

Penerapan Metode *Extreme Gradient Boosting* (XGBOOST) pada Klasifikasi Nasabah Kartu Kredit

Sri Elina Herni Yulianti¹, Oni Soesanto², Yuana Sukmawaty³
^{1,2,3} Program Studi Statistika, Universitas Lambung Mangkurat, Indonesia
e-mail: ¹1811017320018@mhs.ulm.ac.id

Abstrak. Kartu kredit macet merupakan suatu masalah ketidakmampuan pengguna kartu kredit dalam membayar tagihan kartu kredit yang dapat menyebabkan kerugian pada kedua belah pihak yang bersangkutan. Guna menghindari kerugian yang disebabkan oleh kartu kredit macet maka pihak penyedia harus melakukan analisis yang cermat pada calon atau nasabah lama pengguna kartu kredit. Penelitian ini bertujuan untuk mengklasifikasi nasabah kartu kredit yang macet menggunakan teknik *machine learning* yaitu teknik klasifikasi. Salah satu teknik klasifikasi yang digunakan yaitu metode *XGBoost* yang mana metode ini berguna untuk analisis regresi dan klasifikasi berdasarkan *Gradient Boosting Decision Tree* (GBDT), metode *XGBoost* memiliki beberapa *hyperparameter* yang dapat dikonfigurasi untuk meningkatkan kinerja dari model. Metode *hyperparameter tuning* yang digunakan yaitu *grid Search cross validation* yang kemudian divalidasi menggunakan *10-Fold Cross Validation*. *Hyperparameter XGBoost* yang dikonfigurasi antara lain *n_estimators*, *max_depth*, *subsample*, *gamma*, *colsample_bylevel*, *min_child_weight* dan *learning_rate*. Berdasarkan hasil dari penelitian ini membuktikan bahwa penggunaan algoritma dengan *hyperparameter tuning* dapat meningkatkan kinerja algoritma *eXtreme Gradient Boosting* dalam proses klasifikasi nasabah kartu kredit dengan akurasi sebesar 80,039%, presisi sebesar 81,338% dan nilai *recall* sebesar 96,854%.

Kata kunci: *XGBoost*, klasifikasi, akurasi, presisi, *recall*

Abstract. Bad credit card is a problem of inability of credit card users to pay credit card bills that can cause losses to both parties concerned. In order to avoid losses caused by bad credit cards, the provider must conduct a careful analysis of prospective or old customers using credit cards. This study aims to classify bad credit card customers using machine learning techniques, namely classification techniques. One of the classification techniques used is the *XGBoost* method which is useful for regression analysis and classification based on the *Gradient Boosting Decision Tree* (GBDT), the *XGBoost* method has several *hyperparameters* that can be configured to improve the performance of the model. *Hyperparameter tuning* method used is *grid search cross validation* which is then validated using *10-Fold Cross Validation*. *XGBoost* *hyperparameters* configured include *n_estimators*, *max_depth*, *subsample*, *gamma*, *colsample_bylevel*, *min_child_weight* and *learning_rate*. Based on the results of this study proves that the use of algorithms with *hyperparameter tuning* can improve the performance of *eXtreme Gradient Boosting* algorithm in the process of classification of credit card customers with an accuracy of 80.039%, precision of 81.338% and a recall value of 96.854%.

Keywords: *XGBoost*, classification, accuracy, precision, recall

I. PENDAHULUAN

Penggunaan kartu kredit saat ini sangat diminati oleh beberapa kalangan masyarakat salah satu contohnya adalah Indonesia, mengutip dari laporan survei perbankan yang diterbitkan oleh Bank Indonesia ternyata pertumbuhan kredit di Indonesia mengalami kenaikan yang cukup signifikan dalam 1 tahun terakhir. Peningkatan dalam penggunaan kartu kredit ini diperkirakan karena banyaknya kelebihan yang dimiliki oleh kartu kredit. Beberapa kelebihan kartu kredit yaitu dapat mempermudah transaksi jual beli, jangka kartu kredit yang panjang, proses persetujuan dari bank yang cepat dan mudah serta limit pinjaman yang besar, dengan banyaknya kelebihan

