e-ISSN: 3110-0864 | p-ISSN: xxxx-xxxx
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | Page: 252-262
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.163

Prediksi Churn Pelanggan Spotify Menggunakan XGBoost Berbasis Smote untuk Mengatasi Imbalanced Data Perilaku Streaming

Novia Ardani[1], Yakiya Ratna Sari[2*], Zesi Walikhsani[3]

[1,2,3]Teknik Informatika, STMIK Bina Patria, Kota Magelang, Indonesia
Email: [1]viaani863@gmail.com, [2*]ratnasariyakia@email.com, [3]zesiwalikhsani@gmail.com
(* Email Corresponding Author: ratnasariyakia@email.com)
Received: 5 Juli 2026 | Revision: 10 Juli 2026 | Accepted: 16 Juli 2026

Abstrak

Persaingan di industri layanan musik streaming sangat ketat, sehingga platform seperti Spotify perlu memahami perilaku konsumen dengan lebih baik, terutama tentang mengapa pelanggan berhenti menggunakan layanan mereka. Penelitian ini berfokus pada membuat model yang dapat memprediksi pelanggan mana yang mungkin berhenti menggunakan Spotify. Untuk membuat model ini, penelitian ini menggunakan algoritma *Extreme Gradient Boosting* atau XGBoost, yang dikombinasikan dengan metode *Synthetic Minority Over-sampling Technique* atau SMOTE. SMOTE digunakan untuk mengatasi masalah ketidakseimbangan data, di mana jumlah pelanggan yang berhenti menggunakan layanan sangat sedikit dibandingkan dengan yang tetap menggunakan. Data yang digunakan dalam penelitian ini berasal dari *Kaggle* dan terdiri dari 50.000 data pengguna dengan 18 atribut perilaku, seperti berapa lama mereka mendengarkan musik setiap minggu, seberapa sering mereka melewatkan lagu, berapa banyak playlist yang mereka buat, jenis langganan mereka, dan informasi demografis. Proses penelitian meliputi memahami data, menentukan target, membersihkan data, membuat fitur baru, dan membagi data menjadi dua bagian untuk pelatihan dan pengujian. SMOTE digunakan untuk membuat data pelatihan lebih seimbang, sehingga model dapat memprediksi dengan lebih akurat. Penelitian ini juga melakukan pengoptimalan parameter model melalui *RandomizedSearchCV* untuk mendapatkan hasil terbaik. Setelah melakukan evaluasi, model XGBoost dengan SMOTE menunjukkan hasil yang sangat baik, dengan akurasi sebesar 0,9345, presisi sebesar 0,7067, recall sebesar 1,0000, F1-Score sebesar 0,8281, dan AUC-ROC sebesar 0,9610. Hasil *cross-validation* juga menunjukkan bahwa model ini sangat stabil dan dapat digeneralisasi dengan baik. Analisis lebih lanjut menunjukkan bahwa fitur *"inactive_3_months_flag"* adalah prediktor terpenting untuk memprediksi pelanggan yang berhenti menggunakan layanan, diikuti oleh *"months_inactive"* dan *"inactive_listen_ratio"*. Ini berarti bahwa pelanggan yang tidak aktif selama tiga bulan atau lebih memiliki kemungkinan lebih tinggi untuk berhenti menggunakan layanan. Dengan memahami faktor-faktor ini, Spotify dapat mengambil langkah untuk mencegah pelanggan berhenti menggunakan layanan mereka dan meningkatkan kepuasan pelanggan secara keseluruhan.

Kata Kunci: Churn Prediksi, XGBoost, SMOTE, Imbalanced Data, Spotify, Machine Learning, Perilaku Streaming

Abstract

Competition in the music streaming service industry is fierce, necessitating a better understanding of consumer behavior particularly regarding why customers churn for platforms like Spotify. This study focuses on developing a model to predict which customers are likely to stop using Spotify. To build this model, the study employs the Extreme Gradient Boosting (XGBoost) algorithm combined with the Synthetic Minority Over-sampling Technique (SMOTE). SMOTE is used to address the issue of data imbalance, where the number of customers who churn is significantly lower than those who remain. The dataset, sourced from Kaggle, comprises 50,000 user records with 18 behavioral attributes such as weekly listening duration, song-skipping frequency, number of playlists created, subscription type, and demographic information. The research process involves data understanding, target definition, data cleaning, feature engineering, and splitting the data into training and testing sets. SMOTE is applied to balance the training data, thereby enhancing the model's predictive accuracy. The study also optimizes model parameters using RandomizedSearchCV to achieve optimal results. Evaluation reveals that the XGBoost model with SMOTE performs exceptionally well, achieving an accuracy of 0.9345, precision of 0.7067, recall of 1.0000, F1-Score of 0.8281, and AUC-ROC of 0.9610. Cross-validation results further demonstrate the model's stability and strong generalization capabilities. Subsequent analysis identifies the "inactive_3_months_flag" feature as the most critical predictor of customer churn, followed by "months_inactive" and "inactive_listen_ratio." This indicates that customers inactive for three months or longer are more likely to discontinue the service. By understanding these factors, Spotify can implement measures to prevent churn and improve overall customer satisfaction.

***Keywords**: Churn Prediction, XGBoost, SMOTE, Imbalanced Data, Spotify, Machine Learning, Streaming Behavior*

1. PENDAHULUAN

Industri musik streaming di Indonesia dan dunia telah mengalami pertumbuhan yang sangat pesat. Platform seperti Spotify, Apple Music, dan Joox bersaing sangat ketat untuk memperebutkan pangsa pasar yang terus berkembang. Dalam persaingan yang semakin ketat ini, kemampuan untuk mempertahankan pelanggan yang sudah ada menjadi jauh lebih vital dibandingkan menambah pelanggan baru. Keadaan ini

e-ISSN: 3110-0864 | p-ISSN: xxxx-xxxx
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | Page: 252-262
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.163

mendorong urgensi untuk memiliki sistem prediksi *churn* yang terpercaya sebagai landasan dalam pengambilan keputusan strategis perusahaan.

