

# BAB I

## PENDAHULUAN

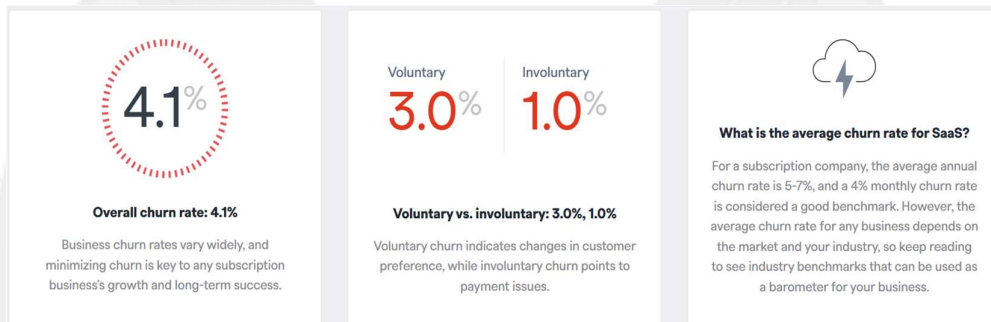
### 1.1 Latar Belakang

Perubahan teknologi telah mengubah cara bisnis dijalankan, terutama di sektor *Business-to-Business* (B2B). Banyak perusahaan B2B yang kini mengadopsi platform berbasis aplikasi untuk memfasilitasi transaksi dengan mitra bisnis mereka. PT XYZ, sebuah perusahaan B2B yang bergerak di bidang distribusi kebutuhan pokok seperti telur, sayur, buah, dan kebutuhan rumah tangga lainnya, adalah salah satu perusahaan yang memanfaatkan teknologi melalui aplikasi *mobile*. Sejak didirikan pada tahun 2017, PT XYZ telah berhasil membangun jaringan pelanggan yang oleh PT XYZ disebut sebagai mitra yang tersebar di wilayah Jabodetabek dan Bandung. Hingga 31 Agustus 2024, PT XYZ memiliki 2400 mitra yang menjual produk kebutuhan pokok kepada teman, tetangga dan komunitas lokal mereka.

Dalam industri B2B, mempertahankan pelanggan atau mengurangi tingkat *churn* (*churn rate*) adalah hal yang esensial untuk mendukung pertumbuhan bisnis yang berkelanjutan. Namun, seiring pertumbuhan perusahaan, tantangan baru muncul, terutama dalam bentuk *customer churn*, yaitu berhentinya mitra dari penggunaan layanan perusahaan. Dalam konteks B2B, *churn* merupakan masalah yang lebih kritis dibandingkan dengan sektor *Business-to-Consumer* (B2C). Ini disebabkan karena hubungan B2B seringkali melibatkan transaksi berulang dengan

volume besar, sehingga hilangnya satu mitra dapat berdampak signifikan pada pendapatan perusahaan (Matuszelański & Kopczewska, 2022).

Menurut laporan Recurly (2024), *churn rate* dapat bervariasi berdasarkan jenis industry, rata-rata *churn rate* untuk perusahaan SaaS B2B yang sehat berkisar antara 5% hingga 7% per tahun. Bahkan dalam situasi sulit seperti pandemi COVID-19, banyak perusahaan SaaS B2B dapat mempertahankan *churn rate* di bawah 10%, menunjukkan kemampuan industri tersebut dalam menjaga stabilitas pelanggan.



**Gambar 1.1 Churn Rate untuk Perusahaan SaaS menurut RecurlyResearch**

Namun, data yang diperoleh dari PT XYZ menunjukkan bahwa tingkat *churn* perusahaan berada di kisaran 20.96% pada tahun 2024. Tingkat *churn* ini secara signifikan lebih tinggi dibandingkan dengan rata-rata industri SaaS, menunjukkan adanya risiko yang cukup besar bagi keberlanjutan bisnis perusahaan. Tingginya *churn rate* ini memicu kebutuhan mendesak untuk memahami faktor-faktor penentu *churn* di PT XYZ serta menerapkan model prediktif untuk memitigasi *churn* dan meningkatkan retensi pelanggan.

