



Perbandingan Xgboost dan Logistic Regression dalam Memprediksi Credit Card Customer Churn

Rearizth Muhammad Daffaa^{1*}, Deris Santika², Fathoni Mahardika³

^{1,2,3}Universitas Sebelas April, Indonesia

Alamat: Jl. Angkrek Situ No.19, Sumedang, Jawa Barat, Indonesia

*Korespondensi penulis: rearizthmd772@gmail.com

Abstract. *With the times, cash transactions that used to use cash are now turning to credit cards. However, the increasing use of credit cards presents challenges, especially in maintaining customer loyalty. Customer churn is the loss of customers within a certain period for various reasons. Logistic Regression is a machine learning algorithm that studies the relationship between a dependent variable and several independent variables and Extreme Gradient Boosting (XGBoost) is a Gradient tree-boosting algorithm that offers out-of-core learning and sparsity awareness. The purpose of this study is to compare the performance between Logistic Regression and Extreme Gradient Boosting (XGBoost) algorithms in predicting customer churn in credit card services using evaluation metrics such as accuracy, precision, recall, and F1-score. Based on the research results, it can be concluded that XGBoost has better performance in all evaluation metrics, both in terms of precision, recall, F1-score, and accuracy. Based on the research, XGBoost shows superior performance compared to Logistic Regression in all evaluation metrics.*

Keywords: *Customer Churn, F1-Score, Logistic Regression, Precision, Recall.*

Abstrak. Dengan perkembangan zaman, transaksi tunai yang awalnya menggunakan uang tunai kini beralih ke kartu kredit. Akan tetapi, meningkatnya penggunaan kartu kredit menghadirkan tantangan, khususnya dalam menjaga loyalitas nasabah. *Customer churn* adalah hilangnya pelanggan dalam periode tertentu karena berbagai alasan. *Logistic Regression* adalah algoritma machine learning yang mempelajari hubungan antara variabel dependen dan beberapa variabel independent dan *Extreme Gradient Boosting (XGBoost)* adalah algoritma *Gradient tree-boosting* yang menawarkan *out-of-core learning* dan *sparsity awareness*. Tujuan dari penelitian ini adalah untuk membandingkan kinerja antara algoritma *Logistic Regression* dan *Extreme Gradient Boosting (XGBoost)* dalam memprediksi *customer churn* pada layanan kartu kredit. Dengan menggunakan metrik evaluasi seperti akurasi, precision, recall, dan F1-score. Berdasarkan hasil penelitian dapat disimpulkan bahwa pada *XGBoost* memiliki performa yang lebih baik di semua metrik evaluasi, baik dari segi precision, recall, F1-score, maupun akurasi. Berdasarkan penelitian, *XGBoost* menunjukkan performa lebih unggul dibandingkan *Logistic Regression* di semua metrik evaluasi.

Kata Kunci: *Customer Churn, F1-Score, Logistic Regression, Precision, Recall.*

1. LATAR BELAKANG

Seiring dengan berkembangnya zaman membawa dampak ke berbagai aspek kehidupan masyarakat salah satunya adalah dalam hal transaksi tunai yang awalnya menggunakan uang cash (tunai), kemudian beralih kepada pembelanjaan dengan menggunakan kartu, dalam hal ini yang dimaksudkan adalah kartu kredit (credit card) (Ardelia & Margaretha Leon, 2021). Kartu kredit ialah fasilitas keuangan yang diberikan oleh lembaga perbankan atau penyedia keuangan lainnya yang dimana setiap pelanggan atau nasabah pengguna kartu kredit dapat melakukan pinjam bayar yang dibantu oleh bank atau lembaga keuangan lainnya setelah

memenuhi persyaratan atau dengan kata lain yakni bank membayar terlebih dahulu biaya belanja nasabah dan nasabah dapat membayar total biaya yang dikeluarkan tersebut secara mencicil (Farid Naufal et al., 2023) [2]. Berdasarkan data dari AKKI (Asosiasi Kartu Kredit Indonesia) penggunaan kartu kredit semakin banyak yang pada tahun 2020 sebanyak 16,9 juta terus bertambah dan menjadi sebanyak 18,89 juta pada tahun 2024 (Akki.or.id, 2024) . Akan tetapi, dengan penggunaan kartu kredit yang semakin banyak tidak lepas dari suatu hambatan khususnya dalam menjaga loyalitas nasabah.