Customer churn dapat didefinisikan sebagai kondisi di mana pelanggan memutuskan untuk berhenti menggunakan layanan atau tidak untuk langganan mereka dalam periode tertentu [1]. Kehilangan pelanggan dapat mempengaruhi secara langsung pendapatan perusahaan karena biaya untuk mendapatkan pelanggan baru jauh lebih tinggi dibandingkan dengan menjaga pelanggan yang sudah ada [2]. Oleh karena itu, kemampuan memprediksi *churn* lebih awal menjadi sangat krusial bagi keberlangsungan bisnis *platform streaming*.

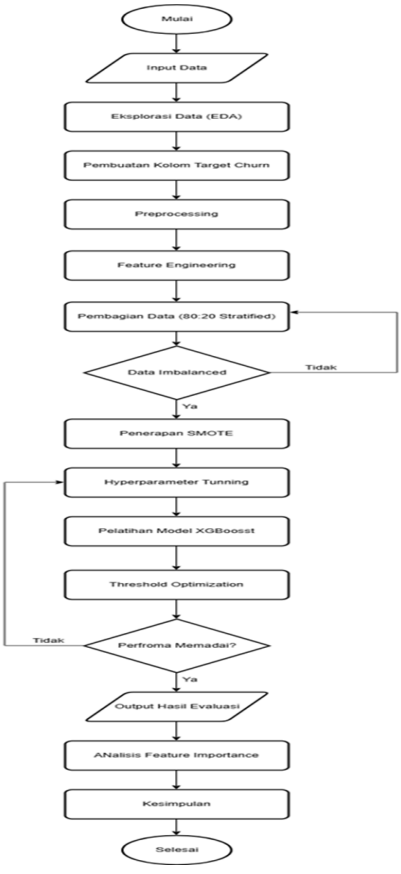
Pendekatan *machine learning* telah banyak diaplikasikan dalam prediksi *churn* telah banyak diaplikasikan dalam prediksi *churn* di berbagai industri. Holilurrahman dan Imron dalam penelitiannya telah menerapkan algoritma *Random Forest* pada prediksi *churn* di industri telekomunikasi Indonesia dan berhasil mencapai akurasi sebesar 95% [3]. Namun, tantangan utama yang sering dihadapi dalam data churn adalah ketidakseimbangan kelas, di mana jumlah pelanggan yang tidak churn jauh lebih banyak dibandingkan dengan pelanggan yang churn. Kondisi ini membuat model lebih condong pada kelas mayoritas, sehingga kemampuannya dalam mendeteksi churn menjadi rendah [4].

XGBoost *(Extreme Gradient Boosting)* adalah algoritma *ensemble* yang memanfaatkan pohon keputusan sebagai dasar yang dikenal memiliki performa tinggi pada berbagai kasus klasifikasi [5]. Yulianti dkk. menunjukkan bahwa XGBoost menghasilkan hasil klasifikasi yang sangat optimal pada data pelanggan kartu kredit [6]. Di sisi lain, SMOTE adalah teknik yang sering diterapkan untuk mengatasi data yang tidak seimbang melalui pembuatan sampel sintetik pada kelas yang kurang [4]. Dalam penelitiannya, Suryana dkk. Membuktikan bahwa kombinasi SMOTE dengan algoritma *boosting* mampu meningkatkan performa prediksi *churn* secara signifikan [7]

Penelitian ini kami mengusulkan kombinasi XGBoost dan SMOTE dalam membangun model prediksi *churn* pelanggan Spotify yang akurat dan andal, khususnya dalam skenario data perilaku *streming* yang bersifat *imbalanced*. Tujuan dari penelitian ini adalah untuk mengembangkan model prediksi churn dengan menggunakan XGBoost, menerapkan SMOTE untuk mengatasi *imbalanced* data, serta mengevaluasi peningkatan performa odel setelah penerapan SMOTE.

2. METODOLOGI PENELITIAN

Tahapan penelitian dilaksanakan secara berurutan dan terstruktur. Secara keseluruhan alur penelitian ini terdiri dari dua belas tahapan utama yang digambarkan pada Gambar 1.



Gambar 1. Alur Tahapan Penelitian Prediksi Churn Pelanggan Spotify

e-ISSN: 3110-0864 | p-ISSN: xxxx-xxxx
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | Page: 252-262
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.163

2.1 Dataset

Dataset yang digunakan dalam penelitian ini adalah *Spotify User Behavior Dataset* yang tersedia secara publik melalui platform *Kaggle*. Dataset ini memuat 50.000 data pengguna dengan 18 atribut sebagaimana ditampilkan pada Tabel 1. Variabel target adalah kolom *subscription_status* yang dikodekan menjadi variabel biner yaitu 1 untuk pelanggan *Inactive (churn)* dan 0 untuk pelanggan *Active* (tidak churn). Dataset menunjukkan karakteristik *imbalanced* dengan distribusi 84,2% pelanggan aktif (42.109 sampel) dan 15,8% pelanggan churn (7.891 sampel), menghasilkan rasio ketidakseimbangan sebesar 5,34:1.