Matuszelański & Kopczewska (2022) dalam penelitian mereka tentang *churn* di sektor *e-commerce* B2C menunjukkan bahwa memprediksi *churn* dengan menggunakan *machine learning* telah terbukti efektif dalam mengidentifikasi

pelanggan berisiko tinggi untuk *churn*. Walaupun penelitian tersebut berfokus pada sektor B2C, prinsip yang sama dapat diterapkan pada sektor B2B karena *churn* merupakan tantangan yang ada di semua industri yang berfokus pada hubungan jangka panjang dengan pelanggan. *Machine learning* memungkinkan perusahaan menganalisis pola perilaku dan tren di antara mitra, serta mengidentifikasi mitra mana yang berpotensi *churn* berdasarkan data historis mereka (Mirkovic et al., 2022).

Dalam penelitian ini, Random Forest dan Regresi Logistik akan digunakan sebagai alat untuk memprediksi *churn*. Kedua algoritma ini dikenal efektif dalam menangani masalah klasifikasi, termasuk *churn*. Random Forest merupakan algoritma ensemble yang bekerja dengan membangun banyak decision trees dan menggabungkan hasilnya untuk meningkatkan akurasi prediksi (Matuszelański & Koczewska, 2022). Sementara itu, Regresi Logistik sering digunakan dalam memprediksi probabilitas suatu peristiwa seperti *churn*, dan berguna ketika hubungan antara variabel independen dan variabel dependen bersifat linier (Russo et al., 2016).

Berbagai penelitian juga telah menunjukkan bahwa faktor-faktor seperti jenis kelamin, usia, status pernikahan, area pengiriman, lama keanggotaan, jumlah transaksi per bulan, nilai belanja per bulan, variasi produk, diskon per bulan, tipe pengiriman, dan persentase keaktifan mitra dapat mempengaruhi keputusan mitra untuk *churn* (Mirkovic et al., 2022; Wadikar, 2020). Misalnya, penelitian oleh Van Haver (2017) menunjukkan bahwa variabel lama keanggotaan dan jumlah transaksi secara signifikan mempengaruhi tingkat retensi pelanggan. Mitra yang lebih lama

menjadi pelanggan cenderung memiliki loyalitas yang lebih tinggi dibandingkan mitra baru yang mungkin lebih mudah *churn*.

Dalam konteks PT XYZ, mempertahankan mitra adalah salah satu kunci kesuksesan perusahaan. Mitra yang loyal memberikan kontribusi yang signifikan terhadap pendapatan jangka panjang perusahaan. Penelitian ini juga menekankan bahwa mempertahankan pelanggan jauh lebih ekonomis dibandingkan memperoleh pelanggan baru, dengan biaya akuisisi pelanggan baru yang bisa mencapai lima kali lipat biaya untuk mempertahankan pelanggan yang sudah ada (Jamjoom, 2021). Oleh karena itu, penting bagi PT XYZ untuk dapat memprediksi mitra mana yang berpotensi *churn* dan melakukan langkah-langkah proaktif untuk mencegah *churn* tersebut, seperti melalui program loyalitas atau penawaran diskon yang lebih menarik (Jahromi et al., 2014).

Melalui penelitian ini, diharapkan PT XYZ dapat mengoptimalkan strategi retensinya dengan memanfaatkan teknologi *machine learning*, sekaligus memperkuat hubungan dengan mitra bisnis mereka. Hasil dari penelitian ini juga diharapkan dapat menjadi acuan bagi perusahaan-perusahaan lain di sektor B2B yang menghadapi tantangan serupa dalam mengelola *churn* di era digital.

## 1.2 Rumusan Masalah

Rumusan masalah dalam penelitian ini adalah sebagai berikut:

- 1) Apa saja faktor-faktor utama yang mempengaruhi *customer churn* di PT XYZ?

- 2) Bagaimana cara membangun model prediksi yang dapat memprediksi *customer churn* di masa depan berdasarkan data transaksi dan profil mitra?
- 3) Faktor mana yang paling signifikan dalam mempengaruhi keputusan pelanggan untuk tetap setia atau berhenti menggunakan layanan PT XYZ?