Customer churn merupakan hilangnya jumlah pelanggan dengan berbagai macam alasan dalam periode waktu tertentu (Kaya et al., 2018). Di dalam konteks bank, customer churn ini terjadi disaat nasabah beralih dari suatu bank ke bank lainnya (Farid Naufal et al., 2023). Berdasarkan itu customer churn bisa memberikan pemahaman bagi bank tentang kepuasan pelanggan dan loyalitas pelanggan. Fenomena ini tidak hanya memengaruhi profitabilitas, tetapi juga menambah beban biaya operasional karena akuisisi pelanggan baru sering kali lebih mahal dibandingkan dengan mempertahankan pelanggan yang sudah ada (Yulianti et al., 2022). Di bank asia terdapat 7 faktor yang mempengaruhi nasabah untuk menutup layanan kartu kredit mereka yaitu kesalahan pelayanan (service failure), harga (price), ketidaknyamanan jarak (inconvenience distance), reputasi, iklan yang efektif, involuntary switching dan switching cost (Sunjaya, 2018). Pada tahun 2017 BRI salah satu bank terbesar di Indonesia kehilangan 17 ribu nasabah yang dimana Hal ini menunjukkan bahwa churn merupakan tantangan yang nyata bagi industri perbankan di Indonesia (Sunjaya, 2018) . Memahami faktor-faktor yang memengaruhi churn dan memprediksinya secara akurat menjadi elemen penting dalam pengelolaan strategi bisnis Perusahaan.

Dalam mendapatkan pemahaman guna memprediksi customer churn data mining menjadi salah satu pendekatan yang sangat penting. 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 (Faruq Mujaddid & Al-Faraby, 2017). Data mining memiliki beberapa teknik dalam kasus prediksi customer churn ini teknik yang akan dipakai adalah klasifikasi. Klasifikasi data bertujuan untuk memprediksi label pada kumpulan data (Mauludin Nur Aziz et al., 2023). Akan tetapi terdapat masalah pada kasus ini dalam klasifikasi yaitu imbalance data. Ini dikarenakan jumlah label didominasi oleh non-churn yang membuat data tidak seimbang yang dimana ini akan mempengaruhi akurasi pada klasifikasi . Salah satu metode untuk menangani imbalance data adalah metode Synthetic Minority Over-sampling Technique

(SMOTE). SMOTE adalah teknik oversampling yang dimana cara kerjanya adalah dengan meningkatkan data minor dengan cara diduplikasi (Faruq Mujaddid & Al-Faraby, 2017).

Logistic Regression dan Extreme Gradient Boosting (XGBoost) adalah algoritma machine learning untuk tugas klasifikasi yang sering dipakai. Metode logistic regression mempelajari hubungan antara variabel dependen dan beberapa variabel independent (Faruq Mujaddid & Al-Faraby, 2017). Sedangkan Extreme Gradient Boosting (XGBoost) adalah algoritma ensemble learning yang digunakan dalam masalah klasifikasi yang dimana XGBoost ini merupakan algoritma gradient tree-boosting yang menawarkan out-of-core learning and sparsity awareness (Fauzan & Murfi, 2018).

Berdasarkan hal itu tujuan dari penelitian ini adalah untuk membandingkan kinerja antara algoritma Logistic Regression dan Extreme Gradient Boosting (XGBoost) dalam memprediksi customer churn pada layanan kartu kredit. Dengan menggunakan metrik evaluasi seperti akurasi, precision, recall, dan F1-score, diharapkan penelitian ini dapat membantu bank dalam memilih algoritma terbaik untuk mengatasi customer churn.

2. KAJIAN TEORITIS

Kajian teoritis dalam penelitian ini meliputi customer churn, machine learning dalam prediksi churn, logistic regression dan extreme gradient boost (XGBoost).

Customer Churn

Customer churn adalah suatu kondisi dimana pelanggan memutuskan berhenti dari layanan suatu perusahaan dan beralih ke layanan perusahaan lain. Customer churn menjadi salah satu masalah krusial yang dihadapi setiap perusahaan terutama dalam persaingan bisnis. Jika kita dapat mengidentifikasi alasan serta waktu pelanggan berhenti menggunakan layanan secara akurat maka memprediksi customer akan sangat membantu perusahaan dalam merancang strategi yang efektif untuk mencegah kehilangan pelanggan (Mauludin Nur Aziz et al., 2023).

Machine Learning dalam Prediksi Churn

Machine Learning adalah teknik pengolahan data yang bertujuan untuk menemukan pola-pola yang hasilnya dapat di analisa dan diterapkan kedalam sebuah sistem. Salah satu dari kemampuan machine learning adalah untuk melakukan klasifikasi data, yaitu proses untuk memprediksi label atau kategori pada kumpulan data (Faruq Mujaddid & Al-Faraby, 2017). Kemampuan machine learning dalam klasifikasi data ini dapat digunakan untuk memprediksi

churn karena dalam prediksi churn terdapat dua label yaitu pelanggan yang churn dan yang tidak churn sehingga ini dapat dijadikan sebagai strategi dalam memprediksi churn.

Logistic Regression

Logistic regression merupakan model regresi biner statistik standar. Metode logistic regression mempelajari hubungan antara variabel dependen dan beberapa variabel independen. Nama regresi logistik digunakan apabila variabel dependen hanya memiliki dua nilai, contohnya 0 dan 1 (Faruq Mujaddid & Al-Faraby, 2017).