Tabel 1. Deskripsi Atribut Dataset Spotify

No	Nama Atribut	Tipe	Keterangan
1	*user_id*	Numerik	Identitas unik pengguna
2	*country*	Kategorik	Negara asal pengguna
3	*age*	Numerik	Usia Pengguna
4	*signup_date*	Tanggal	Tanggal registrasi pengguna
5	*subscription_type*	Kategorik	Jenis langganan
6	*subscription_status*	Kategorik	Status langganan (target)
7	*months_inactive*	Numerik	Lama tidak aktif (bulan)
8	*inactive_3_months_flag*	Biner	Indikator tidak aktif ≥ 3 bulan
9	*ad_interaction*	Kategorik	Interaksi terhadap iklan
10	*ad_conversion_to_subscription*	Kategorik	Konversi iklan menjadi langganan
11	*music_suggestion_rating_1_to_5*	Numerik	Penilaian rekomendasi musik (1-5)
12	*avg_listening_hours_per_week*	Numerik	Rata-rata jam mendengarkan per minggu
13	*favorite_genre*	Kategorik	Genre musik favorit
14	*most_liked_feature*	Kategorik	Fitur yang paling disukai
15	*desired_future_feature*	Kategorik	Fitur yang diharapkan pada masa mendatang
16	*primary_device*	Kategorik	Perangkat utama yang digunakan
17	*playlists_created*	Numerik	Jumlah playlist yang dibuat
18	*avg_skips_per_day*	Numerik	Rata-rata jumlah lagu

2.2 Tahapan Penelitian

Tahapan penelitian ini dilaksanakan melalui 8 langkah utama yang akan dijelaskan sebagai berikut:

1. *Pengumpulan Data*
   Dataset diperoleh dari Kaggle dalam format CSV, kemudian dimuat ke lingkungan Python menggunakan pustaka Pandas untuk proses pengolahan data.
2. *Eksplorasi Data (EDA)*
   Eksplorasi data dilakukan untuk memahami karakteristik dataset. Tahapan ini mencakup analisis distribusi data, identifikasi *missing value*, analisis korelasi antar atribut, serta visualisasi hubungan antara fitur dan variabel target churn [8].
3. *Preprocessing*
   Tahap preprocessing meliputi penghapusan atribut yang tidak relevan, yaitu *user_id*, penanganan *missing value* menggunakan median pada atribut numerik dan modus pada atribut kategorik, rekayasa fitur dari atribut *signup_date* menjadi *tenure_months*, *signup_year*, dan *signup_month*, serta penerapan *label encoding* pada seluruh atribut kategorik [9].
4. *Pembagian Data*
   Dataset dibagi menjadi data latih sebesar 80% dan data uji sebesar 20% menggunakan metode *stratified split* untuk menjaga proporsi kelas pada masing-masing subset data [10].
5. *Normalisasi*
   Normalisasi atribut numerik dilakukan menggunakan metode *StandardScaler [11]*. Proses *fitting* hanya diterapkan pada data latih guna mencegah terjadinya *data leakage* pada data uji.

e-ISSN: 3110-0864 | p-ISSN: xxxx-xxxx
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | Page: 252-262
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.163

6. *Penerapan SMOTE*
   Metode *Synthetic Minority Over-sampling Technique* (SMOTE) diterapkan pada data latih dengan nilai *k neighbors* sebesar 5. Teknik ini digunakan untuk menghasilkan sampel sintetis pada kelas minoritas sehingga distribusi kelas menjadi lebih seimbang [4].
7. *Pelatihan Model*
   Model XGBoost dilatih menggunakan data latih yang telah diseimbangkan melalui SMOTE [12]. Parameter yang digunakan dalam proses pelatihan ditunjukkan pada Tabel 3.
8. *Evaluasi*
   Evaluasi model dilakukan menggunakan data uji dengan memanfaatkan matrix Accuracy, Precision, Recall, F1-Score, dan AUC-ROC. Selain itu, dilakukan pula *5-Fold Stratified Cross Validation* untuk mengukur kestabilan dan konsistensi performa model [13].

2.3 Feature Engineering

Selain 18 fitur asli dari dataset, tiga fitur interaksi baru ditambahkan untuk memperkaya representasi perilaku pengguna sebagaimana pada Tabel 2.

Tabel 2. Fitur Interaksi Baru

Nama Atribut	Formula	Makna
skip_per_hour	*avg_skips_per_day / avg_listening_hours_per_week*	Tingkat skip relatif terhadap durasi mendengarkan
engagement_score	*(listening × 0,4) + (playlist × 0,3) + (rating × 0,3)*	Skor keterlibatan pengguna secara keseluruhan
inactive_listen_ratio	*months_inactive × avg_listening_hours_per_week*	Interaksi antara ketidakaktifan dan kebiasaan mendengarkan

Selain itu, kolom *signup_date* diekstrak menjadi tiga fitur turunan: *tenure_months* (lama berlangganan dalam bulan), *signup_year*, dan *signup_month*. Total fitur akhir yang digunakan dalam pelatihan model berjumlah 21 fitur.

2.4 Hyperparameter XGBoost

Hyperparameter terbaik yang diperoleh dari *RandomizedSearchCV* dengan optimasi F1-Score disajikan pada Tabel 3.

Tabel 3. Hyperparameter Model XGBoost

Nama Atribut	Nilai
n_estimators	500
max_depth	5
learning_rate	0,01
subsample	0,9
colsample_hytree	0,7
min_child_weight	3
gamma	0
reg_alpha	0,5
reg_lambda	1,0

2.5 Metrik Evaluasi

Kinerja model dievaluasi menggunakan beberapa metrik yang umum digunakan pada permasalahan klasifikasi dengan distribusi kelas yang tidak seimbang [14].

1. *Accuracy*
   Digunakan untuk mengukur ketepatan model dalam memprediksi pelanggan yang mengalami churn.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

2. *Precision*
   Digunakan untuk mengukur ketepatan model dalam memprediksi pelanggan yang mengalami churn.

e-ISSN: 3110-0864 | p-ISSN: xxxx-xxxx
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | Page: 252-262
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.163

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

3. *Recall*
   Digunakan untuk menilai kemampuan model dalam mengidentifikasi seluruh pelanggan yang mengalami churn.
$$Recall = \frac{TP}{TP + FN} \quad (3)$$
4. *FI-Score*
   Merupakan rata-rata harmonik antara nilai Precision dan Recall.
$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$
5. *AUC-ROC*
   Digunakan untuk mengukur kemampuan model dalam membedakan kelas positif dan kelas negatif pada berbagai nilai ambang (*threshold*).