### 1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk:

- 1) Mengidentifikasi faktor-faktor penentu *customer churn* yang paling relevan di PT XYZ, terutama dalam konteks B2B dan pengguna ibu rumah tangga.
- 2) Mengembangkan model prediksi *customer churn* berbasis data yang menggabungkan berbagai faktor geodemografis dan perilaku pelanggan
- 3) Mengevaluasi kinerja berbagai algoritma *machine learning*, seperti regresi logistic dan random forest dalam memprediksi *customer churn*.
- 4) Merumuskan strategi retensi pelanggan yang efektif berdasarkan hasil analisis faktor-faktor *customer churn* yang signifikan, dengan fokus pada segmentasi pelanggan yang lebih akurat dan personalisasi interaksi perusahaan dengan pelanggan.

### 1.4 Manfaat Penelitian

Hasil dari penelitian ini diharapkan memberikan manfaat baik bagi PT XYZ maupun akademisi, antara lain:

- 1) Bagi PT XYZ, penelitian ini dapat memberikan wawasan yang mendalam tentang strategi retensi pelanggan yang lebih efektif. Perusahaan dapat menggunakan hasil analisis untuk merancang program loyalitas yang lebih

tepat sasaran, terutama dalam mempertahankan mitra-mitra yang cenderung berisiko *churn*.

- 2) Bagi akademisi, penelitian ini dapat memberikan kontribusi dalam pengembangan metode prediksi *customer churn* berbasis *machine learning* di sektor B2B. Penelitian ini juga menambah referensi tentang bagaimana perusahaan B2B dapat memanfaatkan data transaksi dan analitik untuk mengurangi *customer churn*.
- 3) Bagi masyarakat umum, terutama ibu rumah tangga yang menjadi mitra PT XYZ, penelitian ini diharapkan dapat meningkatkan pengalaman pengguna dan memberikan nilai tambah melalui layanan yang lebih personal dan sesuai dengan kebutuhan mereka.

## **1.5 Ruang Lingkup Penelitian**

Penelitian ini berfokus pada analisis prediktif terhadap *customer churn* di PT XYZ, dengan menggunakan data yang mencakup faktor-faktor demografis, perilaku mitra dan *historical* transaksi mitra selama periode Januari 2023 sampai Agustus 2024. Data yang dianalisis meliputi: jenis kelamin, status pernikahan, area pengiriman, lama keanggotaan, usia mitra, rata-rata jumlah transaksi per bulan, rata-rata nilai belanja per bulan, variasi produk yang dibeli, rata-rata persentase diskon, metode pengiriman, dan persentase keaktifan mitra. Proses analisis dilakukan menggunakan algoritma *machine learning* untuk membangun model prediksi yang dapat mengidentifikasi faktor-faktor paling signifikan dalam menentukan loyalitas. Hasil analisis ini diharapkan dapat memberikan wawasan

yang mendalam bagi PT XYZ dalam menyusun strategi retensi yang lebih personal dan berbasis data.

## 1.6 Metodologi Penelitian

Penelitian ini menggunakan metodologi CRISP-DM (*Cross-Industry Standard Process for Data Mining*) yang dikombinasikan dengan kerangka DCISP-M (*Data, Customer Insights, Strategy, and Performance Measurement*). Langkah-langkah utama dalam metodologi ini adalah:

- 1) Pemahaman Bisnis: Memahami konteks bisnis PT XYZ, yaitu perusahaan B2B berbasis aplikasi yang menjual kebutuhan pokok seperti telur, sayur buah dan kebutuhan sehari – hari lainnya kepada komunitas lokal seperti tetangga sekitar, teman maupun komunitas lain.
- 2) Pemahaman Data: Mengumpulkan data demografis dan transaksi mitra, termasuk jenis kelamin, status pernikahan, area pengiriman, lama keanggotaan, usia, jumlah transaksi, nilai belanja, variasi produk, diskon, tipe pengiriman, dan keaktifan mitra untuk dapat memahami perilaku belanja mitra.
- 3) Persiapan Data dan Pengujian: Membersihkan, mengolah dan melakukan pengujian awal terhadap data untuk mempersiapkan variabel yang akan digunakan dalam pemodelan.
- 4) Pemodelan: Menggunakan algoritma *machine learning* seperti regresi logistik dan random forest untuk membangun model prediksi *customer churn*.



- 5) Evaluasi: Mengevaluasi kinerja model menggunakan metrik akurasi, *precision*, dan *recall* untuk memahami faktor-faktor yang penting maupun signifikan dalam model prediksi ini.
- 6) Metodologi ini memungkinkan PT XYZ untuk tidak hanya memprediksi *customer churn*, tetapi juga menyusun strategi retensi yang lebih efektif berdasarkan data dan wawasan pelanggan yang diperoleh.

