

PENERAPAN *HYPERPARAMETER MACHINE LEARNING* DALAM PREDIKSI GAGAL PINJAM

Dinar Ismunandar^{1*}; Muhammad Rifqi Firdaus²; Yuris Alkhalifi³

Teknologi Informasi¹, Sistem Informasi², Teknologi Komputer³
Universitas Bina Sarana Informatika, Jakarta, Indonesia^{1,2,3}
bsi.ac.id^{1,2,3}

dinar.dim@bsi.ac.id^{1*}; muhammad.mku@bsi.ac.id²; yuris.yak@bsi.ac.id³
(*) Corresponding Author



Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi-NonKomersial 4.0 Internasional.

Abstract—Loans or credit are one of the key factors in advancing the economy. One of them is encouraging business expansion which will have a direct impact on a country's economic growth. Banks and other financing institutions must be able to evaluate the borrower's ability to pay their debts based on the inherent risks to reduce the possibility of default. To this end, machine learning (ML) has emerged as a revolutionary tool in using advanced prediction methods to examine historical data based on customer behavior. This research investigates the application of ML in predicting loan outcomes by optimizing parameters in the Machine Learning algorithm. The ML algorithms examined in this research are Logistic Regression (LR), K-Nearest Neighbor (KNN), Random Forest (RF), Decision Tree (DT), and XGBoost (XGB). Meanwhile, the technique used in hyperparameter tuning is Grid Search Cross Validation (CV). The results show that the algorithm's performance is more optimal than before, it can be seen that the LR algorithm experienced an increase in accuracy of 5%, KNN by 4%, RF by 3%, DT by 3%, and XGB by 2%. By including a default dataset based on customer behavior and optimized algorithm parameters, apart from being able to answer the alignment in previous literature in providing a deeper understanding of loan estimation, this research can also provide an understanding that hyperparameter techniques are worth trying to improve the performance of ML algorithms. So, it will be easier for financing institutions to determine the right loan scenario.

Keywords: *gridsearchCV, hyperparameters, loan default prediction, machine learning.*

Abstrak—Pinjaman atau kredit merupakan salah satu faktor kunci dalam memajukan perekonomian. Salah satunya adalah mendorong ekspansi bisnis yang akan berdampak langsung pada pertumbuhan ekonomi suatu negara. Bank dan lembaga pembiayaan lainnya harus dapat mengevaluasi kemampuan peminjam dalam membayar utangnya berdasarkan risiko yang melekat untuk mengurangi kemungkinan gagal bayar. Untuk itu, *Machine Learning* (ML) hadir sebagai alat revolusioner dalam menggunakan metode prediksi yang canggih untuk memeriksa data historis berdasarkan perilaku nasabah. Penelitian ini menginvestigasi penerapan ML dalam memprediksi hasil pinjaman dengan mengoptimalkan parameter dalam algoritma *Machine Learning*. Algoritma ML yang diteliti dalam penelitian ini adalah *Logistic Regression* (LR), *K-Nearest Neighbor* (KNN), *Random Forest* (RF), *Decision Tree* (DT), dan *XGBoost* (XGB). Sementara itu, teknik yang digunakan dalam *tuning hyperparameter* adalah *Grid Search Cross Validation* (CV). Hasil penelitian menunjukkan bahwa kinerja algoritma lebih optimal dibandingkan sebelumnya, terlihat bahwa algoritma LR mengalami peningkatan akurasi sebesar 5%, KNN sebesar 4%, RF sebesar 3%, DT sebesar 3%, dan XGB sebesar 2%. Dengan menyertakan dataset *default* berdasarkan perilaku nasabah dan parameter algoritma yang telah dioptimasi, selain dapat menjawab keselarasan pada literatur terdahulu dalam memberikan pemahaman yang lebih mendalam mengenai estimasi kredit, penelitian ini juga dapat memberikan pemahaman bahwa teknik hiperparameter layak dicoba untuk meningkatkan performa algoritma ML. Sehingga, lembaga pembiayaan akan lebih mudah menentukan skenario pinjaman yang tepat.

Kata kunci: *gridsearchCV, hyperparameter, prediksi gagal pinjaman, machine learning.*

PENDAHULUAN

Pinjaman atau yang sering disebut dengan kredit sampai saat ini masih menjadi salah satu faktor dalam menggerakkan perekonomian suatu negara (Ulfa & Mulyadi, 2020). Hal ini didasarkan pada mayoritas pengusaha di suatu negara yang telah menggunakan layanan pinjaman untuk memulai atau mempertahankan usahanya dalam menghadapi tantangan global (Zelvia, 2024). Tercatat dalam laporan Otoritas Jasa Keuangan (OJK) di Indonesia Tahun 2023 mengungkapkan bahwa, pelaku usaha Mikro, Kecil dan Menengah (UMKM) dengan penyaluran kepada UMKM perseorangan mencapai Rp15,63 Triliun (Santosa, 2023). Dapat disimpulkan, bahwa di Indonesia saja telah banyak yang menggunakan jasa pinjaman untuk alasan memulai usaha.

