

ANALISIS PERBANDINGAN *LINEAR REGRESSION* DAN *RANDOM FOREST REGRESSION* UNTUK PREDIKSI BATAS KREDIT: PENDEKATAN OPTIMASI *HYPERPARAMETER*

Algies Rifkha Fadillah¹, Mohamad Nurkamal Fauzan², Nuraini Siti Fathonah³

^{1,2,3}Jurusan Informatika, Vokasi, Universitas Logistik Bisnis Internasional
¹algies.rifkha@gmail.com, ²m.nurkamal.f@ulbi.ac.id, ³nurainisf@ulbi.ac.id

Abstrak

Penelitian ini mengeksplorasi penggunaan *Linear Regression* dan *Random Forest Regression* untuk memprediksi batas maksimal pinjaman kredit dalam industri keuangan, di mana penentuan batas kredit sangat penting untuk mengurangi risiko gagal bayar dan menjaga stabilitas lembaga keuangan. Dalam penelitian ini, metode tersebut digunakan untuk membandingkan performa prediksi dengan dua set fitur yang berbeda. Setelah melakukan optimasi *hyperparameter* menggunakan Optuna, hasil menunjukkan bahwa *Random Forest Regression* memberikan akurasi prediksi yang lebih tinggi dibandingkan *Linear Regression*, dengan nilai RMSE terendah sebesar 7.90% dan MAE terendah sebesar 4.72%. Penggunaan 4 fitur menunjukkan sedikit peningkatan dalam akurasi dibandingkan 7 fitur, meskipun tidak signifikan. Hasil ini menyarankan penggunaan *Random Forest Regression* untuk meningkatkan keakuratan dalam menetapkan batas kredit, mengurangi risiko kredit, dan meningkatkan stabilitas lembaga keuangan. Dengan demikian, pengoptimalan *hyperparameter* dengan Optuna dapat meningkatkan performa prediksi model *Regression* dalam konteks ini.

Kata kunci : *linear regression*, *random forest regression*, optimasi *hyperparameter*, batas pinjaman kredit.

1. Pendahuluan

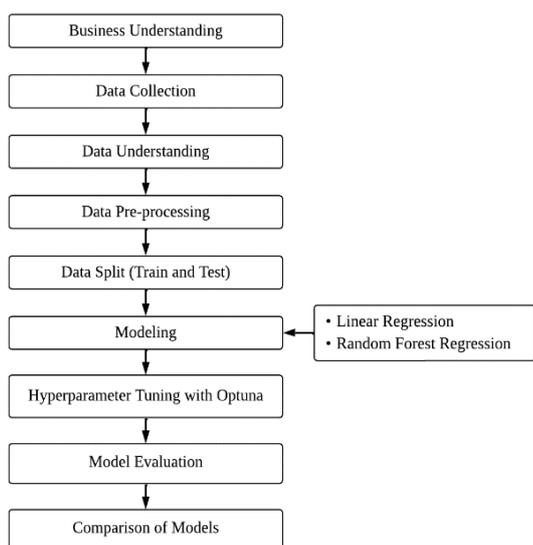
Industri keuangan memiliki peran penting dalam memfasilitasi pertumbuhan ekonomi (Piracha et al., 2022). Salah satu layanan utama yang ditawarkan oleh lembaga keuangan adalah pinjaman kepada nasabahnya (Narayana et al., 2022). Pinjaman ini memainkan peran vital dalam memberikan akses modal kepada individu dan bisnis untuk berinvestasi, mengembangkan usaha, atau memenuhi kebutuhan finansial mendesak (I.O. Artemieva, Igor S. Bondaruk, 2022; Oksana Savko, Juliia Bui, 2022). Namun, dalam memberikan pinjaman, bank harus mempertimbangkan berbagai faktor kompleks untuk memastikan keberlanjutan bisnis mereka serta mengelola risiko secara efektif (Ayhan & Uçar, 2022). Penentuan batas pinjaman maksimum yang tepat adalah salah satu aspek krusial dalam proses ini (Zhang & Wang, 2018). Penentuan yang tidak tepat dapat mengakibatkan risiko kredit yang tinggi, baik berupa gagal bayar oleh nasabah maupun *over-lending* yang berdampak negatif pada kestabilan finansial institusi tersebut (Ganga et al., 2021). Oleh karena itu, berbagai pendekatan analitis dan statistik telah digunakan untuk memprediksi dan menetapkan batas kredit yang optimal. Pendekatan tradisional dalam menentukan batas kredit sering kali tidak memadai dalam menghadapi kompleksitas data dan kebutuhan