inihal kartu kredit cukup diminati oleh masyarakat. Pengguna kartu kredit diimbau untuk menggunakan kartu kredit secara bijak (sesuai kebutuhan) dan selalu melakukan pembayaran sesuai tagihan dengan waktu yang telah ditetapkan, untuk menghindari status kredit macet. Kredit macet adalah ketidakmampuan nasabah dalam membayar atau melunasi pinjamannya kepada penyedia layanan kartu kredit yang akhirnya dapat merugikan bagi kedua belah pihak [12].

Di bidang keuangan, penggunaan kartu kredit kepada nasabah adalah layanan yang memiliki resiko tinggi dalam kerugian atau keuntungan. Lambat laun, kredit macet sering terjadi karena para data analyst di bank yang tidak hati-hati dalam proses pemberian kredit, atau dari nasabah

yang buruk (tidak melakukan pembayaran kartu kredit secara rutin). Guna mencegah kredit buruk atau kredit yang berstatus macet, para data analyst dari bank harus memilih pilihan yang tepat untuk menerima atau menolak aplikasi kredit. Untuk menentukan keandalan nantinya, diperlukan ekspektasi yang tepat, salah satunya dengan memanfaatkan strategi penambangan informasi (data mining) [17].

Data mining adalah istilah yang digunakan untuk menggambarkan pengungkapan informasi dalam kumpulan data. Data mining adalah interaksi yang menggunakan faktual, numerik, kekuatan otak buatan manusia, dan strategi AI untuk memisahkan dan mengenali data bermanfaat dan informasi terkait dari kumpulan data besar. Terkait dengan dengan data mining, terdiri dari beberapa Teknik salah satunya adalah teknik klasifikasi. Klasifikasi adalah cara paling umum untuk menempatkan item tertentu dalam sekelompok karakter. Klasifikasi memiliki beberapa metode, salah satunya adalah metode eXtreme Gradient Boosting (XGBoost). Dalam penggunaan beberapa contoh penelitian menggunakan teknik klasifikasi, perhitungan ini dapat menghasilkan ketepatan yang sangat baik [2]; [8]; [10] dan [14].

Metode XGBoost adalah algoritma pengembangan dari *gradient tree boosting* yang berbasis algoritma *ensemble*, secara efektif bisa menanggulangi kasus *machine learning* yang berskala besar. Metode XGBoost dipilih karena memiliki beberapa fitur tambahan yang berguna untuk mempercepat sistem perhitungan dan mencegah *overfitting*. XGBoost dapat menyelesaikan berbagai contoh klasifikasi, regresi, dan ranking. XGBoost adalah perhitungan pengumpulan pohon yang terdiri dari bermacam-macam pohon sebelumnya (CART). Komponen utama di balik kemakmuran XGBoost adalah kemampuan beradaptasinya dalam berbagai situasi, fleksibilitas ini karena perbaikan dari perhitungan masa lalu.

Berdasarkan latar belakang di atas, maka tujuan penelitian ini adalah menerapkan metode *eXtreme Gradient Boosting (XGBoost)* pada klasifikasi nasabah kartu kredit menggunakan 2 percobaan yaitu tidak menggunakan teknik optimasi pada metode XGBoost dan akan menggunakan teknik optimasi pada metode XGBoost. Tujuan penelitian selanjutnya adalah untuk mengetahui seberapa besar akurasi dari proses klasifikasi nasabah kartu kredit menggunakan metode XGBoost.

II. LANDASAN TEORI

2.1 Klasifikasi

Klasifikasi sesuai [3] adalah cara paling umum untuk menempatkan item tertentu dalam sekelompok karakter, mengingat gagasan artikel yang dirujuk. Berdasarkan urutan keempat bagian tersebut sebagai berikut.

1. Kelas

Kelas adalah variabel terikat sebagai informasi langsung yang membahas 'tanda' yang terkandung dalam artikel setelah karakterisasi. Model: kelas hujan dan kelas panas.

