8]. Kemiskinan di Jakarta



berdampak sosial yang signifikan terhadap kota itu sendiri. Menurut informasi dari Gramedia.com pada 3 November 2022, dampak kemiskinan termasuk peningkatan tingkat putus sekolah pada usia yang seharusnya masih wajib sekolah, masalah kesehatan masyarakat lainnya, serta peningkatan kriminalitas di kota-kota besar, terutama Jakarta [9]. Selain memonitor tingkat kemiskinan, pembahasan mengenai kerentanan terhadap kemiskinan juga sering diperbincangkan karena kemiskinan merupakan masalah yang berubah-ubah dan dapat menimbulkan dampak ekonomi yang signifikan, yang berpotensi mengarahkan banyak rumah tangga ke dalam kondisi kemiskinan [10].

Prediksi atau bisa disebut dengan forecasting, adalah metode perhitungan yang berguna dalam memprakirakan kejadian di masa depan dengan menggunakan data atau informasi dari masa lalu dengan tujuan untuk mengurangi tingkat error Dengan memberikan hasil terbaik [11]. Salah satu dari fungsi prediksi adalah untuk membantu dalam mengambil sebuah keputusan karena prediksi dapat mengurangi kesalahan yang disebabkan oleh kesalahan perencanaan dan dapat membantu dalam membuat keputusan untuk masa depan [12].

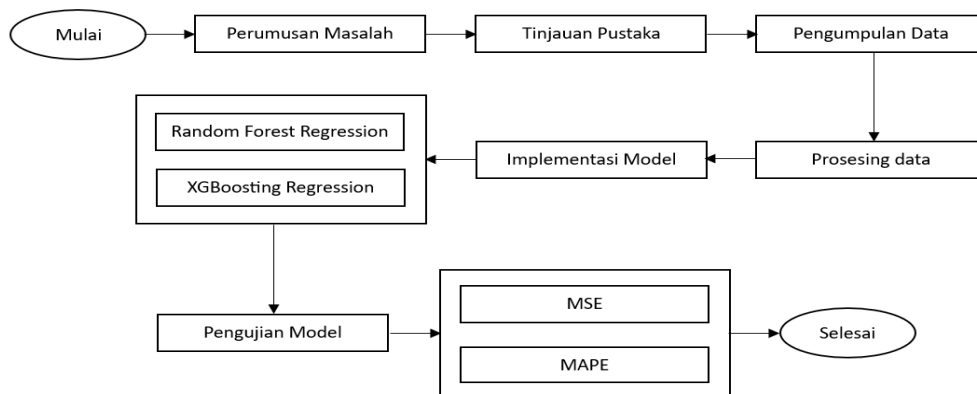
Berdasarkan uraian di atas peneliti mempunyai beberapa referensi dari peneliti terdahulu yang menjadi acuan dalam penelitian ini untuk perbandingan prediksi tingkat kemiskinan di DKI Jakarta yang Menerapkan metode XGBoost Regression dan Random Forest Regression sebagai algoritma dalam memprediksi. Berikut referensi penelitian sebelumnya yang relevan menurut peneliti, studi “Prediksi Jumlah Penduduk Miskin Indonesia menggunakan Metode Single Moving Average dan Double Moving Average” [11] memberikan pengetahuan tentang metode prediksi, dalam “Perbandingan CART dan Random Forest Regression untuk Deteksi Kanker berbasis Klasifikasi Data Microarray” [13] Random Forest Regression menunjukkan akurasi tertinggi 100% untuk data kanker paru-paru dan ovarium, sedangkan CART hanya mencapai 98,26% untuk kanker paru-paru, Life Expectancy Prediction Using Decision Tree, Random Forest Regression, studi “Gradient Boosting, and Regressions” [14] menunjukkan peningkatan signifikan pada XGBoost Regression setelah tuning, dengan RMSE berkurang dari 2,77 menjadi 2,57 dan R2 meningkat dari 0,85 menjadi 0,87. Lalu pada referensi berikutnya ini sebagai acuan untuk pengujian MSE dan MAPE, “Penerapan peramalan penjualan sembako Menggunakan Metode Single Moving Average (Studi Kasus Toko Kelontong Dedeh Retail)” [15], “prediksi penjualan papan bunga menggunakan metode double exponential smoothing” [16].

Dengan adanya beberapa permasalahan di atas dan referensi yang ada, peneliti bertujuan untuk mengetahui keakuratan perbandingan prediksi tingkat kemiskinan di DKI Jakarta dengan menggunakan metode algoritma XGBoost Regression dan Random Forest Regression, menggunakan data yang diperoleh dari Badan Pusat Statistik (BPS) DKI Jakarta. Selain itu Tujuan utama dari prediksi tingkat kemiskinan adalah untuk memberikan informasi yang lebih akurat kepada pembuat kebijakan, organisasi non-pemerintah, dan masyarakat umum untuk mengarahkan upaya-upaya pengentasan kemiskinan dengan lebih efektif. Dengan pemahaman yang lebih baik tentang kondisi kemiskinan yang akan datang, langkah-langkah preventif dan intervensi dapat diambil lebih awal, memungkinkan pengurangan kemiskinan yang lebih signifikan dan berkelanjutan di masa depan.

## 2. METODOLOGI PENELITIAN

### 2.1 Alur Penelitian

Adapun alur penelitian yang dipakai untuk penelitian ini bisa dapat dilihat pada gambar 1 dan akan dijelaskan pada diagram alur.



Gambar 1. Alur Penelitian

### 2.2 Pengumpulan Data

Pengumpulan data merupakan tahap awal dalam proses untuk peramalan tingkat kemiskina di DKI Jakarta. Data tingkat kemiskinan pada tahun 2010 sampai dengan tahun 2023 di daerah DKI Jakarta dalam bentuk csv dan excel ini di dapatkan dari Badan Pusat Statistik (BPS) DKI Jakarta.

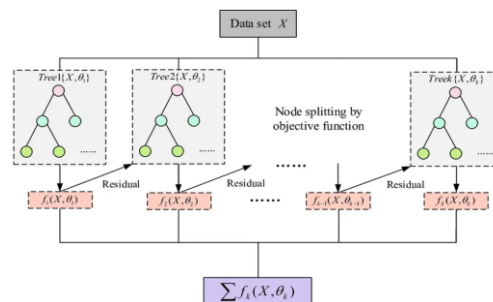


### 2.2.1 Prossesing Data

Prosesing data adalah proses untuk melakukan pengolahan data yang akan dipakai untuk tahap berikutnya yaitu implementasi model. Pada tahap ini, data mentah diubah menjadi data yang siap untuk diproses lebih lanjut untuk menjadi lebih berstruktur dan lebih mudah diproses ke dalam sistem [17]. Data yang sudah dikumpulkan lalu di proses menggunakan tahapan cleansing dan filtering agar data yang dibutuhkan itu lengkap dan tidak terjadi missing value. Yang dimana menganalisa data agar mendapatkan hasil data yang berkualitas tinggi untuk data yang akan digunakan pada saat implementasi model, maka dilakukannya pengubah atau penghapusan untuk yang tidak diperlukan, data yang bukan akurat dan yang memiliki data format atau gile yang salah dalam basis data yang merupakan pada tahapan dari cleansing data [18], sedangkan Filtering data merupakan proses selektif dalam memilih dan menyaring data tertentu dari suatu dataset[19] agar nantinya dapat digunakan dengan baik pada saat memasuki langkah implementasi. Setelah data di cleansing dan fiterling lalu data tersebut di buat pelabelan data. Yang di mana pelabelan data di lakukan secara manual dengan memisahkan data menjadi 2 bagian yaitu data testing dan training [20]. Pertama, model dilatih menggunakan data training untuk memahami pola data. Setelah itu, pengujian data dipakai dalam mengetes kinerja model dengan membuat prediksi untuk data baru yang tidak pernah dilihat pada model sebelumnya.

