

Klasifikasi *Customer Churn* Menggunakan XGBoost dengan Optimasi *GridSearchCV* Berbasis *Shapley Additive Explanations*

I Gusti Ayu Riyana Astarani^{a1}, Luh Arida Ayu Rahning Putri^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹astarani.2308561031@student.unud.ac.id
²rahningputri@unud.ac.id

Abstract

Customer churn is a significant challenge in the banking sector, often leading to revenue loss and requiring predictive strategies to enhance customer retention. This study implements the Extreme Gradient Boosting (XGBoost) algorithm for churn classification, with hyperparameter optimization using the GridSearchCV technique to improve model performance. The dataset comprises 10,000 banking customers with 9 features and 1 target label. To address class imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was applied. Prior to tuning, the XGBoost model achieved an accuracy of 80.8%. After applying optimal parameters, the model's performance improved to 81.5%, along with higher precision and recall values, indicating improved robustness and consistency. For model interpretability, Shapley Additive Explanations (SHAP) were used and visualized through a beeswarm Plot. The analysis identified age, customer activity status, and number of products owned as the most influential features in predicting churn. Based on these findings, this study proposes business recommendations including age-based customer segmentation, enhancing active customer engagement, and optimizing product offerings as strategies to reduce churn.

Keywords: *Customer churn, XGBoost, GridSearchCV, SHAP*

1. Pendahuluan

Perkembangan teknologi informasi telah mendorong terjadinya transformasi digital di berbagai sektor, termasuk industri perbankan yang memiliki peran strategis dalam mendukung pertumbuhan ekonomi global. Persaingan di sektor tersebut semakin meningkat seiring dengan kemunculan bank digital dan perusahaan *financial technology (fintech)* yang menawarkan layanan keuangan berbasis teknologi canggih. Salah satu dampak utama dari dinamika tersebut adalah meningkatnya potensi perpindahan nasabah ke institusi lain, baik bank maupun *fintech* [1]. Dalam dunia bisnis, fenomena ini dikenal dengan istilah *customer churn*. *Customer churn* mengacu pada kecenderungan pelanggan untuk menghentikan penggunaan produk atau layanan suatu perusahaan dan beralih ke penyedia lain [2]. Penelitian yang dilakukan oleh *Forrester Research*, menunjukkan bahwa tingkat *churn* yang tinggi dapat menimbulkan kerugian finansial yang signifikan, bahkan mencapai penurunan pendapatan sebesar 25% [3]. Selain itu, biaya akuisisi pelanggan baru untuk iklan, marketing, dan lain-lain jauh lebih tinggi dibandingkan biaya mempertahankan pelanggan lama, yang umumnya lebih menguntungkan secara jangka panjang [4]. Oleh karena itu, mempertahankan pelanggan menjadi prioritas strategis bagi institusi perbankan.

Identifikasi dini terhadap nasabah berisiko *churn* melalui metode prediktif diperlukan guna mendukung strategi yang efektif. Penerapan *machine learning* telah terbukti efektif dalam memodelkan prediksi *customer churn* melalui analisis pola perilaku nasabah. Salah satu algoritma yang banyak digunakan adalah XGBoost, yang memiliki kemampuan unggul dalam menyelesaikan berbagai tugas klasifikasi, regresi, dan peringkat [5]. Berdasarkan penelitian yang

dilakukan oleh [6], menunjukkan bahwa XGBoost memiliki kinerja prediktif yang lebih baik dibandingkan metode lain seperti *K-Nearest Neighbors* (K-NN) dan *Random Forest*, dengan tingkat akurasi sebesar 79,8% dan nilai *F-score* sebesar 58,2%.

