

# BAB 1

## PENDAHULUAN

### 1.1. Latar Belakang

Perkembangan teknologi berdampak besar dalam persebaran informasi. Yang menjadikan industri telekomunikasi sebagai aspek yang paling penting dalam persebaran informasi tersebut. Perusahaan telekomunikasi bersaing dalam meningkatkan kualitas layanan untuk menjaga kepuasan pengguna dan menjaga konsumen yang telah ada agar tidak beralih ke layanan yang disediakan oleh pesaing. Di dalam bisnis ekonomi, konsep ini dikenal dengan sebutan *Customer Relationship Management* (CRM). CRM adalah strategi bisnis yang menargetkan untuk meyakinkan kepuasan pelanggan. Perusahaan yang berhasil menerapkan CRM kedalam bisnis mereka hampir selalu berhasil dalam peningkatan dalam menjaga pelanggan untuk tidak berpaling.

Industri telekomunikasi memproses dan menghasilkan jumlah data yang sangat besar dan cocok untuk penerapan *data mining*. Pada dasarnya, pertumbuhan pasar telekomunikasi sangat besar dan tumbuh secara eksponensial.

*Churn* di dalam industri telekomunikasi digunakan untuk menganalisis kemungkinan pelanggan untuk merubah layanan atau layanan provider selama jangka waktu tertentu. Secara sukarela maupun tidak sukarela (Mahajan, Misra, Mahajan, 2015). Faktor yang mempengaruhi perpindahan pelanggan bisa terjadi karena beberapa hal seperti pelanggan yang memutuskan untuk berhenti berlangganan, pelanggan yang terpaksa berhenti berlangganan bisa karena masalah finansial atau pindah lokasi yang tidak menyediakan layanan tersebut atau perusahaan memutuskan layanan pelanggan secara sepihak (Lazarov dan Capota, 2007). Mengurangi pengurangan pelanggan memiliki

pengaruh yang signifikan tidak hanya untuk meningkatkan profit untuk komersial bank, tetapi juga meningkatkan inti kompetitif mereka (He, Shi, Wan dan Zhao, 2014).

Dengan berkembangnya teknologi dan ilmu pengetahuan, maka dibutuhkan teknik untuk mengatasi hal tersebut. Salah satu caranya ialah menggunakan algoritma *machine learning*. Algoritma *logistic regression* terbukti menghasilkan nilai performa yang tinggi untuk pengaplikasian prediksi *churn*, yaitu sebesar 95% tertinggi diantara algoritma lain (Petkovski, Aleksandar, Risteska, Biljana, Trivodaliev, 2016). *Logistic regression* berhasil di aplikasikan didalam perusahaan telekomunikasi untuk mendeteksi pelanggan yang berpeluang untuk meninggalkan perusahaan dan penemuan mayoritas faktor-faktor penting penyebab *churn* (Mandák dan Hančlová, 2019). Dengan banyak dan besarnya data yang ada, pendekatan menggunakan *correlation-based feature selection* menghasilkan hasil yang lebih efektif untuk bekerja dengan data yang lebih besar (Yu dan Liu, 2003). Penelitian Yu dan Liu menerapkan metode relevansi dan analisis redundansi untuk seleksi fitur yang menargetkan dataset untuk data yang besar.

Diharapkan peneliti dapat mengetahui hasil performa akan penggunaan *correlation-based feature selection* dan algoritma *logistic regression* yang dipakai untuk memprediksi *churn* terhadap layanan telekomunikasi dan dapat berkontribusi untuk perkembangan akan penelitian selanjutnya.

## 1.2. Perumusan Masalah

1. Bagaimana performa penerapan algoritma *Logistic Regression* untuk memprediksi *Churn* dengan *Correlation-Based Feature Selection-Forward Selection*?
2. Fitur apa saja kah yang berpengaruh dalam menghasilkan performa penerapan algoritma *Logistic Regression*

## 1.3. Batasan Masalah

Penelitian ini mempunyai batasan masalah berikut:

1. Data yang dipakai merupakan data dari perusahaan telekomunikasi Orange yang berjumlah 3333 data dan 19 variabel independent serta 1 variabel dependent (kelas) yang bersumber dari Kaggle.
2. Menggunakan data mentah pelanggan yang mengandung data demografi, data panggilan, data keluhan pelanggan, data tenur, indicator *churn* dan *non-churn*, data riwayat pembelian dan data *billing*.

#### **1.4. Tujuan Penelitian**

Berdasarkan rumusan masalah, tujuan dari penelitian ialah untuk mengetahui tingkat akurasi algoritma *Logistic Regression* dengan beberapa *Correlation-Based Feature Selection* yang dipakai agar dapat di aplikasikan untuk memprediksi pelanggan yang memiliki potensi untuk pindah dari layanan tertentu serta mengetahui akurasi metode yang telah dipakai untuk di aplikasikan ke data baru.

#### **1.5. Luaran Yang Diharapkan**

Dengan penelitian dan implementasi metode *Logistic Regression* dengan *Correlation-Based Feature Selection* ini diharapkan mendapat keluaran berupa hasil perform kinerja model algoritma yang digunakan agar dapat dimanfaatkan untuk pengambilan keputusan yang dapat membantu perusahaan dalam mempertahankan kesetiaan pelanggan yang sudah ada.

#### **1.6. Kontribusi**

Kontribusi dari penelitian ini dilihat dari rumusan masalah dan hasil yang diperoleh dapat memberikan beberapa manfaat sebagai berikut:

- a. Bagi Penulis

Dengan penelitian ini penulis dapat memahami dan meningkatkan pengetahuan dalam analisis Performa kinerja model dalam memprediksi Churn pelanggan menggunakan algoritma *Logistic Regression* dan *Correlation-Based Feature Selection*.

b. Bagi Perkembangan IPTEK dan Perusahaan

Penelitian ini dapat dijadikan bahan referensi untuk perkembangan akan penelitian serta untuk perusahaan bersangkutan terkait prediksi *Churn* dengan menerapkan algoritma *Logistic Regression* dan *Correlation-Based Feature Selection*.

### **1.7. Sistematika Penulisan**

Sistematika penulisan yang akan diterapkan dalam penelitian untuk mengembangkan tugas akhir adalah sebagai berikut:

#### **BAB I PENDAHULUAN**

Bab ini menjelaskan tentang latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, luaran yang diharapkan, kontribusi dan sistematika penulisan.

#### **BAB II TINJAUAN PUSTAKA**

Pada bab ini akan dijelaskan landasan teori yang menjadi dasar analisa dan pemecahan masalah. Tinjauan pustaka ini didapat dari studi pustaka yang memiliki keterkaitan dengan penelitian skripsi ini.

#### **BAB III METODOLOGI PENELITIAN**

Bab ini berisikan metode yang akan digunakan dalam penyusunan laporan.

#### **BAB IV PEMBAHASAN**

Bab ini membahas deskripsi beserta tahapan dari setiap metodologi yang dipakai. Serta uji evaluasi dari analisis algoritma

## BAB V KESIMPULAN

Bab ini menjelaskan tentang penarikan kesimpulan dari keseluruhan hasil yang sudah dibahas dan di uji di bab sebelumnya.

## DAFTAR PUSTAKA