Di balik itu, terdapat laporan lain yang berasal dari OJK mengenai kredit macet yang mencapai di Tahun 2023 secara nasional mencapai Rp1,73 Triliun. Perhitungan tersebut berdasarkan tingkat wanprestasi (TWP) lebih dari 90 hari yang artinya, suatu kredit dikatakan macet jika peminjam gagal membayar utang lebih dari 90 hari sejak tanggal jatuh tempo (Muhamad, 2023). Dari hal tersebut, lembaga pembiayaan atau bahkan perbankan harus melakukan evaluasi yang mendalam terkait risiko dari penilaian portofolio peminjam untuk meminimalisir risiko gagal bayar (Harisanti et al., 2024).

Dalam memecahkan permasalahan tersebut, pembelajaran *Machine Learning* (ML) hadir untuk memberi inovasi dalam membantu pengambilan keputusan dalam berbagai aspek, salah satunya aspek keuangan pada perbankan atau instansi pembiayaan lainnya (Ramadhon et al., 2024). Dalam aspek keuangan ini, ML digunakan untuk mengambil keputusan yang tepat dalam menyetujui atau menolak pinjaman yang dilihat dari perilaku nasabah yang mencerminkan layak atau tidaknya. Akibatnya, prediksi yang dihasilkan dari ML tersebut akan membantu perusahaan dalam menghindari kerugian keuangan dari peminjam yang gagal bayar.

Sebelumnya telah dilakukan penelitian dengan menggunakan algoritma ML yakni, *Logistic Regression*, *KNN*, *Random Forest*, *Decision Tree*, dan *XGBoost* pada sumber data yang sama dengan penelitian ini yaitu penelitian yang dilakukan oleh (Zuama et al., 2024). Namun, hasil dari penelitian tersebut masih belum optimal, bahkan terdapat salah satu model yang mengalami bias terhadap salah satu kelasnya. Hal tersebut, disebabkan karena parameter model yang tidak diatur dengan baik. Akhirnya, penelitian ini dirancang untuk melakukan pengembangan dari penelitian

sebelumnya yaitu dengan penggunaan teknik *Hyperparameter Tuning* di setiap algoritma ML yang digunakan sebelumnya, sehingga menghasilkan parameter yang lebih baik di setiap modelnya dan menghasilkan prediksi yang lebih akurat.

Penelitian ini terinspirasi dari beberapa penelitian sebelumnya yang berhasil mengoptimalkan hasil kinerja algoritma ML dengan menggunakan teknik *Hyperparameter Tuning*. Pertama, pada penelitian mengenai peningkatan kinerja akurasi prediksi diabetes mellitus menggunakan teknik *hyperparameter tuning* dengan metode *GridSearchCV* pada algoritma *Logistic Regression* (Gunawan et al., 2020). Hasilnya, penelitian tersebut memperoleh hasil akurasi yang lebih tinggi dibandingkan model sebelumnya yang tidak menggunakan teknik *hyperparameter tuning* dengan peningkatan akurasi sebesar 11,11%.

Penelitian kedua mengenai klasifikasi nasabah kartu kredit dengan menggunakan teknik *hyperparameter* pada algoritma *XGBoost* (Yulianti et al., 2022). Hasilnya, *hyperparameter tuning* memberikan pengaruh yang signifikan terhadap evaluasi hasil kinerja klasifikasi, dengan peningkatan akurasi, presisi, dan juga *recall* dengan kenaikan 3,40% pada akurasi, 0,04% pada presisi dan 0,42 pada *recall*.

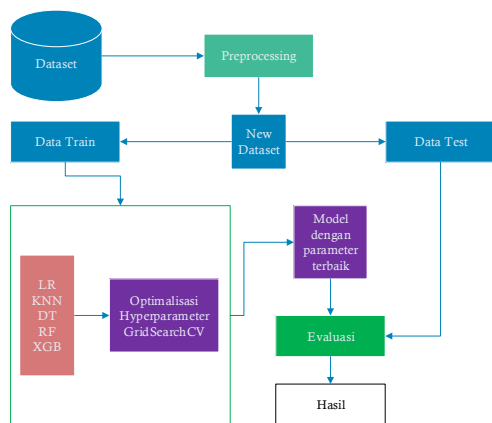
Terakhir pada penelitian yang hasilnya menyimpulkan bahwa optimalisasi dengan menggunakan teknik *hyperparameter tuning* sangat efisien dalam meningkatkan keakuratan prediksi pada setiap algoritma ML. Penelitian tersebut berjudul *GridSearch* dalam optimasi *hyperparameter* model *machine learning* untuk prediksi HIV/Hasil tes AIDS (Belete & Huchaiyah, 2022). Penelitian tersebut membandingkan delapan model ML yang diterapkan untuk memprediksi hasil tes HIV antara penggunaan teknik optimalisasi *hyperparameter tuning* dengan metode *GridSearchCV*, dengan tanpa penggunaan teknik optimalisasi *hyperparameter* melalui metrik standar seperti akurasi, presisi, *recall*, *f1-score*, AUCROC, *Confusion Matrix*, MAE, RMSE dan R2. Hasil evaluasi dari perbandingan model tersebut menunjukkan bahwa teknik optimalisasi *hyperparameter tuning* dengan metode *GridSearchCV* dapat meningkatkan keakuratan dalam prediksi.