akan akurasi yang tinggi (Ayhan & Uçar, 2022). Oleh karena itu, diperlukan metode yang lebih canggih dan adaptif. *Linear Regression* dan *Random Forest Regression* adalah dua metode yang sering digunakan untuk memodelkan hubungan antara variabel *dependen* (misalnya, batas kredit yang diberikan) dan variabel *independen* (misalnya, faktor-faktor risiko nasabah) (P. Meghana et al., 2024). Namun, kedua metode ini memiliki keterbatasan dalam hal optimasi *hyperparameter* yang dapat mempengaruhi akurasi prediksi (Olufemi Ogunleye, 2023). Optuna, sebagai framework optimasi *hyperparameter* berbasis *Bayesian Optimization*, memungkinkan pencarian parameter terbaik secara efisien (Hussain, Almarzooq., U., bin, 2024). Implementasi Optuna telah diterapkan dalam berbagai bidang seperti pengenalan pola, klasifikasi, dan regresi (Akiba et al., 2019). Namun penggunaannya dalam konteks penentuan batas kredit masih relatif baru dan menjanjikan. Penelitian ini bertujuan untuk mengevaluasi performa *Linear Regression* dan *Random Forest Regression* dalam penentuan batas kredit dengan menggunakan Optuna untuk optimasi *hyperparameter*. Tujuan akhir dari penelitian ini adalah untuk menentukan batas maksimal pinjaman kredit yang lebih akurat dan efisien, yang pada akhirnya dapat mengurangi risiko kredit dan meningkatkan kestabilan keuangan

lembaga pemberi pinjaman. Dengan mengoptimalkan *hyperparameter* model penentuan batas kredit menggunakan Optuna, diharapkan dapat dicapai model prediktif yang lebih baik dibandingkan dengan metode konvensional. Penelitian ini juga diharapkan memberikan wawasan baru mengenai penerapan teknik optimasi modern dalam sektor keuangan.

2. Metode Penelitian

Penelitian ini menggunakan metode sendiri yang dirancang untuk mengatasi masalah penentuan batas maksimal pinjaman kredit. Metode ini terdiri dari beberapa tahap sebagai berikut:



Gambar 1. Metode Penelitian

2.1 Business Understanding

Pada tahap pemahaman bisnis, fokus utama adalah memahami masalah utama yang ingin diselesaikan, yaitu penentuan batas maksimal pinjaman kredit yang dapat diberikan kepada nasabah. Penentuan batas ini sangat krusial karena keputusan yang tidak tepat dapat menimbulkan risiko kredit yang tinggi, baik berupa gagal bayar oleh nasabah maupun *over-lending* yang dapat membahayakan stabilitas finansial lembaga keuangan.

2.2 Data Collection

Mengumpulkan data historis nasabah yang mencakup riwayat kredit, pendapatan, dan faktor-faktor risiko lainnya. Metode pengumpulan data yang digunakan dalam penelitian ini meliputi wawancara dan studi pustaka. Dengan kombinasi metode wawancara dan studi pustaka, data yang dikumpulkan mencakup aspek-aspek penting yang diperlukan untuk membangun model prediksi yang akurat. Data ini sangat penting karena menjadi dasar untuk membangun model prediksi batas pinjaman maksimum. Kualitas dan kelengkapan data yang dikumpulkan akan sangat mempengaruhi hasil penelitian.

2.3 Data Understanding

Pada tahap ini, dilakukan pembelajaran dan pemahaman data yang telah dikumpulkan. Ini mencakup eksplorasi awal data untuk mengidentifikasi pola, tren, dan anomali dalam data. Dataset yang digunakan berisi informasi detail mengenai kredit dari 277 nasabah. Setiap aplikasi mencakup nomor aplikasi, nomor kredit, nomor CIF, nama nasabah, pendapatan bulanan, riwayat kredit, tanggal pengajuan, tanggal kredit disetujui, tanggal jatuh tempo, durasi tenor, dan jumlah pinjaman.

Tabel 1. Variabel

Variabel	Keterangan
No Aplikasi	Merupakan nomor aplikasi unik sebagai identifikasi dan tidak digunakan dalam pemodelan.
No Kredit	Merupakan nomor kredit unik sebagai identifikasi dan tidak digunakan dalam pemodelan.
No Cif	<i>Customer Information File</i> adalah nomor unik yang mengidentifikasi nasabah dan tidak digunakan dalam pemodelan.
Nama Nasabah	Merupakan nama nasabah yang mengajukan pinjaman dan tidak digunakan dalam pemodelan
Income Nasabah	Merupakan pendapatan nasabah
Credit History	Merupakan riwayat kredit nasabah yang mencerminkan rekam jejak pembayaran kredit sebelumnya.
Tgl Pengajuan	Merupakan tanggal pengajuan pinjaman oleh nasabah.
Tgl Kredit	Merupakan tanggal disetujuinya kredit.
Tgl Jatuh Tempo	Merupakan tanggal jatuh tempo pinjaman.
Tenor	Merupakan jangka waktu atau lama waktu pinjaman.
Uang Pinjaman	Merupakan jumlah uang yang dipinjamkan kepada nasabah.