Extreme Gradient (XGBoost)

Metode XGBoost merupakan pengembangan dari gradient boosting yang diusulkan oleh Dr. Tianqi Chen dari University of Washington pada tahun 2014. Gradient boosting merupakan algoritma yang dapat menemukan solusi yang optimal untuk berbagai masalah khususnya pada regresi, klasifikasi dan ranking (Yulianti et al., 2022).

3. METODE PENELITIAN

Penelitian ini menggunakan beberapa tahapan yang dimulai dari pengumpulan data, pemrosesan data, pembagian data, penanganan imbalance data, penerapan model, Hyperparameter tuning dengan Grid Search Cross Validation, dan evaluasi model.

Pengumpulan Data

Dataset yang akan digunakan dalam penelitian ini adalah dataset *Credit Card customers* yang didapatkan dari Kaggle. Dataset ini memiliki 19 features dan 1 label dengan jumlah data sebanyak 10.127 yang dimana untuk jumlah labelnya didominasi oleh pelanggan yang tidak churn sebanyak 84% dan sisanya pelanggan yang churn sebanyak 16% (Goyal, 2020). Berikut adalah atribut dan keterangan yang ada di dataset yang ditunjukkan pada tabel dibawah ini.

Tabel 1. Keterangan Dataset

Atribut	Keterangan
Attrition_Flag	Label/kelas yang akan diprediksi
Customer_Age	Usia Nasabah
Gender	Jenis Kelamin
Dependent_count	Jumlah tanggungan
Education_Level	Kualifikasi pendidikan
Marital_Status	Status pernikahan
Income_Category	Kategori pendapatan
Card_Category	Jenis kartu kredit
Months_on_book	Jangka waktu hubungan dengan bank (Bulan)
Total_Relationship_Count	Banyak produk yang dimiliki nasabah
Months_Inactive_12_mon	Jumlah bulan nasabah tidak aktif (12 bulan terakhir)
Contacts_Count_12_mon	Jumlah kontak (12 bulan terakhir)

Atribut	Keterangan
Credit_Limit	Jumlah limit pada kartu kredit
Total_Revolving_Bal	Total saldo bergulir pada kartu kredit
Avg_Open_To_Buy	Rata-rata transaksi kartu kredit
Total_Amt_Chng_Q4_Q1	Perubahan jumlah nominal transaksi Q4 sampai Q1
Total_Trans_Amt	Jumlah nominal transaksi (12 bulan terakhir)
Total_Trans_Ct	Jumlah transaksi (12 bulan terakhir)
Total_Ct_Chng_Q4_Q1	Perubahan jumlah transaksi Q4 sampai Q1
Avg_Utilization_Ratio	Rasio penggunaan kartu rata-rata

Pemrosesan Data

Pada tahapan terdapat beberapa tahap yang akan dilakukan yaitu :

- a. Drop Kolom yang tidak digunakan.

Terdapat 3 kolom yang tidak digunakan yaitu kolom CLIENTNUM, Naive_Bayes_Classifier_Attrition_Flag_Card_Category_Contacts_Count_12_mon_Dependent_count_Education_Level_Months_Inactive_12_mon_2 dan kolom Naive_Bayes_Classifier_Attrition_Flag_Card_Category_Contacts_Count_12_mon_Dependent_count_Education_Level_Months_Inactive_12_mon_1

- b. Handling Missing Value

Di dataset ini semua kolomnya tidak terdapat null, maka tahapan ini bisa dilewati

- c. Transformasi Data

Pada tahapan ini dilakukan normalisasi data dan encoding , untuk normalisasi data menggunakan MinMaxScalar yaitu mensklakan nilai numerik ke rentang 0 hingga 1, lalu untuk encoding yaitu mengubah kolom kategorikal ke numerik dipenelitian ini untuk encoding menggunakan OneHotEncoder (Mohammad Aulia Riftiarraafi & Dira Ernawati, 2024).

	Customer_Age	Dependent_count	Months_on_book	Total_Relationship_Count	Months_Inactive_12_mon	Contacts_Count_12_mon	Credit_Limit
0	0.595745	0.6	0.837209	1.0	0.333333	0.500000	0.355094
1	0.531915	0.0	0.744186	0.4	0.333333	0.500000	0.278426
2	0.404255	0.8	0.372093	0.4	0.333333	0.166667	0.095433
3	0.574468	0.6	0.511628	0.8	0.500000	0.333333	0.071066
4	0.468085	0.4	0.627907	0.4	0.333333	0.666667	0.110821

Gambar 1. Dataset Setelah Transformasi Data

Total_Revolving_Bal	Avg_Open_To_Buy	Total_Amt_Chng_Q4_Q1	...	Income_Category_\$120K +	Income_Category_\$40K - \$60K	Income_Category_\$60K - \$80K	Income_Category_\$80K - \$120K
0.000000	0.381914	0.343244	...	0.0	0.0	0.0	0.0
1.000000	0.235505	0.208419	...	0.0	0.0	0.0	1.0
0.707588	0.081448	0.208419	...	0.0	1.0	0.0	0.0
0.677791	0.060267	0.308213	...	0.0	1.0	0.0	0.0
0.292412	0.126474	0.199882	...	0.0	0.0	0.0	0.0

