

Prediksi Nilai Redaman Jaringan Fiber Optik untuk Menilai Kinerja Jaringan Menggunakan *Random Forest Regression*

Febrianda Putra¹, Susanti², Herwin³, Khusaeri Andesa⁴, Mardainis⁵

febriandaputra11@gmail.com¹, susanti@usti.ac.id², herwin@usti.ac.id³, khusaeri@usti.ac.id⁴, mardainis@usti.ac.id⁵

^{1,2,3,4,5} Universitas Sains dan Teknologi Indonesia

Informasi Artikel

Diterima : 13 Mar 2025
Direvisi : 24 Mar 2025
Disetujui : 30 Apr 2025

Kata Kunci

Random Forest Regression, Fiber Optik, Prediksi, Redaman.

Abstrak

Peningkatan kebutuhan layanan telekomunikasi membutuhkan infrastruktur jaringan fiber optik yang andal. Namun, redaman sinyal menjadi tantangan utama dalam transmisi data yang menyebabkan penurunan kualitas layanan. Penelitian ini bertujuan memprediksi nilai redaman jaringan fiber optik menggunakan algoritma *Random Forest Regression* dengan memanfaatkan data historis dari PT. Telkom Akses Indonesia. Dataset terdiri dari 1225 data Pasang Sambungan Baru (PSB) dengan fitur koordinat ODP, lokasi pelanggan, panjang kabel, dan nilai redaman. Hasil penelitian menunjukkan bahwa model dengan rasio data 90:10 memberikan performa optimal dengan R^2 Score 0.9832, MSE 0.0559, RMSE 0.2363, dan MAE 0.1009. Model ini mampu menjelaskan 98.32% variasi dalam data redaman jaringan. Prediksi nilai redaman untuk periode mendatang menunjukkan tren yang stabil, memungkinkan penilaian kinerja jaringan fiber optik secara proaktif sehingga operator dapat mengantisipasi gangguan sebelum berdampak pada layanan pelanggan.

Keywords

Random Forest Regression, Fiber Optic, Prediction, Attenuation.

Abstract

The increasing demand for telecommunication services requires reliable fiber optic network infrastructure. However, signal attenuation remains a major challenge in data transmission that leads to service quality degradation. This research aims to predict fiber optic network attenuation values using the *Random Forest Regression* algorithm by utilizing historical data from PT. Telkom Akses Indonesia. The dataset consists of 1225 New Connection Installation (PSB) data with features including ODP coordinates, customer location, cable length, and attenuation values. Research results show that the model with a 90:10 data ratio provides optimal performance with an R^2 Score of 0.9832, MSE of 0.0559, RMSE of 0.2363, and MAE of 0.1009. The model successfully explains 98.32% of variation in network attenuation data. Predictions of attenuation values for future periods show a stable trend, enabling proactive assessment of fiber optic network performance so operators can anticipate disruptions before they impact customer service.

A. Pendahuluan

Seiring berkembangnya teknologi, kebutuhan akan layanan telekomunikasi semakin meningkat, terutama dalam mendukung layanan multimedia seperti video, suara, dan data. Salah satu teknologi yang menjadi solusi utama dalam jaringan telekomunikasi adalah fiber optik, yang memiliki keunggulan seperti bandwidth besar, biaya operasional rendah, bobot ringan, serta tahan terhadap gangguan elektromagnetik [1]. Implementasi teknologi ini dalam bentuk *Fiber to the Home* (FTTH) memungkinkan transmisi data berkecepatan tinggi langsung ke rumah pelanggan, sehingga dapat memberikan layanan internet yang cepat, stabil, dan mendukung berbagai kebutuhan pengguna secara simultan [2].

Namun, dalam implementasinya, jaringan fiber optik menghadapi tantangan utama berupa redaman atau *attenuation*, yaitu penurunan daya sinyal sepanjang jalur transmisi dari pengirim ke penerima [3]. Redaman ini dapat berdampak pada efisiensi transmisi data dengan menurunkan *bandwidth*, daya cahaya, serta kapasitas sistem secara keseluruhan [4]. Faktor utama yang memengaruhi redaman mencakup panjang kabel serta kondisi fisik yang rentan terhadap perubahan, seperti pembengkokan kabel atau gangguan eksternal lainnya [5]. Tanpa prediksi yang akurat terhadap nilai redaman, operator jaringan menghadapi risiko gangguan layanan yang tak terduga, penurunan kualitas internet pelanggan, serta peningkatan biaya perbaikan infrastruktur yang mendadak. Oleh karena itu, prediksi nilai redaman menjadi aspek krusial dalam perencanaan dan pemeliharaan jaringan yang lebih efektif serta untuk meminimalkan risiko gangguan pada transmisi data.

