 E-ISSN: 2715-2731


# Analisis Prediktif *Churn* untuk Meningkatkan Tingkat Retensi Pelanggan pada Perusahaan SaaS Menggunakan *Machine Learning*

Marcellina[1,*], Ahmad Mukhlason[2]

[1]Departemen Manajemen Teknologi, Institut Teknologi Sepuluh Nopember, Indonesia
[2]Fakultas Teknologi Elektro dan Informatika Cerdas, Institut Teknologi Sepuluh Nopember, Indonesia

[1]marcellinalim98@gmail.com; [2]mukhlason@is.its.ac.id;

* penulis korespondensi

INFO ARTIKEL

**Sejarah Artikel**
Diterima: 31 Desember 2023
Direvisi: 18 April 2024
Diterbitkan: 30 April 2024

**Kata Kunci**
*Churn*
*Random Forest*
SaaS
*SMOTE*
*XGBoost*

ABSTRAK

*Software-as-a-Service* (SaaS) mengalami peningkatan minat dan pengguna, salah satunya dari industri makanan dan minuman. Tingginya penggunaan dan permintaan SaaS tentu mendorong persaingan ketat sehingga perlu bersaing mempertahankan pengguna aplikasi, salah satunya dengan mengetahui faktor yang dapat menyebabkan seorang pelanggan berhenti menggunakan (*churn)*. Dengan mengetahui faktor tersebut, perusahaan juga dapat memprediksi pelanggan rawan *churn* sehingga dapat dicegah dengan memanfaatkan *machine learning*. Oleh karena itu, dilakukan penelitian untuk mengidentifikasi faktor yang memengaruhi keputusan *churn* di PT XYZ sehingga dapat disusun strategi pencegahan. Algoritma yang akan digunakan adalah *Random Forest* dan *XGBoost*. Terdapat *class imbalance* pada data yang digunakan, sehingga akan diterapkan *SMOTE*. Variabel penelitian adalah jenis usaha, jumlah *outlet*, jumlah penggunaan produk, jumlah tiket, rata-rata tiket bulanan, dan kategori kendala yang dilaporkan klien. Hasil penelitian menunjukkan bahwa model klasifikasi *XGBoost* setelah *SMOTE* memiliki nilai performa terbaik di antara model lainnya, dengan nilai akurasi 0,71, presisi 0,27, *recall* 0,86, F1-*score* 0,41, dan AUC 0,596. *Feature importance* tertinggi adalah seringnya laporan kendala dari klien terutama terkait produk ERP dan POS, kendala penarikan laporan, jenis usaha, dan rata-rata tiket bulanan. Strategi selanjutnya, perusahan dapat memprioritaskan evaluasi kualitas produk serta pelayanan perusahaan, seperti pengenalan dan training penggunaan produk.

## PENDAHULUAN

Seiring dengan perkembangan zaman dan teknologi, perusahaan *Software-as-a-Service* (SaaS) mengalami peningkatan minat dan pengguna. Salah satu industri yang banyak memanfaatkan SaaS adalah industri makanan dan minuman. Berdasarkan laporan Profil Industri Mikro dan Kecil (IMK) tahun 2020, industri makanan merupakan jenis IMK terbanyak dengan 1,52 juta usaha. Selain itu, industri makanan juga menjadi usaha IMK terbanyak yang menggunakan internet dibandingkan industri lainnya. Pemanfaatan internet oleh IMK sendiri sebagian besar terkait pemasaran/penjualan produk, promosi/iklan, dan pembelian bahan baku [1]. Dengan tingginya jumlah IMK yang memanfaatkan internet, berbagai penyedia layanan SaaS bersaing untuk memenuhi kebutuhan usaha para pelaku bisnis.

PT XYZ adalah sebuah perusahaan SaaS yang menawarkan ekosistem teknologi terintegrasi untuk mendukung kebutuhan operasional bisnis F&B, mulai dari proses

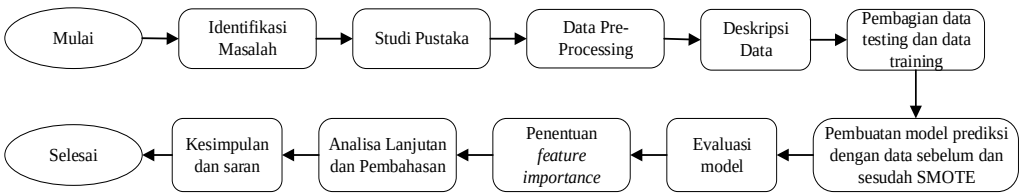
pasokan bahan baku, pemesanan makanan, pembayaran, hingga pemantauan performa bisnis. Beberapa aplikasi yang disediakan oleh PT XYZ adalah aplikasi *enterprise resource planning* (ERP) dan *point of sales* (POS). PT XYZ sendiri memiliki target pasar utama pelaku usaha yang berorientasi *Business to Consumer*, seperti restoran, warung makan, dan kafe. Meskipun PT XYZ terus berkembang, perlu dilakukan tinjauan mendalam untuk menganalisa performa perusahaan secara keseluruhan mengingat betapa pesatnya persaingan dan perubahan dapat terjadi di bidang teknologi.

Maraknya penggunaan dan permintaan SaaS mendorong banyak penyedia untuk bersaing. Agar suatu perusahaan dapat bertahan dengan kondisi persaingan yang ketat, perusahaan perlu mempertahankan pengguna aplikasi. Salah satu cara untuk mempertahankan pelanggan adalah dengan mengetahui faktor apa saja yang dapat menyebabkan seorang pelanggan berhenti menggunakan atau yang disebut sebagai *churn*. Dengan mengetahui faktor tersebut, perusahaan juga dapat memprediksi apakah seorang pelanggan akan *churn* sehingga dapat dicegah dengan memanfaatkan metode *machine learning*. Rata-rata tingkat *churn* di industri adalah sekitar 5%, dan bahkan dapat lebih tinggi untuk perusahan SaaS berskala besar [2].