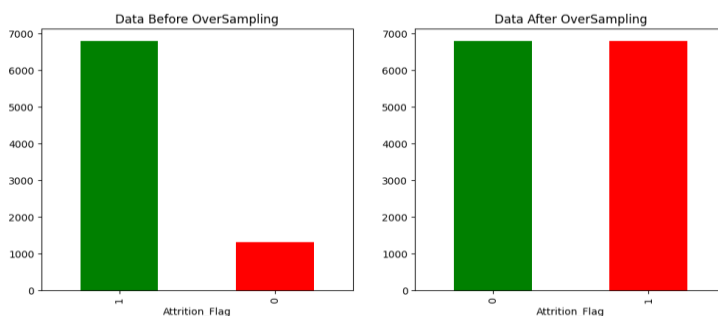
Gambar 2. Lanjutan Dataset Setelah Transformasi Data

Pembagian Data

Sebelum dilakukan modelling dataset dibagi menjadi dua terlebih dahulu yaitu data training dan data test, untuk proporsi data training digunakan sebanyak 80% dan untuk data test sebanyak 20%

Penanganan Imbalance Data

SMOTE (Synthetic Minority Over-sampling Technique) adalah teknik dalam machine learning yang dirancang untuk menangani masalah ketidakseimbangan kelas pada dataset. Ketidakseimbangan ini terjadi ketika jumlah sampel dalam satu kelas (kelas mayoritas) jauh lebih banyak dibandingkan kelas lainnya (kelas minoritas). Ketidakseimbangan ini sering menyebabkan model machine learning bias terhadap kelas mayoritas, menghasilkan akurasi tinggi secara keseluruhan, tetapi performa yang buruk dalam mengenali kelas minoritas yang mungkin memiliki nilai prediktif penting (Persada Pulungan et al., n.d.).



Gambar 3. Sebelum dan setelah SMOTE

Tabel 2. Sebelum dan Sesudah SMOTE

label	Sebelum SMOTE	Sesudah SMOTE
1 / Tidak Churn	6799	6799
0 / Churn	1302	6799

Berdasarkan tabel diatas bisa dilihat data yang sebelum SMOTE untuk label churn terdapat 1302 dan untuk label tidak churn terdapat 6799 dan setelah dilakukan SMOTE baik untuk label churn dan tidak churn datanya menjadi 6799. Alasan digunakannya SMOTE dibanding metode penanganan data tidak seimbang yang lain adalah karena SMOTE memberikan peningkatan yang signifikan terhadap matriks evaluasi seperti f1-score dan ROC-AUC, selain itu SMOTE juga membantu untuk mengurangi risiko overfitting (Last et al., 2017).

Penerapan Model

Pada tahapan ini terdapat 2 model yang akan digunakan yaitu Logistic Regression dan juga Extreme Gradient Boosting.

1) Logistic Regression

Logistic Regression (Regresi Logistik) adalah teknik pembelajaran statistik dan mesin yang digunakan untuk memodelkan hubungan antara variabel dependen biner dan satu atau lebih variabel independent. Logistic Regression bertujuan untuk memprediksi probabilitas kejadian suatu peristiwa atau kategori. Formula dari Logistic Regression terdapat pada persamaan sebagai berikut (Persada Pulungan et al., n.d.).

$$P = \frac{e^{a+bX}}{1 + e^{a+bX}}$$

Dimana:

P: Probabilitas suatu instance/data

a: Konstanta atau bias

bX: Koefisien yang menggambarkan hubungan antara atribut X dan Probabilitas

2) Extreme Gradient Boost

XGBoost adalah algoritma yang mengatasi permasalahan regresi dan klasifikasi berdasarkan Gradient Boosting Decision Tree (GBDT) (Cherif & Kortebi, 2019). XGBoost merupakan pengembangan dari gradient boosting decision tree, yang dirancang untuk membangun boosted trees secara efisien serta mampu beroperasi secara parallel (Chen & Guestrin, 2016). XGBoost dapat diformulasikan sebagai berikut.

$$obj(\theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_k^n \Omega(f_k)$$

Dimana $\sum_{i=1}^n l(y_i, \hat{y}_i)$ adalah differentiable loss function untuk mengukur apakah model tersebut cocok untuk set data pelatihan dan $\sum_k^n \Omega(f_k)$ adalah item yang menentukan kompleksitas model. Ketika kompleksitas model meningkat skor yang sesuai dikurangi nilainya (Muslim Karo Karo, 2020).