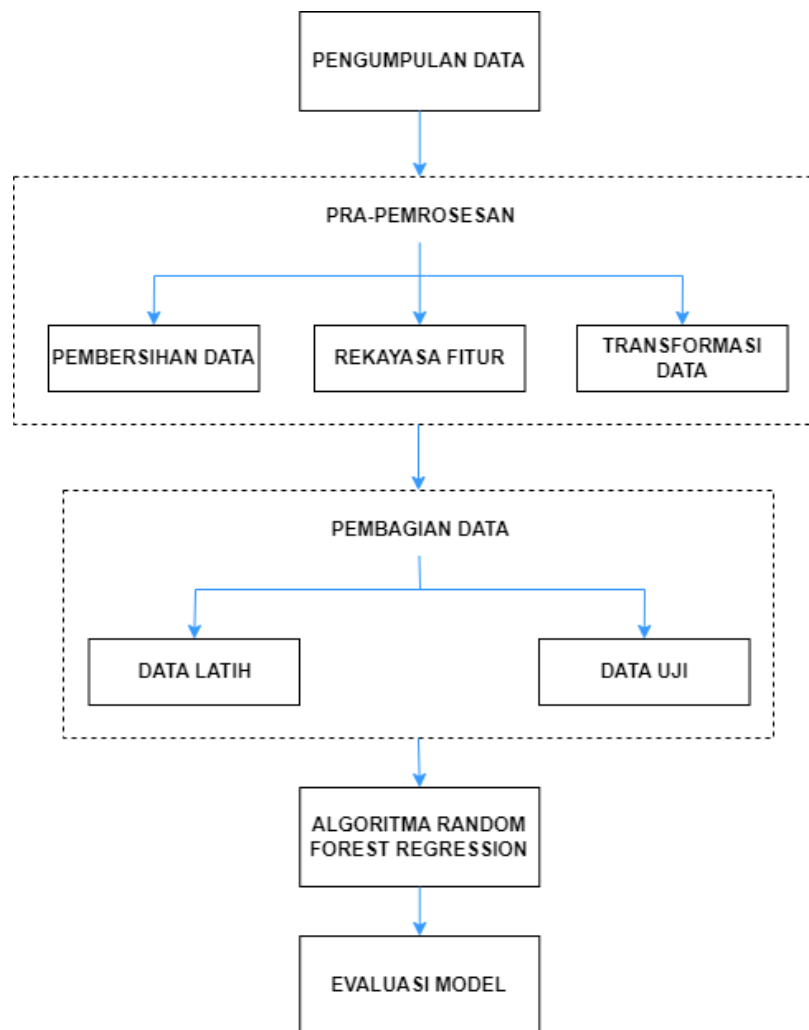
Penelitian terdahulu telah menunjukkan bahwa algoritma *Random Forest Regression* memiliki performa yang unggul dalam analisis data kompleks dan prediksi akurat. Studi oleh [6] membandingkan beberapa metode regresi seperti *Gradient Boosting Trees*, *Linear Regression*, dan *Decision Tree Regression* dalam memprediksi produksi daya pada Pembangkit Listrik Tenaga Surya (PLTS), dan hasilnya menunjukkan bahwa *Random Forest Regression* memberikan performa terbaik dengan R^2 sebesar 0.9679 dan RMSE 0.0438. [7] juga menunjukkan bahwa *Random Forest Regression* lebih akurat dibandingkan dengan metode lain dalam memprediksi harga rumah, dengan RMSE 0.440 dan akurasi 81.5%. Studi lainnya oleh [8] menemukan bahwa algoritma ini memiliki akurasi 80% dan R^2 sebesar 0.8 dalam memprediksi harga kapal layar, dengan variabel seperti region, dimensi kapal, tahun produksi, dan merek sebagai faktor yang paling berpengaruh. Berdasarkan temuan tersebut, *Random Forest Regression* dipandang sebagai metode yang relevan untuk memprediksi nilai redaman jaringan fiber optik karena kemampuannya dalam menangani data kompleks dan hubungan nonlinier antar variabel.

Penelitian ini bertujuan untuk memprediksi nilai redaman jaringan fiber optik menggunakan algoritma *Random Forest Regression* dengan data historis seperti lokasi *Optical Distribution Point* (ODP), lokasi rumah pelanggan, panjang kabel, dan nilai redaman per waktu. Model yang dikembangkan akan dievaluasi menggunakan metrik *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), *R-Squared* (R^2), dan *Mean Absolute Error* (MAE) untuk menilai tingkat akurasi prediksi. Hasil penelitian ini diharapkan dapat memberikan kontribusi dalam pengelolaan jaringan fiber optik dengan meningkatkan efisiensi perencanaan infrastruktur, mengurangi

potensi gangguan, serta mendukung pengambilan keputusan berbasis data dalam pemeliharaan jaringan telekomunikasi.

B. Metode Penelitian

Metodologi penelitian merupakan alur atau tahapan-tahapan yang akan dilakukan pada penelitian ini yang telah disusun sedemikian rupa sesuai dengan objek yang akan diteliti.



Gambar 1. Metode Penelitian

1. Pengumpulan Data

Pengumpulan data adalah langkah pertama yang dilakukan dalam penelitian. Pengumpulan Data sangat penting dilakukan dikarenakan data merupakan sumber utama untuk melakukan penelitian ini. Data yang digunakan dalam penelitian ini yaitu data *private* yang diambil dari PT. Telkom Akses Indonesia bagian Riau Daratan. Data yang diambil merupakan data Pasang Sambungan Baru (PSB) pada bulan Agustus 2024 yang berjumlah 1225 data. Data ini merupakan hasil penggabungan data harian PSB. Data PSB yang digunakan memiliki 15 fitur, di antaranya koordinat ODP (Optical Distribution Point), koordinat rumah pelanggan,

panjang kabel DC (Drop Core), dan 10 nilai pengukuran redaman (Redaman 1 hingga Redaman 10).