2. Indikator

Indikator adalah faktor bebas yang ditunjukkan oleh kualitas informasi yang akan dicirikan dan dilihat dari pengelompokan yang dibuat. Misalnya: Musim, pengaturan umum.

3. Data Latih

Data latih adalah bermacam-macam informasi yang berisi sisi positif dari bagian-bagian masa lalu dan digunakan untuk mempersiapkan model untuk menemukan kelas yang sesuai dengan perkiraan yang sesuai.

4. Data Uji

Informasi pengujian adalah informasi baru yang akan dicirikan oleh model yang baru dibuat dan ketepatan pesanan dapat dinilai.

2.2 Ensemble Learning

Ensemble Learning adalah sekumpulan pengklasifikasi yang dilatih secara individual dimana hasil prediksi dari setiap pengklasifikasi akan digabungkan dengan beberapa cara seperti rata-rata, pemungutan suara terbanyak dan lain lain untuk menghasilkan prediksi akhir. Ensemble learning digunakan untuk meningkatkan kinerja model atau menangani kinerja model yang buruk. Teknik ensemble yang umum digunakan yaitu bagging dan boosting.

2.3 Boosting

Boosting yang dipresentasikan oleh Robert Schapire pada tahun 1998, adalah salah satu strategi pengumpulan yang dapat bekerja pada beberapa hasil pengelompokan yang lemah sehingga mereka dapat berubah menjadi hasil yang kuat. Teknik klasifikasi adalah strategi perhitungan pembelajaran yang dibuat berdasarkan beberapa karakterisasi atau model prakiraan yang kemudian digunakan untuk mengelompokkan informasi baru dengan mempertimbangkan beban yang diantisipasi yang dibuat sebelumnya. Konsep *ensemble* dengan *boosting* bekerja melalui persiapan kumpulan model secara berurutan dan kemudian menggabungkan seluruh model untuk membuat harapan, model berikutnya mendapatkan keuntungan dari kesalahan model sebelumnya [19].

2.4 eXtreme Gradient Boosting (XGBoost)

Metode XGBoost merupakan pengembangan dari *gradient boosting* yang diusulkan oleh Dr. Tianqi Chen dari University of Washington pada tahun 2014 [18]. *Gradient boosting* merupakan algoritma yang dapat menemukan solusi yang optimal untuk berbagai masalah khususnya pada regresi, klasifikasi dan ranking. Konsep dasar dari algoritma ini adalah menyesuaikan parameter pembelajaran secara berulang untuk menurunkan *loss function* (mekanisme evaluasi atas model). XGBoost menggunakan model yang lebih teratur untuk membangun struktur pohon regresi, sehingga dapat memberikan kinerja yang lebih baik dan mampu mengurangi kompleksitas model untuk menghindari *overfitting* [16]. Hasil prediksi akhir dari XGBoost adalah penjumlahan hasil prediksi dari setiap pohon regresi [6]. Algoritma berbasis *decision tree* memiliki kinerja yang baik pada data dengan fitur kategorikal dan tidak terlalu berpengaruh terhadap data dengan kelas tidak seimbang [11].

Pada metode ini diperlukan fungsi objektif yang berguna untuk menilai seberapa bagus model yang didapatkan sesuai dengan data latih [4]. Karakteristik yang terpenting dari fungsi objektif terdiri dari 2 bagian yaitu nilai pelatihan yang hilang dan nilai regularisasi seperti pada persamaan 2.1 berikut ini.

$$obj(\theta) = L(\theta) + \Omega(\theta) \quad (2.1)$$

Dimana L adalah fungsi pelatihan yang hilang, dan Ω adalah fungsi regularisasi, dan θ adalah parameter model terkait. Fungsi pelatihan yang hilang secara umum dapat ditulis seperti pada persamaan 2.2 sebagai berikut.

$$L(\theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) \quad (2.2)$$

Dimana y_i adalah nilai data sebenarnya yang dianggap benar dan \hat{y}_i adalah hasil nilai prediksi dari model, sedangkan n adalah jumlah iterasi nilai dari model.