Salah satu algoritma *machine learning* yang paling banyak digunakan adalah *Random Forest* karena dapat memberikan hasil klasifikasi dengan tingkat akurasi yang cenderung tinggi. Selain *Random Forest*, terdapat algoritma lain seiring dengan pengembangan lebih lanjut dan dibuatnya algoritma baru. *XGBoost* merupakan salah satu algoritma yang telah digunakan dalam beberapa penelitian dan memberikan hasil prediksi *churn* yang lebih baik dibandingkan *Random Forest*, pada penelitian yang dilakukan tentang prediksi *churn* di perusahaan B2B [3] dan penelitian prediksi churn di perusahaan telekomunikasi [4]. Oleh karena itu kedua algoritma berikut digunakan dalam penelitian ini dan diambil model dengan hasil terbaik. Meningkatnya persaingan dan keragaman pasar SaaS menyebabkan kesulitan untuk mencapai konsensus dalam menentukan faktor penting yang dapat menyebabkan terjadinya *churn* di perusahaan SaaS [5]. Oleh karena itu, dengan mempertimbangkan banyaknya jumlah IMK makanan dan minuman, studi ini akan meneliti perusahaan SaaS yang menawarkan layanan dengan target pasar IMK makanan dan minuman.

## METODE

### Tahapan Penelitian




Gambar 1. Tahapan Penelitian

Berdasarkan Gambar 1, Penelitian dimulai dari identifikasi masalah dan pengumpulan data dari PT XYZ. Objek penelitian yang diambil adalah seluruh klien yang menggunakan produk aplikasi POS dan/atau ERP, dengan sampel klien yang mulai penggunaan dari Januari 2021 hingga Januari 2023. Penelitian dilanjutkan dengan studi pustaka dan pemilihan metode penelitian yang akan digunakan. Data diproses terlebih dahulu untuk memastikan kelengkapan data, *missing value*, dan data *outlier*. Dari hasil pengolahan awal, jumlah data yang akan digunakan dalam penelitian sebanyak 1.563

dan diperiksa apakah terdapat *class imbalance*. Tahapan selanjutnya adalah analisis deskriptif dan pembagian data *testing* dan data *training* dengan proporsi 30% dan 70%. Data *training* digunakan untuk membuat model prediktif dengan algoritma *Random Forest* dan *XGBoost* dengan dua kondisi data, yaitu menggunakan semua variabel dan menggunakan variabel hasil *feature selection*. Apabila terdeteksi *class imbalance* di tahapan sebelumnya, akan diterapkan *SMOTE* dan kemudian data *training* setelah *SMOTE* akan digunakan untuk membuat model prediktif dengan algoritma dan kondisi data yang sama. Seluruh model prediktif yang dihasilkan akan dievaluasi berdasarkan nilai *confusion matrix* dan nilai AUC. Model dengan performa terbaik akan digunakan lebih lanjut untuk penentuan *feature importance*. Variabel dengan *importance* tertinggi dianalisa lebih lanjut dan diajukan rekomendasi langkah untuk PT XYZ. Penelitian diakhiri dengan penarikan kesimpulan.

## Variabel Penelitian

Tabel 1 berisi seluruh variabel yang digunakan dalam penelitian dan pembangunan model. Variabel prediktor ditentukan dari hasil penyesuaian terhadap ketersediaan data di PT XYZ, seperti kualitas pelayanan yang juga salah satu faktor yang signifikan di perusahaan telekomunikasi [6] dan di perusahaan manajemen data [7]. Dari hasil penyesuaian data yang tersedia, faktor kualitas pelayanan diteliti berdasarkan produk yang digunakan serta data terkait tiket pelaporan (jumlah tiket, tiket bulanan, dan isu tiket) untuk merepresentasikan kualitas produk. Variabel respon adalah status klien apakah *churn* atau tidak *churn*, dengan klien yang dianggap *churn* adalah yang tercatat memiliki data tanggal pemberhentian penggunaan disertai dengan alasan berhenti.