Dari beberapa penelitian tersebut, dapat disimpulkan bahwa teknik optimalisasi *hyperparameter* memiliki potensi yang baik dalam menghasilkan akurasi prediksi yang lebih optimal pada algoritma ML. Untuk itu, dalam penelitian ini kami akan mengeksplorasi penerapan teknik *hyperparameter tuning* terhadap beberapa algoritma ML dalam memprediksi gagal pinjam yang berbasis pada perilaku nasabah yang

sebelumnya pernah diteliti namun tanpa penerapan teknik *hyperparameter tuning*. Penelitian ini memiliki tujuan untuk memberikan kontribusi signifikan dalam meningkatkan efisiensi dan akurasi prediksi dalam menentukan kelayakan proses pinjaman yang berbasis pada perilaku nasabah, dengan pengoptimalan parameter pada setiap algoritma *machine learning* yang diuji menggunakan teknik *hyperparameter tuning*. Sehingga, dapat menghasilkan suatu model yang lebih baik dengan hasil prediksi yang lebih akurat.

BAHAN DAN METODE

Aliran metodologi yang diusulkan penelitian ini untuk prediksi gagal pinjam yang lebih optimal seperti yang ditunjukkan Gambar 1. Sebelum masuk ke dalam pemodelan, data terlebih dahulu akan melewati tahap *preprocessing* guna menghasilkan data yang optimal. Setelah itu, data yang sudah optimal dibagi menjadi dua menjadi data latih dan data uji. Data latih digunakan untuk pemodelan dari setiap algoritma ML yang digunakan, serta dalam melakukan teknik *hyperparameter tuning* dengan menggunakan metode *GridSearchCV* untuk meningkatkan efisiensi model. Sedangkan data uji, digunakan untuk menguji model yang sudah dilatih sebelumnya dengan menggunakan data latih.



Sumber: (Hasil Penelitian, 2024)
Gambar 1. Diagram Alir Metode yang Diusulkan

A. Dataset

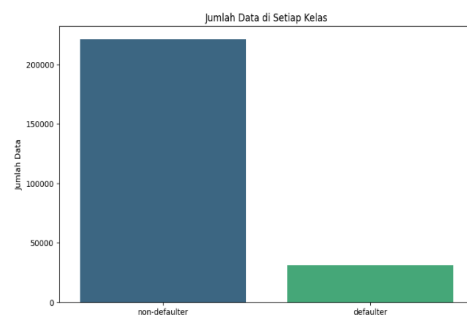
Dataset yang digunakan dalam penelitian ini merupakan data yang bersifat publik dan didapat dari *website* Kaggle (Surana, 2021). Dataset tersebut memiliki 12 variabel yang berjumlah 252.000 data yang terdiri dari dua kelas dengan jumlah masing-masing kelas diantaranya, kelas peminjam yang tidak gagal bayar (*non-defaulter*) berjumlah 221.004 data, dan kelas peminjam yang gagal bayar (*defaulter*) berjumlah 30.996 data. Data-data tersebut berasal dari data historis

peminjaman terkait perilaku nasabah. Untuk rincian lebih lengkap mengenai subset dari data tersebut dapat dilihat dari Tabel 1. dan Gambar 2.

Tabel 1. Deskripsi Data Mentah

No	Kolom	Deskripsi	Tipe
1	<i>income</i>	Pendapatan pengguna	<i>Int</i>
2	<i>age</i>	Usia pengguna	<i>Int</i>
3	<i>experience</i>	Pengalaman profesional pengguna selama bertahun-tahun	<i>Int</i>
4	<i>profession</i>	Profesi	<i>string</i>
5	<i>married</i>	Apakah sudah menikah atau lajang	<i>string</i>
6	<i>house_ownership</i>	Dimiliki atau disewa atau tidak keduanya	<i>string</i>
7	<i>car_ownership</i>	Apakah orang tersebut memiliki mobil	<i>string</i>
8	<i>current_job_years</i>	Pengalaman bertahun-tahun dalam pekerjaan saat ini	<i>Int</i>
9	<i>current_house_years</i>	Berapa tahun lamanya di tempat tinggal saat ini	<i>Int</i>
10	<i>city</i>	Kota tempat tinggal	<i>string</i>
11	<i>state</i>	Negara tempat tinggal	<i>string</i>
12	<i>risk_flag</i>	Gagal membayar pinjaman	<i>int</i>

Sumber: (Zuama, 2024)



Sumber: (Hasil Penelitian, 2024)
Gambar 2. Jumlah Data Mentah di Setiap Kelas

B. Preprocessing

Dalam tahap *preprocessing* atau pemrosesan awal ini dataset diproses terlebih dahulu dengan berbagai proses meliputi, pemilihan fitur yang relevan, memperbaiki data yang bermasalah, serta penyeimbangan kelas agar kualitas data semakin baik. Sehingga, kualitas data yang baik akan berdampak pada hasil akurasi prediksi yang baik. Beda halnya jika data mentah yang masih memiliki kualitas yang rendah dapat mengakibatkan penurunan akurasi prediksi (Azhari et al., 2024).

Dalam penelitian ini, pertama-tama dilakukan pemilihan fitur dengan cara menghapus fitur yang tidak relevan. Pada tahap penghapusan fitur ini, fitur yang tidak dibutuhkan untuk prediksi gagal pinjam di Indonesia di antaranya, fitur '*City*' atau kota tempat tinggal yang hanya berisikan data

dari kota-kota luar Indonesia. Kemudian, fitur 'State' atau negara tempat tinggal yang hanya berisikan data dari negara luar Indonesia. Dan yang terakhir adalah fitur 'profession' atau profesi, di mana fitur sudah tidak dibutuhkan karena sudah terwakili dengan fitur 'income' atau pendapatan dari setiap peminjam.