### 2.2.2 Implementasi Model

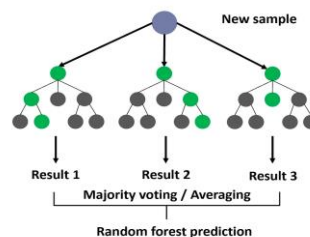
Setelah data di proses dan data sudah disaring maka langkah selanjutnya yaitu implementasi model. Yang dimana implementasi dipenelitian ini menggunakan metode XGBoost Regression dan Random Forest Regression untuk di jadikan suatu perbandingan. Karena peneliti menggunakan metode yang berbeda dalam analisis prediksi seperti XGBoost Regression yang dikenal memiliki performa yang sangat baik dalam berbagai jenis data dan kompetisi prediksi. Metode ini mampu menangani data dengan banyak fitur dan interaksi yang kompleks, serta sering kali memberikan hasil prediksi yang sangat akurat, sedangkan Random Forest Regression memiliki tingkat akurasi yang tinggi dan sering digunakan dalam berbagai aplikasi. Metode ini efektif dalam mengurangi overfitting karena menggunakan banyak pohon keputusan. Berikut pengertian dari kedua metode yang digunakan :



Sumber : <https://python.plainenglish.io/XGBoost-regression-in-depth-9ce4ce5d970d>

Gambar 2. XGBoost Regression

Untuk gambar 2 merupakan tahapan algoritma dari XGBoost Regression yang dimana pengembangan lebih lanjut dari gradient boosting. Membangun struktur pohon regresi menggunakan model yang lebih teratur, yang dapat meningkatkan kinerja dengan mengurangi kompleksitas model [14]. Model ini dikembangkan dengan memakai metode boosting, dari membuat model yang baru berdasarkan model yang lama, sehingga menghasilkan kesalahan prediksi (error) yang lebih kecil dibandingkan dengan model yang di buat sebelumnya agar mendapatkan hasil prediksi yang lebih akurat [21].



Gambar 3. Random Forest Regression

Pada gambar 3 adalah Sekelompok pohon klasifikasi yang dibuat dari pengambilan sampel data bootstrap.Langkah pertama dalam membangun model pohon acak adalah menentukan nilai N yang merupakan jumlah pohon keputusan yang dibentuk [13]. Salah satu keunggulannya yang luar biasa adalah strategi pengacakan variabel input yang digunakan untuk membagi setiap simpul pada pohon keputusan [22].



### 2.2.3 Pengujian Model

Pengujian model adalah proses mengevaluasi dan menilai kinerja suatu model atau algoritma pada kumpulan data yang tidak termasuk dalam fase pelatihan. Tujuannya adalah untuk mengukur kemampuan model dalam menerapkan pengetahuan yang diperoleh selama pelatihan pada data baru yang tidak digunakan selama proses pembelajaran. Proses pengujian model melibatkan penerapan model pada kumpulan data pengujian dan mengevaluasi seberapa baik model melakukan prediksi dan klasifikasi pada data tersebut. Analisis hasil pengujian sangat penting dalam mengevaluasi kehandalan dan efektivitas model sebelum digunakan dalam situasi nyata. Hal ini memungkinkan peneliti untuk membuat keputusan yang lebih tepat sebelum menerapkan model tersebut ke situasi dunia nyata. Dalam pengujian model ini peneliti menggunakan dua pengujian yaitu pengujian MSE dan MAPE, dari hasil pengujian ini peneliti dapat mengevaluasi kinerja dari kedua model. Berikut penjelasan dari kedua pengujian model:

1. Proses pengujian model pada penelitian ini menggunakan metode MSE (Mean Squared Error). MSE mengukur seberapa besar kesalahan kuadrat antara nilai peramalan dan nilai aktual untuk mengestimasi besaran kesalahan dalam peramalan dan membandingkan atau mengevaluasi berbagai metode peramalan adalah dua tujuan umum penggunaan teknik ini. Untuk menghitung MSE, rumus yang digunakan adalah sebagai berikut (1) [15] :

$$MSE = \frac{\sum_{t=1}^n (A_t - F_t)^2}{n} \quad (1)$$

Keterangan :  $A_t$  adalah permintaan aktual pada suatu periode tertentu.  $F_t$  ialah prediksi permintaan pada periode yang bersangkutan.  $n$  untuk jumlah periode  $t$  yang dianalisis. Sedangkan  $t$  mengacu pada periode waktu tertentu dalam analisis tersebut.

2. Pengujian model yang kedua menggunakan model MAPE (Mean Absolute Percentage Error), yaitu metrik untuk menilai kesalahan relatif. MAPE menghitung persentase ketidaksesuaian antara hasil prediksi dan permintaan aktual selama periode tertentu, menyoroti apakah kesalahannya terlalu tinggi atau rendah. Hal ini membuatnya lebih sering digunakan dibandingkan dari pada MAD. Rumus MAPE biasanya dinyatakan sebagai berikut (2) [16]:

$$MAPE = \left( \frac{100}{n} \right) \sum \left| X_t - \frac{F_t}{X_t} \right| \quad (2)$$

Keterangan :  $A_t$  merupakan permintaan aktual pada suatu periode tertentu, sementara  $F_t$  merupakan prediksi permintaan pada periode yang bersangkutan.  $n$  menunjukkan jumlah periode prediksi yang terlibat dalam analisis tersebut.

Metode MAPE ini menggunakan nilai rance sebagai acuan dalam menentukan kategori nilai error dalam bentuk presentase, nilai rance MAPE yang menjadi acuan ini mempunyai beberapa kategori seperti 0%-10% berkategorikan sangat baik, setelah itu 10%-20% berkategorikan baik, sedangkan untuk 20%-50% berkategorikan cukup, dan 50%-100% berkategorikan buruk lalu selebihnya jika melebihi 100% berkategorikan sangat buruk [23].

## 3. HASIL DAN PEMBAHASAN

Dalam penelitian ini, akan dilakukan pengujian untuk membandingkan hasil prediksi dari algoritma XGBoost Regression dan Random Forest Regression untuk tingkat kemiskinan yang ada pada wilayah DKI Jakarta. Tujuan dari penelitian ini ialah guna menentukan mana yang memberikan performa terbaik berdasarkan nilai pengujian pada MSE dan MAPE yang diperoleh dari masing masing algoritma.