2.5 Hyperparameter tuning dengan Grid Search Cross Validation

Pada metode *machine learning*, terdapat beberapa nilai parameter yang diperkirakan dapat meningkatkan kinerja model yang disebut *hyperparameter*. *Hyperparameter* digunakan untuk meningkatkan hasil kinerja algoritma, yang mana cukup mempengaruhi berbagai uji model. *Hyperparameter* dilakukan secara fisik atau dengan menguji sekelompok *hyperparameter* pada batas yang telah ditentukan. Pencarian *hyperparameter* dilakukan secara manual atau dengan menguji kumpulan *hyperparameter* pada parameter yang ditentukan sebelumnya [9]. Salah satu metode *hyperparameter* yang akan diaplikasikan adalah *grid search cross validation (CV)*. Berikut ini parameter yang akan di optimasi dan diperkirakan dapat meningkatkan kinerja model dalam mengklasifikasi yaitu ada 7 parameter, dapat dilihat pada Tabel 1 berikut ini:

Tabel 1. Hasil nilai terbaik untuk hyperparameter tuning

Hyperparameter	Kegunaan Hyperparameter
<i>n_estimators</i>	Banyaknya pohon yang digunakan untuk proses klasifikasi
<i>max_depth</i>	Tingkat dalamnya pohon
<i>min_child_weight</i>	Bobot minimal
<i>eta (learning_rate)</i>	Membantu mempersingkat langkah dalam pembaruan model
<i>gamma</i>	Meminimumkan pengurangan kerugian
<i>subsample</i>	Rasio <i>instance</i> dari data latih
<i>colsample_bylevel</i>	Rasio dari data latih yang digunakan untuk membuat pohon

2.6 Evaluasi Kinerja Algoritma Klasifikasi

Secara umum evaluasi kinerja algoritma klasifikasi dilihat dengan membandingkan nilai prediksi algoritma klasifikasi dengan nilai target variabel data uji sebagai data

sebenarnya. Dan ada beberapa cara untuk mengevaluasi model klasifikasi yang didapatkan yaitu dengan nilai akurasi, presisi dan *recall*. Nilai untuk evaluasi kinerja model dari metode *XGBoost* ini didapatkan dari nilai *confusion matrix*. *Confusion matrix* merupakan suatu alat ukur berbentuk matriks, dari nilai *confusion matrix* didapatkan beberapa nilai evaluasi yaitu nilai akurasi, presisi dan *recall*.

2.7 Kartu Kredit

Secara komunikatif, kartu kredit adalah suatu cara pembayaran angsuran sebagai pengganti uang, yang dapat ditukarkan dengan apa saja yang diinginkan dimana terdapat cabang yang dapat menerima angsuran dengan menggunakan kartu kredit. Satu lagi pemahaman terperinci tentang kartu kredit adalah uang plastik yang diberikan oleh sebuah yayasan yang memungkinkan pemegang kartu untuk memperoleh kredit untuk pertukaran yang mereka selesaikan dan pembayaran dapat dilakukan sebagian atau diselesaikan sepenuhnya pada saat jatuh tempo [5].

III. METODE PENELITIAN

Adapun prosedur pada penelitian ini yaitu:

- a. Melakukan *preprocessing* data
- b. Statistika Deskriptif
Pada tahap ini akan dilakukan analisis deskriptif data berguna untuk memberikan gambaran mengenai data dan menampilkan penyajian data penelitian berguna untuk menyampaikan gambaran mengenai data dalam bentuk tabel dan *bar chart*.
- c. Pembagian data
Pada tahap ini data penelitian akan dibagi menjadi 2 bagian yaitu data latih dan data uji.
- d. Penerapan metode *XGBoost*
Pada tahap ini dilakukan penerapan metode *XGBoost* pada data yang akan diteliti. Juga akan melakukan simulasi klasifikasi menggunakan metode *XGBoost* dengan data yang berjumlah lebih kecil. Tahapan yang akan dilakukan, sebagai berikut:
 - d.1 Melatih model *XGBoost* menggunakan data latih.
 - d.2 Melakukan proses prediksi menggunakan model *XGBoost* yang dilatih sebelumnya.
- e. Melakukan proses *hyperparameter tuning* menggunakan metode *grid search cross validation*.
- f. Melakukan pengukuran kinerja metode *XGBoost* dalam proses klasifikasi dengan akurasi, presisi, *recall*.
- g. Interpretasi hasil dan kesimpulan.