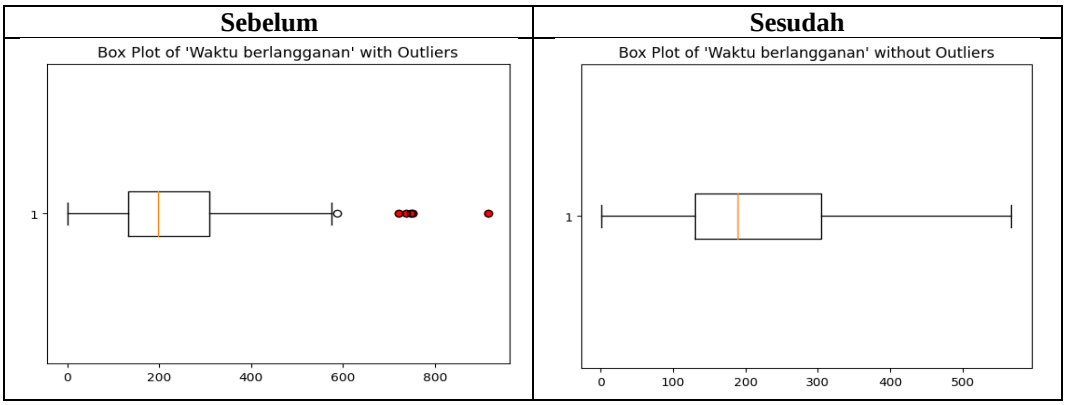
Tabel 1. Daftar Variabel Penelitian

| No | Variabel | Tipe Variabel | Keterangan | Kategori |
|---|---|---|---|---|
| Variabel Prediktor | | | | |
| 1 | Jenis usaha | Nominal | Kategori jenis usaha klien, yaitu perorangan atau perusahaan | UMKM, Enterprise |
| 2 | Jumlah *outlet* | Rasio | Jumlah *outlet* klien terdaftar | - |
| 3 | ERP *usage* | Rasio | Jumlah akun ERP yang digunakan klien | - |
| 4 | POS *usage* | Rasio | Jumlah *station* POS yang digunakan klien | - |
| 5 | Jumlah tiket | Rasio | Jumlah tiket pelaporan/keluhan dari klien secara keseluruhan | - |
| 6 | Tiket bulanan | Rasio | Rata-rata tiket pelaporan/keluhan dari klien setiap bulan | - |
| 7 | Isu tiket | Nominal | Jenis tiket pelaporan/keluhan yang paling banyak diterima dari klien | ERP *Issue*, *General Question*, *Install* POS, *Master* POS, POS *Issue*, dan *None* |
| Variabel Respon | | | | |
| 1 | *Churn*/Tidak *Churn* | Nominal | Apakah klien mengalami *churn*/tidak *churn* | Churn, Tidak Churn |

## HASIL DAN PEMBAHASAN

### Data *Pre-Processing*

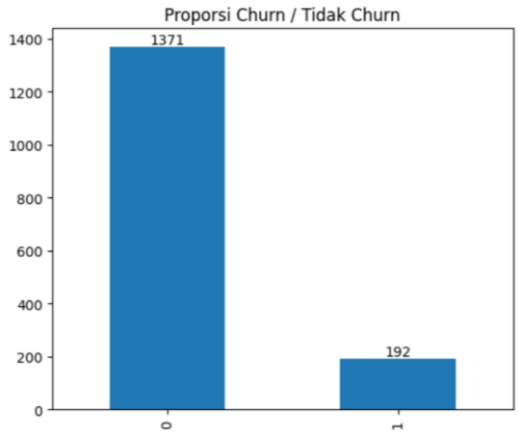
Data *pre-processing* dimulai untuk mendeteksi dan memperbaiki data agar sesuai dengan kebutuhan penelitian. Dari hasil deteksi *missing value*, tidak ditemukan adanya data dengan *missing value*. Langkah berikut adalah deteksi data yang dianggap tidak *valid* atau *outlier*. Pengecekan akan dilakukan pada data waktu berlangganan untuk klien yang mengalami *churn* saja, karena pencatatan data di PT XYZ yang sempat mengalami perubahan, sehingga ada terjadinya kesalahan input atau pemindahan data. Ditemukan sebanyak 7 data klien *churn* yang memiliki waktu berlangganan kurang dari sama dengan 0 hari. Selanjutnya dilakukan uji *outlier* dengan hasil yang ditunjukkan di Gambar 2.




Gambar 2. *Boxplot* sebelum dan sesudah data *outlier* dihilangkan

Selanjutnya dilakukan *one hot encoding* untuk transformasi data kategorikal menjadi numerik biner untuk merepresentasikan keberadaan fitur tersebut pada sampel data. Variabel jenis usaha terbagi menjadi *Enterprise* dan UMKM, sedangkan variabel isu tiket terbagi berdasarkan pengkategorian isu tiket, yaitu ERP *Issue*, *General Question*, *Install* POS, *Master* POS, POS *Issue*, dan *None* untuk klien yang tidak memiliki riwayat pengajuan keluhan/tiket. Dari proses *one hot encoding*, jumlah variabel prediktor menjadi sebanyak 15 *feature*.

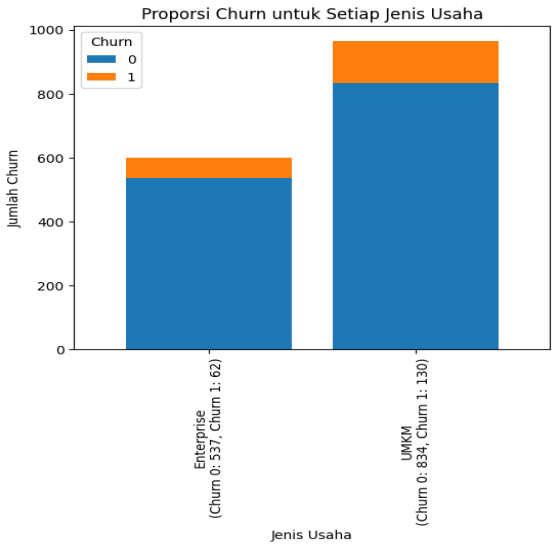
Pengecekan terakhir dilakukan untuk memeriksa apakah data yang digunakan memiliki *class imbalance*, di mana perbandingan antara dua kelas variabel respon memiliki perbedaan yang sangat signifikan. Gambar 3 menunjukkan dari 1.563 klien, 192 mengalami *churn*, yang berarti 12,284% dari klien PT XYZ mengalami *churn*. Perbandingan ini menunjukkan adanya ketidakseimbangan kelas, sehingga akan diimplementasikan *SMOTE* untuk membuat proporsi data klien yang *churn* dan tidak *churn* menjadi seimbang.