2. Pra-pemrosesan

Dataset yang telah dikumpulkan masih berupa data yang tidak terstruktur dan terdapat data yang tidak digunakan, sehingga diperlukan tahapan pra-pemrosesan atau *preprocessing*. Tahapan ini dilakukan agar data yang telah dipersiapkan dapat dirubah menjadi data yang siap untuk dianalisis [9]. Dalam penelitian ini, tahapan pemrosesan data meliputi beberapa langkah, seperti pembersihan data, rekayasa fitur, dan transformasi data.

a. Pembersihan Data

Data yang dikumpulkan sering kali mengandung berbagai masalah. Data yang memiliki *noise*, data yang tidak konsisten, atau tidak relevan dihilangkan melalui pembersihan data [10]. Proses ini meliputi penanganan nilai-nilai yang kosong (*missing values*) pada dataset.

b. Rekayasa Fitur

Data yang telah diberisihkan akan dilakukan rekayasa fitur atau *feature engineering*. Rekayasa Fitur dilakukan untuk mengekstrak fitur yang berguna dari data agar memudahkan model dalam membedakan kelas [11]. Pada penelitian ini, ditambahkan fitur baru berupa *mean_redaman* (rata-rata nilai redaman), *std_redaman* (standar deviasi redaman), dan *slope_redaman* (tren perubahan redaman yang diperoleh melalui regresi linear).

c. Transformasi Data

Data akan dilakukan transformasi yang bertujuan untuk mengubah data asli kedalam data yang dapat diproses oleh model [12]. Transformasi dilakukan dengan mengubah nilai integer ke dalam bentuk desimal agar skala fitur lebih beragam dan sesuai untuk model machine learning.

3. Pembagian Data

Setelah tahapan pra-pemrosesan, maka data akan dibagi menjadi data latih dan data uji. Pembagian data ini dilakukan untuk memastikan model yang digunakan dapat bekerja dengan baik pada data baru [13]. Penelitian ini menggunakan 2 rasio, yaitu 80:20, dan 90:10. Data yang telah dibagi ini akan diuji kedalam model.

Pembagian data dengan dua rasio berbeda ditujukan untuk mengevaluasi pengaruh ukuran data latih terhadap performa model prediksi. Berdasarkan hasil evaluasi, rasio yang memberikan performa terbaik akan dipilih untuk implementasi model akhir.

4. Prediksi *Random Forest Regression*

Penelitian ini menggunakan algoritma *Random Forest Regression* untuk memprediksi nilai redaman jaringan fiber optik berdasarkan bukti empiris dari beberapa penelitian sebelumnya. [6] menunjukkan keunggulan Random Forest Regression dalam memprediksi produksi daya pada PLTS dengan R^2 0.9679 dan RMSE 0.0438, membuktikan kemampuannya menangani data kompleks dan hubungan *non-linear* yang umum pada parameter jaringan telekomunikasi. [7]

membuktikan kemampuan *Random Forest* mengatasi *overfitting* dengan RMSE 0.440 dan akurasi 81.5% dalam prediksi harga rumah, sementara [8] menegaskan keunggulan algoritma ini dengan akurasi 80% dan R^2 0.8 dalam prediksi harga kapal, sekaligus menyoroti kemampuannya mengidentifikasi faktor-faktor berpengaruh, mirip dengan kebutuhan penelitian ini untuk menganalisis parameter penting redaman jaringan. Dengan karakteristik dataset 1225 *record*, *Random Forest* merupakan pilihan tepat karena performa baiknya pada dataset tersebut, yang menjadikannya relevan untuk optimalisasi kinerja sistem jaringan fiber optik.

Prediksi atau *forecasting* adalah proses memperkirakan kemungkinan yang akan terjadi di masa depan berdasarkan data yang tersedia [14]. Proses ini melibatkan penggunaan model tertentu untuk mengolah data historis, sehingga dapat menghasilkan perkiraan yang lebih akurat mengenai kondisi yang akan datang [15].

Random Forest Regression adalah model regresi berbasis *supervised learning* yang menggunakan beberapa pohon keputusan untuk meningkatkan akurasi dan stabilitas prediksi [16]. Metode ini bekerja dengan membangun banyak pohon pada subset data yang dipilih secara acak, lalu menggabungkan hasil prediksi untuk mengurangi *overfitting* dan meningkatkan generalisasi [17]. Berikut adalah tahapan perhitungan menggunakan *Random Forest Regression*:

$$\hat{y} = \frac{1}{N} \sum_{i=1}^N \hat{y}_i \quad (1)$$

Keterangan:

$\hat{y}$ = Nilai prediksi akhir.

N = Jumlah total pohon Keputusan dalam hutan.

$\hat{y}_i$ = Nilai prediksi yang dihasilkan oleh pohon- i.

5. Evaluasi Model

Tahapan selanjutnya dilakukan evaluasi model. Evaluasi model digunakan untuk mengevaluasi sejauh mana model dapat memprediksi dengan baik [18]. Penelitian ini menggunakan metrik evaluasi seperti *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), *R-Squared* (R^2), dan *Mean Absolute Error* (MAE). Pemilihan kombinasi metrik evaluasi ini bertujuan untuk memberikan gambaran komprehensif tentang performa model dari berbagai perspektif - baik dari segi kesalahan absolut (MAE), kesalahan kuadrat (MSE dan RMSE), maupun kemampuan model dalam menjelaskan variasi data (R^2).

a. *Mean Squared Error* (MSE)

Mean Squared Error (MSE) adalah metrik evaluasi yang menghitung rata-rata selisih kuadrat antara nilai aktual dan prediksi, memberikan bobot lebih besar pada kesalahan besar [19]. MSE banyak digunakan dalam model prediktif berbasis *machine learning*, termasuk dalam analisis redaman jaringan fiber optik, untuk mengukur sejauh mana hasil prediksi mendekati nilai aktual [20]. Semakin kecil nilai MSE, semakin akurat model yang digunakan [21]. Berikut rumus MSE:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

Keterangan:

n = Jumlah data.

y_i = Nilai aktual ke- i .