Setelah menghapus fitur yang tidak relevan, selanjutnya merubah jenis data pada fitur yang memiliki jenis data bermasalah, namun masih relevan dengan menggunakan teknik data transformasi dengan menggunakan metode Label Encoder. Pasalnya, untuk meningkatkan akurasi dan efisiensi algoritma yang digunakan dalam sebuah penelitian dapat dilakukan dengan proses encoding seperti metode Label Encoder yang melakukan proses pengkodean untuk mengubah kategori menjadi numerik dan memisahkannya (Jayidan et al., 2024).

Dan yang terakhir, yakni menyeimbangkan kelas pinjaman gagal bayar dan tidak gagal bayar dengan menggunakan teknik oversampling dengan memanfaatkan metode SMOTE (Synthetic Minority Over-sampling Technique). Metode SMOTE merupakan pengembangan dari teknik oversampling. Cara kerja metode tersebut adalah dengan membuat sampel baru yang berasal dari kelas yang memiliki data sedikit (minoritas) dengan tujuan membuat proporsi data menjadi lebih seimbang dengan menggunakan cara sampling ulang dari sampel kelas minoritas (Ridwan et al., 2024).

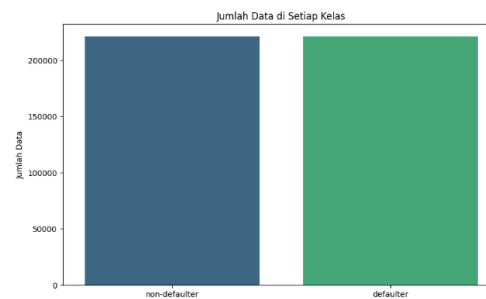
Beberapa pemrosesan awal data sebelumnya, menghasilkan data baru dengan 9 variabel yang berjumlah 442.008 data dengan jumlah masing-masing kelas 221.004 merupakan data peminjam yang tidak gagal bayar (non-defaulter) dan 221.004 merupakan data peminjam yang gagal bayar (defaulter). Untuk rincian lebih lengkap mengenai subset dari data tersebut dapat dilihat dari Tabel 2. dan Gambar 3.

Tabel 2. Deskripsi Data Baru

No	Kolom	Deskripsi	Tipe
1	income	Pendapatan pengguna	Int
2	age	Usia pengguna	Int
3	experience	Pengalaman profesional pengguna selama bertahun-tahun	Int
4	married	Apakah sudah menikah atau lajang	int
5	house_ownership	Dimiliki atau disewa atau tidak keduanya	int
6	car_ownership	Apakah orang tersebut memiliki mobil	int
7	current_job_years	Pengalaman bertahun-tahun dalam pekerjaan saat ini	Int
8	current_housing_use_years	Berapa tahun lamanya di tempat tinggal saat ini	Int

No	Kolom	Deskripsi	Tipe
9	risk_flag	Gagal membayar pinjaman	int

Sumber: (Hasil Penelitian, 2024)



Sumber: (Hasil Penelitian, 2024)

Gambar 3. Jumlah Data Baru di Setiap Kelas

Setelah menghasilkan data baru yang memiliki kualitas lebih baik, dilakukan pembagian data dengan mengimplementasi split data dalam persiapan pembelajaran suatu model. Karena penelitian ini memiliki data yang cukup banyak, dilakukan pembagian sebesar 70% untuk data pelatihan (training data) dan 30% untuk data pengujian (testing data). Data pelatihan dan data pengujian ini menjadi hal yang paling penting dalam membangun model ML yang akurat (Oktavia & Isnain, 2024). Di mana, data pelatihan ini akan digunakan untuk melatih model klasifikasi gagal pinjam, sementara data pengujian akan digunakan untuk menguji performa model serta mengukur seberapa baik model tersebut dalam menggeneralisasi pada data baru.

Sebelum melanjutkan pemodelan, dilakukan normalisasi data terlebih dahulu untuk lebih mengoptimalkan kinerja dari setiap modelnya. Normalisasi dalam penelitian ini menggunakan metode Standard Scaler yang bekerja sebagai standarisasi fitur dengan menghapus rata-rata (mean) dan menskalakan unit varian (standar deviasi) (Sumantri et al., 2023). Metode tersebut dilakukan untuk mencegah adanya data yang memiliki nilai terlalu besar dibanding dengan nilai yang lain. Rumus dari Standard Scaler ditunjukkan pada persamaan (1), di mana x' adalah nilai fitur yang sudah dinormalisasi, x adalah nilai fitur asli, μ adalah mean (rata-rata) dari fitur, dan σ adalah standar deviasi dari fitur.

$$x' = \frac{x - \mu}{\sigma} \tag{1}$$

Dari persamaan (1), Standard Scaler dapat digunakan untuk mengubah skala fitur numerik dalam dataset sehingga memiliki mean 0 dan deviasi standar, serta membantu mengurangi efek outlier dalam dataset (Saputra et al., 2023).

C. Machine Learning

Karena penelitian ini bertujuan untuk memberi kontribusi lebih dengan meningkatkan efisiensi dan akurasi penentuan kelayakan proses pinjaman pada penelitian sebelumnya, algoritma ML yang digunakan dalam penelitian ini menggunakan algoritma yang digunakan pada penelitian sebelumnya, yakni *Logistic Regression* (LR), *K-Nearest Neighbor* (KNN), *Random Forest* (RF), *Decision Tree* (DT), dan *XGBoost* (XGB).