Gambar 3. Proporsi Klien *churn*

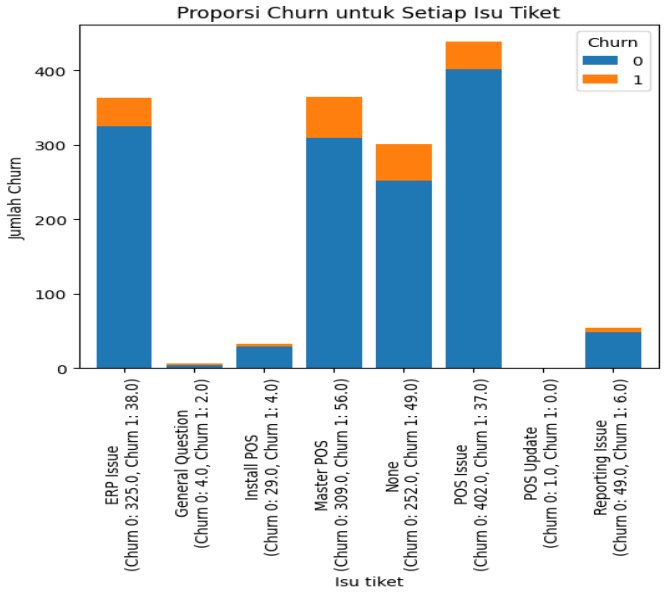
**Analisis Deskriptif**

Berdasarkan jenis usaha klien, PT XYZ memiliki lebih banyak klien dengan jenis usaha UMKM, yaitu sebanyak 964 klien. Pada Gambar 4, dapat dilihat klien dengan jenis usaha UMKM memiliki proporsi klien *churn* sebesar 130 dari 964 klien atau 13,485%, di mana persentase tersebut lebih tinggi dibandingkan klien dengan jenis usaha *Enterprise* dengan 62 dari 599 klien (10,351%).




Gambar 4. Proporsi Klien *Churn* Berdasarkan Jenis Usaha

Variabel isu tiket menunjukkan kategori isu tiket yang paling banyak dilaporkan oleh klien. Dari Gambar 5, dapat dilihat bahwa isu tiket terbanyak yang dilaporkan oleh klien adalah permasalahan terkait produk POS yaitu sebanyak 439 klien. Isu tiket terbanyak berikutnya adalah isu terkait *Master* POS dan permasalahan terkait produk ERP. Kategori dengan proporsi klien *churn* terbesar adalah isu tiket *None*, yaitu sebesar 16,279%. Selanjutnya, kategori isu tiket dengan proporsi klien *churn* terbesar adalah *Master* POS sebesar 15,342% dan ERP *issue* sebesar 10,468%.




Gambar 5. Proporsi Klien *Churn* Berdasarkan Isu Tiket

### *Feature Selection*

Proses *feature selection* dilakukan untuk memilih variabel atau *feature* yang memiliki hubungan kuat dengan variabel respon. Proses ini dilakukan untuk meningkatkan performa model karena semakin banyak *feature* yang dilibatkan dalam pembuatan dan pelatihan model, maka semakin tinggi risiko terjadinya *overfitting*. Metode yang digunakan dalam penelitian ini untuk *feature selec*tion adalah *Lasso* (*Least Absolute Shrinkage and Selection Operator*). Dari15 variabel prediktor yang diperoleh pada proses *one hot encoding*, terpilih 9 *feature* seperti yang dapat dilihat pada Tabel 2.

Tabel 2. Hasil *Feature Selection*

| No | *Feature* |
|---|---|
| 1 | Jenis usaha_*Enterprise* |
| 2 | Tiket bulanan |
| 3 | Isu tiket_ERP *Issue* |
| 4 | Isu tiket_*General Question* |
| 5 | Isu tiket_*Install* POS |
| 6 | Isu tiket_*None* |
| 7 | Isu tiket_POS *Issue* |
| 8 | Isu tiket_POS *Update* |
| 9 | Isu tiket_*Reporting Issue* |

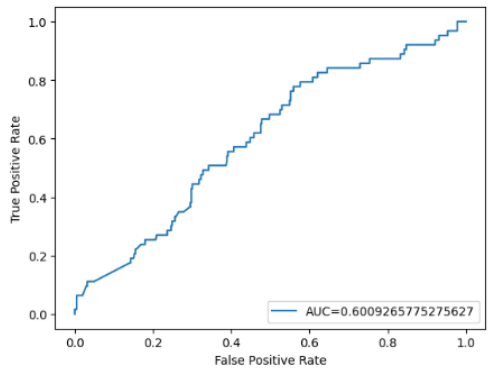
## Hasil dan Evaluasi Model Prediktif

Model prediktif dibangun menggunakan algoritma *Random Forest* dan *XGBoost* untuk kemudian dipilih model dengan performa terbaik. Tabel 3 menunjukkan nilai *confusion matrix* dari hasil *training* model yang melibatkan seluruh *feature* dan setelah *feature selection* dilakukan menggunakan algoritma *Random Forest* dan *XGBoost*. Dari hasil yang ditunjukkan, kedua model memiliki nilai *confusion m*atrix yang sama, dengan nilai presisi, *recall*, dan F1-*score* sebesar 0,00 dan nilai akurasi 0,88. Nilai presisi, *recall*, dan F1-*score* 0 diakibatkan oleh model *training* yang tidak melakukan prediksi kelas positif sama sekali.