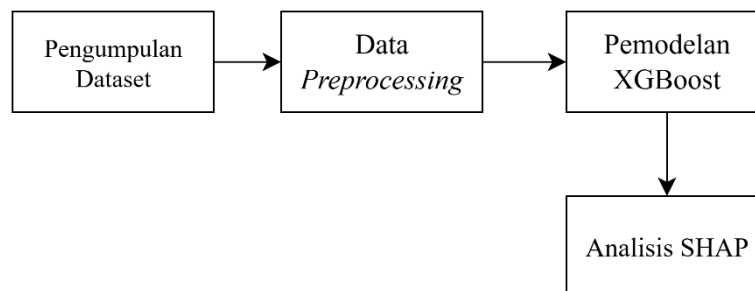
Namun demikian, meskipun algoritma *machine learning* seperti XGBoost menunjukkan performa yang tinggi, tantangan utama yang sering dihadapi adalah rendahnya tingkat interpretabilitas model. Hal ini menjadi perhatian penting di sektor perbankan yang sangat menekankan transparansi dan akuntabilitas dalam proses pengambilan keputusan. Oleh karena itu penelitian ini mengimplementasikan pendekatan SHAP. SHAP merupakan metode interpretabilitas model berbasis teori permainan yang digunakan untuk menjelaskan kontribusi relatif setiap fitur terhadap hasil prediksi. SHAP mampu menjelaskan kontribusi relatif setiap fitur terhadap hasil prediksi model, sehingga meningkatkan pemahaman terhadap proses pengambilan keputusan oleh model secara keseluruhan.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk melakukan klasifikasi *customer churn* di sektor perbankan dengan menggunakan algoritma XGBoost. Selain mengevaluasi kinerja prediksi, penelitian ini juga mengeksplorasi interpretabilitas model menggunakan pendekatan SHAP. Dengan demikian, penelitian ini tidak hanya berfokus pada peningkatan akurasi, tetapi juga pada pemahaman komprehensif terhadap variabel-variabel yang secara signifikan berkontribusi terhadap keputusan *churn*. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi nyata dalam mendukung pengambilan keputusan strategis dalam manajemen relasi nasabah.

2. Metode Penelitian

2.1 Alur dan Desain Penelitian

Penelitian ini dilakukan secara sistematis untuk memperoleh model prediksi *customer churn* menggunakan algoritma XGBoost, serta analisis interpretabilitas berbasis SHAP. Tahapan penelitian secara umum mencakup pengumpulan dataset, data *preprocessing*, pemodelan XGBoost dan interpretasi model menggunakan SHAP. Alur penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

2.2 Pengumpulan Data

Dataset yang digunakan dalam penelitian ini diperoleh dari *platform* kaggle [7] dan mencakup 10.000 entri data nasabah. Dataset tersebut terdiri atas 14 fitur, namun dalam penelitian ini hanya dimanfaatkan 9 atribut utama serta 1 kelas target yang dinilai relevan untuk membangun model prediksi *customer churn* pada sektor perbankan. Rincian atribut yang digunakan dalam penelitian disajikan pada Tabel 1 dan 2.

Tabel 1. Atribut

No	Atribut	Deskripsi
1	<i>CreditScore</i>	Skor kredit pelanggan

No	Atribut	Deskripsi
2	<i>Geography</i>	Negara asal pelanggan (Jerman, Spanyol, Prancis)
3	<i>Gender</i>	Jenis kelamin pelanggan (<i>Male</i> atau <i>Female</i>)
4	<i>Age</i>	Usia pelanggan dalam tahun
5	<i>Tenure</i>	Lama menjadi pelanggan dalam tahun
6	<i>Balance</i>	Saldo rekening pelanggan
7	<i>NumOfProducts</i>	Jumlah produk bank yang digunakan pelanggan
8	<i>HasCrCard</i>	Kepemilikan kartu kredit (1= Ya, 0= Tidak)
9	<i>IsActiveMember</i>	Status keaktifan pelanggan (1= Aktif, 0= Tidak)

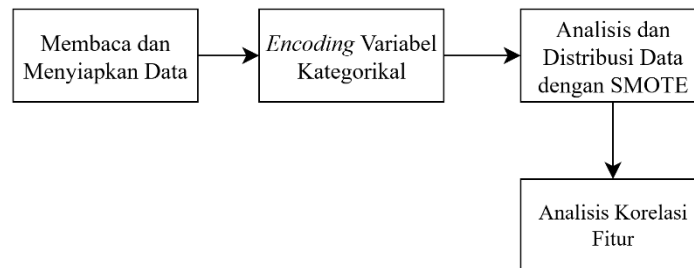
Tabel 2. Kelas Target

No	Atribut	Deskripsi
1	<i>Exited</i>	Status <i>churn</i> pelanggan (1= Keluar, 0= Bertahan).

Atribut yang digunakan dalam penelitian ini meliputi 9 atribut yaitu *CreditScore*, *Geography*, *Gender*, *Age*, *Tenure*, *Balance*, *NumOfProducts*, *HasCrCard*, *IsActiveMember*, serta *Exited* sebagai kelas target. Pemilihan variabel tersebut didasarkan pada penelitian [8] yang menunjukkan bahwa atribut-atribut tersebut secara signifikan memengaruhi perilaku *churn* nasabah di sektor perbankan.