### 3.1 Pengumpulan Data

Data yang dikumpulkan dan diperoleh melalui website Badan Pusat Statistik (BPS) DKI Jakarta, yang mengarah pada tingkat kemiskinan yang ada di DKI Jakarta. Data yang diperoleh dari BPS yaitu data tingkat kemiskinan dari tahun 2010 sampai tahun 2023.

**Tabel 1.** Presentase Garis Kemiskinan di Wilayah DKI Jakarta

Garis Kemiskinan, Jumlah, dan Persentase Penduduk Miskin di Daerah Menurut Kabupaten/Kota di Provinsi DKI Jakarta														
Kab/Kota	Persentase Penduduk Miskin													
	2023	2022	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012	2011	2010
Kep Seribu	13.13	14.11	15.06	14.87	12.09	11.98	12.98	12.58	11.01	11.62	11.01	11.62	11.53	13.07
Jakarta Selatan	3.10	3.52	3.56	3.43	2.73	2.83	3.38	3.27	3.47	3.49	3.47	3.49	3.43	3.80
Jakarta Timur	4.20	4.30	4.28	4.16	3.12	3.14	3.31	3.19	3.10	3.12	3.10	3.12	3.06	3.40
Jakarta Pusat	4.68	4.90	4.94	4.51	3.68	3.59	3.78	3.91	3.70	3.72	3.70	3.72	3.56	3.97
Jakarta Barat	4.09	4.22	4.31	4.25	3.25	3.39	3.25	3.38	3.46	3.47	3.46	3.47	3.44	3.82
Jakarta Utara	6.78	7.24	7.24	6.78	5.04	5.35	5.59	3.57	5.30	5.14	5.30	5.14	5.07	5.62
DKI Jakarta	4.44	4.69	4.72	4.53	3.47	3.57	3.77	3.75	3.72	3.70	3.72	3.70	3.64	4.04

Pada tabel 1 menampilkan data presentase penduduk miskin di wilayah DKI Jakarta dari tahun 2010 hingga tahun 2023 yang diperoleh dari website Badan Pusat Statistik (BPS) DKI Jakarta.



**3.2 Prossessing Data**

Data yang sudah dikumpulkan lalu data diproses melalui Google Colaboratory. Sebelum data di proses pada Google Colaboratory data di filter dan di cleansing terlebih dahulu di dalam excel agar tidak menemukan data yang null. Peneliti juga membutuhkan beberapa library untuk memproses data agar data dapat di proses di dalam Google Colaboratory.

**Tabel 2** Data yang sudah di cleansing

DKI Jakarta	Kep Seribu	Jakarta Selatan	Jakarta Timur	Jakarta Pusat	Jakarta Barat	Jakarta Utara	Tanggal
4.44	13.13	3.10	4.20	4.68	4.09	6.78	01/01/2023
4.69	14.11	3.52	4.30	4.90	4.22	7.24	01/01/2022
4.72	15.06	3.56	4.28	4.94	4.31	7.24	01/01/2021
4.53	14.87	3.43	4.16	4.51	4.25	6.78	01/01/2020
3.47	12.09	2.73	3.12	3.68	3.25	5.04	01/01/2019
3.57	11.98	2.83	3.14	3.59	3.39	5.35	01/01/2018
3.77	12.98	3.38	3.31	3.78	3.25	5.59	01/01/2017
3.75	12.58	3.27	3.19	3.91	3.38	3.57	01/01/2016
3.72	11.01	3.47	3.10	3.70	3.46	5.30	01/01/2015
3.70	11.62	3.49	3.12	3.72	3.47	5.14	01/01/2014
3.72	11.01	3.47	3.10	3.70	3.46	5.30	01/01/2013
3.70	11.62	3.49	3.12	3.72	3.47	5.14	01/01/2012
3.64	11.53	3.43	3.06	3.56	3.44	5.07	01/01/2011
4.04	13.07	3.80	3.40	3.97	3.82	5.62	01/01/2010

Pada tabel 2 ini ialah data yang sudah di cleansing sebelum data dimasukan ke dalam Google Colaboratory data di cleansing dan diolah terlebih dahulu di dalam excel yang dimana semua wilayah dirubah yang sebelumnya berada di kolom menjadi baris dan untuk tahun yang sebelumnya berada di baris dirubah menjadi kolom dan dirubah menjadi tanggal pertahunnya.

	DKI Jakarta	Kepulauan Seribu	Jakarta Selatan	Jakarta Timur	Jakarta Pusat	Jakarta Barat	Jakarta Utara	Tanggal
count	14.000000	14.000000	14.000000	14.000000	14.000000	14.000000	14.000000	14
mean	3.961429	12.618571	3.355000	3.471429	4.025714	3.661429	5.654286	2016-07-02 00:00:00
min	3.470000	11.010000	2.730000	3.060000	3.560000	3.250000	3.570000	2010-01-01 00:00:00
25%	3.700000	11.620000	3.297500	3.120000	3.700000	3.402500	5.140000	2013-04-02 06:00:00
50%	3.735000	12.335000	3.450000	3.165000	3.750000	3.465000	5.325000	2016-07-02 00:00:00
75%	4.340000	13.115000	3.490000	3.970000	4.375000	4.022500	6.490000	2019-10-01 18:00:00
max	4.720000	15.060000	3.800000	4.300000	4.940000	4.310000	7.240000	2023-01-01 00:00:00
std	0.438298	1.327020	0.288304	0.509990	0.501163	0.390539	1.019213	NaN

**Gambar 4.** Tampilan data dalam bentuk statistik deskriptif

Tampilan pada gambar 4 adalah tampilan data yang sudah di import di Google Colaboratory yang sudah menjadi data statistik deskriptif untuk melihat statistic dari data tersebut.

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from xgboost import XGBRegressor
from sklearn.metrics import mean_squared_error, mean_absolute_percentage_error
    
```

**Gambar 5.** Import Library

Gambar 5 ini memperlihatkan untuk import data pada library, memprosesing data hingga tahap implementasi dan evaluasi data. Setelah library sudah di masukan ke Google Colaboratory langkah selanjutnya yaitu filtering.

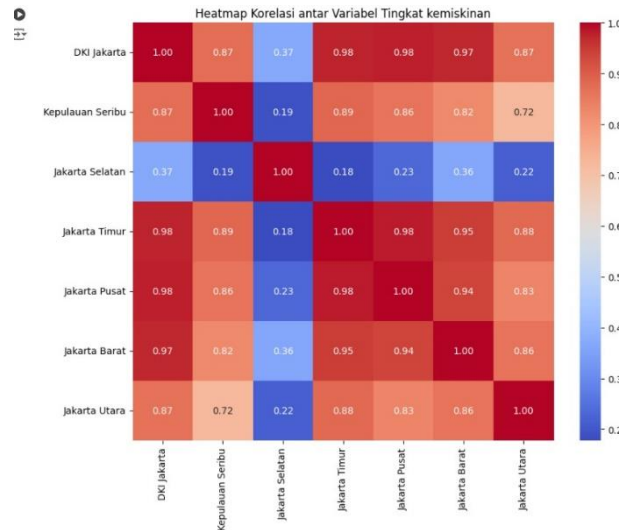
```

target_columns = ["DKI Jakarta", "Kepulauan Seribu", "Jakarta Selatan",
                  "Jakarta Timur", "Jakarta Pusat", "Jakarta Barat", "Jakarta Utara"]
    
