1. Pendahuluan

1.1 Latar belakang

Sektor telekomunikasi Indonesia berkembang dengan pesat, baik dalam layanan maupun dari jumlah pelanggan yang bertambah tiap tahun. Teknologi seluler berevolusi dengan cepat, disertai pertumbuhan operator yang secara simultan tumbuh pada area yang sama menimbulkan persaingan industri telekomunikasi yang keras dan dinamis dengan layanan baru dan teknologi mutakhir. Operator telekomunikasi bersaing untuk mendapatkan pelanggan baru, berbagai cara di tempuh untuk mendapatkan pelanggan baru, promosi agresif dengan layanan yang semakin memanjakan pelanggan meningkatkan persaingan dengan kompetitor. Persaingan ini menimbulkan gejala *churn*, Pengertian *churn* disini adalah ketika semua jasa yang digunakan diputus oleh pelanggan ataupun oleh perusahaan. Perusahaan menyadari bahwa aset yang paling berharga adalah pelanggan mereka, hal ini sejalan dengan perubahan orientasi perusahaan saat ini dari produk ke pelanggan. Untuk itu agar menjaga keuntungan perusahaan maka permasalahann *churn* harus dapat ditangani sedini mungkin. Kerugian akibat kasus *churn* hampir sama dengan kehilangan lima pelanggan baru[2]. Sehingga perusahaan harus mampu memprediksi pelanggan yang akan *churn*. Prediksi dilakukan dengan mengidentifikasi data trend dan perilaku pelanggan sebelum gejala *churn* timbul. Selain untuk mengantisipasi gejala *churn*, prediksi berdasarkan trend data pelanggan juga untuk meningkatkan pelayanan sehingga loyalitas dan kepuasan pelanggan meningkat. Jumlah pelanggan telekomunikasi yang banyak menimbulkan permasalahan tersendiri dalam melakukan *Churn Prediction* maka metode prediksi yang dibangun harus mampu untuk memprediksi apakah seorang pelanggan berpotensi *churn* atau tidak secara efisien dengan jumlah sumber daya perusahaan tersebut.

Data mining adalah salah satu area penting dalam era informasi saat ini dengan jangkauan aplikasi yang luas. Klasifikasi berada dalam domain data mining, dan di identifikasi sebagai bagian penting pembahasan data mining. Dengan menggunakan metode klasifikasi permasalahan *churn* dapat diatasi. Metode klasifikasi tradisional seperti C5.0 *Decison Tree*(DT) tidak menunjukan performansi yang baik untuk kasus kelas minor seperti *churn*, ini karena metode klasifikasi cenderung membiaskan hasil prediksi ke kelas mayor[6]. Di kenalkan satu metode baru yang akurat untuk permasalahan *churn* berupa *Emerging Patterns*(EP), yaitu *itemset* yang frekuensi perubahannya sangat signifikan dari satu dataset ke dataset yang lainnya. EP merepresentasikan *knowledge* pembeda yang sangat kuat antara dataset dan menunjukan kemampuan klasifikasi yang akurat. Dengan mengkombinasikan EP dan C5.0 *Decison Tree*(EPDT) masalah *churn* dapat diatasi. EPDT memanfaatkan *essential Emerging Patterns*(eJEP) yang didefinisikan sebagai minimal *itemset* yang supportnya di satu kelas data nol tetapi dikelas lainnya diatas atau sama dengan nilai *support threshold*(ξ). Dengan memanfaatkan EP maka diharapkan C5.0 *Decision Tree* dapat menangani kasus *Churn Prediction* pada pelanggan telekomunikasi.

1.2 Perumusan masalah

Dengan mengacu pada latar belakang masalah diatas, maka permasalahan yang akan dibahas dan diteliti adalah:

a. Bagaimana menerapkan E*merging Patterns* dan C5.0 *Decision Tree* dalam menangani kasus *Churn Prediction* pelanggan Telekomunikasi.
b. Bagaimana pengaruh *support threshold* dan persentase top item terhadap jumlah JEP dan akurasi prediksi kelas churn pelanggan telekomunikasi dan dapat menggambarkan perbedaan *churner* dengan *non churner* menggunakan *lift curve*.

Hipotesa yang akan dibuktikan bahwa dengan *Emerging Pattern* metode C5.0 *Decision Tree* mampu meningkatkan akurasi prediksi kelas churn pelanggan Telekomunikasi.

1.3 Tujuan

Berdasarkan rumusan masalah di atas, maka tujuan dari tugas akhir ini adalah:

a. Menghasilkan suatu perangkat lunak yang dapat memprediksi pelanggan *churn* .
b. Menganalisa akurasi prediksi antara metode EPDT dan metode C5.0 *Decision Tree* dalam bentuk *recall, precision* dan *lift curve.*
c. Menganalisa pengaruh parameter *support threshold,* persentase top item terhadap jumlah JEP dan akurasi prediksi untuk metode EPDT kasus *Churn Prediction* pelanggan Telekomunikasi.

1.4 Batasan Masalah

Dalam Tugas Akhir ini, yang akan dibahas adalah suatu implementasi EPDT dalam masalah *Churn Prediction* dengan batasan masalahnya sebagai berikut:

a. Pembahasan terfokus pada *Emerging Pattern*, sebagai metode penangan kasus *imbalanced class*.
b. Perangkat lunak yang akan dihasilkan untuk menangani studi kasus pelanggan salah satu perusahaan telekomunikasi.
c. Data inputan mengalami *preprossesing.*
d. Parameter yang digunakan yaitu:
    - Support threshold yaitu minimum support dari sebuah item dalam satu kelas.
    - Persentase top item yaitu pengurangan jumlah item yang digunakan dalam membangun *Pattern Tree.*

1.5 Metodologi penyelesaian masalah

Metode yang digunakan dalam penyelesaian tugas akhir ini adalah menggunakan metode studi pustaka atau studi literatur dan analisis dengan langkah kerja sebagai berikut:

a. Studi literatur mencari referensi yang mendukung pembuatan tugas akhir terutama jurnal-jurnal yang berkaitan dengan pembahasan data mining, *churn predicition*, Clementine 12, Delphi 7, *Emerging Pattern*, *Decision Tree* dan algoritma *single scan* sebagai metode pencarian eJEPs.
b. Mempelajari konsep *imbalancedd class* yaitu kondisi dimana satu kelas direpresentasikan oleh jumlah data yang besar(mayoritas) sedang kelas

yang lainnya hanya diwakili oleh sedikit data(minoritas) khususnya *Churn Prediction* dan memahami data yang digunakan sebagai objek penelitian yang berupa data *artifact*, yaitu data yang sudah diproses dari data pelanggan telekomunikasi Indonesia yang berupa data demografis dan data *traffic*(data Call Detail Record) serta *payment*(data pembayaran), kedalam bentuk satu record satu pelanggan dengan atribut memanjang ke kanan.

c. Merancang dan menganalisis perancangan perangkat lunak dengan menganalisa tahapan-tahapan pembuatan perangkat lunak.
d. Melakukan implementasi perancangan perangkat lunak dengan memanfaatkan Delphi 7 dan Microsoft Access serta tool Clementine 12.
e. Melakukan pengujian perangkat lunak dengan memasukkan data training dan data testing dan mencatat performansi sistem dengan parameter berupa *, gmean, recall, precision* dan *lift curve.*
f. Pengambilan kesimpulan dan penyusunan laporan tugas akhir.