Keterangan:
TP = *True Positive*, yaitu jumlah data churn yang dipredikси churn.
TN = *True Negative*, yaitu jumlah data aktif yang diprediksi aktif.
FP = *False Positive*, yaitu jumlah data aktif yang diprediksi churn.
FN = *False Negative*, yaitu jumlah data churn yang diprediksi aktif.

3. HASIL DAN PEMBAHASAN

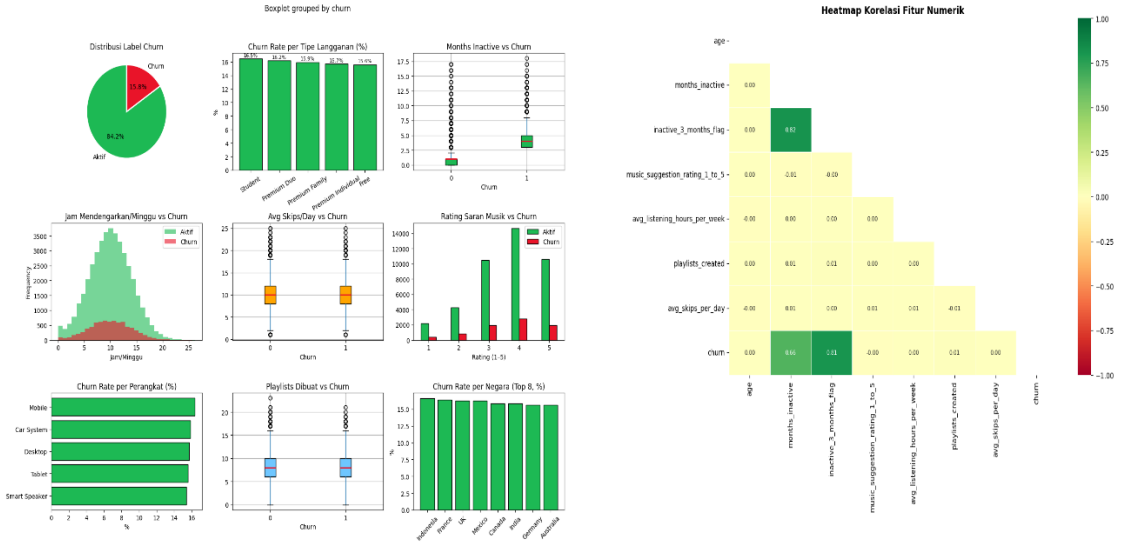
3.1 Analisis Eksplorasi Data

Hasil analisis eksploratif menunjukkan bahwa dataset yang digunakan memiliki distribusi kelas yang tidak seimbang *(imbalanced).* Jumlah pelanggan yang masih aktif tercatat jauh lebih besar dibandingkan jumlah pelanggan yang telah melakukan *churn.* Kondisi ini sejalan dengan temuan penelitian yang dilakukan oleh Suryana dkk., yang menyatakan bahwa pada kasus prediksi *customer churn*, proporsi pelanggan yang menghentikan layanan umumnya lebih kecil dibandingkan pelanggan yang tetap menggunakan layanan. Perbedaan jumlah data antar kelas tersebut berpotensi menimbulkan permasalahan *imbalanced data* yang dapat mempengaruhi kinerja model klasifikasi dalam mengenali kelas minoritas secara optimal [7].

Berdasarkan hasil analisis korelasi pada Gambar 3, atribut *months inactive* memiliki hubungan positif paling kuat terhadap status churn. Hal ini menunjukkan bahwa semakin lama pelanggan tidak menggunakan layanan, semakin besar kemungkinan pelanggan tersebut berhenti berlangganan. Sebaliknya, atribut *avg_listening_hours_per_week* memiliki korelasi negatif yang cukup tinggi, yang mengidentifikasikan bahwa pelanggan dengan intensitas mendengarkan musik yang lebih tinggi cenderung memiliki loyalitas yang lebih baik terhadap layanan Spotify. Selain itu, atribut *playlist_created* juga menunjukkan korelasi negatif terhadap churn, sehingga aktivitas pembuatan playlist dapat dianggap sebagai salah satu indikator keterikatan pengguna terhadap platform.

Jika ditinjau berdasarkan jenis langganan, pelanggan dengan paket *Free* memiliki tingkat churn yang paling tinggi. Sementara itu, pengguna dengan paket *Family* dan *Individual* menunjukkan tingkat churn yang lebih rendah karena adanya komitmen finansial yang lebih besar terhadap layanan yang digunakan.

e-ISSN: 3110-0864 | p-ISSN: xxxx-xxxx
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | Page: 252-262
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.163



(a) (b)