2.3 Preprocessing Data

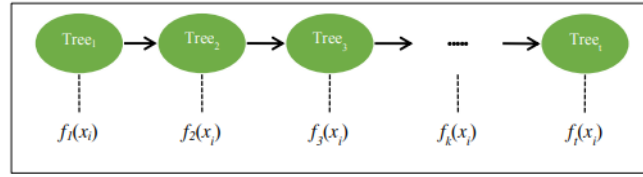
Tahapan *preprocessing* data merupakan langkah awal sebelum proses pelatihan model, yang bertujuan untuk meningkatkan kualitas data agar dapat diolah secara optimal oleh algoritma *machine learning*. Rangkaian tahapan *preprocessing* data dalam penelitian ini disajikan pada Gambar 2.



Gambar 2. Tahapan *Preprocessing* Data

2.4 Pelatihan Model

Pada proses pelatihan model, penelitian ini menggunakan algoritma XGBoost. XGBoost merupakan algoritma *machine learning* yang tergolong dalam metode *ensemble learning* dan dikenal karena efisiensi serta kecepatannya dalam pemrosesan data [9]. Algoritma ini bekerja berdasarkan prinsip *gradient boosting* secara iteratif, di mana setiap pohon keputusan yang dibangun berfokus pada perbaikan kesalahan dari pohon sebelumnya. Proses komputasi dari algoritma XGBoost dapat dilihat pada Gambar 3.



Gambar 3. Skema dari Algoritma XGBoost

Model ini mengoptimalkan fungsi *loss* dengan tujuan meminimalkan kesalahan prediksi melalui perhitungan gradien *loss* pada setiap iterasi menggunakan *binary logistic loss*, yang dinyatakan dalam Persamaan 1:

$$Loss = - \sum_{i=1}^n (y_i \log(p_i) + (1 - y_i) \log(1 - p_i)) \quad (1)$$

Dalam rumus tersebut, y_i merupakan label aktual untuk observasi ke- i sedangkan p_i merupakan probabilitas hasil prediksi model bahwa observasi ke- i termasuk dalam kelas positif. Fungsi ini menghitung selisih antara nilai aktual dan prediksi probabilitas model, di mana nilai *loss* yang lebih rendah menunjukkan prediksi yang lebih akurat.

Dalam metode XGBoost, prediksi untuk setiap sampel dihitung sebagai penjumlahan dari hasil keluaran pohon keputusan yang dibangun pada setiap iterasi, yang dapat dirumuskan dengan persamaan 2:

$$\hat{y}_i = \sum_{k=1}^K f_k(X_i) \quad (2)$$

Berdasarkan persamaan tersebut, prediksi akhir untuk observasi ke- i yang dilambangkan dengan \hat{y}_i diperoleh dari penjumlahan seluruh *output* dari fungsi prediksi $f_k(X_i)$ yang dihasilkan oleh setiap pohon keputusan pada iterasi ke k hingga ke K .

Untuk mengoptimalkan kinerja model berdasarkan formulasi tersebut, model dilatih menggunakan dataset yang telah melalui tahap *preprocessing*, dengan parameter-parameter yang dioptimalkan melalui teknik *grid search cross-validation* (*GridSearchCV*). Metode *GridSearchCV* merupakan pendekatan sistematis dalam menentukan kombinasi *hyperparameter* terbaik dengan mengeksplorasi berbagai nilai parameter secara menyeluruh [10]. Hal ini bertujuan untuk menemukan konfigurasi parameter yang memberikan performa optimal pada model. Untuk memastikan kemampuan generalisasi model terhadap data yang belum pernah dilihat sebelumnya, dataset dibagi menjadi dua subset, yaitu data latih dan data uji, dengan rasio perbandingan 80:20. Selanjutnya, validasi dilakukan menggunakan *k-fold cross-validation* pada data latih untuk meningkatkan keandalan estimasi performa dan membantu pemilihan *hyperparameter* optimal.

2.4 Evaluasi Model

Evaluasi kinerja model dilakukan untuk menilai efektivitas algoritma dalam melakukan klasifikasi data secara akurat. Dalam penelitian ini, beberapa metrik evaluasi yang digunakan meliputi *accuracy*, *precision*, *recall*, dan *F1-score* dalam mengukur aspek performa model secara spesifik.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

2.5 Interpretasi Model dengan SHAP

Untuk memperoleh pemahaman yang lebih mendalam mengenai kontribusi masing-masing variabel terhadap prediksi model, penelitian ini mengadopsi metode interpretabilitas berbasis SHAP. SHAP merupakan pendekatan berbasis teori permainan yang mengkuantifikasi kontribusi fitur terhadap hasil, sehingga memberikan analisis transparan terhadap faktor-faktor yang memengaruhi keputusan model [11].