```

**Gambar 6.** Korelasi antar variable target



Yang ada gambar 6 ialah tampilan untuk melihat korelasi antara data target tahap selanjutnya yaitu tahap pembagian data atau biasa disebut pelebelan data. yang dimana data X itu adalah data tanggal ordinal dan data Y adalah data dari target kolom yang sudah disiapkan. Dalam Pembagian dataset yang dibagi menjadi dua yaitu data trainingnya adalah 80% dan data testingnya adalah 20%.



**Gambar 7.** Filtering Data

Proses selanjutnya yaitu filtering data pada tahap diatas yang dimana target itu dibagi menjadi per daerah yang ada pada DKI Jakarta seperti gambar 7 .

**3.3 Implementasi Model**

Pada tahap implementasi, peneliti menggunakan dua model atau algoritma yang dimana pada penelitian ini memakai algoritma XGBoost Regression dan Random Forest Regression.

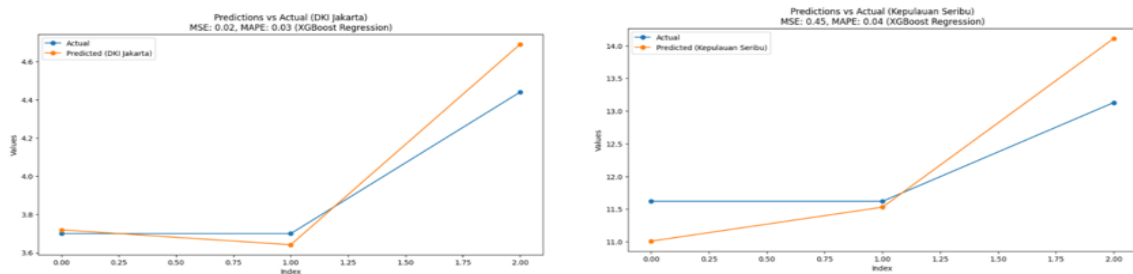
```
[ ] rf_model = RandomForestRegressor(n_estimators=100, random_state=42)
    xgb_model = XGBRegressor(objective='reg:squarederror', n_estimators=100, random_state=42)
```

**Gambar 8.** Persiapan dari kedua Algoritma

Persiapan untuk memanggil dua model algoritma, dengan rf\_model itu adalah variable yang menyimpan model Random Forest Regression dan xgb\_model adalah variable yang menyimpan model XGBoost Regression yang telah dibuat seperti gambar 8.

**3.3.1 Hasil Prediksi Algoritma XGBoost Regression**

Pada tahap ini adalah tahap menampilkan hasil prediksi untuk per 6 bulan dari algoritma XGBoost Regression menggunakan test size 0,2 dan random state 42 yang berbentuk grafik dari beberapa wilayah di DKI Jakarta .



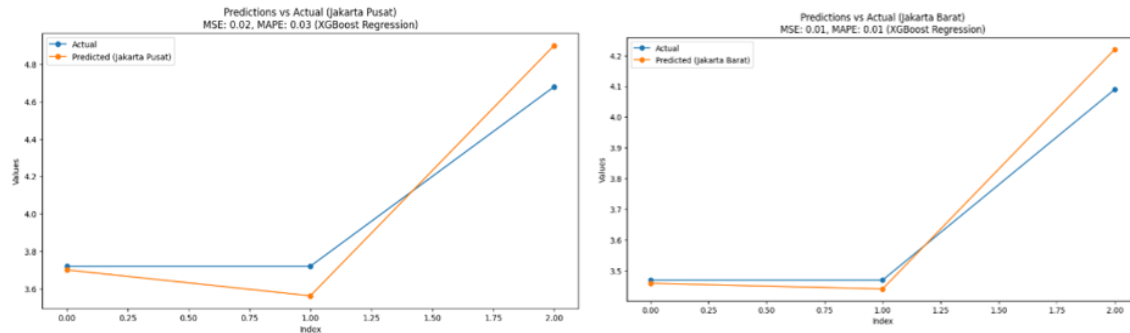
**Gambar 9.** Grafik prediksi DKI Jakarta dan Kepulauan Seribu (XGBoost Regression)

Pada grafik yang berada di gambar 9 ini menampilkan tingkat kemiskinan di wilayah DKI Jakarta yang awalnya mencapai 3,71% mengalami penurunan pada 6 bulan berikutnya mencapai 3,64% dan bulan berikutnya itu mengalami kenaikan yang sangat signifikan yang dimana mencapai 4,69%.

Lalu untuk grafik di Kepulauan Seribu, tingkat kemiskinan untuk bulan 6 pertama mempunyai nilai yang sangat tinggi dari beberapa wilayah yang ada di daerah DKI Jakarta yang dimana memiliki nilai mencapai 11,16%, setelah itu untuk 6 bulan selanjutnya mengalami kenaikan Kembali hingga 0,36% yang mencapai nilai 11,52%



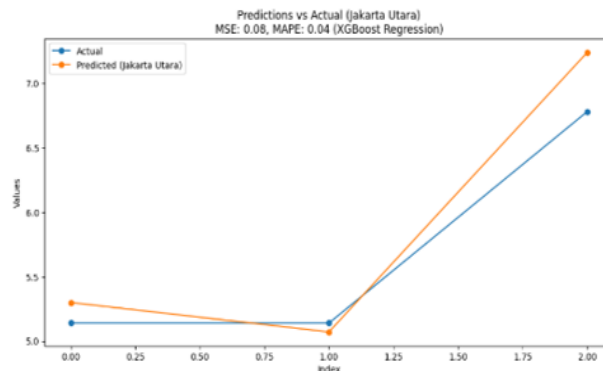
dan untuk 6 bulan berikutnya tingkat kemiskinan di kepulauan seribu mengalami kenaikan yang sangat signifikan hingga mencapai 2,83% yang dimana menjadi 14,35 dari 6 bulan sebelumnya.



**Gambar 10.** Grafik Prediksi Jakarta Selatan dan Jakarta Timur (XGBoost Regression)

Di gambar 10, Grafik untuk prediksi selanjutnya yaitu daerah Jakarta Selatan yang bisa diliha dari gambar di atas sebelah kiri yang dimana untuk 6 bulan awal dari grafik actual sebesar 3,49% dan untuk prediksinya sebesar 3,45% yang dimana mengalami penurunan yang tidak terlalu banyak dan untuk 6 bulan berikutnya kembali naik hanya 0,03%, lalu untuk 6 bulan selanjutnya hanya naik 0,01% saja.

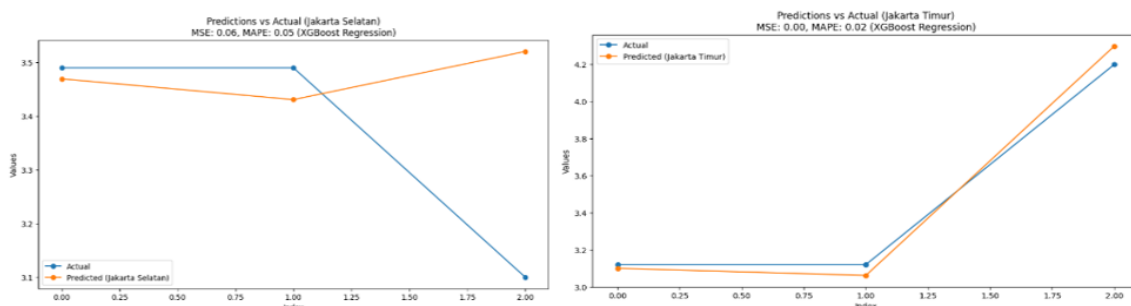
Selanjutnya untuk Grafik prediksi di daerah Jakarta Timur memiliki grafik yang hampir sama dengan grafik di DKI Jakarta yang mengalami kenaikan sangat pesat pada 6 bulan terakhir yang dimana pada 6 bulan pertama di daerah Jakarta Timur itu mencapai 3,10%, lalu mengalami penurunan di 6 bulan selanjutnya dari 3,10% menjadi 3,06% dan 6 bulan berikutnya mengalami kenaikan yang sangat signifikan hingga mencapai 4,24% kenaikan tersebut sangat tinggi dari presentase 6 bulan sebelumnya yang dimana mencapai 1,23%.



**Gambar 11.** Grafik Prediksi Jakarta Pusat dan Jakarta Barat (XGBoost Regression)

Pada gambar 11, grafik prediksi untuk Jakarta Pusat dari 6 bulan pertama yang mencapai 3,70% mengalami penurunan pada 6 bulan berikutnya hingga 0,04% dan untuk 6 bulan selanjutnya grafik berubah menjulang keatas dikarenakan mengalami kenaikan yang sangat pesat dari 0,04% hingga mencapai 1,34%.

