BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Kehadiran *data mining* dilatarbelakangi dengan masalah *data explosion* yang dialami akhir-akhir ini, dimana banyak perusahaan telah mengumpulkan data sekian tahun lamanya (data pembelian, data penjualan, data nasabah, data transaksi, data mahasiswa dan lain-lain). Masalah itu pula yang terjadi pada Universitas Komputer Indonesia (UNIKOM) sebagai perusahaan dalam bidang jasa pendidikan, sebagian data yang tersimpan selama ini hanya sebagai dokumentasi contohnya data calon mahasiswa baru. Pertanyaannya sekarang, apakah data tersebut akan dibiarkan menggunung, tidak berguna lalu dibuang, ataukah dapat me-"nambang"-nya untuk mencari 'emas' dan 'berlian' yaitu informasi yang berguna untuk perusahaan tersebut.

Pertumbuhan yang pesat dari akumulasi data itu ternyata menciptakan kondisi yang sering disebut sebagai *"Rich of Data but Poor of Information"* karena data yang terkumpul itu tidak dapat digunakan untuk pengambilan keputusan. Kumpulan data itu dibiarkan begitu saja seakan-akan menjadi sebuah "kuburan data *(data tombs)"*.

Begitu pula dengan permasalahan yang dihadapi UNIKOM saat ini dalam memanfaatkan data calon mahasiswa baru yang selalu terkumpul tiap tahunnya. Sebagai contoh, pada tahun ajaran 2010/2011 jumlah calon mahasiswa baru UNIKOM yang dinyatakan lulus ujian saringan masuk adalah sebanyak 3813

calon mahasiswa. Namun, ternyata jumlah keseluruhan mahasiswa baru UNIKOM untuk tahun ajaran 2010/2011 adalah sebanyak 3341, itu berarti ada sekitar 472 calon mahasiswa mengundurkan diri dengan cara tidak melakukan registrasi. 12,5 % calon mahasiswa yang mungkin potensial, tidak mampu dipertahankan oleh UNIKOM Bandung. Bayangkan bila dalam waktu 4 tahun jika jumlah calon mahasiswa baru yang mengundurkan diri rata-rata sebanyak 12,5 %, maka hampir 50% calon mahasiswa tidak dapat dipertahankan UNIKOM. Jumlah yang sangat besar mengingat persaingan di dunia pendidikan saat ini yang semakin meningkat.

Berdasarkan hal tersebut, maka jika kemungkinan pengunduran diri seorang calon mahasiswa baru dapat diketahui lebih dini mungkin pihak manajemen UNIKOM dapat melakukan tindakan-tindakan yang diperlukan untuk mempertahankan calon-calon mahasiswa tersebut. Namun, permasalahannya hingga saat ini, UNIKOM belum memiliki standar baku yang dapat dijadikan sebagai alat untuk menganalisis kemungkinan pengunduran diri calon mahasiswa baru tersebut.

Untuk mengatasi permasalahan dalam melakukan pengambilan keputusan dalam proses penjaringan calon mahasiswa baru di Universitas Komputer Indonesia (UNIKOM) diperlukan suatu alat analisis bagi manajemen untuk mengetahui kemungkinan pengunduran diri calon mahasiswa baru. Analisis ini dapat dilakukan dengan memanfaatkan teori penalaran berbasis kasus, yaitu membandingkan kasus calon mahasiswa baru dengan kasus-kasus yang pernah terjadi di tahun-tahun sebelumnya. Alat analisis tersebut berupa sebuah aplikasi

penunjang keputusan berbasis komputer yang mengimplementasikan algoritma klasifikasi *data mining*. Dengan ketersediaan data calon mahasiswa baru yang melimpah, *data mining* dapat menggali informasi yang terkubur dari kumpulan data calon mahasiswa baru tersebut. Aplikasi ini dapat digunakan untuk menganalisis kemungkinan pengunduran diri calon mahasiswa baru berdasarkan data yang telah terkumpul sebelumnya.