Dalam analisis SHAP, beragam visualisasi digunakan untuk mengidentifikasi pola hubungan antara fitur dan hasil prediksi model. Penelitian ini menggunakan SHAP *beeswarm plot*, yang merupakan jenis visualisasi yang menampilkan distribusi nilai dalam bentuk titik-titik yang tersebar di sepanjang sumbu, tanpa saling bertumpuk, menyerupai kawanan lebah (*beeswarm*) yang menyajikan distribusi kontribusi nilai SHAP dari seluruh fitur secara menyeluruh untuk setiap sampel.

Visualisasi ini memudahkan identifikasi fitur-fitur paling berpengaruh serta pola nilai fitur yang berdampak positif atau negatif terhadap prediksi model sehingga dapat memberikan interpretasi yang akurat dan komprehensif yang berkontribusi signifikan dalam memprediksi *churn* pelanggan.

3. Hasil dan Diskusi

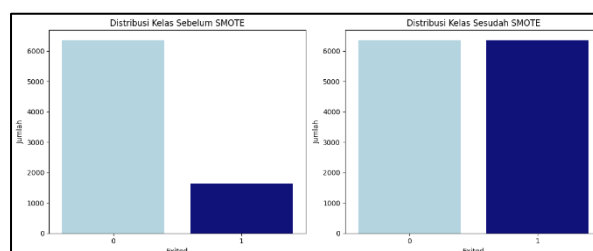
3.1. Preprocessing Data

a. Encoding Variabel Kategorikal

Pada tahap awal *preprocessing* data, variabel kategorikal dikonversi ke dalam representasi numerik agar kompatibel dengan algoritma yang digunakan. Atribut *Gender*, yang terdiri atas dua kategori (*Male* dan *Female*), ditransformasikan menggunakan metode *Label Encoding*, dengan nilai '*Female*' direpresentasikan sebagai 0 dan '*Male*' sebagai 1. Sementara itu, atribut *Geography* yang memiliki tiga kategori (*France*, *Germany*, dan *Spain*) diencoding menggunakan teknik *One-Hot Encoding* untuk menghindari asumsi ordinalitas antar kategori. Dalam penerapannya, digunakan parameter *drop_first=True* untuk mencegah terjadinya *dummy variable trap*, yakni kondisi redundansi informasi antar fitur yang dapat memengaruhi kestabilan estimasi parameter dalam model. Dengan pengaturan ini, hanya dua fitur baru yang dihasilkan yaitu, *Geography_Germany* dan *Geography_Spain*. Kategori '*France*' direpresentasikan secara implisit ketika kedua fitur tersebut bernilai 0.

b. Analisis Distribusi Data

Berdasarkan hasil analisis data, kelas target *Exited* menunjukkan distribusi kelas yang tidak seimbang, di mana jumlah pelanggan yang tidak melakukan *churn* (kelas mayoritas) secara signifikan lebih tinggi dibandingkan dengan pelanggan yang mengalami *churn* (kelas minoritas). Ketidakseimbangan ini dapat mengarah pada bias dalam proses pelatihan model prediktif, di mana algoritma cenderung memprioritaskan akurasi terhadap kelas mayoritas dan mengabaikan performa klasifikasi pada kelas minoritas. Ketidakseimbangan tersebut dapat diamati pada Gambar 4.

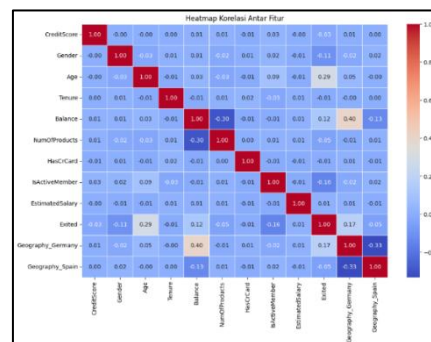


Gambar 4. Distribusi Kelas Sebelum dan Setelah SMOTE

Untuk mengatasi permasalahan ketidakseimbangan kelas pada variabel target, diterapkan teknik *Synthetic Minority Over-sampling Technique (SMOTE)*. SMOTE berfungsi untuk meningkatkan representasi kelas minoritas dengan cara menghasilkan sampel sintesis baru berdasarkan interpolasi antara sampel-sampel yang sudah ada. Setelah penerapan SMOTE, sebagaimana ditampilkan pada Gambar 4, yang menyajikan perbandingan distribusi kelas sebelum dan sesudah penerapan SMOTE, terlihat bahwa distribusi awal menunjukkan dominasi signifikan pada kelas mayoritas (*non-churn*), sedangkan kelas minoritas (*churn*) jauh lebih sedikit. Setelah penerapan SMOTE, distribusi kelas menjadi lebih proporsional, dengan jumlah observasi pada kelas minoritas yang meningkat hingga mendekati jumlah pada kelas mayoritas.