D. Grid Search

Untuk mencapai tujuan dalam mengoptimalkan sebuah algoritma ML, penelitian ini menggunakan teknik *hyperparameter tuning*. Sementara itu, konfigurasi *hyperparameter* yang digunakan pada penelitian ini yakni menggunakan metode *GridSearchCV*. Metode yang memiliki kepanjangan *Grid Search Cross Validation* (CV) ini telah banyak digunakan oleh penelitian-penelitian sebelumnya untuk melatih dan mengoptimalkan model agar bisa memberikan hasil terbaik (Andryan et al., 2022).

Pada *sklearn* pun (Darmawan & Dianta, 2023), *GridSearchCV* dapat digunakan untuk menggunakan mencari grid suatu algoritma dalam mendeteksi *hyperparameter* yang optimal. Meskipun sebelum-sebelumnya banyak juga penelitian ilmiah melakukan pengoptimalan model dengan menggunakan metode lain, namun metode pengoptimalan *hyperparameter* terutama dengan *GridSearchCV* sangat disarankan karena dikategorikan sebagai metode yang teliti dalam menentukan parameter terbaik dengan dilakukan eksplorasi masing masing parameter (Alifah et al., 2024).

GridSearchCV beroperasi dengan menilai Produk *Cartesian* dari serangkaian nilai yang terbatas dan telah ditentukan oleh peneliti sebelum dioperasikan (Rusman et al., 2023). Sementara itu, *GridSearchCV* sendiri tidak akan mengeksplorasi lebih jauh bagian-bagian yang memiliki kinerja baik. Oleh karena itu, terdapat beberapa yang perlu dilakukan secara manual untuk mengidentifikasi secara optimal sebagai berikut (Zöller & Huber, 2021):

1. Memulai dengan ruang yang lebih luas untuk pencarian dan skala fase.
2. Berdasarkan temuan sebelumnya, lingkungan *hyperparameter* yang berkinerja baik akan mempersempit ruang pencarian dan ukuran fase.
3. Ulangi langkah sebelumnya beberapa kali hingga mencapai hasil yang optimal.

E. Evaluasi Performa

Performa model dievaluasi menggunakan metrik-metrik seperti akurasi, presisi, *recall* dan *F1-score*. Tujuan dari evaluasi sendiri yakni untuk membantu menilai sejauh mana kinerja suatu model yang diteliti mampu menggeneralisasi pola dari data pelatihan ke data pengujian (Hasibuan, 2024). Adapun persamaan (2) hingga (5) yang menyajikan metrik standar kinerja setiap model untuk prediksi gagal pinjam pada penelitian ini.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (5)$$

HASIL DAN PEMBAHASAN

Langkah awal yang dilakukan dalam penelitian ini yakni menyiapkan dataset yang merupakan data publik dari *website* Kaggle (Surana, 2021). Kemudian, menyiapkan algoritma ML yang akan digunakan sendiri yakni algoritma LR, KNN, RF, DT, dan XGB yang sebelumnya telah dilakukan penelitian pada penelitian sebelumnya, (Zuama et al., 2024). Hal tersebut dilakukan untuk melihat sejauh mana kontribusi yang dihasilkan dari penelitian ini terhadap penelitian sebelumnya.

Tahap selanjutnya yakni mempersiapkan teknik dan matrik evaluasi. Untuk mengevaluasi performa dari setiap model, diperlukan beberapa data masukan untuk melatih kinerja suatu model tersebut. Sehingga penelitian ini, membagi dataset yang ada menjadi dua bagian dengan besaran 70% untuk training data dan 30% untuk testing data. Dalam proses pembagian data tersebut, penelitian ini menggunakan *library sklearn* dengan metode *split*. Kemudian, sebelum melanjutkan pelatihan, dilakukan pencarian parameter terbaik dengan menggunakan teknik *GridSearchCV* dengan besaran *k-fold* atau $K = 10$. Hal tersebut merupakan besaran pengulangan yang akan dilakukan teknik *hyperparameter* tersebut dalam mencari parameter terbaik dari setiap algoritma ML. Namun, sebelumnya jangkauan parameter yang akan diuji harus ditentukan terlebih dahulu. Jangkauan parameter yang digunakan dalam penelitian ini diambil dengan random seperti yang terlihat di Tabel 3.

Tabel 3. Konfigurasi Jangkauan *Hyperparameter* pada Setiap Algoritma *Machine Learning*

Model	Jangkauan <i>Hyperparameter</i> yang dicari	Parameter terbaik
LR	C = np.logspace (-4,4,20) max_iter = [1000, 10000] penalty = ['l1', 'l2']	C = 0.00026366508987303583 max_iter = 1000 penalty = 'l2'
KNN	n_neighbors = list(range(1,30)) p=[1,2] weights = ['uniform', 'distance']	n_neighbors = 7 p: 1 weights: 'distance'
RF	min_samples_leaf = [1, 2, 3] min_samples_split = np.arange(3, 8) max_features = ['sqrt', 'auto', 'log2']	min_samples_leaf = 1 min_samples_split = 7 max_features = 'log2'
DT	splitter = ['best', 'random'] min_samples_leaf = [3, 5, 8, 10, 15, 50] min_samples_split = [2, 5, 8, 10, 15, 50]	Splitter = 'best' min_samples_leaf = 3 min_samples_split = 15
XGB	learning_rate = [0.1,0.3,0.5] n_estimators = [700, 900, 1000] max_depth = [4, 6, 9]	learning_rate = 0.3 max_depth = 9 n_estimators = 1000