Hyperparameter tuning dengan Grid Search Cross Validation

Hyperparameter adalah sekumpulan nilai parameter yang digunakan dalam pembelajaran mesin yang dianggap dapat meningkatkan kinerja model. Performa algoritme dapat ditingkatkan dengan menggunakan hyperparameter, yang cukup untuk mempengaruhi pengujian model yang berbeda. Proses pengujian hyperparameter dapat dilakukan dengan

mengevaluasi sekumpulan nilai pada rentang yang telah ditentukan atau melalui eksplorasi langsung. Pendekatan ini dapat dilakukan secara manual, dengan memilih nilai-nilai hyperparameter tertentu, atau dengan menguji kombinasi nilai hyperparameter yang telah ditentukan sebelumnya. Grid search cross validation (CV) akan menjadi salah satu teknik hyperparameter yang digunakan (Yulianti et al., 2022). Untuk meningkatkan kinerja klasifikasi model, tiga parameter berikut yang ditunjukkan pada tabel di bawah ini akan diestimasi dan dioptimalkan:

Tabel 3. Hyperparameter Logistic Regression

Hyperparameter	Kegunaan Hyperparameter	Grid Search Value	Nilai Hyperparameter Terbaik
penalty	Menentukan jenis regularisasi yang diterapkan untuk mencegah overfitting	l1,l2,elasticnet	l2
C	Merepresentasikan parameter regulasi yang mengontrol kekuatan regularisasi	0.001,0.01,0.1,1	1
max_iter	Menentukan jumlah maksimum iterasi yang dilakukan solver selama proses optimisasi untuk mencapai konvergensi	100, 150, 200	100

Tabel 4. Hyperparameter XGBoost

Hyperparameter	Kegunaan Hyperparameter	Grid Search Value	Nilai Hyperparameter Terbaik
max_depth	Menentukan kedalaman maksimum setiap decision tree dalam ensemble.	2,3,4	3
n_estimators	Menentukan jumlah decision tree yang akan dibangun oleh algoritma dalam ensemble	100,200	200
min_samples_leaf	Menentukan jumlah minimum sampel yang diperlukan untuk membentuk sebuah leaf node.	60,75,100	60

Berdasarkan tabel diatas untuk hyperparameter Logistic Regression didapatkan parameter penelati dengan nilai l2 ,parameter C dengan nilai 1 dan max_iter dengan nilai 100.Lalu untuk hyperparameter XGBoost nilai hyperparameter terbaik yang didapatkan adalah max_depth dengan nilai 3,n_estimator sebanyak 200,dan min_samples_leaf sebanyak 60.

Evaluasi Model

Pada Tahapan ini evaluasi dilakukan untuk membandingkan performa kedua algoritma, Logistic Regression dan Extreme Gradient Boosting (XGBoost). Akurasi digunakan untuk mengukur tingkat kecocokan antara hasil prediksi dengan nilai actual (Farid Naufal et al., n.d.) [2]. Precision adalah rasio antara jumlah hasil positif yang diprediksi dengan benar terhadap total hasil yang diprediksi sebagai positif (Bintoro et al., n.d.). Recall mengacu pada rasio hasil

relevan yang diidentifikasi secara akurat oleh Algoritma(Bintoro et al., n.d.).F1-score adalah rata-rata dari precision dan recall(Bintoro et al., n.d.). Accuracy, precision, racall, dan F1-score dirumuskan sebagai berikut .

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+FP+FN} * 100\%$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Precision} = \frac{TP}{TP+FN}$$

$$\text{F1 -score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

4. HASIL DAN PEMBAHASAN

Logistic Regression

Dalam penelitian ini peneliti mengimplementasikan GridSearchCV dalam mencari parameter terbaik dan didapatkan parameter terbaik untuk digunakan pada logistic regression ini adalah {'model__C': 1,'model__max_iter': 100,'model__penalty': 'l1' }.Hasil laporan klasifikasi akan dijelaskan pada gambar berikut ini.

Model Performance in Data Testing Logistic Regression				
	precision	recall	f1-score	support
0	0.51	0.80	0.62	325
1	0.96	0.86	0.90	1701
accuracy			0.85	2026
macro avg	0.73	0.83	0.76	2026
weighted avg	0.89	0.85	0.86	2026

Gambar 4. Classification Report Logistic Regression

Berdasarkan gambar diatas,didapat hasil klasifikasi yang dimana untuk label 0 atau nasabah yang churn mendapatkan nilai precision 51%,recall 80%,dan f1-score dengan nilai 62%. Lalu untuk label 1 atau nasabah yang tidak churn mendapatkan nilai precision 96%,recall 86%,dan f1-score mendapatkan 90%.Dan akurasi untuk model logistic regression ini adalah 85%.

XGBoost

Sama halnya dengan logistic regression dimodel ini peneliti mengimplementasikan GridSearchCV dalam mencari parameter terbaik dan didapatkan parameter terbaik untuk digunakan pada XGBoost ini adalah {'model__max_depth': 3, 'model__min_samples_leaf':

60,'model__n_estimators': 200}.Hasil laporan klasifikasi akan dijelaskan pada gambar berikut ini.