Untuk grafik pada daerah Jakarta Barat 6 bulan pertama adalah 3,45% setelah itu 6 bulan selanjutnya mengalami penurunan dari yang sebelumnya 3,45% menjadi 3,44% dan untuk 6 bulan berikutnya mengalami kenaikan yang sangat tinggi hingga 0,78% yang menjadi 4,22%. Grafik terakhir pada gambar 12 adalah grafik untuk daerah Jakarta Utara yang dimana untuk tingkat kemiskinan di 6 bulan pertama ada pada 5,29% dan menurun pada 6 bulan berikut hingga 0,12% yang menjadi 5,07%, lalu untuk 6 bulan terakhir mengalami kenaikan yang sangat derastis hingga 2,1% yang menjadi 7,23%.



**Gambar 12.** Grafik Prediksi Jakarta Utara (XGBoost Regression)



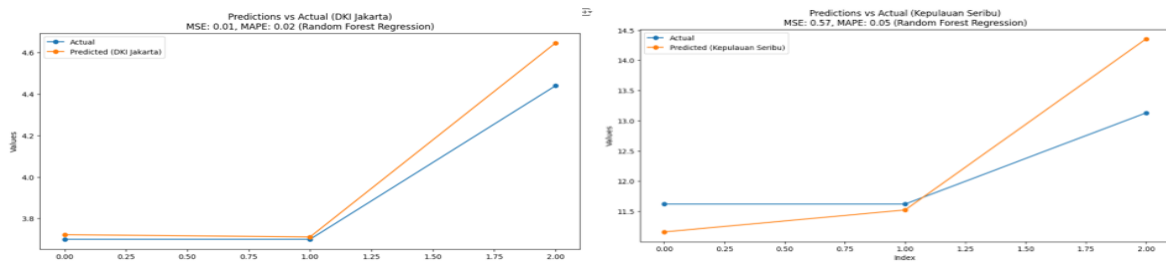
**Tabel 3.** Hasil Prediksi per 6 bulan XGBoost Regression

Tampilan data aktual dan prediksi per 6 bulan dari XGBoost Regression			
Bulan	Target	Aktual	Prediksi
6 bulan Pertama	DKI Jakarta	3.70	3.719889
6 bulan Kedua	DKI Jakarta	3.70	3.641960
6 bulan Ketiga	DKI Jakarta	4.44	4.690413
6 bulan Pertama	Kepulauan Seribu	11.62	11.010991
6 bulan Kedua	Kepulauan Seribu	11.62	11.530025
6 bulan Ketiga	Kepulauan Seribu	13.13	14.110274
6 bulan Pertama	Jakarta Selatan	3.49	3.469469
6 bulan Kedua	Jakarta Selatan	3.49	3.439469
6 bulan Ketiga	Jakarta Selatan	3.10	3.520477
6 bulan Pertama	Jakarta Timur	3.12	3.100061
6 bulan Kedua	Jakarta Timur	3.12	3.061568
6 bulan Ketiga	Jakarta Timur	4.20	4.298836
6 bulan Pertama	Jakarta Pusat	3.72	3.700071
6 bulan Kedua	Jakarta Pusat	3.72	3.561665
6 bulan Ketiga	Jakarta Pusat	4.68	4.900460
6 bulan Pertama	Jakarta Barat	3.47	3.459571
6 bulan Kedua	Jakarta Barat	3.47	3.441266
6 bulan Ketiga	Jakarta Barat	4.09	4.220270
6 bulan Pertama	Jakarta Utara	5.14	5.299187
6 bulan Kedua	Jakarta Utara	5.14	5.071409
6 bulan Ketiga	Jakarta Utara	6.78	7.239448

Tampilan tabel 3 adalah data aktual dan prediksi per 6 bulan dari model XGBoost Regression menunjukkan perbandingan antara nilai sebenarnya dengan nilai yang di prediksi dari model XGBoost Regression dalam rentan waktu setiap 6 bulan dalam memprediksi nilai di masa depan berdasarkan pada data sebelumnya.

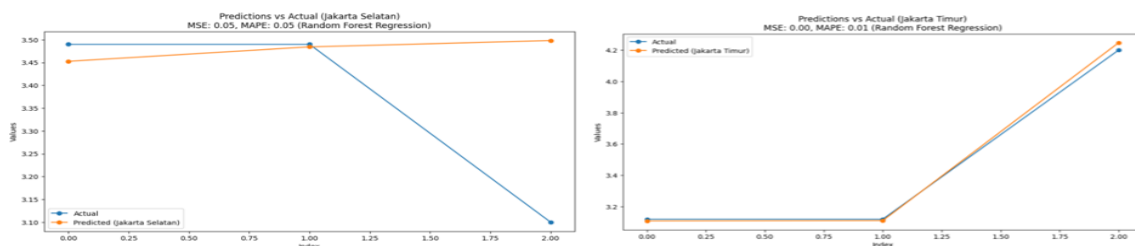
**3.3.2 Hasil prediksi algoritma Random Forest Regression**

Selanjutnya adalah hasil prediksi menggunakan algoritma Random Forest Regression yang menggunakan test size dan random state nya sama seperti Regression yaitu 0,2 dan 42 yang berbentuk grafik untuk per 6 bulan pada wilayah DKI Jakarta.



**Gambar 13.** Grafik prediksi DKI Jakarta dan Kepulauan Seribu (Random Forest Regression)

Pada gambar 13 menunjukkan grafik di wilayah DKI Jakarta tingkat kemiskinan yang awalnya mencapai 3,72% mengalami penurunan pada 6 bulan selanjutnya yang mencapai 3,71% dan bulan 6 bulan berikutnya mengalami kenaikan yang sangat tinggi yang dimana mencapai 4,64%. Untuk grafik di Kepulauan Seribu memiliki tingkat kemiskinan yang sangat tinggi dari beberapa wilayah di DKI Jakarta hingga mencapai 11,16%, lalu untuk 6 bulan kedepannya kembali terjadi kenaikan menjadi 11,52%, untuk 6 bulan selanjutnya tingkat kemiskinan di kepulauan seribu mengalami kenaikan yang sangat derastis yang mencapai 14,35%.



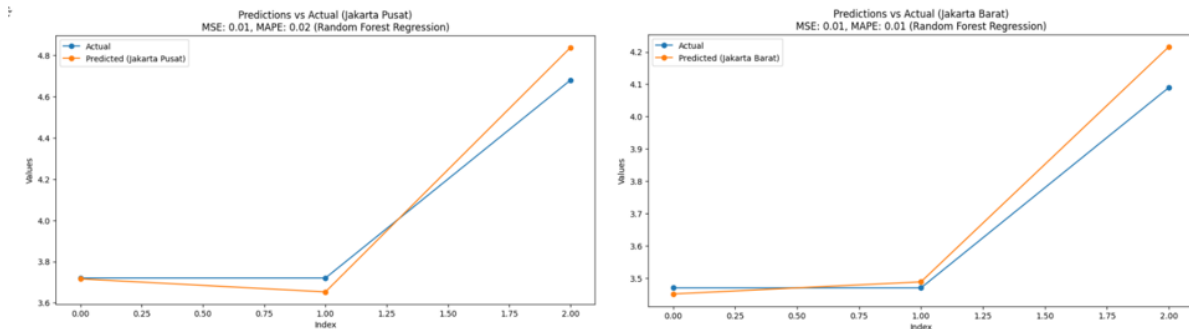
**Gambar 14.** Grafik Prediksi Jakarta Selatan dan Jakarta Timur (Random Forest Regression)





Gambar 14 ini menunjukkan Grafik prediksi berikutnya dari daerah Jakarta Selatan yang bisa dilihat dari gambar di atas sebelah kiri yang sebesar 3,45% yang dimana untuk 6 bulan berikutnya kembali naik menjadi 3,48% dan untuk 6 bulan selanjutnya hanya naik 0,01% saja menjadi 3,49%.

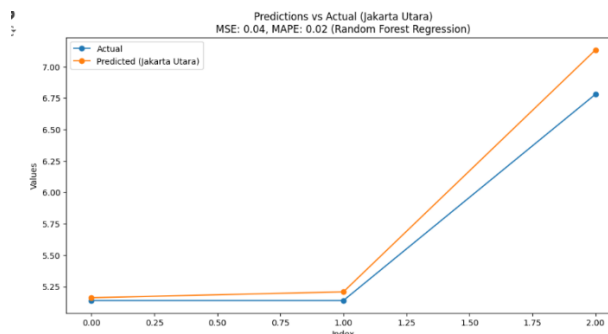
Lalu grafik prediksi pada Jakarta Timur yang dimana pada 6 bulan pertama di daerah Jakarta Timur itu mencapai 3,10%, lalu mengalami kenaikan pada 6 bulan selanjutnya dari 3,10% menjadi 3,11% dan 6 bulan berikutnya mengalami kenaikan yang sangat signifikan hingga mencapai 4,24% kenaikan tersebut sangat tinggi dari presentase 6 bulan sebelumnya yang dimana mencapai 1,13%.



**Gambar 15.** Grafik Prediksi Jakarta Pusat dan Jakarta Barat (Random Forest Regression)

Selanjutnya pada gambar 15, grafik prediksi untuk Jakarta Pusat dari 6 bulan pertama yaitu 3,71% mengalami penurunan pada 6 bulan berikutnya menjadi 3,65% dan untuk 6 bulan selanjutnya grafik berubah naik keatas dikarenakan mengalami kenaikan yang sangat signifikan hingga mencapai 4,83%.

Untuk grafik berikutnya daerah Jakarta Barat yang dimana 6 bulan pertama 3,45% setelah itu 6 bulan selanjutnya mengalami kenaikan dari yang sebelumnya 3,45% menjadi 3,48% dan untuk 6 bulan terakhir mengalami kenaikan kembali yang sangat signifikan hingga 0,73% yang menjadi 4,21%.