Sumber: (Hasil Penelitian, 2024)

Keterangan: LR (*Logistic Regression*), KNN (*K-Nearest Neighbor*), RF (*Random Forest*), DT (*Decision Tree*), dan XGB (*XGBoost*)

B. Hasil Ekperimen

Setelah persiapan eksperimen selesai, di tahap ini dilakukan evaluasi menggunakan metrik-metrik seperti akurasi, presisi, *recall*, dan *F1-score*. Untuk memastikan perbandingan yang akurat, penelitian ini melakukan perbandingan dengan eksperimen sebelumnya, yakni dengan pengujian model tanpa konfigurasi *hyperparameter* (Zuama et al., 2024) dan pengujian model dengan menggunakan konfigurasi *hyperparameter* pada penelitian ini. Hasil dari pengujian tersebut dapat dilihat pada Tabel 4 dan Tabel 5.

Tabel 4. Hasil Pengujian Model *Machine Learning* tanpa *Hyperparameter*

Model	Class	Precision	Recall	F1-score	Accuracy
LR	Non-Defaulter	0.88	1.00	0.93	0.50
	Defaulter	0.00	0.00	0.00	

KNN	Non-Defaulter	0.98	0.86	0.92	0.86
	Defaulter	0.46	0.86	0.60	
RF	Non-Defaulter	0.97	0.89	0.93	0.88
	Defaulter	0.51	0.81	0.63	
DT	Non-Defaulter	0.98	0.87	0.92	0.87
	Defaulter	0.47	0.85	0.61	
XGB	Non-Defaulter	0.97	0.91	0.94	0.89
	Defaulter	0.54	0.78	0.64	

Sumber: (Zuama, 2024)

Keterangan: LR (*Logistic Regression*), KNN (*K-Nearest Neighbor*), RF (*Random Forest*), DT (*Decision Tree*), dan XGB (*XGBoost*)

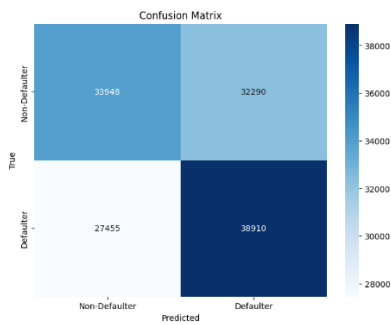
Tabel 5. Hasil Pengujian Model *Machine Learning* dengan *Hyperparameter*

Model	Class	Precision	Recall	F1-score	Accuracy
LR	Non-Defaulter	0.55	0.51	0.53	0.55
	Defaulter	0.55	0.59	0.57	
KNN	Non-Defaulter	0.93	0.86	0.90	0.90
	Defaulter	0.87	0.94	0.90	
RF	Non-Defaulter	0.96	0.86	0.91	0.91
	Defaulter	0.88	0.96	0.92	
DT	Non-Defaulter	0.93	0.86	0.89	0.90
	Defaulter	0.87	0.94	0.90	
XGB	Non-Defaulter	0.95	0.87	0.91	0.91
	Defaulter	0.88	0.95	0.92	

Sumber: (Hasil Penelitian, 2024)

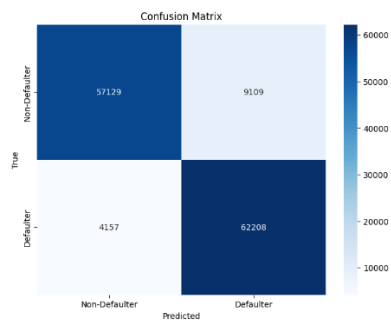
Keterangan: LR (*Logistic Regression*), KNN (*K-Nearest Neighbor*), RF (*Random Forest*), DT (*Decision Tree*), dan XGB (*XGBoost*)

Perbandingan hasil dari Tabel 4 dan Tabel 5 terlihat bahwa teknik *hyperparameter* sangat efektif untuk menghasilkan tingkat performa suatu model lebih baik. Terlihat dari setiap modelnya yang mengalami kenaikan performa menunjukkan bahwa kinerja algoritma pada penelitian ini lebih optimal dibandingkan sebelumnya. Kenaikan-kenaikan performa dari algoritma-algoritma tersebut yakni, pada algoritma LR mengalami peningkatan akurasi sebesar 5%, KNN sebesar 4%, RF sebesar 3%, DT sebesar 3%, dan XGB sebesar 2%. Sehingga, rata-rata kenaikan performa dari setiap algoritma yang diteliti tersebut sebesar 3,4%. Sementara, untuk dapat memvisualisasikan hasil kinerja prediksi dari kelima model tersebut penelitian ini menggunakan *confusion matrix*. Setiap nilai entri *matrix* tersebut menunjukkan jumlah prediksi yang dibuat dari masing-masing model ketika mengklasifikasikan kelas dengan benar atau salah. *Confusion matrix* pada penelitian ini dapat dilihat dari Gambar 4 hingga Gambar 8.