Pembuatan aplikasi ini memanfaatkan teknik klasifikasi *data mining* dengan membuat sebuah *classifier* berupa pohon keputusan. Algoritma yang digunakan dalam pembuatan pohon keputusan adalah C4.5 yang dibuat oleh J. Ross Quinlan pada tahun 1992. Metode ini dipilih karena proses analisis bersifat mengklasifikasikan. Algoritma C4.5 digunakan dalam penelitian ini mengingat algoritma ini merupakan algoritma paling berpengaruh dan terbaik dalam komunitas penelitian oleh *IEEE International Conference Data mining* (ICDM) pada Desember 2006, disamping sejumlah kemampuan yang dimilikinya dalam menganalisis data. Hasil dari penelitian tugas akhir ini adalah untuk mengetahui apakah model pohon keputusan yang dihasilkan algoritma C4.5 dalam aplikasi dapat digunakan untuk menganalisis kemungkinan pengunduran diri calon mahasiswa baru. Dengan adanya penelitian ini diharapkan proses penjaringan calon mahasiswa baru yang dilakukan UNIKOM dapat dilakukan dengan lebih efisien dan efektif. Adapun judul penelitian Tugas Akhir ini adalah IMPLEMENTASI ALGORITMA C4.5 DALAM PEMBUATAN APLKASI UNTUK MENGANALISIS KEMUNGKINAN PENGUNDURAN DIRI CALON MAHASISWA BARU.

1.2 Rumusan Masalah

Adapun masalah yang ingin diselesaikan dalam penelitian ini adalah bagaimana cara membuat aplikasi untuk membantu manajemen UNIKOM dalam melakukan analisis kemungkinan pengunduran diri seorang calon mahasiswa dengan didasarkan pada kasus-kasus yang sudah ada. Dalam penelitian ini model pohon keputusan yang akan digunakan untuk mendeteksi kemungkinan pengunduran diri seorang calon mahasiswa adalah dengan menerapkan algoritma C4.5.

1.3 Maksud dan Tujuan Penelitian

Maksud dari Penelitian ini adalah untuk mengimplementasikan algoritma pembentukan pohon keputusan C4.5 dalam sebuah aplikasi yang dapat mendeteksi kemungkinan pengunduran diri seorang calon mahasiswa di UNIKOM Bandung.

Sedangkan tujuan dari penelitian adalah:

1. Dengan adanya aplikasi ini diharapkan kemungkinan seorang calon mahasiswa akan mengundurkan diri dari UNIKOM dapat diketahui secara dini
2. Untuk membantu pihak manajemen unikom dalam melakukan proses penjaringan calon mahasiswa baru.
3. Untuk membentuk klasifikasi data calon mahasiswa baru yang melakukan registrasi atau tidak registrasi.
4. Untuk mendukung keputusan dalam proses penerimaan mahasiswa baru UNIKOM.

1.4 Batasan Masalah

Pada tugas akhir ini, permasalahan dibatasi pada :

1. Pembentukan pohon keputusan dengan menggunakan algoritma C4.5.
2. Proses yang dilibatkan dalam aplikasi yang dibangun adalah berupa proses *data mining.*
3. Aplikasi dibangun menggunakan Borland Delphi 7.
4. Keluaran yang dihasilkan aplikasi ini berupa informasi pengklasifikasian data calon mahasiswa baru dengan variabel tujuan melakukan registrasi dan tidak registrasi berdasarkan atribut-atribut dalam data mahasiswa baru berupa pohon keputusan yang dibentuk menggunakan algoritma C4.5.
5. Penyajian laporan daftar aturan yang terbentuk berdasarkan *tree* yang telah dibentuk disimpan ke dalam format Excel.
6. Pencocokkan kasus baru dengan kasus lama dilakukan secara *manual* oleh Analis.
7. Klasifikasi yang dihasilkan berdasarkan pohon keputusan adalah mengelompokkan data calon mahasiswa baru berdasarkan variabel tujuan yang terdiri dari registrasi, tidak registrasi, dan tidak terklasifikasi.
8. Data yang digunakan merupakan data sampel yang diambil secara acak dari tabel CalonMhsBaru. Tujuannya untuk membandingkan hasil pengklasifikasian yang dihitung secara manual dan secara otomatis oleh aplikasi yang dibangun.
9. *Database* yang digunakan adalah *database* yang menggunakan Interbase sebagai *Database Management System* (DBMS) dan telah melalui proses *data*

preprocessing yang hanya menyertakan informasi-informasi yang dibutuhkan saja dengan membuang informasi yang tidak dibutuhkan.

10. Analisis pemodelan yang digunakan adalah berdasarkan aliran data terstruktur. Alat yang digunakan dalam menggambarkan model fungsional dan aliran informasi yaitu menggunakan Diagram Konteks dan *Data Flow Diagram* (DFD).