Dataset ini memiliki 277 data dan 12 kolom dengan tipe data yang beragam. Tidak ada nilai *null* dalam *dataset* ini, dan tidak ditemukan baris duplikat.

Statistik deskriptif menunjukkan bahwa nomor aplikasi, nomor kredit, dan nomor CIF memiliki nilai yang unik. Pendapatan nasabah bervariasi dari 1.6 juta hingga 7.45 juta dengan rata-rata 4.18 juta. Tenor kredit berkisar dari 12 hingga 36 bulan, dengan rata-rata sekitar 27.21 bulan. Jumlah pinjaman bervariasi dari 3 juta hingga 10 juta, dengan rata-rata sekitar 7.18 juta. Analisis ini memberikan pemahaman awal yang penting tentang distribusi data aplikasi kredit, termasuk variasi dalam pendapatan nasabah, durasi tenor, dan jumlah pinjaman

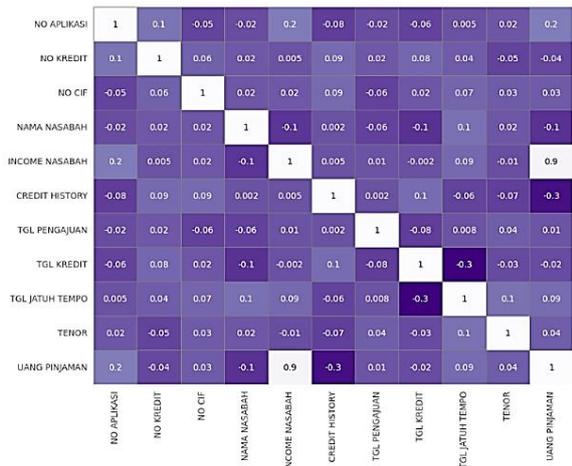
2.4 Data Pre-processing

Pra-pemrosesan data melibatkan pembersihan dan transformasi data. Langkah ini mencakup:

- Pembersihan data: menghilangkan data yang tidak lengkap atau memiliki nilai anomali (*outliers*) yang dapat mengganggu analisis.
- Transformasi data: normalisasi atau

standarisasi data agar setiap variabel memiliki skala yang sama, sehingga model dapat bekerja dengan baik dan memberikan hasil yang akurat.

Setelah melakukan transformasi data, langkah selanjutnya adalah melakukan analisis korelasi. Korelasi mengukur sejauh mana dua variabel berhubungan. Dalam analisis ini, korelasi antara berbagai kolom dalam DataFrame yang telah ditransformasi divisualisasikan menggunakan heatmap.



Gambar 2. Correlation matrix

Dari Gambar 3, fitur No_Aplikasi, No_Kredit, No_CIF, Nama_Nasabah dipilih untuk dihapus. Income_Nasabah, Credit_History, Tgl_Pengajuan, Tgl_Kredit, Tgl_Jatuh_Tempo, Tenor, dan Uang_Pinjaman merupakan tujuh atribut yang dipilih yang akan digunakan dalam prediksi. Penelitian ini akan menggunakan *Linear Regression* dan *Random Forest Regression*, dengan menggunakan dua set fitur yang berbeda: satu set dengan 7 fitur dan satu set dengan 4 fitur.

- a) Set dengan 7 fitur: Income_Nasabah, Credit_History, Tgl_Pengajuan, Tgl_Kredit, Tgl_Jatuh_Tempo, Tenor, dan Uang_Pinjaman.
- b) Set dengan 4 fitur: Income_Nasabah, Credit_History, Tenor, dan Uang_Pinjaman.

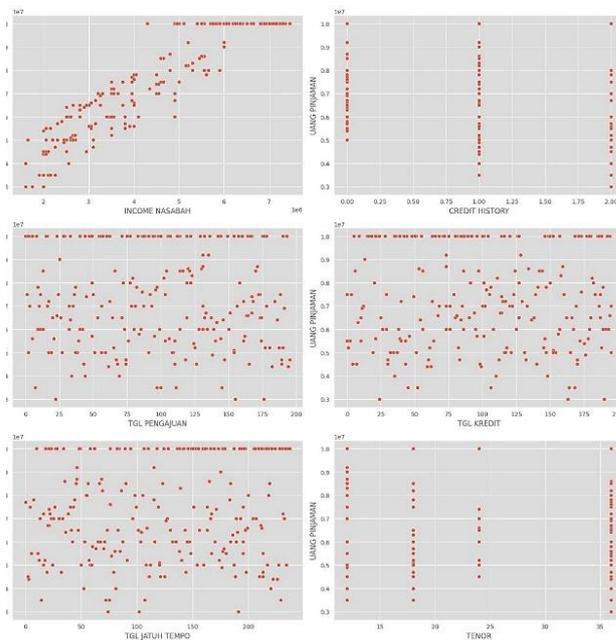
2.5 Data Split

Data yang telah dipra-pemroses kemudian dibagi menjadi dua set, yaitu set pelatihan dan set pengujian. Pembagian dilakukan dengan rasio 80:20, dimana 80% data digunakan untuk melatih model dan 20% sisanya digunakan untuk menguji kinerja model.