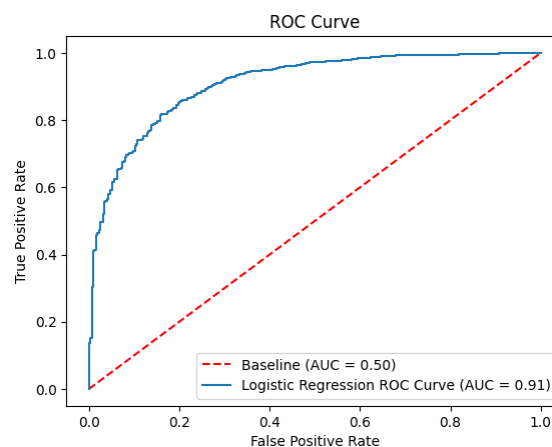
Model Performance in Data Testing XGBoost				
	precision	recall	f1-score	support
0	0.88	0.94	0.91	325
1	0.99	0.98	0.98	1701
accuracy			0.97	2026
macro avg	0.94	0.96	0.95	2026
weighted avg	0.97	0.97	0.97	2026

Gambar 5. Classification Report XGBoost

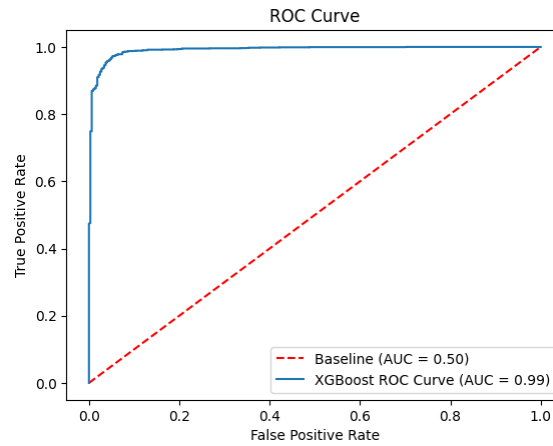
Berdasarkan gambar diatas,didapat hasil klasifikasi yang dimana untuk label 0 atau nasabah yang churn mendapatkan nilai precision 88%,recall 94%,dan f1-score dengan nilai 91%. Lalu untuk label 1 atau nasabah yang tidak churn mendapatkan nilai precision 99%,recall 98%,dan f1-score mendapatkan 98%.Dan akurasi untuk model XGBoost ini adalah 97%.

ROC dan AUC

ROC atau receiver operating characteristics adalah kurva yang menggambarkan keakuratan model klasifikasi(Siringoringo et al., 2021). ROC curve dibentuk berdasarkan hasil perhitungan dari confusion matrix, dengan mengacu pada hubungan antara False Positive Rate (FPR) dan True Positive Rate (TPR). Nilai performa algoritma dapat dibandingkan dengan menggunakan luas area di bawah kurva, yang dikenal sebagai AUC (Area Under Curve), untuk menentukan keunggulan relatif dari setiap model(Kristiawan & Widjaja, 2021). Nilai ROC dari logsitic regression adalah 91% dan XGBoost sebanyak 99%.Perbandingan hasil dari kedua algoritma tersebut dapat dilihat pada gambar dibawah ini.



Gambar 6. ROC Logsitic Regression



Gambar 7. ROC XGBoost

Pada gambar diatas menunjukkan kurva ROC AUC dalam memprediksi pengguna kartu kredit churn. Pada curva XGBoost atau pada gambar 7 menunjukkan jika XGBoost dapat membedakan antara nasabah yang churn dan tidak dengan kemampuan yang hampir sempurna dengan nilai Area Under Curve (AUC) sebesar 0.99. Di kurva tersebut juga menunjukkan jika XGBoost mendaptkan False Positive Rate (FPR) yang rendah tapi disisi lain mempertahankan True Positive Rate (TPR) yang tinggi.Sementara itu untuk model Logistic Regression atau yang ditunjukkan pada gambar 6 mendapatkan nilai Area Under Curve (AUC) sebesar 0.91 yang lebih rendah dibandingkan dengan XGBoost meskipun begitu ,model ini tetap memberikan prediksi yang baik ,dengan melampaui baseline (AUC = 50).

Analisis Performa Model

Tabel 5 dibawah adalah ringkasan hasil ringkasan evaluasi Logistic Regression dan XGBoost berdasarkan beberapa metrik evaluasi utama, termasuk confidence interval untuk memperkuat keandalan hasil.