IV. HASIL DAN PEMBAHASAN

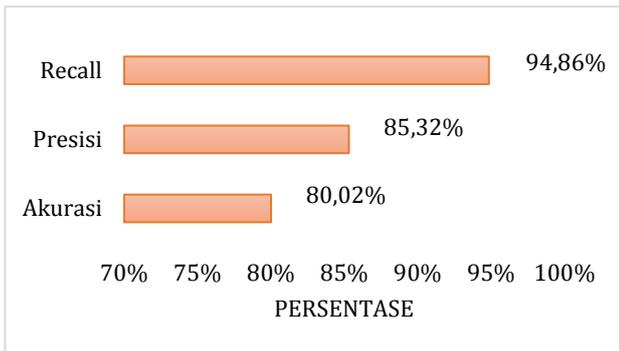
4.1 Hasil klasifikasi data tanpa hyperparameter tuning

Penerapan metode *XGBoost* pada klasifikasi data nasabah kartu kredit macet dan tidak macet pada penelitian ini menggunakan dataset *default of credit card clients* yang didapatkan dari situs *UCI Machine Learning Repository*. Dataset ini berisi 30000 data nasabah dengan 24 variabel dan memiliki 2 kelas yaitu sebanyak 23370 nasabah kartu kredit tidak macet dilambangkan dengan angka 0 dan sebanyak 6630 nasabah kartu kredit macet dilambangkan

dengan angka 1. Data yang digunakan pada penelitian ini hanya pada 7 variabel yaitu terdiri dari 6 variabel bebas dan 1 variabel terikat. Perbandingan persentase kelas pada dataset ini yaitu kelas tidak macet sebesar 78% dan kelas macet sebesar 22%. Pada tahapan preprocessing hal pertama yang dilakukan adalah mengeluarkan variabel yang tidak diperlukan pada penelitian. Selanjutnya, melakukan deteksi missing value dan data outlier.

Pada saat untuk melakukan modeling, dataset harus dibagi menjadi data latih dan data uji. Data latih yang digunakan sebanyak 80% dari jumlah nasabah dan data uji yang digunakan sebanyak 20% dari jumlah nasabah. Pada penelitian ini akan dilakukan 2 percobaan yaitu percobaan pertama klasifikasi menggunakan *XGBoost* tanpa hyperparameter tuning dan percobaan kedua klasifikasi menggunakan *XGBoost* dengan hyperparameter tuning dengan metode grid search CV.

Pada penelitian ini klasifikasi menggunakan metode *XGBoost* dilakukan dengan bantuan library *XGBoost* pada python. Pada percobaan klasifikasi menggunakan *XGBoost* tanpa hyperparameter tuning, seluruh data latih digunakan untuk melatih model *XGBoost*, sedangkan data uji digunakan untuk mengukur kinerja dari model yang dihasilkan berupa akurasi, presisi dan recall. Adapun hasil dari kinerja yang dihasilkan pada model *XGBoost* tanpa hyperparameter tuning dapat dilihat pada Gambar 1.



Gambar 1. Hasil evaluasi kinerja tanpa *hyperparameter tuning*

Dari grafik diatas dapat diketahui bahwa akurasi atau persentase ukuran yang menentukan kemiripan antara hasil prediksi nasabah kartu kredit yang macet menggunakan algoritma *XGBoost* dengan data uji yang sebenarnya diukur menghasilkan persentase sebesar 80,02%. Presisi adalah persentase tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan sistem, dan hasil presisi dari algoritma *XGBoost* sebesar 85,32%. Recall adalah persentase prediksi nasabah kartu kredit macet dibandingkan dengan keseluruhan data nasabah kartu kredit yang macet, maka persentasenya sebesar 94,86%. Setelah dilakukan penelitian terkait klasifikasi nasabah kartu kredit menggunakan metode *XGBoost* tanpa hyperparameter tuning hasil yang didapatkan cukup memuaskan termasuk dalam kategori good classification.