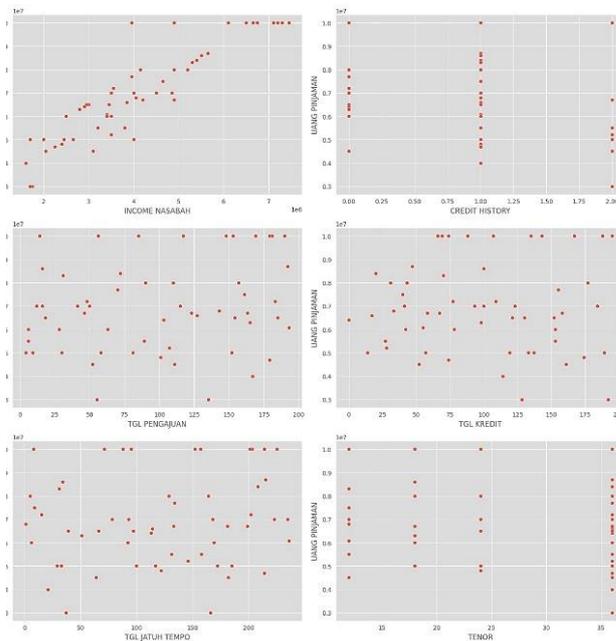
2.5.1 Set dengan 7 fitur

Scatter plot pada Gambar 4 dan Gambar 5 menunjukkan hubungan antara fitur Income_Nasabah, Credit_History, Tgl_Pengajuan, Tgl_Kredit, Tgl_Jatuh_Tempo, dan Tenor yang

dipilih dengan variabel target (Uang_Pinjaman) pada data set pelatihan dan pengujian. Plot ini memvisualisasikan pola, distribusi, dan korelasi antara fitur-fitur dan variabel target.



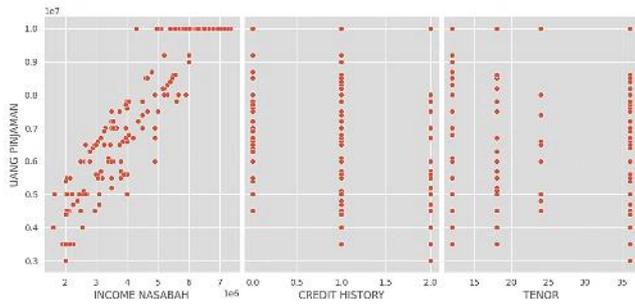
Gambar 3. Linearitas data train pada set 7 Fitur



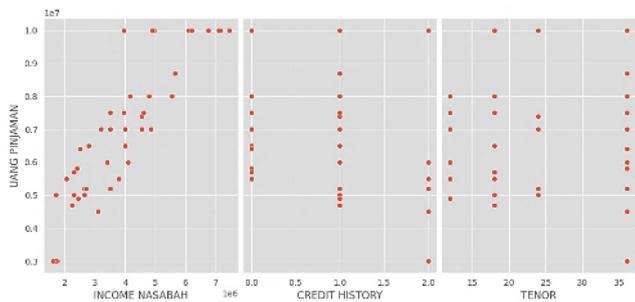
Gambar 4. Linearitas data test pada set 7 Fitur

2.5.2 Set dengan 4 fitur

Scatter plot pada Gambar 6 dan Gambar 7 menunjukkan hubungan antara fitur Income_Nasabah, Credit_History, dan Tenor yang dipilih dengan variabel target (Uang_Pinjaman) pada data set pelatihan dan pengujian. Plot ini memvisualisasikan pola, distribusi, dan korelasi antara fitur-fitur dan variabel target.



Gambar 5. Linearitas data *train* pada set 4 Fitur



Gambar 6. Linearitas data *test* pada set 4 Fitur

2.6 Modeling

Pada tahap ini, dua model digunakan untuk perbandingan yaitu *Random Forest Regression* dan *Linear Regression*. Kedua model ini dilatih menggunakan data latih untuk memprediksi batas maksimal pinjaman. *Random Forest Regression* adalah model *ensemble* yang menggunakan banyak pohon keputusan, sementara *Linear Regression* adalah model linier yang sederhana dan mudah diinterpretasikan (Karl, 2023; V., Ignatenko., Anton, Surkov., Sergej, 2024).