c. Analisis Korelasi Fitur

Analisis korelasi pada dataset setelah dilakukan proses *oversampling* bertujuan untuk mengevaluasi sejauh mana hubungan antar fitur memengaruhi probabilitas terjadinya *churn*. Dalam studi ini, korelasi antar variabel dihitung menggunakan koefisien *pearson* dan divisualisasikan melalui *correlation matrix heatmap*. Visualisasi tersebut disajikan pada Gambar 5.



Gambar 5. Heatmap Korelasi Antar Fitur

Berdasarkan hasil visualisasi, beberapa korelasi positif yang signifikan ditemukan antara lain hubungan positif yang cukup kuat antara *Balance* dan variabel target *Exited*, yang mengindikasikan bahwa nasabah dengan saldo lebih tinggi memiliki kecenderungan yang lebih besar untuk melakukan *churn*. Hasil tersebut menunjukkan bahwa nasabah bernilai tinggi secara finansial justru lebih rentan meninggalkan layanan, yang dapat memberikan implikasi strategis bagi *customer relationship management*. Selain itu, variabel *Age* juga menunjukkan korelasi positif terhadap *Exited*, mengindikasikan bahwa aspek demografis seperti usia turut berkontribusi dalam menentukan risiko *churn*. Hal ini menggarisbawahi pentingnya pendekatan segmentasi usia dalam strategi retensi. Sementara itu, korelasi negatif antara *NumOfProducts* dan *Balance* menunjukkan bahwa nasabah dengan saldo tinggi cenderung menggunakan lebih sedikit produk perbankan, yang dapat menjadi sinyal lemahnya keterikatan (*engagement*) dengan institusi keuangan.

3.2 Evaluasi Model dan Interpretasi Hasil

Evaluasi terhadap performa model dilakukan untuk mengukur sejauh mana algoritma XGBoost mampu memprediksi *churn* nasabah secara akurat. Tahap ini dilakukan dalam dua bagian yaitu, sebelum dan sesudah penerapan hyperparameter tuning. Pengukuran performa model didasarkan pada metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Pada tahap awal, model XGBoost diimplementasikan dengan parameter *default*. Hasil evaluasi performa model disajikan pada Tabel 3.

Tabel 3. Hasil Evaluasi Model Tanpa *Hyperparameter Tuning*

Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Accuracy</i>
0	0.90	0.85	0.88	0.80
1	0.52	0.62	0.57	0.80

Berdasarkan hasil pada Tabel 3, model XGBoost memiliki tingkat akurasi sebesar 80,8%. *Precision* tinggi pada kelas 0 menunjukkan bahwa model sangat baik dalam mengidentifikasi nasabah yang tidak *churn*. Namun, pada kelas 1 (nasabah *churn*), *precision* dan *recall* masih tergolong rendah, masing-masing 0.52 dan 0.62. Hal ini menunjukkan bahwa model cenderung kurang optimal dalam mengenali nasabah yang benar-benar akan *churn*, sehingga diperlukan penyempurnaan parameter untuk meningkatkan performa prediksi pada kelas minoritas tersebut.

Sebagai upaya peningkatan performa model, dilakukan proses *hyperparameter tuning* menggunakan teknik *GridSearchCV* dengan *3-fold cross-validation*. *GridSearchCV* mengevaluasi total 243 kombinasi parameter, namun efektifnya sebanyak 216 kombinasi valid. Hasil nilai *hyperparameter* terbaik disajikan pada Tabel 4.

Tabel 4. Ruang Pencarian *Hyperparameter*

<i>Hyperparameter</i>	<i>Grid Search Values</i>	Nilai <i>Hyperparameter</i> Terbaik
<i>max_depth</i>	3, 5, 7	7
<i>learning_rate</i>	0.01, 0.1, 0.2	0.2
<i>n_estimators</i>	50, 100, 150	150
<i>subsample</i>	0.7, 0.8, 1.0	1.0
<i>colsample_bytree</i>	0.7, 0.8, 1.0	1.0

Adapun parameter dan nilai parameter terbaik yang dapat meningkatkan kinerja algoritma dapat dilihat pada Tabel 4. 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 tabel 5.