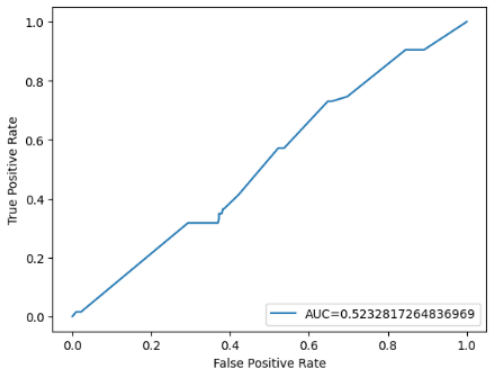
Tabel 3. *Confusion Matrix* Model Sebelum *SMOTE*

| | **Tanpa *Feature Selection*** | | **Dengan *Feature Selection*** | |
|---|---|---|---|---|
| **Metrik** | ***Random Forest*** | ***XGBoost*** | ***Random Forest*** | ***XGBoost*** |
| Akurasi | 0,88 | 0,88 | 0,88 | 0,88 |
| Presisi | 0,00 | 0,00 | 0,00 | 0,00 |
| *Recall* | 0,00 | 0,00 | 0,00 | 0,00 |
| F1-*score* | 0,00 | 0,00 | 0,00 | 0,00 |

Gambar 6 dan Gambar 7 menunjukkan kurva *Receiver Operating Characteristic* (ROC) dari kedua model yang telah dibangun. Pada Gambar 6 dan Gambar 7, kedua model sebelum diterapkan *SMOTE* dan tanpa *feature selection* memiliki nilai AUC masing-masing 0,601 dan 0,523. AUC bernilai > 0,5 dan ≤ 0,6 menunjukkan model sangat buruk untuk digunakan karena memiliki tingkat keakuratan yang sangat rendah dan gagal untuk mengklasifikasikan kelas negatif dan kelas positif (*failed discrimination*) [8]. Sedangkan untuk nilai AUC > 0,6 dan ≤ 0,7 dianggap sebagai *poor discrimination* [8]. Dari hasil evaluasi yang diperoleh, kedua model dinilai kurang baik untuk digunakan.

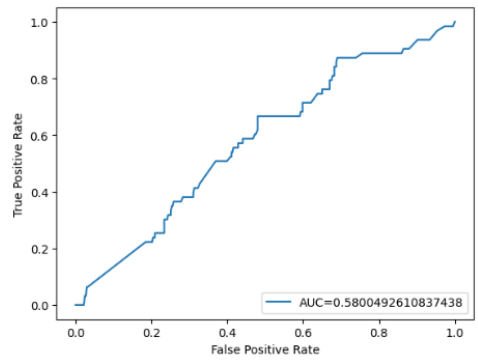



Gambar 6. Kurva ROC *Random Forest* Sebelum *SMOTE* tanpa *Feature Selection*

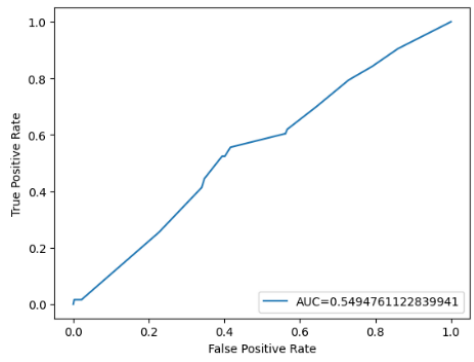



Gambar 7. Kurva ROC *XGBoost* Sebelum *SMOTE* tanpa *Feature Selection*

Gambar 8 dan Gambar 9 menunjukkan kurva ROC dari kedua model yang telah dibangun sebelum diterapkan *SMOTE* dan menggunakan *feature selection*. Dari kedua kurva tersebut, dapat dilihat nilai AUC kedua model masing-masing adalah 0,580 dan 0,549 yang menandakan model sangat buruk untuk digunakan [8]. Hasil ini sesuai dengan nilai *confusion matrix* pada Tabel 3 yang menunjukkan performa model yang buruk.

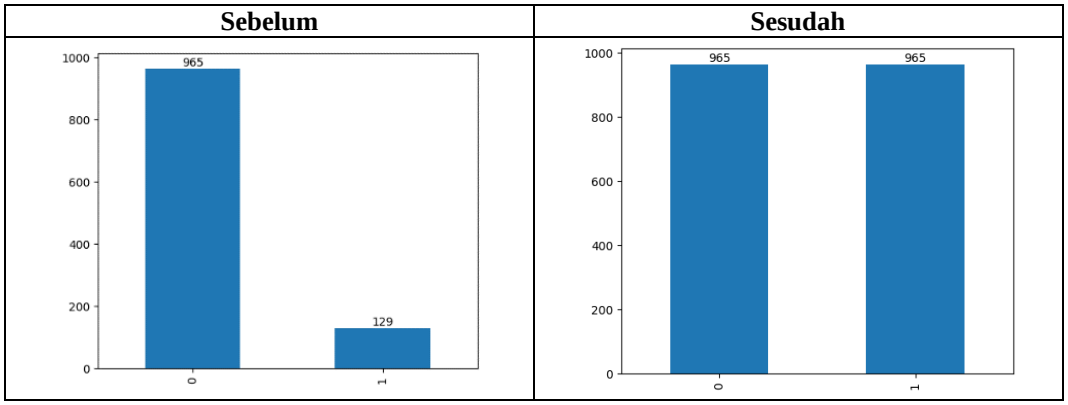



Gambar 8. Kurva ROC *Random Forest* Sebelum *SMOTE* dengan *Feature Selection*




Gambar 9. Kurva ROC *XGBoost* Sebelum *SMOTE* dengan *Feature Selection*

Dari hasil pengujian data yang telah ditunjukkan pada Gambar 3, dapat dilihat bahwa terdapat ketidakseimbangan jumlah data antara kelas positif (*Churn*) dan negatif (Tidak *Churn*). Adanya ketidakseimbangan ini dapat memengaruhi model yang dihasilkan, oleh karena itu dilakukan proses *SMOTE* untuk menambah data penelitian agar jumlah data kelas positif menjadi sama banyak dengan kelas negatif. Setelah dilakukan *SMOTE* terhadap data *training*, diperoleh hasil jumlah data seperti yang dapat dilihat pada Gambar 10.