Tabel 5. Ringkasan Hasil Evaluasi

Matrik Evaluasi	Logistic Regression	XGBoost
Akurasi	85% (CI: 83.17%, 86.23%)	97% (CI: 96.25%, 97.73%)
Precision (Label 0 / Nasabah Churn)	51%	88%
Recall (Label 0 / Nasabah Churn)	80%	94%
F1-score (Label 0 / Nasabah Churn)	62%	91%
Precision (Label 1 / Nasabah tidak Churn)	96%	99%
Recall (Label 1 / Nasabah tidak Churn)	86%	98%
F1-score (Label 1 / Nasabah tidak Churn)	90%	98%
AUC-ROC	90.88% (CI: 88.94%, 92.53%)	99.13% (CI: 98.65%, 99.52%)

Hasil Evaluasi menunjukkan jika Logistic Regression cenderung lebih baik dalam memprediksi kelas nasabah yang tidak churn akan tetapi dalam memprediksi kelas nasabah yang churn performanya lebih rendah sedangkan untuk XGBoost performanya lebih unggul baik dalam memprediksi kelas nasabah yang tidak churn ataupun nasabah yang churn dibandingkan dengan Logistic Regression. Alasan kenapa XGBoost lebih unggul karena XGBoost dapat menangkap pola non-linear yang lebih kompleks, sedangkan Logistic Regression hanya mampu menangkap hubungan linier lalu penggunaan regularisasi dan ensemble learning pada XGBoost yang dapat mengurangi risiko overfitting, serta memanfaatkan metode boosting untuk meningkatkan akurasi prediksi.

Penyertaan confidence interval memperkuat hasil penelitian yang dimana confidence interval untuk AUC-ROC pada Logistic Regression adalah 90.88% (CI: 88.94%, 92.53%), sementara pada XGBoost adalah 99.13% (CI: 98.65%, 99.52%), mengindikasikan bahwa XGBoost memiliki kemampuan diskriminasi yang lebih konsisten. Selain itu confidence interval untuk akurasi Logistic Regression adalah 85% (CI: 83.17%, 86.23%), sedangkan untuk XGBoost adalah 97% (CI: 96.25%, 97.73%). Interval yang lebih sempit pada XGBoost menunjukkan stabilitas model yang lebih tinggi.

Berdasarkan matriks evaluasi seperti akurasi, precision, recall, f1-score dan juga AUC menunjukkan jika XGBoost memiliki performa yang lebih baik dibandingkan dengan logistic regression yang dimana XGBoost mendapatkan nilai akurasi 97% sedangkan logistic regression mendapatkan akurasi sebesar 85%.

Perbandingan dengan penelitian sebelumnya dilakukan oleh Yulianti S. E. H., Soesanto, O, dan Sukmawaty .Y yang berjudul “Penerapan Metode Extreme Gradient Boosting(XGBOOST) pada Klasifikasi Nasabah Kartu Kredit” menunjukkan jika hasil dari XGBoost mendapatkan akurasi sebesar 80,02%, presisi sebesar 85,32% dan nilai recall sebesar 94,86%. Nilai-nilai ini menggambarkan keseimbangan yang baik antara metrik evaluasi di mana akurasi, presisi, dan recall memiliki nilai yang relatif tinggi dan konsisten. Sedangkan disisi lain pengimplementasian logistic regression juga pernah dilakukan oleh Aziz A. M. N, Mauludin, A, Hafiz R. D., Sintalana, V. A, dan Rismayadi A. A. yang berjudul “Prediksi Customer Churn Menggunakan Logistic Regression dan Decision Tree” dalam penelitian itu domain dalam adalah customer churn dalam Perusahaan telekomunikasi akan tetapi sifat masalah dan pendekatan analitis serupa, sehingga memungkinkan perbandingan yang relevan. Dan mendapatkan 80% nilai akurasi, 69% nilai precision, 57% nilai recall dan 62% nilai F1 Score. Berdasarkan penelitian-penelitian sebelumnya juga dapat

dinyatakan jika XGBoost memiliki performa yang lebih unggul dibandingkan dengan logistic regression.

Secara keseluruhan dalam penelitian ini memberikan wawasan mengenai algoritma mana yang akan dipakai untuk mengatasi nasabah yang akan berhenti berlayanan kartu kredit dan diharapkan dapat membantu pihak-pihak seperti bank atau Lembaga keuangan lainnya dalam pemilihan algoritma machine learning.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian dapat disimpulkan bahwa pada XGBoost memiliki performa yang lebih baik di semua metrik evaluasi, baik dari segi precision, recall, F1-score, maupun akurasi. dibandingkan dengan logistic regression yang dimana XGBoost mendapatkan akurasi sebesar 97% lalu pada label 0 atau customer yang churn mendapatkan precision 88%, recall 94%, dan f1-score dengan nilai 91% dan pada label 1 atau customer yang tidak churn mendapatkan nilai precision 99%, recall 98%, dan f1-score mendapatkan 98% dan nilai AUC sebesar 99%. Sebagai saran untuk penelitian selanjutnya dapat mencoba untuk menggunakan algoritma lainnya atau menambahkan hyperparameter tuning Grid search cross validation yang lebih kompleks sehingga dapat meningkatkan performa dari model machine learning.