$\hat{y}_i$ = Nilai prediksi ke- i .

b. *Root Mean Squared Error (RMSE)*

Root Mean Squared Error (RMSE) merupakan akar dari MSE, yang memberikan interpretasi lebih intuitif terhadap tingkat kesalahan dalam satuan yang sama dengan data asli [22]. RMSE lebih sensitif terhadap kesalahan besar, sehingga sering digunakan dalam penelitian prediksi redaman jaringan untuk mengukur seberapa jauh prediksi model dari nilai aktual [23]. Nilai RMSE yang lebih kecil menunjukkan model dengan akurasi lebih baik dalam memperkirakan nilai redaman [24]. Berikut rumus RMSE:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

Keterangan:

n = Jumlah data.

y_i = Nilai aktual ke- i .

$\hat{y}_i$ = Nilai prediksi ke- i .

c. *R-Squared (R^2)*

R-Squared (R^2) digunakan untuk mengukur sejauh mana model dapat menjelaskan variabilitas dalam data, dengan nilai berkisar antara 0 hingga 1 [25]. Nilai R^2 yang tinggi menunjukkan bahwa model mampu menangkap pola hubungan antar variabel dengan baik, sedangkan nilai yang rendah mengindikasikan bahwa model kurang efektif dalam menjelaskan variasi data [26]. Dalam prediksi redaman jaringan, R^2 digunakan untuk menilai keandalan model regresi dalam memprediksi nilai redaman berdasarkan data historis. Berikut rumus R^2 :

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

Keterangan:

SS_{res} = Jumlah kuadrat kesalahan residu (*Sum of Squares of Residuals*).

SS_{tot} = Jumlah kuadrat total (*Total Sum of Squares*).

d. *Mean Absolute Error (MAE)*

Mean Absolute Error (MAE) menghitung rata-rata selisih absolut antara nilai aktual dan prediksi, memberikan gambaran langsung tentang tingkat akurasi model [27]. Berbeda dengan MSE dan RMSE, MAE tidak memberikan bobot lebih besar pada kesalahan besar, sehingga lebih stabil terhadap outlier [19]. Dalam penelitian redaman jaringan, MAE digunakan sebagai metrik yang mudah diinterpretasikan untuk mengevaluasi seberapa dekat hasil prediksi dengan nilai sebenarnya. Berikut rumus dari MAE:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{actual,i} - y_{pred,i}| \quad (5)$$

Keterangan:

- n = Jumlah total data.
 $y_{actual,i}$ = Nilai aktual pada data ke- i.
 $y_{pred,i}$ = Nilai prediksi pada data ke- i.
 $| \cdot |$ = Melambungkan nilai absolut.

C. Hasil dan Pembahasan

1. Pengumpulan Data

Pada tahap ini, data yang digunakan dalam penelitian ini yaitu data *private* yang diambil dari PT. Telkom Akses Indonesia bagian Riau Daratan. Data yang diambil merupakan data Pasang Sambungan Baru (PSB) pada bulan Agustus 2024. Data ini merupakan hasil penggabungan data harian PSB. Data ini berjumlah 1225 data dan memiliki 15 kolom. Dataset dapat dilihat pada Tabel 1.

Tabel 1. Dataset

ID	Longitude ODP	Latitude ODP	Longitude Rumah	Latitude Rumah	Panjang Kabel DC	Redaman 1	Redaman 2
0					200.0	-22.000	-22.070
1					150.0	-23.180	-23.370
2					150.0	-16.230	-16.290
3					250.0	-20.700	-20.700
4					150.0	-16.860	-16.920
...	...	...	...	...	...	...	...
1220					250.0	-22.598	-22.598
1221					100.0	-22.078	-22.148
1222					150.0	-17.722	-17.826
1223					250.0	-15.348	-15.348
1224					200.0	-21.428	-21.428

Redaman 3	Redaman 4	Redaman 5	Redaman 6	Redaman 7	Redaman 8	Redaman 9	Redaman 10
-21.940	-22.000	-21.940	-22.000	-22.000	-22.000	-22.000	-22.000
-23.560	-23.370	-23.100	-23.560	-23.370	-23.280	-23.180	-23.560
-16.290	-16.290	-16.230	-16.290	-16.290	-16.230	-16.230	-16.290
-20.750	-20.650	-20.750	-20.700	-20.700	-20.810	-20.750	-20.750
-16.860	-16.960	-16.920	-16.770	-16.820	-16.770	-16.860	-16.820
...	...	...	...	...	...	...	...
-22.598	-22.518	-22.598	-22.598	-22.518	-22.518	-22.366	-22.366
-22.008	-22.008	-22.008	-21.872	-21.938	-21.872	-21.938	-21.872
-17.748	-17.800	-17.852	-18.042	-18.210	-18.268	-18.298	-18.268
-15.348	-15.422	-15.200	-15.200	-15.230	-15.230	-15.230	-15.200
-21.428	-21.488	-21.428	-21.488	-21.488	-21.488	-21.428	-21.428

Tabel 1 merupakan dataset yang memiliki beberapa fitur, yaitu:

- Longitude* ODP adalah koordinat garis bujur dari titik *Optical Distribution Point* (ODP).
- Latitude* ODP adalah koordinat garis lintang dari titik *Optical Distribution Point* (ODP).
- Longitude* Rumah adalah koordinat garis bujur dari lokasi rumah pelanggan.
- Latitude* Rumah adalah koordinat garis lintang dari lokasi rumah pelanggan.
- Panjang Kabel DC adalah panjang kabel *drop core* (DC) yang digunakan dari ODP ke rumah pelanggan (dalam meter).
- Redaman 1 - Redaman 10 adalah nilai redaman optik yang diukur dalam sepuluh kali pengujian, dinyatakan dalam desibel (dB).