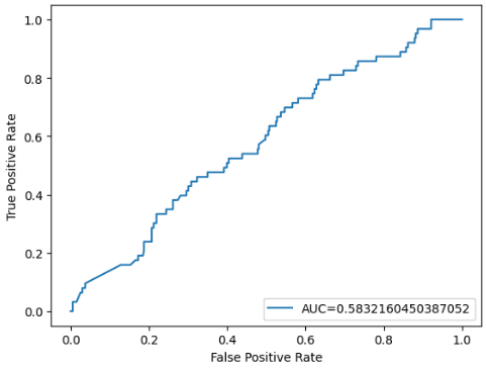

Gambar 10. Proporsi Klien *Churn* Sebelum dan Sesudah *SMOTE*

Tabel 4 menunjukkan nilai *confusion matrix* untuk model menggunakan data setelah proses *SMOTE*. Apabila dibandingkan dengan hasil pada Tabel 3, nilai presisi, *recall*, dan F1-*score* meningkat secara signifikan meskipun nilai akurasi menurun. Penurunan nilai akurasi yang diikuti oleh kenaikan nilai metrik lainnya menunjukkan adanya "*accuracy paradox*" yang diakibatkan oleh ketidakseimbangan data penelitian [9]. Nilai akurasi yang tinggi tidak berarti model mampu secara tepat membedakan data antar kelas, terutama kelas positif yang jumlah datanya jauh lebih sedikit. Secara keseluruhan performa model setelah proses *SMOTE* lebih baik dibandingkan tanpa penerapan *SMOTE*.

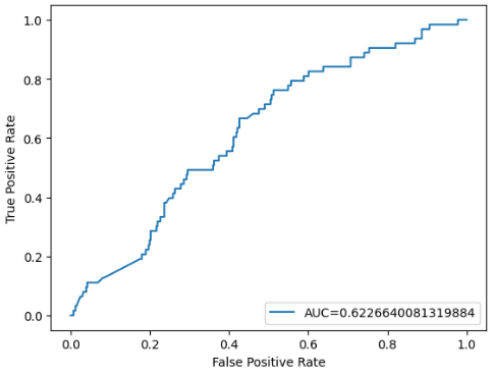
Tabel 4. *Confusion Matrix* Model Setelah *SMOTE*

| | **Tanpa *Feature Selection*** | | **Dengan *Feature Selection*** | |
|---|---|---|---|---|
| **Metrik** | ***Random Forest*** | ***XGBoost*** | ***Random Forest*** | ***XGBoost*** |
| Akurasi | 0,79 | 0,72 | 0,61 | 0,71 |
| Presisi | 0,34 | 0,44 | 0,21 | 0,27 |
| *Recall* | 0,86 | 0,89 | 0,87 | 0,86 |
| F1-*score* | 0,49 | 0,59 | 0,34 | 0,41 |

Gambar 11 dan Gambar 12 menunjukkan kurva ROC kedua model yang dibangun setelah penerapan *SMOTE* tanpa dilakukan *feature selection*. Apabila dibandingkan dengan Gambar 6 dan Gambar 7, penerapan *SMOTE* terhadap data memberikan peningkatan performa model untuk *XGBoost* dan penurunan untuk *Random Forest*. Untuk model setelah penerapan *SMOTE* tanpa adanya *feature selection*, model dengan algoritma *XGBoost* memiliki tingkat akurasi yang lebih tinggi. Meskipun begitu, nilai AUC kedua model sebesar 0,583 dan 0,623 masih berada di rentang *failed discrimination* dan *poor discrimination*, yang berarti model buruk untuk digunakan [8].

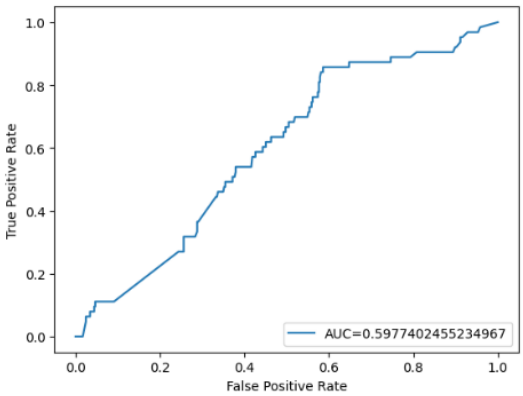



Gambar 11. Kurva ROC *Random Forest* Setelah *SMOTE* tanpa *Feature Selection*

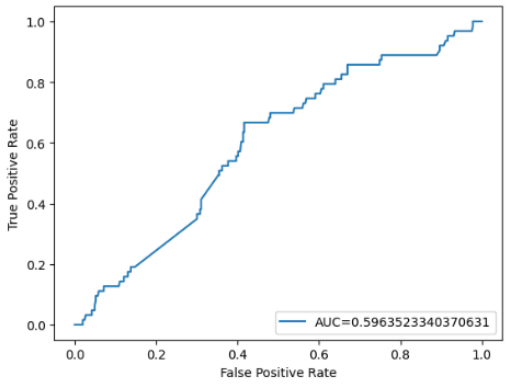



Gambar 12. Kurva ROC *XGBoost* Setelah *SMOTE* tanpa *Feature Selection*

Gambar 13 dan Gambar 14 menunjukkan kurva ROC dari kedua model setelah diterapkan *SMOTE* dan ada penerapan *feature selection*. Apabila dibandingkan dengan Gambar 8 dan Gambar 9, nilai AUC kedua model mengalami peningkatan yang berarti *SMOTE* memberikan peningkatan performa untuk model yang menerapkan *feature selection*. Nilai AUC kedua model masih termasuk *failed discrimination* [8]. Mempertimbangkan nilai *confusion matrix*, model XGBoost memiliki performa yang lebih seimbang dibandingkan *Random Forest*, baik tanpa maupun dengan adanya *feature selection* karena memiliki F1-*score* yang lebih tinggi dan nilai AUC yang sedikit lebih rendah (untuk model dengan *feature selection)*. Untuk analisis lebih lanjut, akan dilakukan perbandingan nilai *confusion matrix* model *XGboost* menggunakan data *testing*.