2.7 Hyperparameter Tuning with Optuna

Optuna digunakan untuk melakukan tuning *hyperparameter* secara otomatis dan efisien (Maryam, Hassanali., Mohammadreza, Soltanaghaei., Taghi, Javdani, Gandomani., Farsad, Zamani, 2024). *Hyperparameter tuning* melibatkan pencarian set *hyperparameter* terbaik yang meningkatkan performa model (M.S., Nidhya., Nidhi, Saraswat., Prashant, 2024). Dalam Optuna, *study* dibuat dan fungsi tujuan didefinisikan untuk mengevaluasi performa model berdasarkan set *hyperparameter* yang berbeda.

2.8 Model Evaluation

Setelah model dilatih dan di-tuning, langkah selanjutnya adalah evaluasi performa model menggunakan data uji. Metrik evaluasi seperti *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan *R-squared* (R^2) digunakan untuk mengukur akurasi prediksi model. Kedua model, *Random Forest Regression* dan *Linear Regression*, dibandingkan untuk menentukan model yang paling efektif.

2.9 Comparison of Modelgambar

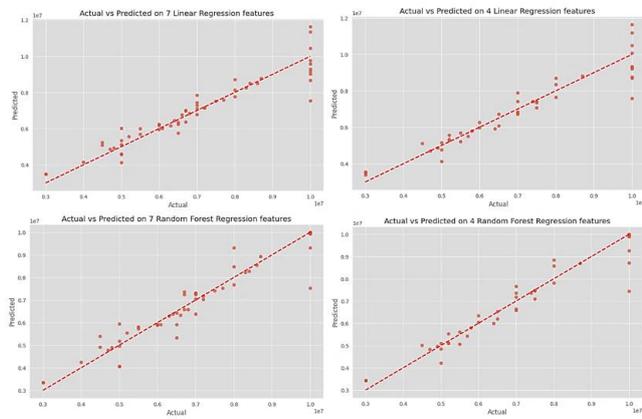
Pada bagian ini, dilakukan perbandingan kinerja dua model *Regression*, yaitu *Linear Regression* dan *Random Forest Regression*, dengan menggunakan dua set fitur berbeda, yaitu set pertama dengan 7 fitur dan set kedua dengan 4 fitur. Evaluasi model dilakukan menggunakan tiga metrik: *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan *R-squared* (R^2).

Perbandingan ini dapat memberikan wawasan mengenai dampak pemilihan fitur dan model terhadap kinerja prediksi, serta membantu dalam membuat keputusan yang tepat dalam pemodelan prediktif.

3. Hasil Dan Pembahasan

Bagian ini menyajikan hasil dari perbandingan kinerja dua model *Regression*, yaitu *Linear Regression* dan *Random Forest Regression*, yang dilatih menggunakan dua set fitur yang berbeda. Evaluasi dilakukan menggunakan metrik *Linear*, *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan *R-squared* (R^2). Hasil evaluasi ini akan dianalisis dan diinterpretasikan untuk mengidentifikasi model dan set fitur yang memberikan kinerja prediksi terbaik.

3.1 Linearitas



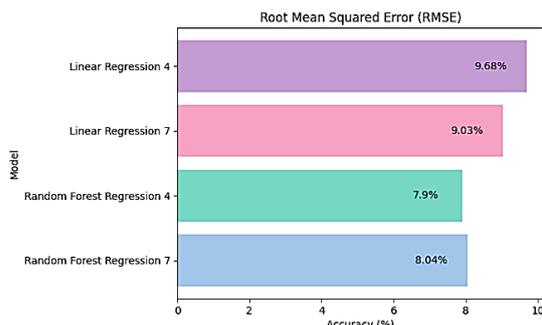
Gambar 7. Linearitas Hasil Pemodelan

Berdasarkan pada Gambar 8 yang menampilkan grafik Linear, perbandingan kinerja antara *Linear Regression* dan *Random Forest Regression* dengan menggunakan dua set fitur yang berbeda dapat dilakukan. *Linear Regression* dengan 7 fitur menunjukkan prediksi yang lebih akurat dibandingkan dengan *Linear Regression* yang hanya menggunakan 4 fitur, terlihat dari titik-titik data yang lebih dekat dengan garis diagonal.

Sementara itu, *Random Forest Regression* juga menunjukkan kinerja yang lebih baik ketika menggunakan 7 fitur dibandingkan dengan 4 fitur. Titik-titik data pada model dengan 7 fitur lebih terkonsentrasi di sekitar garis diagonal, mengindikasikan prediksi yang lebih akurat.