2. Pra-pemrosesan

Pada tahap prapemrosesan, dilakukan beberapa langkah untuk memastikan data siap digunakan dalam proses pemodelan. Pertama, dilakukan pembersihan data yang memiliki nilai kosong. Lalu dilakukan transformasi data dengan mengubah nilai integer ke dalam bentuk desimal agar skala fitur lebih seragam dan sesuai untuk model *machine learning*. Selanjutnya, dilakukan *feature engineering* dengan menambahkan fitur baru, yaitu *mean_redaman*, *std_redaman*, dan

slope_redaman, yang dihitung berdasarkan nilai redaman dari beberapa titik pengukuran sebelumnya. *Mean* digunakan untuk mengetahui rata-rata redaman dalam suatu jalur, standar deviasi digunakan untuk mengukur sebaran nilai redaman, sedangkan *slope* diperoleh melalui regresi *linear* untuk memahami tren perubahan redaman dari waktu ke waktu.

3. Pembagian Data

Dikarenakan jumlah data hanya 1225 data, maka penelitian ini menggunakan 2 kali rasio pembagian data, 90:10, dan 80:20. Untuk pembagian data 90:10, data dengan nilai 90% dari total jumlah data akan digunakan sebagai data latih, dan nilai 10% digunakan sebagai data uji. Untuk pembagian data 80:20, data dengan nilai 80% dari total jumlah data akan digunakan sebagai data latih, dan nilai 20% digunakan sebagai data uji. Diantara dua ratio tersebut akan digunakan salah satu yang terbaik untuk dimasukkan kedalam model prediksi.

4. Prediksi *Random Forest Regression*

```
# Preparing Features and Target
X = df[['REDAMAN 1', 'REDAMAN 2', 'REDAMAN 3', 'REDAMAN 4', 'REDAMAN 5',
        'REDAMAN 6', 'REDAMAN 7', 'REDAMAN 8', 'REDAMAN 9',
        'mean_redaman', 'std_redaman', 'slope_redaman',
        'PANJANG KABEL DC', 'LONGTITUDE ODP', 'LATITUDE ODP',
        'LONGTITUDE RUMAH', 'LATITUDE RUMAH']]
y = df['REDAMAN 10'] # Assume we want to predict the last redaman value

# Splitting data into train and test sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.1, random_state=42)

# Model Training
model = RandomForestRegressor(n_estimators=100, random_state=42)
model.fit(X_train, y_train)

# Prediction
y_pred = model.predict(X_test)

# Evaluation
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
rmse = np.sqrt(mse) # RMSE
mae = mean_absolute_error(y_test, y_pred)

print(f'Mean Absolute Error (MAE): {mae}')
print(f'Mean Squared Error (MSE): {mse}')
print(f'Root Mean Squared Error (RMSE): {rmse}')
print(f'R^2 Score: {r2}')
```

Gambar 2. Source Code *Random Forest Regression*

Pada gambar 2 menunjukkan skrip program yang digunakan untuk membangun model *Random Forest Regression* dalam memprediksi nilai redaman jaringan fiber optik. Pada skrip ini, data terlebih dahulu dibagi menjadi data latih (*training set*) dan data uji (*test set*) dengan perbandingan 90:10 menggunakan

metode *train_test_split*. Selanjutnya, model dilatih menggunakan 100 pohon keputusan ($n_estimators=100$) dengan parameter *random_state=42* untuk memastikan hasil yang konsisten. Setelah proses pelatihan selesai, model digunakan untuk melakukan prediksi pada data uji.

5. Evaluasi Model

Evaluasi model dilakukan untuk menilai kinerja *Random Forest Regression* dalam memprediksi nilai redaman jaringan fiber optik dengan dua skenario pembagian data: 90:10 dan 80:20. Hasil evaluasi menunjukkan bahwa model dengan *splitting* 90:10 memiliki performa lebih baik dibandingkan 80:20, dengan tingkat kesalahan yang lebih rendah dan akurasi lebih tinggi.

Tabel 2. Metrik Evaluasi

Metrik	90:10	80:20
<i>Mean Absolute Error</i> (MAE)	0.1009	0.1479
<i>Mean Squared Error</i> (MSE)	0.0559	0.6424
<i>Root Mean Squared Error</i> (RMSE)	0.2363	0.8015
R^2 Score	0.9832	0.8278