4.2 Hasil Kalsifikasi Data dengan *Hyperparameter Tuning*

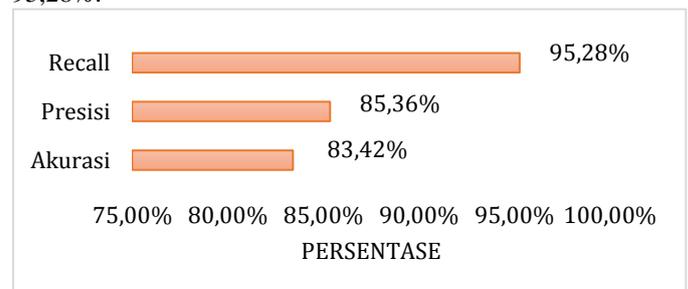
Pada penelitian ini juga dilakukan klasifikasi menggunakan algoritma *XGBoost* dengan proses *hyperparameter tuning* yaitu optimasi parameter yang

berguna untuk meningkatkan kinerja model dalam mengklasifikasi. Terdapat 7 parameter yang diperkirakan dapat meningkatkan kinerja model dalam mengklasifikasi menggunakan metode *XGBoost*. *Hyperparameter tuning* yang dilakukan pada 7 parameter ini menggunakan metode *grid search CV*. *Grid search CV* dikategorikan sebagai metode yang teliti, karena dalam menentukan parameter terbaik dilakukan eksplorasi masing masing parameter dengan mengatur jenis nilai prediksi terlebih dahulu. Konfigurasi *hyperparameter* optimal dari *grid search* dipilih berdasarkan nilai akurasi *cross validation* tertinggi dari kandidat *hyperparameter*. Hasil nilai *hyperparameter* terbaik sebagai berikut:

Tabel 2. Hasil nilai terbaik untuk hyperparameter tuning

<i>Hyperparameter</i>	<i>Grid Search Values</i>	Nilai <i>Hyperparameter</i> Terbaik
<i>n_estimator</i>	400, 300, 200, 100	100
<i>max_depth</i>	8, 7, 6, 5, 4	6
<i>min_child_weight</i>	0, 1, 2, 3, 4, 5, 6, 7	7
<i>eta (learning_rate)</i>	0,3, 0,2, 0,1, 0,05, 0,025	0,025
<i>gamma</i>	0, 0,1, 0,2, 0,3, 0,4, 1, 1,5 ,2	2
<i>subsample</i>	1, 0,75, 0,5, 0,15	1
<i>colsample_bylevel</i>	0,1, 0,2, 0,25, 1,0	1

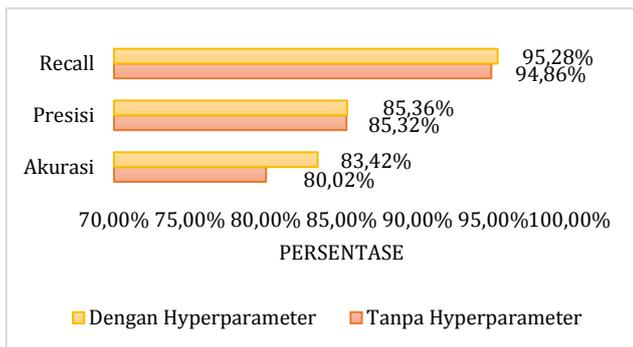
Adapun parameter dan nilai parameter terbaik yang dapat meningkatkan kinerja algoritma dapat dilihat pada Tabel 2. Konfigurasi *hyperparameter* ini kemudian digunakan untuk melatih ulang model *XGBoost* menggunakan keseluruhan dari data latih. Model *XGBoost* dengan *hyperparameter tuning* kemudian dievaluasi menggunakan data uji untuk mengukur kinerja berupa akurasi, presisi dan recall. Adapun untuk hasil kinerja yang dihasilkan pada model *XGBoost* dengan *hyperparameter tuning* disajikan dalam bentuk grafik pada Gambar 2. Dapat dilihat pada Gambar 2 model *XGBoost* dengan *hyperparameter tuning* menghasilkan akurasi sebesar 83,42%, presisi sebesar 85,36% dan recall sebesar 95,28%.