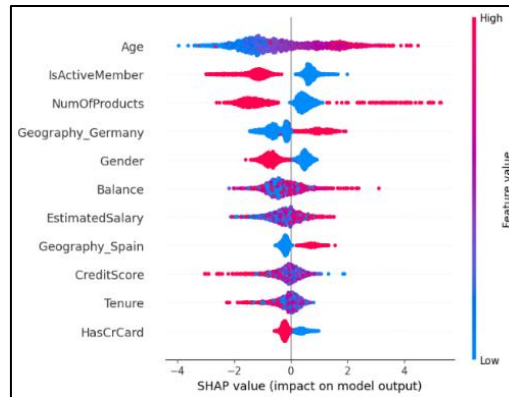
Tabel 5. Hasil Evaluasi Setelah *Hyperparameter Tuning*

Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Accuracy</i>
0	0.90	0.87	0.88	0.81
1	0.54	0.61	0.57	0.81

Setelah penerapan parameter optimal, performa model mengalami sedikit peningkatan. Akurasi naik menjadi 81,5% dari sebelumnya 80,8%. *Precision* pada kelas 1 meningkat dari 0.52 menjadi 0.54, menunjukkan model menjadi lebih akurat dalam mengidentifikasi nasabah *churn*. Meskipun *recall* pada kelas 1 sedikit menurun dari 0.62 menjadi 0.61, namun *f1-score* tetap stabil pada nilai 0.57, yang mengindikasikan adanya *trade-off* yang seimbang antara *precision* dan *recall*. Berdasarkan hasil tersebut, penerapan *hyperparameter tuning* menghasilkan model yang lebih konsisten serta memiliki ketahanan yang lebih baik dalam memprediksi *churn* sehingga dapat disimpulkan bahwa *hyperparameter tuning* terbukti mampu meningkatkan kinerja algoritma dalam proses klasifikasi. Hal ini sejalan dengan penelitian [12], yang menyatakan bahwa *hyperparameter tuning* merupakan tahapan yang direkomendasikan sebelum melakukan proses klasifikasi, karena dapat meningkatkan kinerja model secara signifikan.

3.3. Hasil Analisis SHAP

Dalam memahami kontribusi setiap fitur terhadap prediksi model XGBoost, dilakukan analisis interpretabilitas menggunakan metode SHAP dengan memanfaatkan visualisasi *beeswarm plot*.



Gambar 6. Visualisasi *Beeswarm Plot*

Gambar 6 menyajikan visualisasi nilai SHAP menggunakan *beeswarm plot* untuk seluruh fitur yang digunakan dalam model. Sumbu vertikal menunjukkan fitur-fitur yang diurutkan berdasarkan kontribusi rata-rata terhadap *output* prediksi, sementara sumbu horizontal merepresentasikan nilai SHAP, yang mengindikasikan besarnya dan arah pengaruh masing-masing fitur terhadap probabilitas *churn*.

Setiap titik dalam plot mewakili satu data observasi (nasabah), dengan warna merah menunjukkan nilai fitur yang tinggi, dan warna biru menunjukkan nilai fitur yang rendah. Nilai-nilai SHAP pada sumbu horizontal (misalnya dari -4 hingga 4) menunjukkan besarnya kontribusi suatu fitur terhadap prediksi untuk individu tertentu. Apabila nilai SHAP bernilai 0, maka fitur tersebut tidak memberikan pengaruh terhadap keputusan model. $SHAP > 0$ artinya fitur tersebut mendorong model untuk memprediksi kelas positif (dalam hal ini, *churn* atau keluar), sedangkan $SHAP < 0$ menunjukkan bahwa fitur mengurangi kemungkinan *churn* dan mendorong prediksi ke arah kelas negatif (tetap menjadi nasabah).

Berdasarkan hasil visualisasi, terdapat tiga fitur utama yang memberikan kontribusi paling signifikan terhadap keputusan model XGBoost dalam memprediksi *churn*, yakni *Age*, *NumOfProducts*, dan *IsActiveMember*. Analisis terhadap ketiga fitur tersebut dijelaskan sebagai berikut:

- Fitur *Age* menempati posisi tertinggi sebagai prediktor *churn*. Hasil analisis SHAP menunjukkan bahwa nasabah dengan usia yang lebih tinggi (ditunjukkan oleh nilai SHAP positif yang besar berwarna merah) memiliki kecenderungan lebih tinggi untuk meninggalkan layanan perbankan. Hasil ini sejalan dengan analisis korelasi yang menemukan hubungan positif antara *Age* dan *Exited*, menguatkan bahwa faktor demografis seperti usia perlu diperhatikan dalam segmentasi risiko *churn*. Strategi retensi berbasis usia sangat relevan dalam konteks ini. Berdasarkan yang dilakukan oleh [13] yang menyatakan bahwa nasabah berusia lanjut cenderung menunjukkan perilaku konservatif dan lebih cepat berpindah layanan ketika merasa tidak puas. Dengan demikian, usia menjadi indikator penting dalam segmentasi risiko *churn* dan penentuan strategi retensi yang bersifat demografis.
- Fitur *IsActiveMember* menggambarkan keterlibatan nasabah dalam aktivitas perbankan. Dari visualisasi SHAP, terlihat bahwa nasabah yang tidak aktif (bernilai rendah, warna biru) cenderung memiliki nilai SHAP positif yang tinggi, menunjukkan kontribusi besar terhadap peningkatan kemungkinan *churn*. Hal ini menunjukkan bahwa keaktifan nasabah merupakan indikator kuat dari loyalitas dan keberlanjutan hubungan pelanggan dengan

institusi perbankan. Hasil ini sejalan dengan pendekatan relationship marketing yang menekankan pentingnya interaksi berkelanjutan untuk mempertahankan pelanggan.

- c. Fitur *NumOfProducts* juga menunjukkan dampak yang kuat dalam mempengaruhi *output* model. Nasabah yang memiliki lebih sedikit produk (terlihat dari warna biru dengan nilai SHAP positif) cenderung lebih rentan mengalami *churn*. Sebaliknya, semakin banyak produk yang dimiliki nasabah (warna merah), nilai SHAP cenderung negatif, menunjukkan korelasi negatif terhadap *churn*. Hal ini mengindikasikan bahwa keterikatan nasabah terhadap bank meningkat seiring dengan jumlah produk yang digunakan sehingga penggunaan produk menjadi faktor protektif terhadap *churn*.

Hasil visualisasi SHAP *beeswarm plot* tidak hanya memperjelas fitur mana yang paling berkontribusi terhadap prediksi *churn*, tetapi juga selaras dengan hasil korelasi fitur. Konsistensi ini memperkuat validitas model serta memberikan wawasan yang mendalam untuk pengambilan keputusan berbasis data, khususnya dalam merancang strategi retensi pelanggan yang lebih efektif.

Dari hasil interpretasi ini, ketiga variabel tersebut dapat dijadikan sebagai indikator kunci dalam perumusan strategi penanganan *churn*. Bank dapat mengimplementasikan pendekatan berbasis data untuk mengidentifikasi segmen pelanggan yang rentan *churn* berdasarkan usia lanjut, minimnya penggunaan produk, dan rendahnya aktivitas. Intervensi preventif, seperti personalisasi layanan, penawaran *bundling* produk, dan promosi aktivitas, dapat diarahkan secara lebih tepat sasaran pada segmen-segmen ini guna meningkatkan retensi. Dengan memahami kontribusi fitur utama secara lebih komprehensif, strategi retensi dapat difokuskan pada peningkatan keterlibatan nasabah serta pengembangan penawaran produk yang relevan.

4. Kesimpulan

Penelitian ini berhasil mengembangkan dan mengevaluasi model prediksi *customer churn* pada sektor perbankan menggunakan algoritma XGBoost yang dioptimalkan melalui teknik *GridSearchCV*. Model awal menunjukkan performa yang cukup baik, dengan akurasi sebesar 80,8%. Pada kelas *non-churn*, model mencapai *precision* sebesar 90%, *recall* 85%, dan *f1-score* 88%. Sementara itu, pada kelas *churn*, *precision* hanya sebesar 52%, *recall* 62%, dan *f1-score* 57%. Untuk meningkatkan performa, dilakukan *hyperparameter tuning* menggunakan *grid search* yang mengevaluasi 243 kombinasi parameter dengan teknik *3-fold cross-validation*. Konfigurasi terbaik diperoleh dengan parameter *max_depth* = 7, *learning_rate* = 0.2, *n_estimators* = 150, serta *subsample* dan *colsample_bytree* masing-masing sebesar 1.0. Hasil tuning ini memberikan peningkatan performa model, dengan akurasi menjadi 81%. Pada kelas *churn*, *precision* meningkat menjadi 54%, *recall* sebesar 61%, dan *f1-score* tetap 57%. Adapun pada kelas *non-churn*, *precision* mencapai 90%, *recall* 87%, dan *f1-score* 88%. Untuk memahami kontribusi setiap fitur terhadap prediksi model, dilakukan analisis interpretabilitas menggunakan SHAP. Hasil visualisasi menunjukkan bahwa *Age*, *IsActiveMember*, dan *NumOfProducts* merupakan tiga fitur paling berpengaruh terhadap kecenderungan nasabah untuk melakukan *churn*. Secara umum, nasabah yang memiliki risiko tinggi untuk *churn* ditandai oleh usia yang lebih lanjut, kepemilikan produk yang terbatas, serta tingkat aktivitas rendah dalam menggunakan layanan perbankan. Penelitian ini memberikan dasar yang kuat bagi penyusunan strategi bisnis yang lebih tepat sasaran, seperti segmentasi nasabah berdasarkan usia, peningkatan keterlibatan nasabah aktif, serta pengembangan portofolio produk yang sesuai dengan kebutuhan pelanggan. Untuk pengembangan lebih lanjut, disarankan penerapan teknik *deep learning*, integrasi analitik secara *real-time*, serta pemanfaatan indikator ekonomi makro sebagai fitur tambahan. Pendekatan ini diharapkan dapat meningkatkan akurasi prediksi dan memperkuat relevansi model dalam menghadapi dinamika industri perbankan.

