

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Pelanggan adalah bagian penting bagi perkembangan dan keberhasilan suatu perusahaan. pelanggan tidak hanya mencerminkan kesetiaan terhadap suatu produk atau layanan tetapi penting dalam stabilitas pendapatan. Dalam lingkungan bisnis yang kompetitif, mempertahankan pelanggan menjadi tantangan strategis bagi Perusahaan, khususnya disektor telekomunikasi. Upaya memperoleh pelanggan baru memerlukan biaya yang jauh lebih besar yaitu 5-10 kali dibandingkan mempertahankan pelanggan. Oleh sebab itu, pemahaman terhadap perilaku dan loyalitas pelanggan menjadi aspek penting dalam mendukung keberlanjutan bisnis [1]

Sektor telekomunikasi menjadi salah satu industry dengan Tingkat perpindahan pelanggan (*customer churn*) yang tinggi. Tingkat churn tahunan di sektor ini dapat mencapai sekitar 30-35% yang berpotensi memengaruhi stabilitas pendapatan perusahaan [2]. Fenomena ini memberikan Gambaran penting mengenai perlunya analisis terhadap faktor-faktor yang memengaruhi loyalitas pelanggan. Beberapa aspek seperti lama berlangganan (*tenure*), tipe kontrak, kualitas layanan, serta metode pembayaran memiliki kontribusi signifikan terhadap tingkat loyalitas pelanggan. Pemahaman terhadap faktor-faktor tersebut dapat membantu perusahaan merumuskan strategi retensi yang lebih efektif[2], [3]

Pendekatan dengan pembelajaran mesin (*machine learning*) untuk menganalisis data pelanggan secara mendalam dan memprediksi loyalitas berdasarkan pola perilaku. Beberapa penelitian terdahulu, seperti oleh prabadevi et al.(2023), membandingkan performa beberapa algoritma *machine learning* dan hasil studi tersebut menunjukkan bahwa SGB memiliki tingkat akurasi tertinggi yaitu 83,9%, disusul oleh RF dengan 82,9%, dan LR dengan 82,6%. Sementara itu, KNN mencatat akurasi sebesar 78,1%, meskipun memiliki keterbatasan dalam aspek interpretabilitas model yang berperan dalam pengambilan keputusan bisnis[4].

Penelitian yang dilakukan oleh S.Wu, et al.(2021) mengembangkan kerangka analitik pelanggan yang mengintegrasikan prediksi churn dan segmentasi pelanggan pada industri teekomunikasi. Studi ini menggunakan beberapa algoritma pembelajaran mesin dan menerapkan Teknik SMOTE. Hasil penelitian menunjukkan bahwa AdaBoost memberikan performa terbaik pada dataset 1 dengan F-1Score 63,11% dan random forest unggul pada dataset 2 dengan F1-score 77,20%. Pendekatan ini menegaskan pentingnya pemilihan algoritma dan penanganan data yang tepat dalam meningkatkan akurasi prediksi churn[5].

Penelitian oleh Yunus Agung Kristanto (2022) menyoroti pentingnya kualitas layanan dalam meningkatkan loyalitas pelanggan. Penelitian tersebut mengidentifikasi tingkat layanan memiliki dampak besar pada kesetiaan pelanggan sebagai variabel mediasi. Meningkatkan mutu layanan terbukti dapat memperkuat hubungan antara perusahaan dan pelanggan serta berdampak positif pada loyalitas[6].

Dalam penelitian ini menerapkan dua algoritma, yaitu XGBoost dan Random Forest, untuk memprediksi loyalitas pelanggan. Kedua model ini dipilih karena memiliki kemampuan tinggi dalam menangani data besar dan kompleks, serta unggul dalam aspek stabilitas dan akurasi prediksi. XGBoost memiliki keunggulan dalam mengatasi ketidakseimbangan data serta mekanisme regularisasi L1 dan L2 yang mampu mengurangi overfitting [7]. Sementara itu, Random Forest efektif dalam menangani interaksi antar variabel dan memberikan fitur importance yang bermanfaat dalam interpretasi model.

penelitian ini juga menerapkan Metode Sequential Feature Selection (SFS) digunakan sebagai teknik seleksi fitur utama karena mampu menilai kombinasi fitur terbaik berdasarkan peningkatan performa model. Selanjutnya, Analisis Varians (ANOVA) diterapkan sebagai analisis statistik pendukung untuk menguji signifikansi awal fitur terhadap variabel target. Sementara itu, metode SHAP (Shapley Additive Explanations) digunakan sebagai teknik interpretasi model guna menjelaskan kontribusi setiap fitur [8], [9] Dengan pendekatan tersebut, diharapkan

penelitian dapat memberikan informasi yang berguna dan strategis bagi Perusahaan telekomunikasi untuk membuat strategi retensi pelanggan.

## 1.2 Rumusan Masalah

Berdasarkan penjelasan di bagian sebelumnya, penelitian ini bertujuan untuk menjawab pertanyaan berikut:

1. Faktor apa saja yang signifikan memengaruhi loyalitas pelanggan, serta bagaimana hasil interpretasi fitur dapat membantu dalam memahami kontribusi setiap faktor terhadap prediksi loyalitas pelanggan?
2. Bagaimana hasil prediksi model ini dapat digunakan untuk merancang strategi bisnis yang efektif dalam meningkatkan loyalitas pelanggan?

## 1.3 Tujuan Penelitian

Tujuan dilakukan penelitian ini yaitu menemukan fitur-fitur penting yang berkontribusi terhadap loyalitas pelanggan. Teknik seleksi fitur yang digunakan adalah Sequential Feature Selection (SFS), sedangkan teknik interpretasi model yang digunakan adalah SHAP (Shapley Additive Explanations). Selain itu, metode ANOVA sebagai analisis statistik pendukung untuk menilai signifikansi awal fitur terhadap variabel target. Diharapkan mampu memberikan pemahaman lebih mendalam mengenai faktor-faktor yang mempengaruhi loyalitas pelanggan. penelitian ini juga bertujuan untuk memberikan rekomendasi yang bisa digunakan untuk meningkatkan loyalitas pelanggan.

## 1.4 Batasan Masalah

Penelitian ini dibatasi pada hal-hal berikut:

1. Penelitian ini hanya menggunakan dataset Telco Customer Churn yang bersumber dari platform Kaggle dan diunggah oleh BlastChar.
2. Model prediksi yang digunakan adalah Xgboost dan Random Forest dengan focus pada task klasifikasi churn pelanggan.

3. Analisis fitur dilakukan menggunakan metode SFS sebagai teknik seleksi fitur utama, ANOVA sebagai analisis pendukung, serta SHAP untuk interpretasi model.
4. Penelitian ini menggunakan data historis pelanggan tanpa mempertimbangkan perubahan perilaku pelanggan di masa depan atau dinamika pasar yang dapat memengaruhi tingkat churn.