Gambar 2. Hasil evaluasi kinerja dengan *hyperparameter tuning*

Guna mengetahui pengaruh *hyperparameter tuning* pada kinerja *XGBoost* dalam melakukan klasifikasi pada

dataset *credit card*, maka dilakukan perbandingan antara kinerja model *XGBoost* tanpa *hyperparameter tuning* dengan kinerja model *XGBoost* dengan *hyperparameter tuning*. Dari hasil *confusion matrix* klasifikasi *XGBoost* tanpa *hyperparameter tuning* dan hasil *confusion matrix* klasifikasi *XGBoost* dengan *hyperparameter tuning* pada Tabel 2, dapat dilihat bahwa *hyperparameter tuning* memberikan pengaruh terhadap evaluasi hasil kinerja klasifikasi, yaitu pada akurasi mengalami peningkatan sebesar 3,4 %, pada presisi mengalami peningkatan sebesar 0,04% dan untuk nilai recall juga mengalami peningkatan sebesar 0,421%. Pada Gambar 3 dapat dilihat grafik perbandingan antara kinerja model *XGBoost* tanpa *hyperparameter tuning* dengan kinerja model *XGBoost* dengan *hyperparameter tuning*.



Gambar 3. Perbandingan hasil evaluasi kinerja algoritma metode *XGBoost*

Berdasarkan hasil penelitian dapat disimpulkan bahwa *hyperparameter tuning* terbukti mampu meningkatkan kinerja algoritma dalam proses klasifikasi. Sejalan dengan penelitian [7] dan [13] , 2 penelitian ini juga memberikan bukti bahwa proses *hyperparameter tuning* yang digunakan sebelum melatih data dapat meningkatkan kinerja algoritma, khususnya untuk teknik klasifikasi.

Dari uraian diatas dapat dikatakan bahwa *hyperparameter tuning* merupakan hal yang disarankan menjadi salah satu tahapan sebelum melakukan klasifikasi. Karena sesuai dengan hasil penelitian dan penelitian terdahulu yang dilakukan bahwa *hyperparameter tuning* dapat meningkatkan kinerja model pada proses klasifikasi.

V. KESIMPULAN

Berdasarkan hasil penelitian pada pembahasan didapatkan kesimpulan bahwa Hasil klasifikasi menggunakan metode *XGBoost* dengan parameter yang default pada dataset nasabah pengguna kartu kredit menghasilkan model yang dikatakan cukup baik yaitu akurasi model sebesar 80,02% , untuk presisi sebesar 85,32%, *recall* sebesar 94,86% dan dapat dikategorikan sebagai *good classification*. Untuk percobaan kedua menggunakan teknik optimasi yaitu proses *hyperparameter tuning* menggunakan 7 *hyperparameter* dengan memvalidasi data, maka didapatkan hasil *hyperparameter tuning* yang diperoleh akurasi model sebesar 83,42%, presisi sebesar 85,36%, *recall* sebesar 95,28% dan hasil klasifikasi termasuk kategori *good*

classification. Dengan hasil yang didapatkan ini terbukti bahwa *hyperparameter tuning* merupakan solusi terbaik jika ingin meningkatkan kinerja algoritma *XGBoost* dalam klasifikasi.