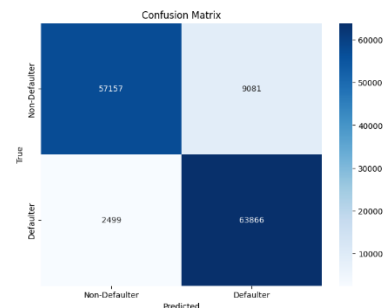
Sumber: (Hasil Penelitian, 2024)
Gambar 4. Confusion Matrix LR

Dari Gambar 4, dapat dilihat bahwa nilai *True Negative* (TN), *False Positive* (FP), *False Negatives* (FN), dan *True Positives* (TP) dari model LR memiliki jumlah 33.948, 32.290, 27.455 dan 38.910. Dari nilai hasil perhitungan prediksi tersebut jika dihitung dengan rumus (2) menghasilkan akurasi sebesar 0,55 atau 55%.



Sumber: (Hasil Penelitian, 2024)
Gambar 5. Confusion Matrix KNN

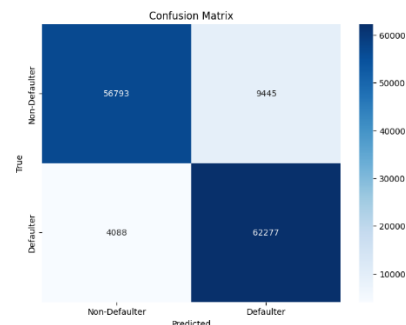
Dari Gambar 5, dapat dilihat bahwa nilai TN, FP, FN, dan TP dari model KNN memiliki jumlah 57.129, 9.109, 4.157 dan 62.208. Dari nilai hasil perhitungan prediksi tersebut jika dihitung dengan rumus (2) menghasilkan akurasi sebesar 0,90 atau 90%.



Sumber: (Hasil Penelitian, 2024)
Gambar 6. Confusion Matrix RF

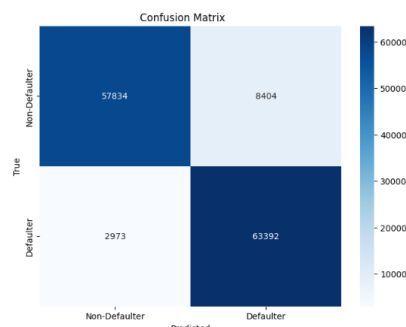
Dari Gambar 6, dapat dilihat bahwa nilai TN, FP, FN, dan TP dari model RF memiliki jumlah

57.157, 9.081, 2.499 dan 63.866. Dari nilai hasil perhitungan prediksi tersebut jika dihitung dengan rumus (2) menghasilkan akurasi sebesar 0,91 atau 91%.



Sumber: (Hasil Penelitian, 2024)
Gambar 7. Confusion Matrix DT

Dari Gambar 7, dapat dilihat bahwa nilai TN, FP, FN, dan TP dari model DT memiliki jumlah 56.793, 9.445, 4.088 dan 62.277. Dari nilai hasil perhitungan prediksi tersebut jika dihitung dengan rumus (2) menghasilkan akurasi sebesar 0,90 atau 90%.



Sumber: (Hasil Penelitian, 2024)
Gambar 8. Confusion Matrix XGB

Dari Gambar 8, bisa dilihat bahwa nilai TN, FP, FN, dan TP dari model XGB memiliki jumlah 57.834, 8.404, 2.973 dan 63.392. Dari nilai hasil perhitungan prediksi tersebut jika dihitung dengan rumus (2) menghasilkan akurasi sebesar 0,91 atau 91%.

KESIMPULAN

Kesimpulan yang dapat diambil dari penelitian ini yaitu penggunaan teknik *hyperparameter* pada setiap algoritma ML dapat meningkatkan performa algoritma tersebut. Karena, penelitian prediksi gagal pinjam berdasarkan perilaku nasabah ini dapat membuktikan bahwa performa algoritma ML dengan *hyperparameter* lebih baik dari pada penelitian sebelumnya yang tidak menggunakan

hyperparameter. Pembuktian tersebut terlihat dari kenaikan performa algoritma-algoritma tersebut seperti, algoritma LR yang mengalami peningkatan akurasi sebesar 5%, KNN sebesar 4%, RF sebesar 3%, DT sebesar 3% dan XGB sebesar 2%, dengan rata-rata kenaikan performa sebesar 3,4%. Namun, hal ini dapat ditingkatkan lagi dengan cara pengoptimalan *hyperparameter* yang berbeda seperti, menggunakan teknik *tuning hyperparameter* yang berbeda, serta menentukan jangkauan konfigurasi dalam pencarian *hyperparameter* yang lebih bervariasi.