**Gambar 16.** Grafik Prediksi Jakarta Utara ( Random Forest Regression)

Grafik yang terakhir ini pada gambar 16 adalah grafik untuk daerah Jakarta Utara yang dimana tingkat kemiskinan di 6 bulan pertama ada pada 5,16% setelah itu mengalami kenaikan pada 6 bulan berikutnya hingga 0,04% yang menjadi 5,20% dan untuk 6 bulan terakhir mengalami kenaikan yang sangat signifikan hingga mencapai 1,93%.

**Tabel 4.** Hasil presentase prediksi 6 bulan Random Forest Regression

Tampilan data aktual dan prediksi per 6 bulan dari Random Forest Regression			
Bulan	Target	Aktual	Prediksi
6 bulan Pertama	DKI Jakarta	3.70	3.7221
6 bulan Kedua	DKI Jakarta	3.70	3.7112
6 bulan Ketiga	DKI Jakarta	4.44	4.6465
6 bulan Pertama	Kepulauan Seribu	11.62	11.1605
6 bulan Kedua	Kepulauan Seribu	11.62	11.5226
6 bulan Ketiga	Kepulauan Seribu	13.13	14.3553
6 bulan Pertama	Jakarta Selatan	3.49	3.4527
6 bulan Kedua	Jakarta Selatan	3.49	3.4843
6 bulan Ketiga	Jakarta Selatan	3.10	3.4982
6 bulan Pertama	Jakarta Timur	3.12	3.1089
6 bulan Kedua	Jakarta Timur	3.12	3.1110
6 bulan Ketiga	Jakarta Timur	4.20	4.2470



Tampilan data aktual dan prediksi per 6 bulan dari Random Forest Regression			
Bulan	Target	Aktual	Prediksi
6 bulan Pertama	Jakarta Pusat	3.72	3.7162
6 bulan Kedua	Jakarta Pusat	3.72	3.6527
6 bulan Ketiga	Jakarta Pusat	4.68	4.8383
6 bulan Pertama	Jakarta Barat	3.47	3.4513
6 bulan Kedua	Jakarta Barat	3.47	3.4886
6 bulan Ketiga	Jakarta Barat	4.09	4.2161
6 bulan Pertama	Jakarta Utara	5.14	5.1622
6 bulan Kedua	Jakarta Utara	5.14	5.2087
6 bulan Ketiga	Jakarta Utara	6.78	7.1326

Tabel 4 adalah tampilan data actual dan prediksi untuk per 6 bulan pada algoritma Random Forest Regression tampilan tersebut mengelompokan data per 6 bulan untuk menunjukkan hasil sebenarnya dan hasil yang sudah di prediksi oleh model dalam jangka waktu setiap 6 bulan untuk masing-masing kota yang ada di wilayah DKI Jakarta.

### 3.4 Pengujian MAPE

Pada tabel 5, terdapat perbandingan nilai MAPE antara dua model, yaitu XGBoost (XGB) dan Random Forest (RF), untuk beberapa target di wilayah DKI Jakarta. Untuk DKI Jakarta, Random Forest memiliki MAPE sebesar 2.3381, lebih rendah dibandingkan XGBoost yang memiliki MAPE sebesar 2.5821. Di Kepulauan Seribu, XGBoost lebih baik dengan MAPE sebesar 4.4938 dibandingkan Random Forest yang memiliki MAPE sebesar 4.5652. Pada wilayah Jakarta Selatan, Jakarta Timur, Jakarta Pusat, Jakarta Barat, dan Jakarta Utara. Nilai MAPE Random Forest Regression lebih rendah daripada XGBoost Regression, memperlihatkan untuk model Random Forest memberikan hasil prediksi yang lebih akurat dilokasi tertentu.

**Tabel 5.** Pengujian menggunakan MAPE

Perbandingan Pengujain Menggunakan MAPE		
Target	MAPE (XGB)	MAPE (Random Forest)
DKI Jakarta	2.582050730158256	2.3381339202056197
Kepulauan Seribu	4.493756198258355	4.565244459164017
Jakarta Selatan	5.28287064563796	5.086052983854935
Jakarta Timur	1.6217056211534446	1.2770569016887916
Jakarta Pusat	3.1675780743056556	2.6999152667659136
Jakarta Barat	1.4378963973797412	1.420603991391247
Jakarta Utara	3.7359938949125433	3.265003981138813

Dapat dilihat dari hasil tersebut model Random Forest (RF) mendapatkan hasil yang lebih baik untuk sebagian besar lokasi, kecuali untuk Kepulauan Seribu di mana model XGBoost (XGB) memiliki MAPE yang lebih rendah. Secara umum, Random Forest cenderung memberikan prediksi yang lebih akurat berdasarkan nilai MAPE yang lebih rendah di sebagian besar wilayah yang dibandingkan. Hal ini menunjukkan bahwa dalam konteks data yang diberikan, model Random Forest lebih andal dibandingkan XGBoost untuk sebagian besar target lokasi.

### 3.5 Pengujian MSE

Pada tabel 6, terdapat perbandingan nilai MSE antara dua model, yaitu (XGB) dan Random Forest (RF), untuk beberapa target lokasi. Untuk DKI Jakarta, Random Forest memiliki MSE sebesar 0.019578, lebih rendah dibandingkan yang memiliki MSE sebesar 0.022158. Di Kepulauan Seribu, lebih baik dengan MSE sebesar 0.446641 dibandingkan Random Forest yang memiliki MSE sebesar 0.489093. Di Jakarta Selatan, Jakarta Timur, Jakarta Pusat, Jakarta Barat, dan Jakarta Utara, Random Forest memiliki nilai MSE yang lebih rendah dibandingkan, yang memperlihatkan dari model Random Forest mendapatkan hasil prediksi yang lebih akurat di lokasi-lokasi tersebut.