Ketika membandingkan kedua model *Regression* tersebut, baik pada set fitur dengan 7 maupun 4 fitur, *Random Forest Regression* secara konsisten menunjukkan kinerja yang lebih baik dibandingkan *Linear Regression*. Titik-titik data pada grafik *Random Forest Regression* lebih dekat dengan garis diagonal dibandingkan dengan titik-titik data pada grafik *Linear Regression*, menunjukkan bahwa model ini memberikan prediksi yang lebih akurat

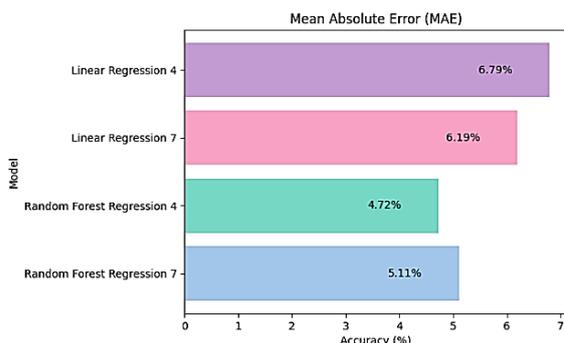
3.2 Root Mean Squared Error (RMSE)



Gambar 8. Diagram Root Mean Squared Error (RMSE)

Pada Gambar 9 menampilkan Diagram yang menunjukkan *Root Mean Squared Error (RMSE)* untuk empat model *Regression*. Model *Random Forest Regression* dengan 4 fitur memiliki RMSE terendah (7.9%), diikuti oleh model dengan 7 fitur (8.04%). Model *Linear Regression* menunjukkan RMSE lebih tinggi, dengan model 7 fitur mencapai 9.68% dan model 4 fitur 9.03%. Ini menunjukkan bahwa model *Random Forest Regression* lebih akurat dibandingkan model *Linear Regression* dalam dataset ini.

3.3 Mean Absolute Error (MAE)

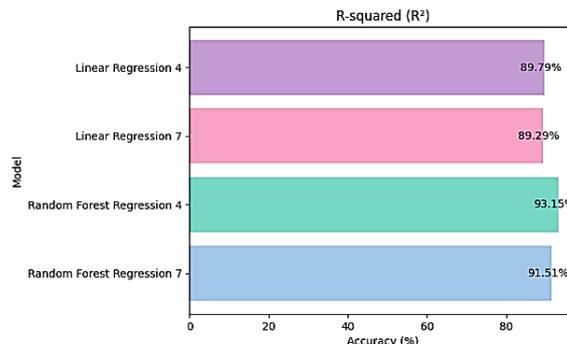


Gambar 9. Diagram Mean Absolute Error (MAE)

Pada Gambar 10 menampilkan Diagram *Mean Absolute Error (MAE)* untuk empat model *Regression*. Model *Random Forest Regression* dengan 4 fitur memiliki MAE terendah (4.77%), diikuti oleh model dengan 7 fitur (5.17%). Model *Linear Regression* menunjukkan MAE lebih tinggi, dengan model 7 fitur memiliki MAE sebesar 6.79% dan model 4 fitur sebesar 6.19%. Ini mengindikasikan bahwa model *Random Forest*

Regression lebih baik dalam meminimalkan kesalahan absolut rata-rata dibandingkan dengan model *Linear Regression* dalam dataset ini.

3.4 R-squared (R²)



Gambar 10. Diagram R-squared (R²)

Pada Gambar 11 menunjukkan akurasi model *Random Forest Regression* dengan 4 fitur memiliki akurasi tertinggi, mencapai lebih dari 93.15%. Model *Random Forest Regression* lainnya juga menunjukkan akurasi tinggi, sekitar 91.51%. Di sisi lain, model *Linear Regression* menunjukkan akurasi yang lebih rendah dibandingkan dengan model *Random Forest Regression*, dengan akurasi sekitar 89% untuk kedua parameter yang diuji. Grafik ini mengindikasikan bahwa model *Random Forest Regression* umumnya lebih efektif dalam memprediksi data pada dataset ini dibandingkan dengan model *Linear Regression*.

