

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi yang demikian pesat mendukung kompetisi dalam dunia perdagangan. Salah satu tujuan utama aktivitas pemasaran seringkali dilihat dari pencapaian loyalitas pelanggan melalui strategi pemasaran [1]. Hal ini dikarenakan pelanggan dianggap sebagai salah satu aset terpenting bagi bisnis di berbagai perusahaan, hasil pengamatan menyatakan bahwa biaya mempertahankan pelanggan yang loyal lebih murah daripada merekrut pelanggan baru. Berbagai strategi pemasaran digunakan oleh perusahaan agar pelanggan tidak menghentikan pembelian. Apalagi di era digital seperti saat ini mendapatkan pelanggan baru sangatlah sulit. Dengan adanya pasar yang kompetitif, pelanggan memiliki banyak pilihan penyedia layanan, mereka dapat dengan mudah beralih layanan. Dalam dunia bisnis pelanggan tersebut disebut sebagai *customer churn*. *Customer churn* terjadi ketika pelanggan merasa tidak puas, biaya lebih tinggi, kualitas rendah, kurangnya fitur, dan masalah privasi [2].

Dalam penelitian [3] dilaporkan bahwa perusahaan industri *telecommunication* mengalami masalah substansial terkait dengan perpindahan pelanggan karena adanya persaingan yang ketat, kondisi dinamis dan adanya inovasi penawaran terbaru yang lebih menarik. Saat ini, perusahaan *telecommunication* memiliki banyak informasi tentang pelanggan mereka [4]. Data-data tersebut berjenis numerik dan diskrit yang berjumlah besar tersimpan secara elektronik pada basis data perusahaan sehingga dapat dimanfaatkan untuk menemukan pola-pola perilaku dan karakteristik konsumen. Oleh sebab itu perusahaan *telecommunication* membutuhkan pengolahan data yang tepat agar dapat menghasilkan pengetahuan-pengetahuan yang bermanfaat untuk memprediksi pelanggan mana yang *churn* dan mana yang tidak [5]. Salah satu proses analisa data yang tepat untuk mengolah data dengan jumlah besar adalah klasifikasi. Klasifikasi adalah pembagian sesuatu menurut kelas, dalam ilmu *science*, klasifikasi adalah proses pengelompokan objek berdasarkan ciri-ciri persamaan dan perbedaan. K-NN (*K-nearest neighbor*), *Bayesian Classifiers*, *Naïve Bayes Classifiers*, *Decision Tree*, *Neural Network*, Algoritma Genetika, *Support*

Vector Machine (SVM) ,dan *Memory Based Reasoning* merupakan metode yang sering digunakan untuk mengklasifikasi data secara luas [6].

Metode *decision tree c4.5* cukup banyak digunakan dalam penelitian, seperti penelitian yang dilakukan oleh Panji Bimo Nugraha Setio, dkk dengan judul “Klasifikasi dengan Pohon Keputusan Berbasis *Algoritma C4.5*”, dalam penelitian ini dilakukan pembahasan mengenai *algoritma* ID3 dan C4.5, berdasarkan pembahasan tersebut, disimpulkan bahwa *algoritma* C4.5 mengkonstruksi ulang pohon keputusan (*decision tree*) pada *algoritma* ID3 dengan memperbaiki skema pada ID3 yaitu atribut dengan data kosong dan pemangkasan cabang [7]. Pada penelitian lain yang dilakukan oleh Yohana Tri Utami, dkk dengan judul “Penerapan Metode C4.5 untuk Prediksi *Churn Rate* Pengguna Jasa *Telecommunication*”. Disimpulkan bahwa metode *decision tree* C4.5 dalam melakukan klasifikasi pada set data pengguna jasa *telecommunication* menghasilkan nilai *accuracy* sebesar 87%, *precision* sebesar 87,5% dan *recall* sebesar 96% yang dianggap memperoleh hasil yang baik. Atribut yang dianggap paling menarik dari hasil pengujian ini adalah atribut harga yang selanjutnya diidentifikasi sebagai prediksi pola perpindahan pelanggan (*customer churn*) [2].

Berdasarkan permasalahan yang ada maka penulis akan melakukan penelitian yang berjudul “Klasifikasi *Customer Churn* pada Perusahaan *Telecommunication* dengan Menggunakan Metode *Decision Tree C4.5*”. Penelitian ini bertujuan untuk mengukur nilai *accuracy* metode *decision tree* C4.5 dalam mengklasifikasi *customer churn* pada perusahaan *telecommunication*.

1.2 Rumusan Masalah

Dari latar belakang yang telah dijelaskan, adapun rumusan masalah pada penelitian ini sebagai berikut :

1.2.1 Permasalahan

Data-data pelanggan yang dikumpulkan oleh perusahaan *telecommunication* berjenis numerik dan diskrit. Maka dibutuhkan metode yang dapat memproses data-data dengan fitur berjenis numerik dan diskrit. Sehingga dapat menghasilkan pengetahuan-pengetahuan yang bermanfaat untuk memprediksi pelanggan mana yang *churn* dan mana yang tidak.

1.2.2 Metode Usulan

Metode *decision tree c4.5* pada teknik klasifikasi memiliki kelebihan dapat mengolah data dengan nilai numerik dan diskrit sehingga dapat memperoleh *accuracy* yang baik, serta dapat dengan mudah mengidentifikasi karakteristik konsumen pada *customer churn* di perusahaan *telecommunication*.

1.2.3 Pertanyaan Penelitian

Berapa nilai *accuracy* yang dihasilkan dari metode *decision tree c4.5* dalam pengklasifikasian *customer churn* pada perusahaan *telecommunication* ?

1.3 Tujuan dan Manfaat

Adapun tujuan dan manfaat penelitian ini sebagai berikut :

1.3.1 Tujuan

Mengetahui nilai *accuracy* dari metode *decision tree c4.5* dalam pengklasifikasian *customer churn* pada perusahaan *telecommunication*.

1.3.2 Manfaat

Diharapkan dapat membantu pelaku usaha jasa *telecommunication* untuk mengetahui pola *customer churn*, sehingga dapat membuat promosi atau strategi yang efisien untuk mempertahankan pelanggannya.

1.4 Batasan Masalah

Agar penulisan penelitian ini mudah dipahami dan tidak terlalu luas, adapun batasan-batasan masalah sebagai berikut :

1. Dataset yang digunakan adalah *customer churn prediction* yang telah disediakan oleh <https://www.kaggle.com/datasets/blstchar/telco-customer-churn?resource=download>.
2. Dataset yang digunakan memiliki 7043 data, 21 fitur dan 1 sebagai class.

1.5 Sistematika Proposal

Adapun sistematika yang digunakan dalam penulisan skripsi adalah sebagai berikut :

BAB I PENDAHULUAN

Bab ini akan menjelaskan latar belakang masalah, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian, metode penelitian dan sistematika penulisan laporan.

BAB II KAJIAN PUSTAKA

Bab ini membahas landasan teori utama yang diperlukan dalam menyelesaikan penelitian dan penelitian sebelumnya yang telah digunakan untuk menyelesaikan dengan metode atau masalah yang sama.

BAB III METODE USULAN

Bab ini membahas tentang metode yang diusulkan, arsitektur sistem, data penelitian, tahapan penelitian, serta rencana pengujian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini membahas hasil implementasi dari sistem yang sudah dibuat. Serta akan dilakukannya proses uji coba sistem sehingga mendapatkan hasil analisa dari uji coba tersebut.

BAB V PENUTUP

Bab ini berisi kesimpulan dari hasil analisa dan saran-saran untuk pengembangan selanjutnya.