**Tabel 6.** Pengujian menggunakan MSE

Perbandingan Pengujain Menggunakan MSE		
Target	MSE (XGB)	MSE (Random Forest)
DKI Jakarta	0.02215705901272968	0.01957760600848653
Kepulauan Seribu	0.44664172081054865	0.4890930472070263
Jakarta Selatan	0.06024282230206159	0.05793821709026316



Perbandingan Pengujain Menggunakan MSE		
Target	MSE (XGB)	MSE (Random Forest)
Jakarta Timur	0.004526792465221291	0.0032859960879253382
Jakarta Pusat	0.024689947786246016	0.01974914519083045
Jakarta Barat	0.005968196773778786	0.0058228934047416196
Jakarta Utara	0.08037906441268126	0.06797929738623235

Dari uraian diatas dapat kita lihat untuk model Random Forest (RF) memberikan hasil yang lebih baik untuk sebagian besar lokasi, kecuali untuk Kepulauan Seribu di mana model XGBoost (XGB) memiliki MSE yang lebih rendah. Secara umum, Random Forest cenderung memberikan prediksi yang lebih akurat berdasarkan nilai MSE yang lebih rendah di sebagian besar wilayah yang dibandingkan. Hal ini menunjukkan bahwa dalam konteks data yang diberikan, model Random Forest lebih andal dibandingkan XGBoost untuk sebagian besar target lokasi.

#### 4. KESIMPULAN

Dari Hasil yang sudah di dapatkan pada penelitian ini, peneliti dapat menyimpulkan bahwa metode Random Forest Regression memiliki performa yang lebih baik dibandingkan XGBoost Regression dalam memprediksi tingkat kemiskinan di DKI Jakarta. Untuk DKI Jakarta Random Forest Regression memiliki MSE sebesar 0.019578, yang lebih rendah dibandingkan XGBoost Regression sebesar 0.022158, dan MAPE sebesar 2.3381, yang juga lebih rendah dibandingkan MAPE XGBoost Regression sebesar 2.5821. Di Kepulauan Seribu, XGBoost Regression memberikan hasil yang lebih baik dengan MSE sebesar 0.446641 dan MAPE sebesar 4.4938 dibandingkan dengan Random Forest Regression yang memiliki MSE sebesar 0.489093 dan MAPE sebesar 4.5652. Namun, di wilayah-wilayah seperti Jakarta Selatan, Jakarta Timur, Jakarta Pusat, Jakarta Barat, dan Jakarta Utara, Random Forest Regression secara konsisten menunjukkan nilai MSE dan MAPE yang lebih rendah dibandingkan XGBoost Regression. Hal ini memperlihatkan bahwa metode Random Forest Regression lebih baik dan memberikan hasil prediksi yang lebih akurat dari sebagian besar lokasi. Hasil ini menunjukkan bahwa pemilihan metode yang tepat sangat penting dalam analisis data kemiskinan untuk menghasilkan prediksi yang akurat.

#### REFERENCES

- [1] Syaharuddin, E. Pujiana, I. P. Sari, V. M. Mardika, and M. Putri, "Analisis Algoritma Back Propagation Dalam Prediksi Angka Kemiskinan Di Indonesia," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 4, no. 2, pp. 198–207, 2020, [Online]. Available: <https://jurnal.unimed.ac.id/2012/index.php/cess/article/view/13601/pdf>
- [2] Devi Arfiani, Berantas Kemiskinan, Digital 20. ALPRIN, 2020. [Online]. Available: [https://www.google.co.id/books/edition/Berantas\\_Kemiskinan\\_xnn7DwAAQBAJ?hl=id&gbpv=1](https://www.google.co.id/books/edition/Berantas_Kemiskinan_xnn7DwAAQBAJ?hl=id&gbpv=1)
- [3] R. B. Budhijana, "Analisis Pengaruh Pertumbuhan Ekonomi, Index Pembangunan Manusia (IPM) dan Pengangguran Terhadap Tingkat Kemiskinan Di Indonesia Tahun 2000-2017," *J. Ekon. Manaj. dan Perbank. (Journal Econ. Manag. Banking)*, vol. 5, no. 1, p. 36, 2020, doi: 10.35384/jemp.v5i1.170.
- [4] D. A. Mukmin, R. Irsyada, and H. A. Audytra, "Penerapan Metode Moving Average Pada Sistem Informasi Prediksi Angka Kemiskinan," *Multidiscip. Appl. Quantum Inf. Sci.*, vol. 1, no. 1, pp. 43–49, 2022, doi: 10.32665/almantiq.v1i1.330.
- [5] R. B. Praja, M. Muchtar, and P. R. Sihombing, "Analisis Pengaruh Indeks Pembangunan Manusia, Laju Pertumbuhan Penduduk, dan Tingkat Pengangguran Terbuka terhadap Kemiskinan di DKI Jakarta," *Ecoplan*, vol. 6, no. 1, pp. 78–86, 2023, doi: 10.20527/ecoplan.v6i2.656.
- [6] Eka Wati, "Analisis Pengaruh Tingkat Pengangguran Terbuka, Indeks Pembangunan Manusia, PDRB Terhadap Kemiskinan di Provinsi DKI Jakarta Tahun 2017-2022," *Nucl. Phys.*, vol. 13, no. 1, pp. 104–116, 2023.