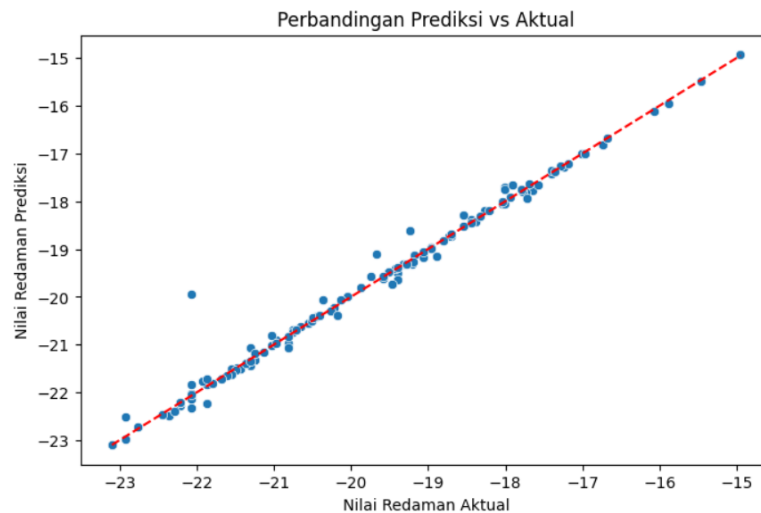
Pada skenario 90:10, nilai *Mean Absolute Error* (MAE) sebesar 0.1009 menunjukkan bahwa rata-rata selisih antara nilai prediksi dan nilai aktual relatif kecil. *Mean Squared Error* (MSE) sebesar 0.0559 dan *Root Mean Squared Error* (RMSE) sebesar 0.2363 juga mengindikasikan bahwa model mampu melakukan prediksi dengan tingkat kesalahan yang rendah. Selain itu, nilai R^2 Score sebesar 0.9832 menandakan bahwa model dapat menjelaskan 98.32% variasi dalam data, yang berarti model mampu menangkap pola redaman jaringan dengan sangat baik.

Sebaliknya, pada skenario 80:20, performa model menurun secara signifikan. MAE meningkat menjadi 0.1479, menunjukkan adanya peningkatan rata-rata kesalahan prediksi. MSE juga melonjak menjadi 0.6424, sedangkan RMSE meningkat menjadi 0.8015, yang mengindikasikan bahwa tingkat kesalahan prediksi menjadi lebih besar. Nilai R^2 Score juga mengalami penurunan menjadi 0.8278, menunjukkan bahwa model hanya dapat menjelaskan 82.78% variasi dalam data, yang berarti tingkat akurasi model dalam memahami pola redaman jaringan menurun ketika jumlah data latih berkurang. Maka akan digunakan pembagian data dengan rasio 90:10 dalam memprediksi nilai redaman.

Untuk memahami sejauh mana akurasi model dalam memprediksi nilai redaman, dilakukan beberapa langkah evaluasi. Visualisasi berikut menggambarkan bagaimana model bekerja dalam membandingkan hasil prediksi dengan nilai aktual serta melakukan estimasi nilai redaman untuk titik-titik selanjutnya.

a. Visualisasi *Scatter Plot* Perbandingan Prediksi vs Aktual

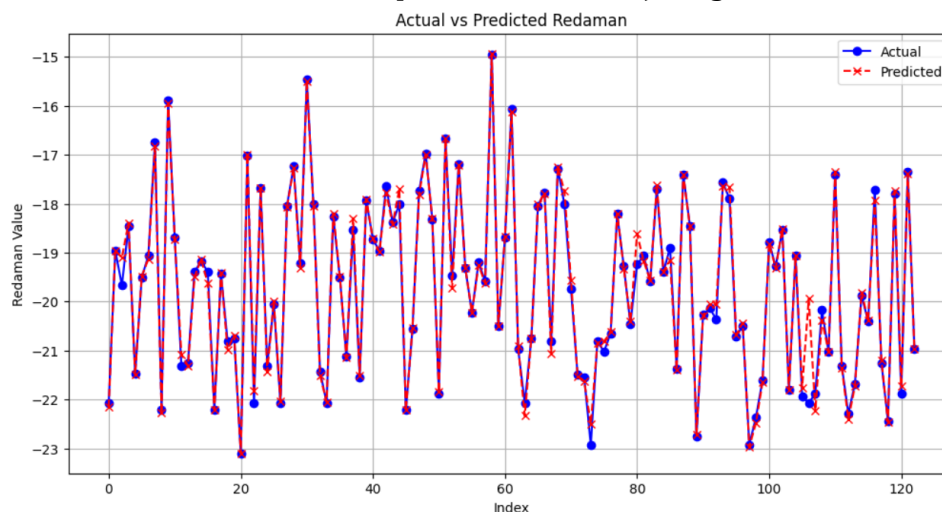
Scatter plot ini menunjukkan hubungan antara nilai redaman aktual dan nilai redaman prediksi. Semakin mendekati garis merah $y = x$, semakin baik model dalam melakukan prediksi.



Gambar 3. Scatter plot perbandingan prediksi vs aktual

b. Visualisasi Garis Tren Aktual vs Prediksi

Grafik ini menampilkan perbandingan nilai aktual dan prediksi dalam bentuk garis tren. Perbedaan antara kedua garis menunjukkan tingkat kesalahan model dalam memprediksi redaman jaringan.



Gambar 4. Garis tren prediksi vs aktual

c. Prediksi Nilai Redaman 11 – 20

Untuk menguji performa model dalam memprediksi nilai redaman jaringan fiber optik, dilakukan pengujian dengan menggunakan satu sampel data sebagai input. Sampel data ini terdiri dari nilai redaman sebelumnya (REDAMAN 1 hingga REDAMAN 9) serta informasi tambahan mengenai panjang kabel dan lokasi. Model *Random Forest Regression* kemudian digunakan untuk memperkirakan nilai redaman berikutnya secara iteratif.