REFERENSI

- Alifah, R. N., Najib, M. K., Nurdiati, S., Sari, A. P., Herlambang, K., Ginting, D. T. P. B., & Sya'adah, S. N. (2024). Perbandingan Metode Tree Based Classification untuk Masalah Klasifikasi Data Body Mass Index. *Indonesian Journal of Mathematics and Natural Sciences*, 47(1), 49–65.
- Andryan, M. R., Fajri, M., & Sulistyowati, N. (2022). Komparasi Kinerja Algoritma Xgboost Dan Algoritma Support Vector Machine (Svm) Untuk Diagnosis Penyakit Kanker Payudara. *JIKO (Jurnal Informatika Dan Komputer)*, 6(1), 1–5.
- Azhari, M., Maulana, H., & Riza, F. (2024). Data Mining Dalam Analisis Faktor Drop Out Mahasiswa Menerapkan Algoritma Decision Tree. *Jurnal Media Informatika Budidarma*, 8(2), 1209–1217.
- Belete, D. M., & Huchaiah, M. D. (2022). Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results. *International Journal of Computers and Applications*, 44(9), 875–886.
- Darmawan, Z. M. E., & Dianta, A. F. (2023). Implementasi optimasi hyperparameter GridSearchCV pada sistem prediksi serangan jantung menggunakan SVM. *Teknologi: Jurnal Ilmiah Sistem Informasi*, 13(1), 8–15.
- Gunawan, M. I., Sugiarto, D., & Mardianto, I. (2020). Peningkatan Kinerja Akurasi prediksi penyakit diabetes mellitus menggunakan metode grid Search Pada algoritma logistic regression. *JEPIN (Jurnal Edukasi Dan Penelitian Informatika)*, 6(3), 280–284.
- Harisanti, N. N., Meliala, M. E. B., & Oktafia, R. (2024). Analisis Pembiayaan Perbankan (Studi Kasus) Pada Bank Syariah PT. Bank Muamalat Indonesia. *Jurnal Rumpun Manajemen Dan Ekonomi*, 1(1), 52–63.
- Hasibuan, F. H. (2024). Klasifikasi Data Material Pending Pada Perusahaan dengan Metode SVM. *Innovative: Journal Of Social Science Research*, 4(1), 5080–5090.
- Jayidan, Z., Siregar, A. M., Faisal, S., & Hikmayanti, H. (2024). Improving Heart Disease Prediction Accuracy Using Principal Component Analysis (PCA) In Machine Learning Algorithms. *Jurnal Teknik Informatika (Jutif)*, 5(3), 821–830.
- Muhamad, N. (2023, August 23). *Gen Z dan Milenial Jadi Penyumbang Kredit Macet Pinjol Terbesar pada Juni 2023*. Katadata.Co.Id.
- Oktavia, I., & Isnain, A. R. (2024). Analisis Sentimen Opini Terhadap Tools Artificial Intelligence (AI) Berdasarkan Twitter Menggunakan Algoritma Naive Bayes. *Jurnal Media Informatika Budidarma*, 8(2), 777–787.
- Ramadhon, R. N., Ogi, A., Agung, A. P., Putra, R., Febrihartina, S. S., & Firdaus, U. (2024). Implementasi Algoritma Decision Tree untuk Klasifikasi Pelanggan Aktif atau Tidak Aktif pada Data Bank. *Karimah Tauhid*, 3(2), 1860–1874.
- Ridwan, R., Hermaliani, E. H., & Ernawati, M. (2024). Penerapan: Penerapan Metode SMOTE Untuk Mengatasi Imbalanced Data Pada Klasifikasi Ujaran Kebencian. *Computer Science (CO-SCIENCE)*, 4(1), 80–88.
- Rusman, J., Haryati, B. Z., & Michael, A. (2023). Optimisasi Hiperparameter Tuning pada Metode Support Vector Machine untuk Klasifikasi Tingkat Kematangan Buah Kopi. *J-Icon: Jurnal Komputer Dan Informatika*, 11(2), 195–202.
- Santosa, A. (2023, July 8). *Siaran Pers: Pembiayaan UMKM Lewat Pinjaman Online terus Berkembang, Pinjaman Masyarakat masih Terkendali*. Ojk.Go.Id.
- Saputra, M., Sidabuke, J. P., Sinulingga, R. P., & Tamba, R. B. (2023). Analisis Metode Algoritma K-Nearest Neighbor (KNN) Dan Naive Bayes Untuk Klasifikasi Diabetes Mellitus. *Jurnal Tekinkom (Teknik Informasi Dan Komputer)*, 6(2), 723–729.
- Sumantri, G., Novianto, M. D., & Prihastuti, P. P. (2023). Implementasi Fuzzy C-Means dalam Pengelompokan Provinsi di Indonesia untuk Pemerataan Kualitas Pendidikan. *Prosiding Seminar Pendidikan Matematika Dan Matematika*, 8.
- Surana, S. (2021). *Loan Prediction Based on Customer Behavior*. Kaggle.Com.
- Ulfa, M., & Mulyadi, M. (2020). Analisis dampak kredit usaha rakyat pada sektor Usaha Mikro terhadap penanggulangan kemiskinan di Kota Makassar. *Aspirasi: Jurnal Masalah-Masalah Sosial*, 11(1), 17–28.
- Yulianti, S. E. H., Soesanto, O., & Sukmawaty, Y. (2022). Penerapan Metode Extreme Gradient

- Boosting (XGBOOST) pada Klasifikasi Nasabah Kartu Kredit. *Journal of Mathematics: Theory and Applications*, 21–26.
- Zelvia, R. (2024). Peran Pembiayaan Kredit Usaha Rakyat (KUR) Terhadap Perkembangan Umkm (Studi Pada PT. Bank Syariah Indonesia, Tbk Kantor Cabang Kalianda). *Ekonodinamika: Jurnal Ekonomi Dinamis*, 6(1).
- Zöller, M.-A., & Huber, M. F. (2021). Benchmark and survey of automated machine learning frameworks. *Journal of Artificial Intelligence Research*, 70, 409–472.
- Zuama, R. A., Ichsan, N., Pohan, A. B., Azis, M. S., & Lase, M. (2024). An implementation of machine learning on loan default prediction based on customer behavior. *Jurnal Info Sains: Informatika Dan Sains*, 14(01), 157–164.