Daftar Pustaka

- [1] N. Gurung, M. R. Hasan, M. S. Gazi, and F. R. Chowdhury, "AI-Based Customer *Churn* Prediction Model for Business Markets in the USA: Exploring the Use of AI And Machine

- Learning Technologies in Preventing Customer *Churn*," *Journal of Computer Science and Technology Studies*, vol. 6, no. 2, pp. 19–29, 2024.
- [2] A. Sulisty Wibowo, "Analisis *Churn* Nasabah Bank Dengan Pendekatan *Machine Learning* Dan Pengelompokan Profil Nasabah Dengan Pendekatan Clustering," *Aritekin Jurnal*, vol. 2, no. 1, pp. 30–41, 2024.
- [3] Forrester Research, "How customer relationship management impacts *churn* rate reduction," May 2025. [Online]. Available: <https://growett.com/blogs/How-customer-relationship-management-impacts-churn-rate-reduction.html>. [Accessed: May 31, 2025].
- [4] M. Marcellina, "Analisis Prediktif *Churn* untuk Meningkatkan Tingkat Retensi Pelanggan", Institut Teknologi Sepuluh Nopember, 2024.
- [5] N. Selayanti, D. A. Putri, T. Trimono, and M. Idhom, "Prediksi harga penutupan saham BBRI dengan model hybrid LSTM-XGBoost", "*Jurnal Teknik Informatika dan Multimedia*", vol. 5, no. 1, pp. 52–64, 2025.
- [6] A. H. Primandari, "Implementasi Metode Random Forest dan XGBoost pada Klasifikasi *Customer Churn*", Universitas Islam Indonesia, 2020.
- [7] M. Zangari, *Customer Churn from a Bank*, Kaggle, 2020. [Online]. Available: <https://www.kaggle.com/datasets/murilozangari/customer-churn-from-a-bank>. [Accessed: May 31, 2025].
- [8] TA. B. Wirayuda, "*Customer Churn Prediction* Pada Streaming Musics Platform Menggunakan Ensemble Learning", *eProceedings of Engineering*, Bandung, 2025, vol. 12, pp. 1982–1989.
- [9] B. A. Maulana and N. Hidayati, "*Churn* prediction in credit customers using Random Forest and XGBoost methods", *Indonesian Journal of Data and Science*, vol. 6, no. 1, pp. 81–89, 2025.
- [10] S. C. Lu, C. L. Swisher, C. Chung, D. Jaffray and C. Sidey-Gibbons, "On the Importance Of Interpretable Machine Learning Predictions to Inform Clinical Decision Making in Oncology," *Front. Oncol*, vol. 13, p. 1129380, 2023.
- [11] F. Fakhriza, D. Subekti, and P. W. Cahyo, "Optimalisasi *Algoritma Random Forest Feature Selection* Dan *Hyperparameter Tuning* Klasifikasi Genre Musik", *JIKA (Jurnal Informatika)*, vol. 9, no. 1, pp. 16–25, 2025.
- [12] A.-G. Văduva, S.-V. Oprea, A.-M. Niculae, A. Bâra, and A.-I. Andreescu, "Improving *Churn* Detection in the Banking Sector: A Machine Learning Approach with Probability Calibration Techniques", *Electronics*, vol. 13, no. 22, p. 4527, 2024.