Gambar 13. Kurva ROC *Random Forest* Setelah *SMOTE* dengan *Feature Selection*




Gambar 14. Kurva ROC *XGBoost* Setelah *SMOTE* dengan *Feature Selection*

Tabel 5. *Confusion Matrix XGBoost Data Testing Setelah SMOTE*

| **Metrik** | **Tanpa Feature Selection** | **Feature Selection** |
|---|---|---|
| Akurasi | 0,72 | 0,61 |
| Presisi | 0,16 | 0,18 |
| *Recall* | 0,24 | 0,52 |
| F1-*score* | 0,19 | 0,27 |

Berdasarkan hasil yang ditampilkan pada Tabel 5, nilai *confusion matrix* model prediktif yang dijalankan menggunakan data testing menunjukkan model dengan

*feature selection* memiliki performa lebih baik. Hasil ini berlawanan dengan hasil pada Tabel 4, di mana baik model algoritma *Random Forest* maupun *XGBoost* yang tidak menerapkan *feature selection* memiliki performa lebih baik. Hasil yang berlawanan ini menunjukkan ada kemungkinan terjadi *overfitting*, di mana model tanpa *feature selection* dapat dengan baik memprediksi data yang digunakan untuk melatih model (data *training*), tetapi ketika digunakan untuk data baru (data *testing*), hasilnya tidak sebaik itu yang berarti model kurang tergeneralisasi.

*Overfitting* sendiri dapat diakibatkan oleh berbagai faktor, seperti ukuran data *training* yang terbatas, adanya *noise (*data *training* mengandung banyak informasi yang kurang relevan*)*, ataupun *complex classifiers* [10]. Salah satu contoh *noise* adalah *feature* Jenis usaha_UMKM dan Jenis usaha_*Enterprise*. Karena kategori jenis usaha hanya UMKM atau *Enterprise*, penyertaan salah satu *feature* sudah cukup mewakili untuk variabel jenis usaha. Dari hasil dan analisis yang telah dilakukan, model *XGBoost* setelah *SMOTE* dengan data yang telah diproses dengan *feature selection* akan digunakan untuk analisis lebih lanjut dan penentuan *feature importance*.

### *Feature Importance*

Tabel 6 menampilkan 5 *feature* dengan tingkat kepentingan tertinggi dan sudah dapat menginterpretasikan 87% dari total *feature* yang ada.

Tabel 6. *Feature Importance* 5 Teratas

| ***Feature*** | ***Importance*** |
|---|---|
| Isu tiket_ERP *Issue* | 0,24 |
| Isu tiket_POS *Issue* | 0,23 |
| Isu tiket_*Reporting Issue* | 0,16 |
| Jenis usaha_*Enterprise* | 0,12 |
| Tiket bulanan | 0,12 |

Dari lima *feature importance* yang telah diperoleh, secara garis besar dapat dibagi menjadi dua aspek utama, yaitu tiket laporan kendala dan jenis usaha klien, secara khusus jenis klien UMKM. *Feature* dengan *importance* tertinggi adalah tiket dengan isu permasalahan di produk ERP, produk POS, dan *reporting*. Tiket dengan kendala yang dikategorikan sebagai kendala ERP adalah adanya *error* atau *bug* di produk ERP, permasalahan di pengiriman email otomatis, laporan terkait akses akun pengguna, ataupun pertanyaan terkait pengaturan menu dan bahan baku menu. Tiket yang dikategorikan sebagai kendala POS adalah *error* atau *bug* di produk POS, produk POS yang hanya berhenti di tampilan *loading*, pengaturan cabang restoran ataupun daftar menu, tampilan menu untuk pelanggan, hingga permasalahan saat melakukan proses pembayaran. Tiket yang dikategorikan sebagai kendala *reporting* adalah permasalahan saat ingin mengakses laporan dari produk ERP, dari produk POS, ataupun laporan penutupan *shift*.

Selain isu tiket, rata-rata tiket bulanan juga menjadi *feature importance* ke-5. Semakin sering klien melaporkan kendala, terutama yang berkaitan dengan isu produk ERP, POS, maupun *reporting*, maka semakin tinggi kemungkinan klien merasa tidak puas dengan pengalaman penggunaan ataupun kualitas produk karena sering mengalami kendala atau *error* dalam pemakaian. Selain adanya kendala, pelaporan isu dapat terjadi karena klien mengalami kesulitan saat melakukan penggunaan ataupun pengaturan produk, yang menunjukkan kurang maksimalnya pemahaman klien terhadap produk. Kesulitan dalam penggunaan produk pun dapat

menimbulkan rasa tidak puas pada klien. Dengan hadirnya kompetitor, muncul potensi klien untuk beralih ke produk lain yang diharapkan lebih jarang mengalami kendala. Dari hasil *feature importance*, PT XYZ mengetahui isu tiket apa yang teridentifikasi perlu diprioritaskan untuk diteliti dan diperbaiki.

