

BAB I

PENDAHULUAN

I.1 Latar Belakang Masalah

Seiring dengan berkembangnya zaman, informasi yang dibutuhkan juga semakin berkembang dan telah menjadi salah satu hal yang paling penting dalam kehidupan manusia. Menyadari pentingnya informasi, muncullah konsep penambangan data yang merupakan solusi untuk memperoleh informasi yang berguna dari sekumpulan data. Kumpulan data yang sangat banyak jumlahnya ini, jika tidak digunakan untuk aplikasi yang berguna, akan menimbulkan suatu kondisi yang sering disebut sebagai "*rich of data but poor of information*". Tidak jarang kumpulan data yang dibiarkan begitu saja akan menyebabkan "kuburan data" (*data tombs*).

Penambangan data saat ini terutama digunakan oleh perusahaan-perusahaan retail, komunikasi, keuangan, dan organisasi pemasaran. Penambangan data tersebut dapat digunakan untuk menentukan hubungan diantara faktor-faktor internal (seperti : kinerja karyawan atau keterampilan pegawai) dan faktor-faktor eksternal (seperti : indikator ekonomi, kompetisi, dan demografis konsumen). Penambangan data juga membuat perusahaan-perusahaan tersebut dapat menentukan dampak pada penjualan, kepuasan konsumen, dan laba perusahaan. Akhirnya penambangan data membuat perusahaan tersebut dapat menarik ringkasan informasi untuk melihat detail data yang ada. Dengan menggunakan penambangan data, pihak pengusaha dapat menggunakan data karyawan untuk

menentukan perkiraan besar gaji, kriteria penerimaan karyawan, dan hal lain secara tepat.

Metode yang akan digunakan adalah metode klasifikasi dengan algoritma statistik Bayes. Metode klasifikasi mungkin merupakan salah satu teknik penambangan data yang paling umum dan sering digunakan. Beberapa contoh aplikasi metode ini antara lain pengenalan pola dan gambar, diagnosa medis, aplikasi pendeteksian kesalahan dalam industri, dan pengelompokan finansial tren pasar. Perkiraan dan prediksi dapat dipandang sebagai jenis-jenis klasifikasi. Ketika seseorang memperkirakan umur kita atau menebak jumlah kelereng dalam toples, sebenarnya ini termasuk dalam klasifikasi. Prediksi dapat diartikan dengan mengelompokkan sebuah nilai atribut ke dalam suatu kumpulan kelas yang ada/memungkinkan. Selain itu, juga sering dipandang sebagai meramalkan suatu nilai yang kontinu, sementara klasifikasi meramalkan suatu nilai diskret.

Untuk mengukur prestasi kerja seorang karyawan dalam perusahaan ini, maka diperlukan beberapa kriteria penilaian. Kriteria penilaian yang digunakan antara lain : presensi, produktifitas, lama bekerja, tanggung jawab, dan kemampuan kerja. Dari kriteria-kriteria yang ada tersebut, prestasi kerja karyawan dapat dikategorikan menjadi 3 kategori, yaitu prestasi baik, prestasi cukup, dan prestasi kurang. Kriteria yang paling mempengaruhi adalah presensi kehadiran kerja. Pengukuran prestasi kerja karyawan ini untuk selanjutnya dapat digunakan untuk proses pengambilan keputusan, seperti untuk menentukan berapa besar bonus

yang layak diberikan kepada karyawan tergantung dari kelas prestasinya.

Selama ini, pengukuran prestasi kerja hanya dilakukan secara manual oleh kepala bagian masing-masing, sehingga cenderung menyebabkan pemborosan waktu. Maka dari itu, DMBC diharapkan mampu membantu dalam proses pengklasifikasian karyawan sehingga dapat meningkatkan efisiensi waktu.

I.2 Rumusan Masalah

Dari latar belakang masalah tersebut di atas dapat dirumuskan masalah sebagai berikut :

1. Bagaimana membangun suatu aplikasi data mining yang dapat digunakan untuk memperoleh pola yang menarik/pengetahuan yang baru dari data?
2. Bagaimana menggunakan klasifikasi Bayesian untuk memperoleh pengetahuan yang menarik tentang perkiraan kelas suatu tuple yang memiliki nilai-nilai atribut tertentu?
3. Bagaimana menyajikan pengetahuan yang menarik yang diperoleh dari data yang ditambang dalam penyajian yang berguna?

I.3 Batasan Masalah

- a. Sebagai data percobaan (*training data*) akan diambil sample dari data karyawan pada PT. Iskandar Indah Printing Textile Surakarta.

- b. Penambahan data dilakukan dengan menggunakan algoritma *naive Bayesian classification*.
- c. Sistem akan diimplementasikan dengan menggunakan tools Microsoft Visual Basic .NET 2003.
- d. Sistem manajemen basis data (DBMS) yang digunakan hanya SQL Server 2000.
- e. Sistem hanya dapat dijalankan pada sistem operasi Windows.

I.4 Kebutuhan Perangkat Lunak

Adapun tools yang digunakan untuk mengimplementasikan penambahan data menggunakan metode klasifikasi dengan algoritma Bayes ini antara lain :

- a. Nama : MSSql Server 2000
Sumber : Microsoft
Sebagai database yang dibutuhkan dalam mengoperasikan perangkat lunak yang akan dibuat.
- b. Nama : Windows 2000/XP
Sumber : Microsoft
Sebagai sistem operasi komputer.
- c. Nama : Visual Basic .NET 2003
Sumber : Microsoft
Sebagai perangkat lunak yang digunakan untuk mengimplementasikan penambahan data metode klasifikasi ini.

I.5 Metode Pengumpulan Data

Adapun informasi/data yang diperlukan dikumpulkan dengan metode pengumpulan data sebagai berikut :

a. Metode Kepustakaan

Yaitu : membaca buku-buku yang berkaitan dengan penambangan data dan teknik-teknik yang digunakan untuk menambang data.

b. Metode Observasi

Yaitu : cara pengumpulan data/informasi yang dilakukan dengan mengadakan penelitian terhadap data-data karyawan yang ada.

c. Metode Interview / Wawancara

Yaitu : cara pengumpulan data dengan melakukan tanya jawab pada pegawai yang ditunjuk untuk memberikan informasi atau data yang ada kaitannya.

I.6 Sistematika Penulisan Tugas Akhir

Tugas akhir ini akan disusun dengan format penulisan sebagai berikut :

BAB I. PENDAHULUAN

Bab ini berisi latar belakang masalah, batasan masalah, metode pengumpulan data, dan sistematika penulisan skripsi.

BAB II. LANDASAN TEORI

Bab ini berisi uraian sistematis dari teori yang ada pada literatur maupun penjabaran tinjauan pustaka yang mendasari pemecahan masalah.

BAB III. ANALISIS DAN DESAIN SISTEM

Bab ini berisi uraian hasil penelitian/ perancangan, analisis, dan pembahasannya secara terpadu yang disajikan dalam bentuk penjelasan, baik secara kualitatif maupun kuantitatif.

BAB IV. IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini berisi penjelasan tentang implementasi sistem dan hasil pengujian sistem terhadap data yang ada.

BAB V. KESIMPULAN DAN SARAN

Bab ini berisi ringkasan hasil penelitian dan ide-ide mengenai langkah-langkah untuk perbaikan dan pengembangan penelitian yang telah dilakukan.