Gambar 2. (a) Hasil EDA dan (b) Hasil Correlation Heatmap

3.2 Preprocessing dan Feature Enginering

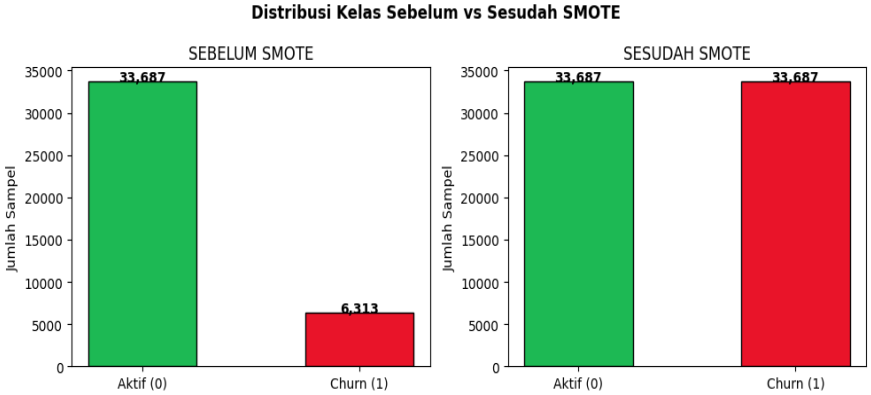
Preprocessing menghasilkan 21 fitur akhir dari 18 fitur asli setelah penambahan tiga fitur interaksi dan tiga fitur turunan dari *signup_date*. Delapan kolom kategorik berhasil dikonversi menggunakan *Label Encoding*, yaitu *country*, *subscription_type*, *ad_interaction*, *ad_conversion_to_subscription*, *favorite_genre*, *most_liked_feature*, *desired_future_feature*, dan *primary_device*. Data dibagi menjadi 40.000 sampel latih (6.313 churn, 15,8%) dan 10.000 sampel uji (1.578 churn, 15,8%), mempertahankan proporsi kelas yang konsisten melalui *stratified split.*

3.3 Hasil penerapan SMOTE

Kondisi data latih sebelum penerapan SMOTE menunjukkan adanya *class imbalance*, yang ditandai dengan jumlah sampel pada kelas minoritas yang jauh melebihi jumlah sampel pada kelas minoritas. Setelah metode SMOTE diterapkan dengan nilai *k_neighbors* sebesar 5, jumlah sampel pada kelas aktif sehingga menghasilkan rasio kelas sebesar 1:1

Proses penyeimbangan data menggunakan SMTOTE hanya dilakukan pada data latih dan tidak melibatkan data uji. Hal ini bertujuan untuk mencegah kebocoran informasi antara data pelatihan dan data pengujian sehingga evaluasi dapat mencerminkan performa model yang sebenarnya. Pendekatan tersebut sesuai dengan rekomendasi yang disampaikan oleh Chieka dan Kurniasih [15].

Distribusi kelas sebelum dan sesudah penerapan SMOTE ditunjukkan pada Gambar 4



Gambar 3. Distribusi Kelas Sebelum dan Sesudah SMOTE

3.4 Hasil Hyperparameter Tuning

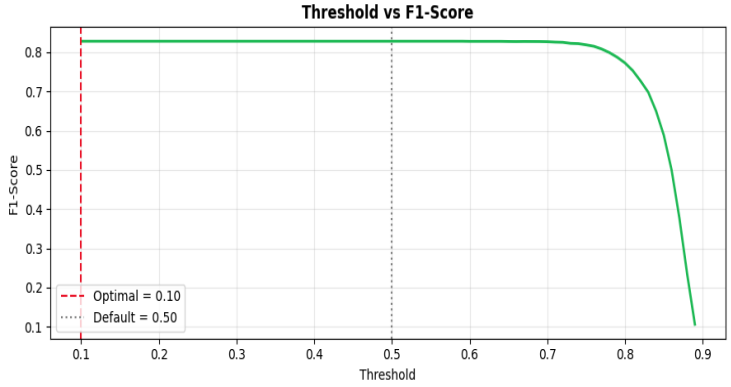
RandomizedSearchCV dengan 30 *iterasi* dan 5-*fold cross-validation* menghasilkan best CV *F1-Score* sebesar 0,9632 dengan kombinasi *hyperparameter* terbaik yaitu, *n_estimators*=500, *max_depth*=5, *learning_rate*=0,01, *subsample*=0,9, *colsample_bytree*=0,7, *min_child_weight*=3, *gamma*=0, *reg_alpha*=0,5,

e-ISSN: 3110-0864 | p-ISSN: xxxx-xxxx
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | Page: 252-262
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.163

dan r*eg_lambda*=1,0. Nilai *learning_rate* yang rendah (0,01) dikombinasikan dengan jumlah estimator yang banyak (500) menghasilkan model yang belajar secara bertahap dan hati-hati, meminimalkan risiko overfitting.

3.5 Hasil Threshold Optimization

Optimasi *threshold* menggunakan pencarian *F1-Score* pada rentang 0,10 hingga 0,90 menghasilkan threshold optimal sebesar 0,10. Nilai threshold yang rendah ini menunjukkan bahwa model cenderung memberikan probabilitas prediksi yang rendah namun konsisten untuk kelas churn setelah penerapan SMOTE, sehingga *threshold* perlu diturunkan agar deteksi *churn* menjadi maksimal. Dengan *threshold* 0,10, model berhasil mencapai *Recall* sebesar 1,0000, artinya seluruh pelanggan yang benar-benar churn berhasil terdeteksi. Hasil dari *threshold* ditunjukkan pada Gambar 5



Gambar 4. Threshold Optimization

3.6 Hasil Evaluasi Model

Perbandingan performa model XGBoost sebelum dan sesudah penerapan SMOTE pada data uji disajikan pada Tabel 4.

Tabel 4. Perbandingan Performa Model Pada Data Uji (N=10.000)

Matrix	XGBoots Tanpa SMOTE (thr=0,50)	XGBoots + SMOTE (thr=0,50)	XGBoots + SMOTE S (thr=0,10 optimal)
Accuracy	0,9338	0,9345	0,9345
Precision	0,7069	0,9345	0,7067
Recall	0,7069	1,0000	1,0000
F1-Score	0,7069	0,8281	0,8281
AUC-ROC	0,9605	0,9610	0,9610