Tabel 3. Data Input untuk Prediksi Nilai Redaman

REDAMAN	Nilai
---------	-------

REDAMAN 1	-22.598
REDAMAN 2	-22.598
REDAMAN 3	-22.598
REDAMAN 4	-22.518
REDAMAN 5	-22.598
REDAMAN 6	-22.598
REDAMAN 7	-22.518
REDAMAN 8	-22.518
REDAMAN 9	-22.366
PANJANG KABEL DC	250.0
LONGTITUDE ODP	
LATITUDE ODP	
LONGTITUDE RUMAH	
LATITUDE RUMAH	

Setelah model dilatih menggunakan data historis, prediksi nilai redaman dilakukan secara iteratif berdasarkan data input di atas. Model *Random Forest Regression* memperkirakan nilai redaman untuk REDAMAN 11 hingga REDAMAN 20 dengan hasil sebagai berikut:

Tabel 4. Hasil prediksi nilai redaman 11 - 20

REDAMAN	Nilai Prediksi
REDAMAN 11	-22.4017
REDAMAN 12	-22.4122
REDAMAN 13	-22.4287
REDAMAN 14	-22.4312
REDAMAN 15	-22.4323
REDAMAN 16	-22.4237
REDAMAN 17	-22.4341
REDAMAN 18	-22.4357
REDAMAN 19	-22.4373
REDAMAN 20	-22.4341

Hasil ini menunjukkan bahwa nilai redaman cenderung stabil dengan sedikit fluktuasi dari REDAMAN 11 hingga REDAMAN 20. Tren yang diperoleh dari model dapat digunakan sebagai dasar dalam menilai kinerja jaringan fiber optik di masa depan.

D. Simpulan

Algoritma *Random Forest Regression* terbukti efektif dalam memprediksi nilai redaman jaringan fiber optik dengan akurasi tinggi (R^2 Score 0.9832), menunjukkan superioritas model ini dalam menangani data kompleks dengan hubungan nonlinier antara variabel input dan output. Rasio pembagian data 90:10 menghasilkan performa optimal dibandingkan 80:20, yang mengindikasikan bahwa model machine learning ini membutuhkan volume data training yang cukup besar untuk menghasilkan generalisasi yang baik. Model berhasil memprediksi tren redaman untuk periode mendatang (Redaman 11-20) dengan pola yang stabil, memungkinkan penilaian kinerja jaringan fiber optik di masa depan berdasarkan data prediktif, bukan hanya reaktif. Implementasi machine learning dalam prediksi nilai redaman memungkinkan operator telekomunikasi menilai kinerja jaringan secara proaktif, mengidentifikasi potensi penurunan performa sebelum berdampak

pada layanan pelanggan, sehingga meningkatkan keandalan infrastruktur FTTH secara keseluruhan.

E. Referensi

- [1] M. Nurus, O. Nurdiawa, and M. Martanto, "Analisis Jaringan Akses Fiber to The Home Menggunakan Teknologi Gigabit Passive Optical Network," *Jurnal Janitra Informatika dan Sistem Informasi*, vol. 3, no. 2, pp. 56–66, Oct. 2023, doi: 10.25008/janitra.v3i2.168.
- [2] P. A. Saputra *et al.*, "Pengaruh Teknik Splice Fiber Optic Terhadap Redaman (DB) Pada PT. Pandawa Karya Wijaya Putra Mitra ICON NET," 2025. [Online]. Available: <https://jurnal.bsi.ac.id/index.php/ijec/>
- [3] M. Rasyid Ali, A. Rifai, I. Nisaa, P. studi Manajemen Industri, and A. Telekomunikasi Bogor, "Fiber Optic Network Service Quality Measurement," 2022.
- [4] M. A. Rahmatulloh, D. Hanto, M. Yantidewi, A. Rianaris, and R. A. Firdaus, "Analisis Redaman Fiber Optik dengan Menggunakan Pemodelan Software Optisystem Optik Fiber Attenuation Analysis Using Optisystem Software Modeling," 2023. [Online]. Available: <https://jurnal.unismuhpalu.ac.id/index.php/JKS>
- [5] J. Arya, D. Purnama, I. Lammada, T. Elektro, S. Karawang, and J. H. R. Waluyo, "Analisa Performansi Redaman Serat Optik Pada Otb (Optical Termination Box) Menggunakan Optical Power Meter Di Pt Aquila Wijaya Teknik," 2024.
- [6] Ardianto, Agus Budi Raharjo, and Diana Purwitasari, "Random Forest Regression Untuk Prediksi Produksi Daya Pembangkit Listrik Tenaga Surya," 2022.
- [7] E. Fitri, "Analisis Perbandingan Metode Regresi Linier, Random Forest Regression dan Gradient Boosted Trees Regression Method untuk Prediksi Harga Rumah," *JOURNAL OF APPLIED COMPUTER SCIENCE AND TECHNOLOGY (JACOST)*, vol. 4, no. 1, pp. 2723–1453, 2023, doi: 10.52158/jacost.491.
- [8] Z. Lin, X. Huang, and J. Chen, "Prediction Model of Sailing Ship Price Based on Random Forest Regression," 2023.
- [9] I. Saputra *et al.*, "Analisis Sentimen Pengguna Marketplace Bukalapak dan Tokopedia di Twitter Menggunakan Machine Learning," *Faktor Exacta*, vol. 13, no. 4, p. 200, Feb. 2021, doi: 10.30998/faktorexacta.v13i4.7074.
- [10] F. Putra, H. F. Tahiyat, R. M. Ihsan, R. Rahmaddeni, and L. Efrizoni, "Penerapan Algoritma K-Nearest Neighbor Menggunakan Wrapper Sebagai Preprocessing untuk Penentuan Keterangan Berat Badan Manusia," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, pp. 273–281, Jan. 2024, doi: 10.57152/malcom.v4i1.1085.
- [11] E. Dariato, "Analisa dan Perancangan Machine Learning Untuk Mendeteksi Kegagalan Job di Apache Spark," *Arcitech: Journal of Computer Science and Artificial Intelligence*, vol. 2, no. 1, p. 1, Jun. 2022, doi: 10.29240/arcitech.v2i1.4124.
- [12] A. Khoirala Ermy Pily, Oktavianda, F. Aprilia, Rahmaddeni, and L. Efrizoni, "Komparasi Algoritma K-Nearest Neighbors dan Naïve Bayes dalam Klasifikasi Penyakit Diabetes Gestasional," *Indonesian Journal of Computer Science Attribution*, vol. 13, no. 1, pp. 2024–1195, 2024.