1.5 Metodologi Penelitian

Metodologi yang digunakan dalam penulisan tugas akhir ini adalah sebagai berikut :

1. Tahap pengumpulan data

Metode pengumpulan data yang digunakan dalam penelitian ini adalah sebagai berikut :

a. Studi Pustaka

Pengumpulan data dengan cara mengumpulkan literatur, jurnal, *paper* dan bacaan-bacaan yang ada kaitannya dengan judul penelitian.

b. Observasi

Teknik pengumpulan data dengan mengadakan penelitian dan peninjauan langsung terhadap permasalahan yang diambil.

c. Wawancara

Wawancara adalah tahap pengumpulan data dengan cara tanya jawab langsung dengan pihak Manajemen UNIKOM dan staf terkait terhadap permasalahan yang diambil.

2. Tahap pembuatan perangkat lunak.

Teknik analisis data dalam pembuatan perangkat lunak menggunakan paradigma perangkat lunak secara *waterfall*, yang meliputi beberapa proses diantaranya:

a. *System Engineering*

   Merupakan bagian terbesar dalam pengerjaan suatu proyek, dimulai dengan mencari dan menetapkan berbagai kebutuhan dari semua elemen yang diperlukan oleh suatu sistem.

b. *Analysis*

   Merupakan tahap menganalisis hal-hal yang diperlukan dalam pelaksanaan proyek pembuatan perangkat lunak.

c. *Design*

   Tahap penerjemahan dari data yang dianalisis kedalam bentuk yang mudah dimengerti oleh *user*.

d. *Coding*

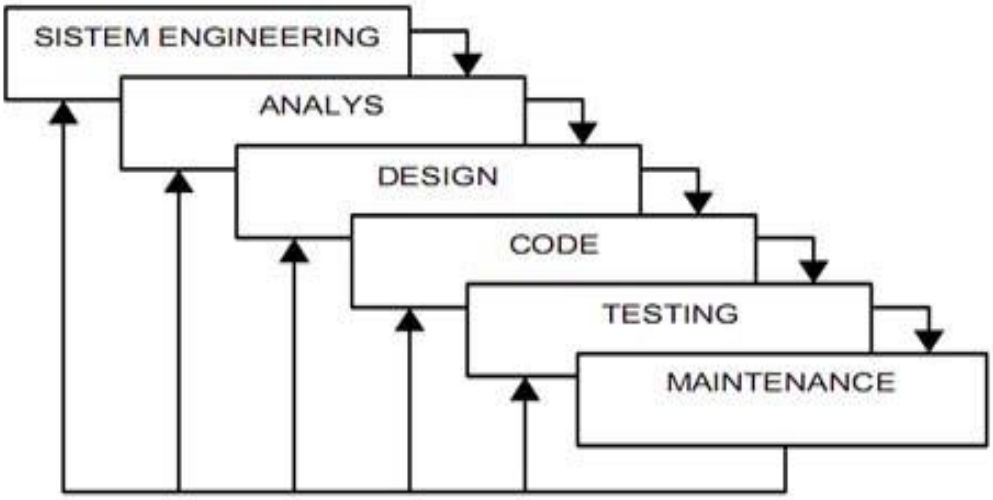
   Tahap penerjemahan data atau pemecahan masalah yang telah dirancang kedalam bahasa pemrograman tertentu.

e. *Testing*

   Merupakan tahap pengujian terhadap perangkat lunak yang dibangun apakah sudah sesuai dengan kebutuhan atau keinginan user

f. *Maintenance*

Tahap akhir dimana suatu perangkat lunak yang sudah selesai dapat mengalami perubahan–perubahan, penambahan, atau perbaikan sesuai dengan permintaan *user*.



Gambar I.1 Metode Waterfall

1.6 Sistematika Penulisan

Adapun sistematika penulisan tugas akhir ini adalah sebagai berikut :

BAB I PENDAHULUAN

Bab ini menguraikan tentang latar belakang permasalahan, rumusan masalah, maksud dan tujuan, batasan masalah, metodologi penelitian, sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Membahas berbagai konsep dan dasar-dasar teori yang menunjang dalam kaitan dengan topik pembuatan aplikasi.

BAB III ANALISIS DAN PERANCANGAN

Menganalisis masalah yang dihadapi dalam membuat aplikasi dan tentang perancangan dalam pembuatan sistem.

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Berisi tahapan-tahapan yang dilakukan untuk menerapkan sistem yang telah dirancang.

BAB V KESIMPULAN DAN SARAN

Berisi rangkuman atau kesimpulan dari penelitian tugas akhir dan saran yang diperlukan untuk pengembangan aplikasi yang telah dibangun.