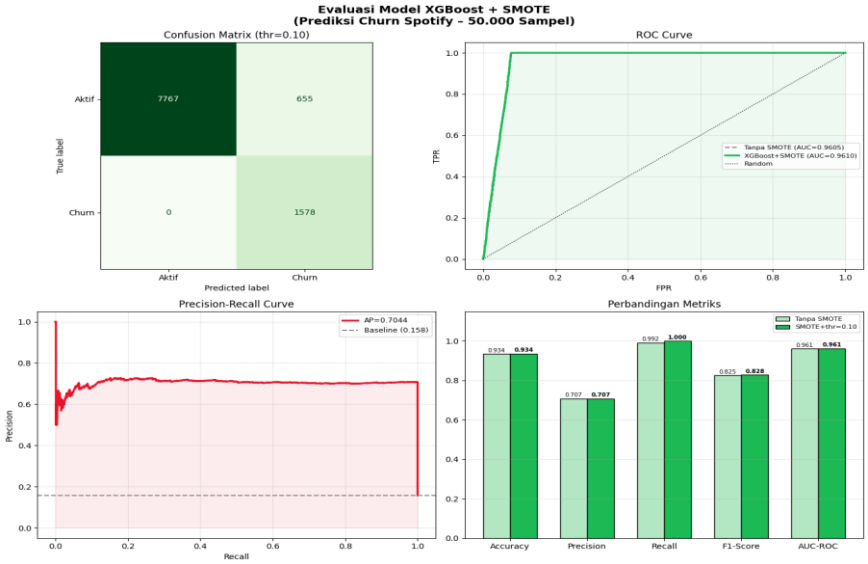
Hasil menunjukkan bahwa model *XGBoost + SMOTE* dengan *threshold* optimal mencapai *Recall* sempurna sebesar 1,0000, artinya tidak ada satu pun pelanggan yang akan churn yang luput dari deteksi model. Nilai *AUC-ROC* sebesar 0,9610 mengindikasikan kemampuan diskriminasi model yang sangat baik. Nilai *Accuracy* sebesar 0,9345 menunjukkan bahwa 93,45% prediksi keseluruhan adalah benar. *Precision* sebesar 0,7067 berarti dari seluruh pelanggan yang diprediksi churn, 70,67% di antaranya memang benar-benar churn.

Classification report menunjukkan bahwa untuk kelas Aktif (0), model mencapai *Precision* 1,00 dan *Recall* 0,92 dengan *F1-Score* 0,96. Untuk kelas Churn (1), model mencapai *Precision* 0,71 dan *Recall* 1,00 dengan *F1-Score* 0,83. *Macro average F1-Score* mencapai 0,89 dan *weighted average F1-Score* mencapai 0,94, mencerminkan performa yang sangat baik secara keseluruhan.

Temuan ini sejalan dengan penelitian Suryana dkk. yang menyatakan bahwa kombinasi SMOTE dan boosting secara konsisten meningkatkan performa prediksi churn dibandingkan model tanpa penyeimbangan data [7]. Dibandingkan dengan model baseline tanpa SMOTE, penerapan SMOTE berhasil meningkatkan Recall dari 0,9918 menjadi 1,0000 dan AUC-ROC dari 0,9605 menjadi 0,9610.

Hasil evaluasi pengujian model ditampilkan pada gambar 6

e-ISSN: 3110-0864 | p-ISSN: xxxx-xxxx
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | Page: 252-262
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.163



Gambar 5. Hasil Evaluasi Model XGBoost+SMOTE

3.7 Hasil Cross-validation

Tabel 5. Hasil Cross-Validation 5-Fold

Nama Atribut	Formula	Makna
AUC-ROC	*0,9924*	±0,0006
F1-Score	*0,9632*	±0,0019
Recall	*1,0000*	±0,0000

Nilai standar deviasi yang sangat kecil pada seluruh metrik mengindikasikan bahwa model bersifat sangat stabil dan konsisten. Nilai AUC-ROC rata-rata sebesar 0,9924 pada *cross-validation* menunjukkan kemampuan generalisasi model yang sangat tinggi. *Recall* sempurna (1,0000 ± 0,0000) di seluruh fold membuktikan bahwa model secara konsisten mampu mendeteksi seluruh kasus churn pada setiap subset data latih [7].

3.8 Analisis Feature Importance

Analisis feature importance yang diperoleh berdasarkan nilai *gain score* pada algoritma XGBoost menunjukkan bahwa atribut *inactive_3_months_flag* merupakan fitur yang memberikan kontribusi terbesar dalam proses prediksi churn dengan *gain score* sebesar 0,63, jauh melampaui fitur-fitur lainnya. Dominasi fitur ini menunjukkan bahwa status tidak aktif selama tiga bulan atau lebih merupakan sinyal paling kritis yang digunakan model dalam mengidentifikasi pelanggan yang akan berhenti berlangganan.

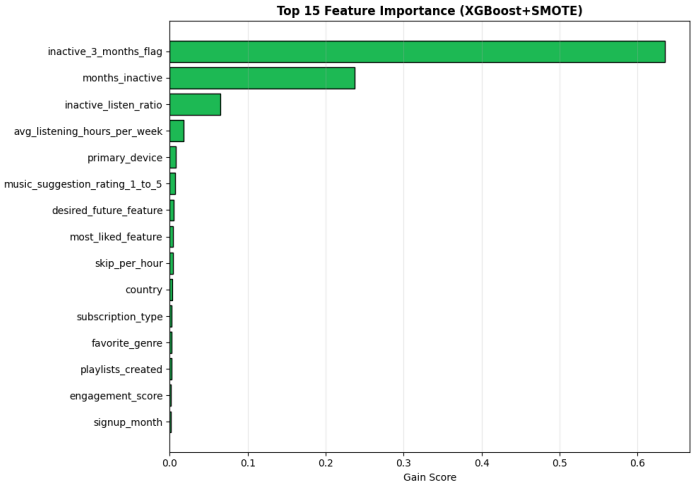
Fitur tersebut diikuti oleh *months_inactive* pada posisi kedua dengan *gain score* sekitar 0,24, dan *inactive_listen_ratio* pada posisi ketiga dengan *gain score* sekitar 0,07. Kehadiran *inactive_listen_ratio* yang merupakan fitur hasil rekayasa dari perkalian *months_inactive* dan *avg_listening_hours_per_week* di posisi ketiga memvalidasi bahwa penambahan fitur interaksi baru memberikan nilai tambah yang nyata bagi model. Secara keseluruhan, dominasi tiga fitur berbasis ketidakaktifan di posisi teratas mengkonfirmasi bahwa perilaku ketidakaktifan pengguna merupakan determinan utama churn pada platform streaming musik.