- [13] F. Putra, R. Muhammad Ihsan, H. Fulaila Tahiyat, L. Efrizoni, and Rahmadden, "Evaluation Of Gojek Application Performance Through User Review Word Classification Using The SVM Method," 2024.
- [14] S. Wulandari, P. Sugiartawan, I. Made, and D. Setiawan, "Prediksi Luas Sebaran Hama Wareng pada Tanaman Padi dengan RNN Time Series," *Jurnal Sistem Informasi dan Komputer Terapan Indonesia (JSIKTI)*, vol. 5, no. 1, pp. 21–32, 2022, doi: 10.33173/jsikti.174.
- [15] C. Chandra and S. Budi, "Analisis Komparatif ARIMA dan Prophet dengan Studi Kasus Dataset Pendaftaran Mahasiswa Baru," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 6, no. 2, Aug. 2020, doi: 10.28932/jutisi.v6i2.2676.
- [16] Diana Tri Susetianingtias, Eka Patriya, and Rodiah, "Model Random Forest Regression Untuk Peramalan Penyebaran Covid-19 Di Indonesia," *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 2, no. 2, pp. 84–95, Sep. 2022, doi: 10.51454/decode.v2i2.48.
- [17] V. Hafizh, C. Putra, M. Al-Husaini, A. P. Wahyu, and A. R. Raharja, "MALCOM: Indonesian Journal of Machine Learning and Computer Science Design of an Intelligent Monitoring System Based on Internet of Things (IoT) with Random Forest Regression Algorithm for Height Detection in Cherry Tomato Plants Perancangan Sistem Monitoring Cerdas Berbasis Internet of Things (IoT) dengan Algoritma Random Forest Regression untuk Deteksi Ketinggian pada Tanaman Tomat Cherry," vol. 5, no. 1, pp. 10–25, 2025, doi: 10.57152/malcom.v5i1.1612.
- [18] D. N. Agustia and R. R. Suryono, "Comparison Of Naïve Bayes, Random Forest, And Logistic Regression Algorithms For Sentiment Analysis Online Gambling Komparasi Algoritma Naïve Bayes, Random Forest, Dan Logistic Regression Untuk Analisis Sentimen Judi Online," vol. 10, no. 1, 2025.
- [19] Hommy D. E. Sinaga and Novica Irawati, "Perbandingan Double Moving Averagedengan Double Exponential Smoothingpada Peramalan Bahan Medis Habis Pakai," 2018.
- [20] A. Sulaiman and A. Juarna, "Peramalan Tingkat Pengangguran Di Indonesia Menggunakan Metode Time Series Dengan Model Arima Dan Holt-Winters," *Jurnal Ilmiah Informatika Komputer*, vol. 26, no. 1, pp. 13–28, 2021, doi: 10.35760/ik.2021.v26i1.3512.
- [21] A. Adnan Rusdy, "Buletin Sistem Informasi dan Teknologi Islam Penerapan Metode Regresi Linear pada Prediksi Penawaran dan Permintaan Obat Studi Kasus Aplikasi Point of Sales INFORMASI ARTIKEL ABSTRAK," vol. 3, no. 2, pp. 121–126, 2022.
- [22] S. muntari and W. Agustina, "Implementasi Data Mining Untuk Memprediksi Harga Cabai Merah Kota Pagar Alam Menggunakan Metode K-Nearest Neighbor," 2024.
- [23] M. ' Ruf *et al.*, "Algoritma Adaptive Neuro Fuzzy Inference System Untuk Perkiraan Intensitas Curah Hujan," 2021.
- [24] Laila Sari, Syaiful Zuhri Harahap, and Irmayanti Ritonga, "Memprediksi Data Saham Bank Mandiri Menggunakan Metode Algoritma Regresi Linear Dengan Bantuan Rapid Miner," 2024.
- [25] Muh. Asharif Suleman and Zulfi Idayanti, "Faktor-Faktor Yang Mempengaruhi Keberhasilan Proses Pembelajaran diSDMI," 2024.

- [26] M. Akil, H. Umar, and B. Sitohang, "Analysis Of Factors Influencing The Decision To Purchase Tour Packages Using The Decision Trees, Random Forest And K-Nearest Neighbors Classification Models," *Journal of Social and Economics Research*, vol. 6, no. 2, 2024, [Online]. Available: <https://idm.or.id/JSER/index.php/JSER>
- [27] Ilham Amansyah, Jamaludin Indra, Euis Nurlaelasari, and Ayu Ratna Juwita, "Prediksi Penjualan Kendaraan Menggunakan Regresi Linear: Studi Kasus pada Industri Otomotif di Indonesia," 2024.