3.5 Tuning Hyperparameter pada Model Regression

Dilakukan *tuning hyperparameter* untuk dua model *Regression* yaitu *Linear Regression* dan *Random Forest Regression* dengan dua set fitur yang berbeda. Proses *tuning* ini menggunakan Optuna untuk menemukan kombinasi *hyperparameter* terbaik yang meminimalkan *Mean Squared Error (MSE)*. Berikut ditampilkan pada Tabel 2 hasil dan durasi *tuning hyperparameter* untuk masing-masing model:

Tabel 2. Tuning Hyperparameter

Model	Best Parameters	Duration of hyperparameter tuning
Linear Regression (7 Fitur)	'fit_intercept': True	4.846s
Linear Regression (4 Fitur)	'fit_intercept': True	0.84s
Random Forest Regression (7 Fitur)	'n_estimators': 123, 'max_depth': 13, 'min_samples_split': 4, 'min_samples_leaf': 4	50.538s
Random Forest Regression (4 Fitur)	'n_estimators': 21, 'max_depth': 18, 'min_samples_split': 2, 'min_samples_leaf': 2	29.231s

a. Linear Regression

- **Linear Regression dengan 7 fitur:** parameter terbaik yang ditemukan adalah 'fit_intercept' disetel ke *True*, yang berarti model menyertakan intersep dalam persamaan *Regression*. Proses *tuning hyperparameter* untuk model *Linear Regression* memakan waktu 4.846 detik.
- **Linear Regression dengan 4 fitur:** parameter terbaik juga fit_intercept disetel ke *True*, dengan durasi *tuning* yang lebih singkat, yaitu 0.84 detik.

b. Random Forest Regression

- **Random Forest Regression dengan 7 fitur:** Terdapat beberapa parameter terbaik yang ditemukan. 'n_estimators: 123', yang berarti jumlah pohon dalam hutan adalah 123. 'max_depth: 13', yang berarti kedalaman maksimum setiap pohon adalah 13. 'min_samples_split: 4', yang berarti jumlah minimum sampel yang diperlukan untuk membagi node adalah 4. 'min_samples_leaf: 4', yang berarti jumlah minimum sampel yang diperlukan untuk menjadi daun adalah 4. Proses *tuning hyperparameter* untuk model *Random Forest Regression* memakan waktu 4.846 detik.
- **Random Forest Regression dengan 4 fitur:** Terdapat beberapa parameter terbaik yang ditemukan. 'n_estimators: 21', yang berarti jumlah pohon dalam hutan adalah 21. 'max_depth: 18', yang berarti kedalaman maksimum setiap pohon adalah 18. 'min_samples_split: 2', yang berarti jumlah minimum sampel yang diperlukan untuk membagi node adalah 2. 'min_samples_leaf: 2', yang berarti jumlah minimum sampel yang diperlukan untuk menjadi daun adalah 2. Proses *tuning hyperparameter* untuk model *Random Forest Regression* memakan waktu 29.231 detik.

Proses *tuning hyperparameter* untuk model *Linear Regression* dengan 7 fitur membutuhkan waktu lebih lama dibandingkan dengan model yang hanya menggunakan 4 fitur. Di sisi lain, model *Random Forest Regression* menunjukkan bahwa *tuning hyperparameter* untuk set fitur yang lebih banyak (7 fitur) juga memerlukan waktu yang lebih lama dibandingkan dengan set fitur yang lebih sedikit (4 fitur).

Secara keseluruhan, hasil *tuning hyperparameter* menunjukkan bahwa jumlah fitur dan kompleksitas model memiliki dampak signifikan terhadap durasi *tuning*. Model dengan lebih banyak fitur dan lebih banyak *hyperparameter* cenderung memerlukan waktu *tuning* yang lebih lama.

4. Kesimpulan dan Saran

Penelitian ini bertujuan untuk mengevaluasi performa dua metode regresi, yaitu *Linear Regression* dan *Random Forest Regression*, dalam menentukan batas maksimal pinjaman kredit. Dua set fitur digunakan dalam analisis ini: satu set dengan 7 fitur dan satu set dengan 4 fitur. Hasil penelitian menunjukkan bahwa *Random Forest Regression* secara konsisten mengungguli *Linear Regression* dalam hal akurasi prediksi. Model *Random Forest Regression* dengan 4 fitur berhasil mencapai nilai *Root Mean Squared Error (RMSE)* terendah sebesar 7,9% dan *Mean Absolute Error (MAE)* 4,72%, serta akurasi prediksi tertinggi yang melebihi 93,15%. Meskipun perbedaan antara penggunaan 4 fitur dan 7 fitur tidak signifikan, set fitur dengan 7 fitur sudah cukup memberikan hasil yang baik, terutama pada model *Random Forest Regression*.

Optimasi hyperparameter menggunakan Optuna terbukti efektif dalam meningkatkan performa model *Regression*. Implementasi ini tidak hanya meningkatkan akurasi prediksi, tetapi juga efisien dalam hal waktu, memungkinkan pencarian parameter optimal dilakukan lebih cepat tanpa mengurangi kualitas hasil prediksi. Berdasarkan temuan ini, penggunaan model *Random Forest Regression* disarankan untuk diterapkan dalam penentuan batas maksimal pinjaman kredit karena menunjukkan performa yang lebih baik dibandingkan *Linear Regression*. Selain itu, *optimasi hyperparameter* dengan Optuna juga sangat direkomendasikan dalam pemodelan prediktif karena kemampuannya untuk meningkatkan akurasi model secara signifikan.