REFERENCE

- [1] Chen, T., & Guestrin, C. 2016. Xgboost: A Scalable Tree Boosting System. In Proceedings Of The 22nd Acm Sigkdd International Conference On Knowledge Discovery And Data Mining (Pp. 785-794).
- [2] Cherif, Iyad Lahsen, And Abdesselem Kortebi. 2019. "On Using Extreme Gradient Boosting (Xgboost) Machine Learning Algorithm For Home Network Traffic Classification." IFIP Wireless Days 2019-April: 1–6.
- [3] Gorunescu, F. 2011 *Data Mining Concepts, Models And Techniques*.
- [4] Hanif, I. 2020 'Implementing *Extreme Gradient Boosting (Xgboost)* Classifier To Improve Customer Churn Prediction'.
- [5] Kristianti, D. S. 2014 'Kartu Kredit Syariah dan Perilaku Konsumtif Masyarakat', *AHKAM: Jurnal Ilmu Syariah*, 14(2), Pp. 287–296.
- [6] Li, S. And Zhang, X. 2020 'Research On Orthopedic Auxiliary Classification and Prediction Model Based On *Xgboost* Algorithm', *Neural Computing And Applications*, 32(7), Pp. 1971–1979.
- [7] Mujilawati, S. Dkk. 2021. 'Optimasi Hyperparameter Tensorflow Dengan Menggunakan Optuna Di Python: Study Kasus Klasifikasi Dokumen Abstrak Skripsi', 5, Pp. 1084–1089. Doi: 10.30865/Mib.V5i3.3090.
- [8] Nokeri, T. C. 2021 *Data Science Revealed, Data Science Revealed*. Doi: 10.1007/978-1-4842-6870-4.
- [9] Muslim, Ichwanul, Karo Karo, 2020. "Implementasi Metode *Xgboost* dan Feature Importance Untuk Klasifikasi Pada Kebakaran Hutan Dan Lahan." 1(1): 10–16.
- [10] Nokeri, Tshepo Chris. 2021. *Data Science Revealed Data Science Revealed*.
- [11] Nyoman, Ngakan Et Al. 2020. "Prediksi Kecelakaan Lalu Lintas Di Bali Dengan *Xgboost* Pada Python." 8(3): 188–96.
- [12] Pavan, R. Et Al. 2021 'Bayesian Optimization and Gradient Boosting To Detect Phishing Websites', *2021 55th Annual Conference On Information Sciences And Systems, CISS 2021*, Pp. 2–6.
- [13] Pritasari, Mardhiyah Putri Ayu Siregar Riki Ruli A Palupiningsih. 2020. "Klasifikasi Untuk Memprediksi Pembayaran Kartu Kredit Macet." *Jurnal Teknologia* 3(1): 91–101.
- [14] Ramadhan, M. M. Et Al. 2017 'Parameter Tuning In Random Forest Based On Grid Search Method For Gender Classification Based On Voice Frequency', *Destech Transactions On Computer Science And Engineering*, (Cece), Pp. 625–629.
- [15] Siringoringo, Rimbun, Resianta Perangin-Angin, and Mufria J Purba. 2021. "Segmentasi dan Peramalan Pasar Retail Menggunakan *Xgboost* dan Principal Component Analysis." 5(1): 42–47.
- [16] Sunata, Haliem Azrullah, Firman Jodi Rianto, Y.

- 2020 'Komparasi Tujuh Algoritma Identifikasi Fraud ATM Pada PT. Bank Central Asia Tbk 1,2,3', 7(3), Pp. 441–450.
- [17] Tama, B. A. *Et Al.* 2020 'An Enhanced Anomaly Detection In Web Traffic Using A Stack Of Classifier Ensemble', *IEEE Access*, 8, Pp. 24120–24134.
- [18] Wijayanti, Reny. 2018. "Analisa Klasifikasi Kartu Kredit Menggunakan Algoritma Naive Bayes 1,2." *Jurnal Mantik Penusa* 10(2): 80–86.
- [19] Zhang, D. And Gong, Y. 2020 'The Comparison Of Lightgbm And Xgboost Coupling Factor Analysis And Prediagnosis Of Acute Liver Failure', *IEEE Access*, 8.
- [20] Zhou, Z.-H. 2012 *Ensemble Methods Foundations And Algorithms*, *SEAFSI Quarterly (South East Asia Iron And Steel Institute)*.