DAFTAR REFERENSI

- Akki.or.id. (2024). *Jumlah kartu kredit*. <https://akki.or.id/statistik/>
- Ardelia, N., & Margaretha Leon, F. (2021). Faktor-faktor yang mempengaruhi konsumen dalam menggunakan dan berhutang dengan kartu kredit di Jabodetabek. *Jurnal Akuntansi, Keuangan, dan Manajemen*, 2(2), 83–101. <https://doi.org/10.35912/jakman.v2i2.136>
- Bintoro, P., Zulkifli, Z., Anggista Putri, N., Pringsewu, A., Kesehatan, F., & Aisyah Pringsewu, U. (n.d.). Analisis perbandingan klasifikasi virus cacar monyet dengan pendekatan algoritma machine learning. *SNESTIK Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika*. <https://doi.org/10.31284/p.snestik.2024.5849>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). <https://doi.org/10.1145/2939672.2939785>
- Cherif, I. L., & Kortebi, A. (2019). On using eXtreme Gradient Boosting (XGBoost) machine learning algorithm for home network traffic classification. *IFIP Wireless Days, 2019(April)*, 1–6. <https://doi.org/10.1109/WD.2019.8734193>

- Farid Naufal, M., Fernando Susanto, A., Nathaneil Kansil, C., Huda, S., & Kunci, K. (n.d.). Analisis perbandingan algoritma machine learning untuk prediksi potensi hilangnya nasabah bank. *Bank Application of Machine Learning*, 22(1).
- Faruq Mujaddid, M., & Al-Faraby, S. (n.d.). Analisis churn prediction menggunakan metode logistic regression dan SMOTE (Synthetic Minority Over-sampling Technique) pada perusahaan telekomunikasi.
- Fauzan, M. A., & Murfi, H. (2018). The accuracy of XGBoost for insurance claim prediction. *International Journal of Advance Soft Computing Applications*, 10(2). <https://www.claimsjournal.com/news/national/2013/11/21/240353.htm>
- Goyal, S. (2020). *Credit card customers*. Kaggle. <https://www.kaggle.com/datasets/sakshigoyal7/credit-card-customers>
- Kaya, E., Dong, X., Suhara, Y., Balcisoy, S., Bozkaya, B., & Pentland, A. S. (2018). Behavioral attributes and financial churn prediction. *EPJ Data Science*, 7(1). <https://doi.org/10.1140/epjds/s13688-018-0165-5>
- Kristiawan, K., & Widjaja, A. (2021). Perbandingan algoritma machine learning dalam menilai sebuah lokasi toko ritel. *Jurnal Teknik Informatika dan Sistem Informasi*, 7(1). <https://doi.org/10.28932/jutisi.v7i1.3182>
- Last, F., Douzas, G., & Bacao, F. (2017). Oversampling for imbalanced learning based on K-means and SMOTE. *Information Sciences*, 1–19. <https://doi.org/10.1016/j.ins.2018.06.056>
- Mauludin Nur Aziz, A., Mauliddin, A., Armanda Sintalana, V., Daryl Hafiz, R., Akbar Rismayadi, A., Studi Sistem Informasi, P., & Adhirajasa Reswara Sanjaya, U. (2023). Prediksi customer churn menggunakan logistic regression dan decision tree. *Jurnal Ilmiah Teknologi Informasi dan Komputer*, 4(1).
- Mohammad Aulia Riftiarraafi, & Ernawati, D. (2024). Penerapan algoritma C4.5 untuk klasifikasi customer churn pada perusahaan perbankan. *Sammajiva: Jurnal Penelitian Bisnis dan Manajemen*, 2(1), 178–190. <https://doi.org/10.47861/sammajiva.v2i1.808>
- Muslim Karo Karo, I. (2020). Implementasi metode XGBoost dan feature importance untuk klasifikasi pada kebakaran hutan dan lahan. *Journal of Software Engineering, Information and Communication Technology*, 1(1).
- Persada Pulungan, M., Purnomo, A., Kurniasih, A., & Tinggi Ilmu Manajemen dan Ilmu Komputer ESQ, S. (n.d.). Penerapan SMOTE untuk mengatasi imbalance class dalam klasifikasi kepribadian MBTI menggunakan Naive Bayes classifier. <https://doi.org/10.25126/jtiik.2024117989>
- Siringoringo, R., Perangin-angin, R., & Purba, M. J. (2021). Segmentasi dan peramalan pasar retail menggunakan XGBoost dan principal component analysis. *Methomika: Jurnal Manajemen Informatika dan Komputerisasi Akuntansi*, 5(1), 42–47. <https://doi.org/10.46880/jmika.Vol5No1.pp42-47>
- Sunjaya, M. I. (2018). Tingkat churn tabungan pada industri perbankan. *Ekspektra: Jurnal Bisnis dan Manajemen*, 2, 157–169. <https://doi.org/10.25139/ekt.v2i2.1101>

Yulianti, E. H., Soesanto, O., & Sukmawaty, Y. (2022). Penerapan metode Extreme Gradient Boosting (XGBoost) pada klasifikasi nasabah kartu kredit. *JOMTA: Journal of Mathematics: Theory and Applications*, 4(1).