*Feature importance* ke-4 tertinggi adalah jenis usaha, yang mana berdasarkan Gambar 4 klien UMKM memiliki kecenderungan untuk *churn* yang lebih tinggi dibandingkan *Enterprise*. Hasil ini menunjukkan adanya terdapat faktor di luar penelitian yang berpotensi memiliki kontribusi lebih besar terhadap kecenderungan klien UMKM untuk melakukan churn dibandingkan keempat *feature* tertinggi lainnya. Beberapa contoh faktor yang tidak diteliti pada penelitian ini meliputi harga dan diskon yang ditawarkan, produk selain ERP dan POS, jumlah kegiatan atau transaksi klien, ataupun jumlah tagihan yang dibayarkan. Faktor-faktor tersebut tidak diteliti karena adanya batasan penelitian untuk lebih fokus meninjau faktor-faktor yang pernah diteliti oleh penelitian terdahulu yang menjadi referensi [6][7]. Selain itu, penelitian ini lebih banyak meneliti faktor yang berkaitan dengan kualitas dan penggunaan produk, secara khusus produk ERP dan POS karena kedua produk tersebut merupakan produk utama dan yang paling lama telah ditawarkan oleh PT XYZ.

Berdasarkan hasil analisis yang telah dilakukan, terdapat beberapa implikasi manajerial yang dapat diterapkan oleh PT XYZ, seperti perlu perhatian pada klien yang sering melaporkan adanya kendala (rata-rata tiket bulanan tinggi), terutama dengan isu terkait permasalahan di produk ERP dan produk POS. Perusahaan juga perlu meninjau kembali pelayanan tim dan kualitas produk, karena banyaknya laporan kendala dari klien dapat menunjukkan adanya permasalahan di produk, misalnya seperti *bug* ataupun panduan/pelatihan penggunaan produk yang kurang memadai bagi klien. Perlu dipastikan bahwa klien sebagai pengguna dan operator produk telah memahami mekanisme dan cara penggunaan produk.

Selain itu, PT XYZ juga perlu memperhatikan klien dengan jenis usaha UMKM karena saat ini klien PT XYZ dengan jenis usaha UMKM lebih banyak dibandingkan *Enterprise*, dan proporsi klien UMKM yang *churn* lebih tinggi seperti yang ditunjukkan pada Gambar 4. PT XYZ dapat melakukan peninjauan lebih lanjut terkait faktor yang memengaruhi klien UMKM menjadi *churn*. Faktor yang dapat dianalisis lebih lanjut dapat berupa *feature* yang diteliti pada penelitian ini, seperti jenis isu tiket atau jumlah tiket, ataupun variabel lainnya, seperti harga dan promo yang ditawarkan kepada klien.

## KESIMPULAN

Berdasarkan hasil analisis, model dengan performa terbaik adalah model algoritma *XGBoost* dengan penerapan *SMOTE* dan *feature selection* berdasarkan nilai *confusion matrix* akurasi 0,71, presisi 0,27, *recall* 0,86, F1-*score* 0,41, dan AUC 0,596. Dari model tersebut, diperoleh hasil faktor yang paling memengaruhi tingkat *churn* PT XYZ adalah klien yang sering melaporkan kendala, terutama kendala terkait produk ERP, kendala terkait produk POS, kendala terkait penarikan laporan kegiatan operasional, jenis usaha klien, dan rata-rata tiket yang dilaporkan klien setiap bulannya. Klien yang sering melaporkan adanya kendala ataupun klien dengan jenis usaha UMKM memiliki risiko *churn* lebih tinggi. PT XYZ dapat melakukan evaluasi produk untuk memastikan bahwa kendala yang sering

dilaporkan oleh klien tidak dialami berulang kali. PT XYZ juga dapat melakukan peninjauan ulang terhadap penjelasan dan pelatihan penggunaan produk ke klien untuk memastikan bahwa klien memahami produk dan mekanisme penggunaan produk dengan baik. Selain itu, PT XYZ dapat melakukan peninjauan lebih dalam untuk klien dengan jenis usaha UMKM untuk mengidentifikasi faktor yang menyebabkan klien UMKM memiliki risiko *churn* lebih tinggi.

## REFERENSI

[1] Badan Pusat Statistik. *Profil Industri Mikro dan Kecil 2020*. Jakarta: BPS, 2022.
[2] Kumar, Saravana, "Churn in SaaS and Five Ways to Reduce It", *Forbes*, October 6, 2022. [Online]. Available: Forbes, https://www.forbes.com. [Accessed December 8, 2023].
[3] Ghaffari, Benjamin and Osman, Yasin, "Customer Churn Prediction Using Machine Learning: A Study in the B2B Subscription Based Service Context", Master thesis, Blekinge Institute of Technology, Karlskrona. 2021.
[4] Singh, Yash, et al, "Prediction of Customer Churn Using Machine Learning", *International Research Journal of Modernization in Engineering Technology and Science*, vol.4, pp. 318-330, 2022.
[5] Ge, Y., He, S., Xiong, J., and Brown, D. E., "Customer Churn Analysis for a Software-as-a-service Company", *2017 Systems and Information Engineering Design Symposium (SIEDS)*, pp. 106-111, 2017.
[6] Rautio, Anton, "Churn Prediction in SaaS using Machine Learning", Master thesis, Tampere University, Tampere. 2019.
[7] Apostolov, Nikola, "Reducing Churn in Data Management SaaS Companies: A Case Study", Master thesis, University of Twente, Enschede. 2020.
[8] Kleinbaun, D. G., and Klein, M., *Logistic Regression: A Self Learning Text (3rd ed)*. New York: Springer-Verlag, 2010.
[9] Valverde-Albacete, Francisco and Pelaez-Moreno, Carmen, "100% Classification Accuracy Considered Harmful: The Normalized Information Transfer Factor Explains the Accuracy Paradox", *PLoS One*, Vol.9, 2014.
[10] Ying, Xue. "An Overview of Overfitting and Its Solutions", *Journal of Physics: Conference Series*, 1168, 2019.