- [7] deris desmawan reza maulana, "Analisis Faktor-Faktor Yang Mempengaruhi Tingkat Kemiskinan Di Pulau Jawa tahun 2018-2022," vol. 7, pp. 29433–29440, 2023, [Online]. Available: <http://repository.upnjatim.ac.id/11258/>
- [8] N. A. Tanamal, "Analisa Faktor Kemiskinan Warga Kampung Ujung Cipinang Besar Selatan (Manusia Kuburan) Jakarta Timur," *J. Rev. Pendidik. dan Pengajaran*, vol. 5, no. 1, pp. 89–97, 2022, doi: 10.31004/jrpp.v5i1.4745.
- [9] S. T. W. Ariyanto, A. Tjalla, and M. Mahdiyah, "Analisis Pengaruh Meningkatnya Jumlah Kemiskinan di Jakarta Dalam 20 Tahun Terakhir Terhadap Jumlah Kriminalitas di Wilayah Hukum Polda Metro Jaya," *J. Litbang Polri*, vol. 26, no. 2, pp. 50–55, 2023, doi: 10.46976/litbangpolri.v26i2.200.
- [10] M. Fuad, F. I., and M. U. Basuki, "Analisis Faktor-Faktor Yang Mempengaruhi Kerentanan Kemiskinan Relatif Di Kota Jakarta Barat Tahun 2018," *Diponegoro J. Econ.*, vol. 9, no. 2, p. 168, 2020, [Online]. Available: <https://ejournal2.undip.ac.id/index.php/dje>
- [11] F. Kusuma, M. Ahsan, and S. Syahminan, "Prediksi Jumlah Penduduk Miskin Indonesia menggunakan Metode Single Moving Average dan Double Moving Average," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 3, no. 2, p. 105, 2021, doi: 10.36499/jinrpl.v3i2.4594.
- [12] A. Alfani W.P.R., F. Rozi, and F. Sukmana, "Prediksi Penjualan Produk Unilever Menggunakan Metode K-Nearest Neighbor," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 6, no. 1, pp. 155–160, 2021, doi: 10.29100/jipi.v6i1.1910.
- [13] Riska Chairunisa, Adiwijaya, and Widi Astuti, "Perbandingan CART dan Random Forest untuk Deteksi Kanker berbasis Klasifikasi Data Microarray," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 5, pp. 805–812, 2020, doi: 10.29207/resti.v4i5.2083.



- [14] G. Chairunisa et al., "Life Expectancy Prediction Using Decision Tree, Random Forest, Gradient Boosting, and XGBoost Regressions," *J. Sintak*, vol. 2, no. 2, pp. 71–82, 2024, doi: 10.62375/jsintak.v2i2.249.
- [15] N. Kurnia Informatika, I. Komputer, U. Singaperbangsa, and K. Abstrak, "Penerapan Peramalan Penjualan Sembako Menggunakan Metode Single Moving Average (Studi Kasus Toko Kelontong Dedeh Retail)," *J. Ilm. Wahana Pendidik.*, vol. 8, no. 17, pp. 307–316, 2022, [Online]. Available: <https://doi.org/10.5281/zenodo.7076573>
- [16] M. H. Elison, R. Asrianto, and Aryanto, "Prediksi Penjualan Papan Bunga Menggunakan Metode Double Exponential Smoothing," *J. Ris. Sist. Inf. dan Teknol. Inf.*, vol. 2, no. 3, pp. 45–56, 2020, doi: 10.52005/jursistekni.v2i3.60.
- [17] H. Faisal, A. Febriandirza, and F. N. Hasan, "Analisis Sentimen Terkait Ulasan Pada Aplikasi PLN Mobile Menggunakan Metode Support Vector Machine," *KESATRIA J. Penerapan Sist. Inf. (Komputer Manajemen)*, vol. 5, no. 1, pp. 303–312, 2024.
- [18] D. Darwis, N. Siskawati, and Z. Abidin, "PENERAPAN ALGORITMA NAIVE BAYES UNTUK ANALISIS SENTIMEN REVIEW DATA TWITTER BMKG NASIONAL," *J. Tekno Kompak*, vol. 15, no. 1, p. 131, 2021, doi: 10.33365/jtk.v15i1.744.
- [19] A. Nugraha, Y. H. Chrisnanto, and R. Yuniarti, "Prediksi Sentimen Pada Sosial Media Twitter Mengenai Produk Smartphone Menggunakan Algoritma K-NN Classification," *Sensasi*, pp. 251–258, 2019.
- [20] I. Iqbal Wibowo and A. Febriandirza, "Analisis Sentimen Ulasan Pengguna Game Pubg Di Google Play Store Menggunakan Algoritma Naïve Bayes," *J. Sist. Komput. dan Inform.*, vol. 5, no. 3, pp. 590–599, 2024, doi: <http://dx.doi.org/10.30865/json.v5i3.7264>.
- [21] G. A. Mursianto, I. M. Falih, M. Irfan, T. Sakinah, and D. S. Prasvita, "Perbandingan Metode Klasifikasi Random Forest dan XGBoost Serta Implementasi Teknik SMOTE pada Kasus Prediksi Hujan," *J. Senamika*, vol. 2, no. 2, pp. 41–50, 2021.
- [22] N. Nur, F. Wajidi, S. Sulfayanti, and W. Wildayani, "Implementasi Algoritma Random Forest Regression untuk Memprediksi Hasil Panen Padi di Desa Minanga," *J. Komput. Terap.*, vol. 9, no. 1, pp. 58–64, 2023, doi: 10.35143/jkt.v9i1.5917.
- [23] S. Pratista, A. Nazir, I. Iskandar, E. Budianita, and I. Afrianty, "Perbandingan Prediksi Obat Berdasarkan Pemakaian Menggunakan Algoritma Single Moving Average dan Support Vector Regression," *J. Media Inform. Budidarma*, vol. 7, no. 4, pp. 1860–1868, 2023, doi: 10.30865/mib.v7i4.6859.