

BAB V

KESIMPULAN DAN SARAN

V.1 Kesimpulan

Ada beberapa kesimpulan yang dapat ditarik dari tugas akhir ini, yaitu :

1. Perangkat lunak DMBC telah berhasil dibuat dan dapat digunakan untuk memperoleh suatu pengetahuan baru tentang klasifikasi prestasi kerja seorang karyawan berdasarkan nilai atribut yang dimilikinya.
2. Penambahan data dengan metode klasifikasi dengan menggunakan algoritma naive Bayesian telah berhasil diimplementasikan pada DMBC.
3. Pengetahuan yang baru diperoleh dengan melakukan penghitungan probabilitas dari nilai-nilai atribut tuple baru yang diinputkan oleh user. Hasil penghitungan nilai probabilitas tuple tersebut selanjutnya digunakan untuk mengklasifikasikan tuple tersebut kedalam kelas-kelas yang ada. Pengetahuan yang diperoleh dari hasil klasifikasi ini dapat digunakan untuk membantu proses pengambilan keputusan.

V.2 Saran

Beberapa saran untuk tugas akhir ini antara lain:

1. Tambahkan grafik agar visualisasi pengetahuan yang diperoleh dapat lebih membantu dalam pengambilan keputusan.
2. Tambahkan fungsi/*feature* baru yang belum ada dalam sistem ini, seperti fungsi untuk menampilkan seluruh tuple yang ada dari tabel yang dipilih.



DAFTAR PUSTAKA

Bradley, Julia Case, and Millspangh, Anita C., *Advanced Programming using Visual Basic .NET*, McGraw - Hill, 2003.

Dilly, Ruth, *Data Mining an Introduction*, Queens University Belfast, 1995.

Dunham, Margaret H., *Data Mining Introductory and Advanced Topics*, Prentice Hall, 2003.

Han, Jiawei, and Kamber, Micheline, *Data Mining Concepts and Techniques*, Morgan-Kaufman, 2001.

Yaeger, Deitel-Deitel, *Simply Visual Basic .NET 2003. An Application-Driven Tutorial Approach*, Prentice Hall, 2004.

MSDN Library - October 2004, Microsoft, 2004.

LAMPIRAN 1



SKPL

SPEKIFIKASI KEBUTUHAN PERANGKAT LUNAK

Penambangan Data dengan Algoritma Bayesian (DMBC - Data Mining Bayesian Classification)

Untuk :


PT. Iskandar Indah Printing Textile Surakarta

Dipersiapkan oleh:

Liliana Iskandar / 03493

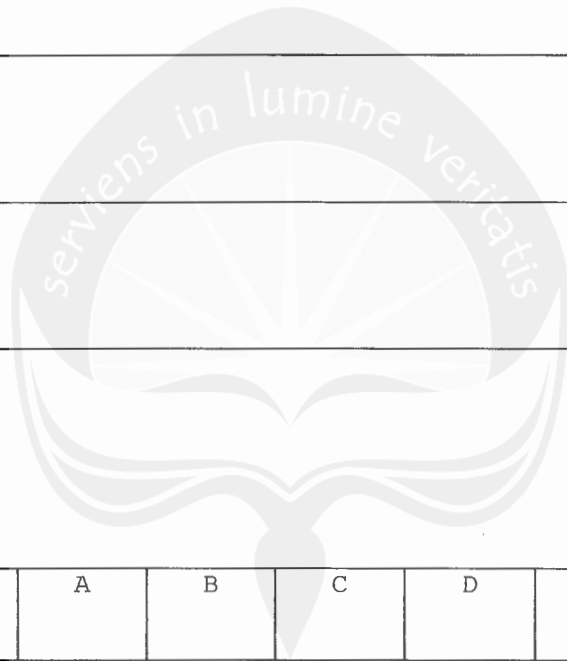
Program Studi Teknik Informatika - Fakultas Teknologi Industri

Universitas Atma Jaya Yogyakarta

	Program Studi Teknik Informatika	Nomor Dokumen		Halaman
		SKPL-DMBC		1/25
		Revisi		Tgl : 26-11-2005

DAFTAR PERUBAHAN

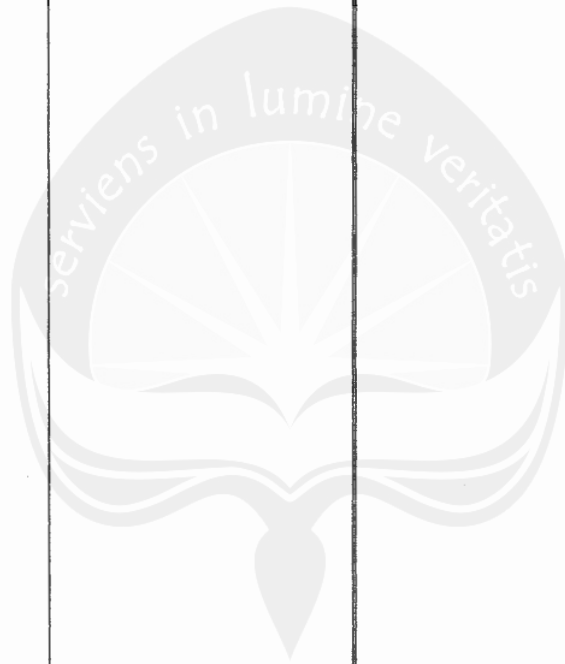
Revisi	Deskripsi
A	
B	
C	
D	
E	
F	



INDEX TGL	-	A	B	C	D	E	F	G
Ditulis oleh								
Diperik sa oleh								
Disetuj ui oleh								