## 1.7 Sistematika Penulisan

Penulisan tesis ini terdiri dari lima bab utama yang dirancang untuk menjelaskan secara sistematis tahapan penelitian, analisis, dan hasil yang diperoleh dalam rangka mengembangkan model prediksi *customer churn* di PT XYZ. Berikut adalah sistematika penulisan yang digunakan:

### **Bab 1: Pendahuluan**

Bab ini memberikan pengantar tentang latar belakang masalah yang dihadapi oleh PT XYZ dalam mempertahankan retensi pelanggan, terutama mengingat tantangan bisnis B2B berbasis aplikasi yang melibatkan mitra ibu rumah tangga. Selain itu, bab ini menguraikan rumusan masalah yang menjadi fokus penelitian, tujuan penelitian, manfaat penelitian bagi perusahaan dan dunia akademis, ruang lingkup penelitian yang mengidentifikasi data-data utama yang digunakan, serta metodologi yang mengombinasikan pendekatan CRISP-DM dan DCISP-M untuk analisis prediktif loyalitas pelanggan. Bab ini juga mencakup deskripsi singkat tentang bagaimana penelitian ini akan memberikan wawasan bagi PT XYZ dalam merancang strategi retensi pelanggan yang lebih efektif.



## **Bab 2: Tinjauan Pustaka**

Bab ini membahas teori-teori dan penelitian terdahulu yang relevan dengan *customer churn*, analisis prediktif, dan penggunaan *machine learning* dalam memprediksi perilaku pelanggan. Fokus utama tinjauan pustaka mencakup kajian literatur tentang faktor-faktor yang memengaruhi *customer churn* dalam konteks B2B dan aplikasi berbasis platform, peran strategi dalam meningkatkan retensi, serta metode prediksi yang sering digunakan seperti regresi logistik dan random forest. Penelitian ini juga akan mereview model yang digunakan untuk menjelaskan hasil model dan memudahkan interpretasi faktor-faktor penentu *customer churn*.

## **Bab 3: Metode Penelitian**

Pada bab ini dijelaskan secara rinci metode penelitian yang digunakan, termasuk langkah-langkah dalam pendekatan CRISP-DM dan DCISP-M yang diterapkan dalam penelitian ini. Proses pengumpulan data, baik data primer maupun sekunder, juga dijelaskan, termasuk jenis data yang digunakan seperti jenis kelamin, status pernikahan, area pengiriman, lama keanggotaan, usia, jumlah transaksi per bulan, nilai belanja per bulan, variasi produk, besaran diskon, tipe pengiriman, dan persentase keaktifan mitra. Selain itu, bab ini menjelaskan teknik analisis data yang diterapkan menggunakan algoritma *machine learning* dan validasi model.

## **Bab 4: Hasil dan Pembahasan Penelitian**

Bab ini menyajikan hasil analisis data yang telah dikumpulkan dan diolah sesuai dengan metodologi yang dipilih. Hasil dari pemodelan *machine learning* yang diterapkan untuk memprediksi *customer churn* akan disajikan dalam bentuk metrik akurasi, *precision*, *recall*, serta penjelasan lebih lanjut mengenai faktor-

faktor signifikan yang memengaruhi *churn*. Selain itu, bab ini juga membahas interpretasi hasil analisis dan implikasinya terhadap strategi bisnis PT XYZ. Analisis ini mencakup evaluasi kinerja model prediksi dan strategi retensi yang dapat diterapkan oleh perusahaan.

## **Bab 5: Kesimpulan dan Saran**

Bab terakhir ini menyajikan kesimpulan dari hasil penelitian, yang mencakup ringkasan temuan utama mengenai faktor-faktor penentu *customer churn* dan efektivitas model prediksi yang dikembangkan. Berdasarkan hasil analisis, saran-saran strategis untuk PT XYZ juga disampaikan, terutama terkait implementasi program loyalitas yang lebih terarah dan personal. Selain itu, bab ini juga memberikan rekomendasi untuk penelitian lebih lanjut, termasuk peluang pengembangan model prediksi dengan variabel tambahan atau metode analisis yang lebih kompleks.