BAB I PENDAHULUAN

1.1 Latar Belakang

Dalam era digital yang semakin maju, bisnis dan perusahaan di berbagai sektor telah mengalami transformasi dalam pengumpulan, penyimpanan, dan analisis data. Data telah menjadi aset berharga yang dapat memberikan wawasan penting untuk pengambilan keputusan bisnis yang lebih cerdas. Dalam konteks ini, pemahaman yang lebih dalam tentang perilaku pelanggan menjadi kunci penting dalam mempertahankan dan meningkatkan keberhasilan suatu perusahaan.[1]

Salah satu tantangan yang dihadapi oleh banyak perusahaan adalah tingkat churn atau perpindahan pelanggan. Churn terjadi ketika pelanggan berhenti menggunakan produk atau layanan yang diberikan oleh perusahaan. Fenomena ini dapat memiliki dampak signifikan terhadap pendapatan dan pertumbuhan perusahaan. Oleh karena itu, prediksi churn telah menjadi fokus utama bagi banyak perusahaan yang ingin mempertahankan basis pelanggan mereka.[2]

Dalam rangka mengatasi tantangan ini, teknik Machine Learning (ML) telah menjadi alat yang penting dalam mengembangkan model prediksi churn yang akurat. Salah satu tahap penting dalam pengembangan model prediksi adalah pemilihan fitur atau atribut yang paling relevan untuk meningkatkan performa model. Di sinilah teknik Feature Selection masuk ke dalam permainan. Feature Selection adalah proses identifikasi dan pemilihan subset fitur yang paling berpengaruh dalam meningkatkan performa model.[3]

Dalam skripsi ini, akan diteliti penerapan teknik Feature Selection dalam konteks meningkatkan akurasi prediksi churn menggunakan metode Machine Learning. Metode seleksi fitur akan diukur berdasarkan akurasi prediksi dari beberapa model klasifikasi, yaitu *Naïve Bayes*, *Support Vector Machine (SVM)*, dan *Random Forest*. *Naïve Bayes* adalah salah satu metode klasifikasi berbasis probabilitas dengan asumsi bahwa semua fitur saling independen, sedangkan SVM merupakan metode yang bekerja dengan menemukan hiperplane optimal untuk

pemisahan kelas, dan *Random Forest* adalah model berbasis pohon keputusan yang menggunakan pendekatan ansambel.[4]

Dengan menggunakan teknik Feature Selection, penelitian ini bertujuan untuk mengidentifikasi fitur-fitur yang paling relevan dalam memprediksi churn dan memperbaiki akurasi model prediksi. Hasil dari penelitian ini dapat memberikan wawasan berharga bagi perusahaan-perusahaan yang ingin meningkatkan strategi retensi pelanggan mereka. Selain itu, penelitian ini juga dapat memberikan kontribusi pada pemahaman tentang pentingnya pemilihan fitur dalam mengembangkan model prediksi churn yang akurat dan efisien.[5]

Melalui pendekatan yang komprehensif dan eksperimen yang terstruktur, diharapkan penelitian ini akan menghasilkan pemahaman yang lebih dalam tentang potensi teknik Feature Selection dalam meningkatkan akurasi prediksi churn menggunakan metode klasifikasi *Naïve Bayes*, SVM, dan *Random Forest*. Dengan demikian, penelitian ini dapat memberikan panduan praktis bagi perusahaan dalam menghadapi tantangan churn dan mempertahankan keberhasilan dalam lingkungan bisnis yang kompetitif saat ini.[6]

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, terdapat beberapa permasalahan yang perlu diatasi dalam penelitian ini:

1. Bagaimana mengatasi permasalahan dimensi data yang dapat mempengaruhi kinerja model prediksi churn?
2. Bagaimana mengidentifikasi subset atribut yang paling berpengaruh dalam prediksi churn?
3. Apakah penerapan teknik Feature Selection dapat meningkatkan akurasi prediksi churn menggunakan metode klasifikasi *Naïve Bayes*, SVM, dan *Random Forest*?
4. Sejauh mana peningkatan atau penurunan akurasi yang dapat dicapai dengan penerapan teknik Feature Selection dalam prediksi churn menggunakan metode klasifikasi tersebut?

1.3 Batasan Masalah

Agar penelitian ini memiliki fokus yang jelas, beberapa batasan masalah diidentifikasi:

1. Penelitian ini akan berfokus pada penerapan teknik *Feature Selection* pada metode klasifikasi *Naïve Bayes*, *SVM*, dan *Random Forest* dalam prediksi *churn*.
2. Data yang digunakan dalam penelitian ini terbatas pada data pelanggan yang relevan dengan prediksi *churn* dalam suatu perusahaan atau layanan.
3. Teknik *Feature Selection* yang akan digunakan akan difokuskan pada metode *Chi2*.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk menerapkan 3eputu *Feature Selection* dalam meningkatkan akurasi prediksi *churn* menggunakan metode klasifikasi *Naïve Bayes*, *SVM*, dan *Random Forest*. Dalam konteks ini, penelitian bertujuan untuk:[13]

1. Menganalisis pengaruh penerapan 3eputu *Feature Selection* terhadap akurasi prediksi *churn*.
2. Membandingkan kinerja beberapa metode klasifikasi *Naïve Bayes*, *SVM*, dan *Random Forest* sebelum dan sesudah menggunakan *Feature Selection* yang diterapkan pada prediksi *churn*.

Mengidentifikasi kelebihan dan kelemahan metode klasifikasi yang digunakan dalam konteks prediksi *churn*.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini meliputi:

1. Memberikan kontribusi pada pemahaman tentang penggunaan teknik *feature selection* dalam konteks prediksi *churn*.
2. Memberikan panduan praktis kepada perusahaan dalam meningkatkan akurasi prediksi *churn* dan mengurangi *churn rate*.

Mengidentifikasi faktor-faktor yang mempengaruhi keberhasilan penerapan teknik *feature selection* dalam prediksi *churn*.

1.6 Sistematika Penulisan

Berisi sistematika penulisan skripsi yang memuat uraian secara garis besar isi skripsi untuk tiap-tiap bab. Peneliti harus dapat mendeskripsikan (menggambarkan) apa saja isi masing-masing Bab yang akan disusun. Jelaskan secara singkat isi dari bab I, bab II, bab III, bab IV, dan bab V.

Contoh :

BAB I PENDAHULUAN, berisi Latar belakang masalah, rumusan masalah, ...

BAB II TINJAUAN PUSTAKA, berisi tinjauan pustaka, dasar-dasar teori yang digunakan, ...

BAB III METODE PENELITIAN, didalamnya terdapat tinjauan umum tentang objek penelitian, analisis masalah, solusi yang ditawarkan, rancangan, ...

BAB IV HASIL DAN PEMBAHASAN, bab ini merupakan tahapan yang penulis lakukan dalam mengembangkan aplikasi, testing hingga penerapan aplikasi di objek penelitian, ...

BAB V PENUTUP, berisi kesimpulan dan saran yang dapat peneliti rangkum selama proses penelitian, ...