Fitur *avg_listening_hours_per_week* menempati posisi keempat, menunjukkan bahwa durasi mendengarkan musik per minggu juga berkontribusi dalam membedakan pelanggan yang loyal dari yang berisiko *churn*. Pelanggan dengan durasi mendengarkan yang rendah cenderung memiliki keterikatan yang lebih lemah terhadap platform sehingga lebih rentan untuk berhenti berlangganan.

Atribut *music_suggestion_rating_1_to_5* juga memberikan kontribusi dalam model meskipun dengan gain score yang lebih kecil. Temuan ini menunjukkan bahwa kualitas rekomendasi musik yang diterima pengguna memiliki pengaruh terhadap keputusan pelanggan untuk tetap menggunakan layanan berlangganan. Dengan demikian, sistem rekomendasi yang mampu memberikan saran musik yang relevan berpotensi meningkatkan retensi pelanggan pada platform streaming musik

Visualisasi tingkat kepentingan fitur ditampilkan pada Gambar 7

e-ISSN: 3110-0864 | p-ISSN: xxxx-xxxx
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | Page: 252-262
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.163



Gambar 6. Feature Importance Model XGBoost

Temuan tersebut sejalan dengan penelitian Mukhlason yang menyatakan bahwa fitur-fitur yang berkaitan dengan perilaku pengguna layanan memiliki pengaruh yang lebih besar terhadap prediksi churn dibandingkan dengan karakteristik demografis pengguna [1].

3.9 Pembahasan

Berdasarkan hasil pengujian, penerapan algoritma XGBoost yang dipadukan dengan metode SMOTE terbukti efektif dalam melakukan prediksi *churn* pelanggan pada dataset perilaku streaming yang mengalami permasalahan ketidakseimbangan kelas. Model yang dikembangkan berhasil mencapai nilai Recall sebesar 1,0000, *Accuracy* sebesar 0,9345, *F1-Score* sebesar 0,8281, dan AUC-ROC sebesar 0,9610 pada data uji yang terdiri dari 10.000 sampel. Peningkatan nilai *Recall* yang signifikan setelah penerapan SMOTE menunjukkan bahwa masalah imbalanced data perlu ditangani secara khusus sebelum proses pelatihan model dilakukan [4].

Menurut Siringoringo, penerapan SMOTE memungkinkan model untuk mempelajari pola-pola pada kelas minoritas secara lebih baik, sehingga peluang kelas tersebut untuk teridentifikasi menjadi lebih tinggi dibandingkan ketika menggunakan data yang tidak diseimbangkan [4]. Hal ini terbukti pada penelitian ini, di mana SMOTE berhasil menyeimbangkan rasio kelas dari 5,34:1 menjadi 1:1 pada data latih, yang secara langsung berdampak pada kemampuan model mendeteksi seluruh pelanggan yang benar-benar akan churn tanpa ada yang terlewat.

Jika dibandingkan dengan penelitian Wardana dkk. yang menggunakan algoritma XGBoost tanpa teknik penyeimbangan data pada kasus *churn* pelanggan Indibiz [16], model yang dikembangkan pada penelitian ini menunjukkan peningkatan performa yang lebih baik terutama pada metrik *Recall* setelah penerapan SMOTE. Perbandingan antara model XGBoost tanpa SMOTE dan dengan SMOTE pada penelitian ini juga menunjukkan perbedaan yang nyata, di mana penerapan SMOTE meningkatkan nilai *Recall* dari 0,9918 menjadi 1,0000 dan AUC-ROC dari 0,9605 menjadi 0,9610. Hasil tersebut menegaskan bahwa proses penanganan *imbalanced* data merupakan tahapan *preprocessing* yang sangat penting, khususnya pada kasus prediksi churn, karena kelas minoritas merupakan kelas yang paling bernilai untuk dideteksi secara akurat dari sudut pandang bisnis [7].

4. KESIMPULAN

Churn pelanggan Spotify menggunakan algoritma XGBoost berbasis SMOTE untuk mengatasi permasalahan ketidakseimbangan kelas pada data perilaku *streaming*. Berdasarkan hasil eksperimen dan analisis yang telah dilakukan terhadap dataset sebanyak 50.000 sampel pengguna dengan 21 fitur, dapat ditarik tiga kesimpulan utama. Pertama, penerapan SMOTE terbukti efektif dalam menangani permasalahan imbalanced data dengan menyeimbangkan *rasio* kelas dari 5,34:1 menjadi 1:1 pada data latih. Kondisi ini secara langsung meningkatkan kemampuan model dalam mendeteksi pelanggan yang berisiko churn, dibuktikan dengan tercapainya nilai *Recall* sempurna sebesar 1,0000 yang berarti seluruh pelanggan yang akan berhenti berlangganan berhasil teridentifikasi oleh model tanpa ada yang terlewat. Kedua, model XGBoost yang dioptimalkan menggunakan *RandomizedSearchCV* dengan 30 *iterasi* menghasilkan *hyperparameter* terbaik dengan best *CV F1-Score* sebesar 0,9632. Pada data uji, model mencapai nilai *Accuracy* sebesar 0,9345, *Precision* sebesar 0,7067, *Recall* sebesar 1,0000, *F1-Score* sebesar 0,8281, dan AUC-ROC sebesar 0,9610. Hasil *cross-validation* 5-fold menunjukkan rata-rata AUC-ROC sebesar 0,9924 ± 0,0006 dengan standar deviasi yang sangat kecil, mengindikasikan bahwa model bersifat stabil dan memiliki kemampuan generalisasi