Penelitian lebih lanjut dapat dilakukan dengan menambahkan data baru atau menggunakan teknik *machine learning* lainnya untuk lebih meningkatkan performa prediksi. Eksplorasi terhadap fitur tambahan yang mungkin berpengaruh terhadap penentuan batas maksimal pinjaman juga dapat dilakukan, serta studi tentang penerapan teknik optimasi *modern* lainnya seperti *Gradient Boosting* atau *Deep Learning* untuk memberikan wawasan tambahan dan potensi peningkatan performa prediktif.

Daftar Pustaka:

- Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2623–2631. <https://doi.org/10.1145/3292500.3330701>
- Ayhan, T., & Uçar, T. (2022). Determining customer limits by data mining methods in credit allocation process. *International Journal of*

- Electrical and Computer Engineering*, 12(2), 1910–1915.
<https://doi.org/10.11591/ijece.v12i2.pp1910-1915>
- Ganga, A., Vijayalakshmi, S., & Kalpana, R. A. (2021). Deciding bank client's credit limit and home loan using machine learning. *Advances in Parallel Computing*, 38, 288–293.
<https://doi.org/10.3233/APC210053>
- Hussain, Almarzooq., U., bin, W. (2024). Automating hyperparameter optimization in geophysics with Optuna: A comparative study. *Geophysical Prospecting*, 72(5), 1778–1788.
<https://doi.org/10.1111/1365-2478.13484>
- I.O. Artemieva, Igor S. Bondaruk, I. O. M. (2022). Credit as a necessary economic tool of economic relationships between financial companies and banks. *Ekonomični Gorizonti*, 4(22), 67–78.
[https://doi.org/https://www.doi.org/10.31499/2616-5236.4\(22\).2022.267019](https://doi.org/https://www.doi.org/10.31499/2616-5236.4(22).2022.267019)
- Karl, L. (2023). Simple Linear Regression. *International Series in Management Science/Operations Research*, 57–90.
https://doi.org/10.1007/978-3-031-21480-6_4
- M.S., Nidhya., Nidhi, Saraswat., Prashant, K. (2024). Leveraging Adaptive Hyper Parameter Tuning for Automated Machine Learning. *International Conference on Optimization Computing and Wireless Communication (ICOCWC)*, 1–7.
<https://doi.org/10.1109/ICOCWC60930.2024.10470671>
- Maryam, Hassanali., Mohammadreza, Soltanaghaei., Taghi, Javdani, Gandomani., Farsad, Zamani, B. (2024). Software development effort estimation using boosting algorithms and automatic tuning of hyperparameters with Optuna. *Journal of Software: Evolution and Process*.
<https://doi.org/https://doi.org/10.1002/smr.2665>
- Narayana Darapaneni, Akshay Kumar, Archanna Dixit, M. S. (2022). Loan Prediction Software for Financial Institutions. *Interdisciplinary Research in Technology and Management (IRTM)*, 1–8.
<https://doi.org/https://doi.org/10.1109/IRTM54583.2022.9791797>
- Oksana Savko, Juliia Bui, Y. S. (2022). Lending activity development of financial companies in the service sector of individual and legal entities. *Infrastruktura Rinku*, 65.
<https://doi.org/https://www.doi.org/10.32843/infrastruct65-6>
- Olufemi Ogunleye, J. (2023). *Predictive Data Analysis Using Linear Regression and Random Forest*. IntechOpen.
<https://doi.org/10.5772/intechopen.107818>
- P. Meghana et al. (2024). Analysis of Neural Network Algorithm in Comparison to Multiple Linear Regression and Random Forest Algorithm. *2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETISIS)*, 437–443.
<https://doi.org/10.1109/ICETISIS61505.2024.10459496>
- Piracha, S., Hiba, H., Farooq, U., Fatima, U., Aslam, F., Ishaq, A., & Sial, H. (2022). New Insights and Different Economic Factors in Banking Sectors and Current Challenges. *Scholars Bulletin*, 8(2), 75–78.
<https://doi.org/10.36348/sb.2022.v08i02.005>
- V., Ignatenko., Anton, Surkov., Sergei, K. (2024). Random forests with parametric entropy-based information gains for classification and regression problems. *PeerJ*, 10.
<https://doi.org/10.7717/peerj-cs.1775>
- Zhang, M., & Wang, M. (2018). Review of the Management Theory and Method of Loan Limitation. *Atlantis Press, 130*(Ifmeita 2017), 1–5. <https://doi.org/10.2991/ifmeita-17.2018.1>

Halaman ini sengaja dikosongkan