e-ISSN: 3110-0864 | p-ISSN: xxxx-xxxx
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | Page: 252-262
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.163

yang sangat baik terhadap data baru. Ketiga, analisis *feature importance* mengungkapkan bahwa fitur *inactive_3_months_flag* dengan *gain score* sebesar 0,63 merupakan prediktor churn yang paling dominan, diikuti oleh *months_inactive* dan *inactive_listen_ratio*. Temuan ini memberikan wawasan berharga bagi Spotify bahwa strategi retensi pelanggan sebaiknya diprioritaskan pada pemantauan perilaku ketidakaktifan pengguna sebagai sinyal peringatan dini *churn*. Penelitian selanjutnya diharapkan dapat memperluas kajian dengan membandingkan efektivitas metode SMOTE terhadap teknik penyeimbangan data lainnya, seperti ADASYN dan *Borderline-SMOTE*. Selain itu, penggunaan metode interpretabilitas, seperti SHAP, dapat dipertimbangkan untuk meningkatkan transparansi dan pemahaman terhadap keputusan yang dihasilkan model. Penelitian berikutnya juga disarankan untuk menganalisis keseimbangan antara *Precision* dan *Recall* pada berbagai nilai *threshold* agar model dapat dioptimalkan sesuai dengan tujuan dan kebutuhan bisnis yang spesifik.

REFERENCES

[1] M. Marcellina and A. Mukhlason, "Analisis Prediktif Churn untuk Meningkatkan Tingkat Retensi Pelanggan pada Perusahaan SaaS Menggunakan Machine Learning," *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, vol. 6, no. 1, pp. 21–32, Apr. 2024, doi: 10.28926/ilkomnika.v6i1.598.

[2] M. Rizki Kurniawan *et al.*, "PREDIKSI CUSTOMER CHURN PADA PERUSAHAAN TELEKOMUNIKASI MENGGUNAKAN ALGORITMA C4.5 BERBASIS PARTICLE SWARM OPTIMIZATION," 2023.

[3] M. Imron and P. Korespondensi, "Implementasi Model Prediksi Churn Pelanggan Menggunakan Algoritma Random Forest Pada Website Industri Telekomunikasi," 2025.

[4] R. Siringoringo, "KLASIFIKASI DATA TIDAK SEIMBANG MENGGUNAKAN ALGORITMA SMOTE DAN k-NEAREST NEIGHBOR," 2018.

[5] Nurlaelatul Maulidah, "Prediksi Peningkatan Jumlah Nasabah Deposito Berjangka Menggunakan Algoritma KNN, Decision Tree, Random Forest Dan Xgboost," vol. vol.13, no.2, Aug. 2023, doi: 10.22441/incomtech.v13i3.16921.

[6] S. Yulianti, O. Soesanto, and Y. Sukmawaty, "Penerapan Metode Extreme Gradient Boosting (XGBOOST) pada Klasifikasi Nasabah Kartu Kredit," *Journal of Mathematics Theory and Application*, vol. 4, pp. 21–26, Aug. 2022, doi: 10.31605/jomta.v4i1.1792.

[7] Suryana Nana, Pratiwi Pratiwi, and Prasetio Tri Rizki, "Penanganan Ketidakseimbangan Data pada Prediksi Customer Churn Menggunakan Kombinasi SMOTE dan Boosting," *LPPM Universitas Bina Sarana Informatika*, vol. Vol 6, No.1, 2021.

[8] Digital Skola Content Team, "Apa itu Exploratory Data Analysis (EDA): Definisi dan Teknik Visualisasi Data." Accessed: Jul. 06, 2026. [Online]. Available: https://digitalskola.com/blog/data-science/exploratory-data-analysis-eda

[9] Sckit Learn, "Preprocessing data." Accessed: Jul. 06, 2026. [Online]. Available: https://scikit-learn.org/stable/modules/preprocessing.html

[10] Chopra Abhishek, "Stratified Split: Why Data Splitting is Crucial in Machine Learning." Accessed: Jul. 06, 2026. [Online]. Available: https://www.linkedin.com/pulse/ml-6-stratified-split-why-data-splitting-crucial-machine-chopra-zdvmc

[11] "StandardScaler, MinMaxScaler and RobustScaler techniques - ML." Accessed: Jul. 06, 2026. [Online]. Available: https://www.geeksforgeeks.org/machine-learning/standardscaler-minmaxscaler-and-robustscaler-techniques-ml/

[12] XGBoosting, "XGBoost for Imbalanced Classification with SMOTE", Accessed: Jul. 06, 2026. [Online]. Available: https://xgboosting.com/xgboost-for-imbalanced-classification-with-smote/

[13] Shalev Ofir, "Recall, Precision, F1, ROC, AUC, and everything."

e-ISSN: 3110-0864 | p-ISSN: xxxx-xxxx
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | Page: 252-262
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.163

[14] N. Agian, S. Dinata, G. Abdurrahman, and N. Q. Fitriyah, "Perbandingan Optimasi Algoritma Random Forest Menggunakan Teknik Boosting Terhadap Kasus Klasifikasi Churn Pelanggan Di Industri Telekomunikasi." [Online]. Available: www.kaggle.com.

[15] C. Mulia *et al.*, "Teknik SMOTE untuk Mengatasi Imbalance Class dalam Klasifikasi Bank Customer Churn Menggunakan Algoritma Naïve bayes dan Logistic Regression," 2023.

[16] A. Wardana, W. Rahayu, and K. Mustaqim, "Model Prediksi Churn Pelanggan Indibiz Menggunakan Regresi Logistik dan Extreme Gradient Boosting (XGBoost)," *Jurnal Informatika Polinema*, vol. 12, pp. 425–430, May 2026, doi: 10.33795/jip.v12i3.9377.