Daftar Halaman Perubahan

Halaman	Revisi	Halaman	Revisi



Daftar Isi

1	Pendahuluan	7
1.1	Tujuan.....	7
1.2	Lingkup Masalah.....	7
1.3	Definisi, Akronim dan Singkatan.....	7
1.4	Referensi	7
1.5	Deskripsi umum (Overview)	7
2	Deskripsi Kebutuhan	9
2.1	Perspektif produk.....	9
2.2	Fungsi Produk.....	10
2.3	Karakteristik Pengguna	11
2.4	Batasan-batasan.....	11
3	Kebutuhan khusus.....	12
3.1	Kebutuhan antarmuka eksternal.....	12
3.1.1	Antarmuka pemakai	12
3.1.2	Antarmuka perangkat keras	12
3.1.3	Antarmuka perangkat lunak.....	12
3.2.1.1.1	Entitas data.....	13
3.2.1.1.1.1	Proses.....	13
3.2.1.1.1.2	Topologi.....	13
3.2.1.1.2	DFD Level 1 DMBC.....	14
3.2.1.1.2.1	Entitas Data	14
3.2.1.1.2.2	Proses.....	14
3.2.1.1.2.3	Topologi.....	15
3.2.1.1.3.1	Entitas Data	15
3.2.1.1.3.2	Proses.....	15
3.2.1.1.3.3	Topologi.....	16
3.2.1.2	Deskripsi proses	17
3.2.1.2.1	Proses Connect	17
3.2.1.2.1.1	Entitas data masukan.....	17
3.2.1.2.1.2	Algoritma atau formula dari proses.....	18
3.2.1.2.1.3	Entitas data terlibat.....	18
3.2.1.2.2	Proses Pemilihan Database dan tabel	18
3.2.1.2.2.1	Entitas data masukan.....	18
3.2.1.2.2.2	Algoritma atau formula dari proses.....	18
3.2.1.2.2.3	Entitas data terlibat.....	18
3.2.1.2.3	Proses Input Nilai Atribut.....	19
3.2.1.2.3.1	Entitas data masukan.....	19
3.2.1.2.3.2	Algoritma atau formula dari proses.....	19
3.2.1.2.3.3	Entitas data terlibat.....	19
3.2.1.2.4	Proses Pemilihan Field Atribut untuk Klasifikasi	19
3.2.1.2.4.1	Entitas data masukan.....	19
3.2.1.2.4.2	Algoritma atau formula dari proses.....	19
3.2.1.2.4.3	Entitas data terlibat.....	19
3.2.1.2.5	Proses Penghitungan probabilitas terjadinya kelas	20
3.2.1.2.5.1	Entitas data masukan.....	20
3.2.1.2.5.2	Algoritma atau formula dari proses.....	20
3.2.1.2.5.3	Entitas data terlibat.....	20
3.2.1.2.6	Proses Penghitungan Probabilitas Tuple per kelas.....	20
3.2.1.2.6.1	Entitas data masukan.....	20

3.2.1.2.6.2	Algoritma atau formula dari proses.....	20
3.2.1.2.6.3	Entitas data terlibat.....	21
3.2.1.2.7	Proses penghitungan kemungkinan kelas.....	21
3.2.1.2.7.1	Entitas data masukan.....	21
3.2.1.2.7.2	Algoritma atau formula dari proses.....	21
3.2.1.2.7.3	Entitas data terlibat.....	21
3.2.1.2.8	Proses penghitungan probabilitas tuple total.....	21
3.2.1.2.8.1	Entitas data masukan.....	21
3.2.1.2.8.2	Algoritma atau formula dari proses.....	22
3.2.1.2.8.3	Entitas data terlibat.....	22
3.2.1.2.9	Proses penghitungan probabilitas posterior.....	22
3.2.1.2.9.1	Entitas data masukan.....	22
3.2.1.2.9.2	Algoritma atau formula dari proses.....	22
3.2.1.2.9.3	Entitas data terlibat.....	22
4.	Spesifikasi data DMBC.....	23
4.1	Konstruksi data DMBC.....	23
5.	Kamus data.....	23
6.	Entity Relationship Diagram.....	25



Daftar Gambar

Gambar 1. DFD Level 0 DMBC	14
Gambar 2. DFD Level 1 DMBC.....	15
Gambar 3. DFD Level 2 Proses Klasifikasi Tuple.....	16
Gambar 4. Entity Relationship Diagram(ERD).....	25



1 Pendahuluan

1.1 Tujuan

Dokumen Spesifikasi Kebutuhan Perangkat Lunak (SKPL) tersebut merupakan dokumen spesifikasi kebutuhan perangkat lunak DMBC (Data Mining using Bayesian Classification) untuk mendefinisikan kebutuhan perangkat lunak, yang meliputi antarmuka eksternal, dan atribut, serta mendefinisikan fungsi perangkat lunak, juga mendefinisikan batasan perancangan perangkat lunak.

1.2 Lingkup Masalah

Perangkat Lunak DMBC dikembangkan dengan tujuan untuk melakukan penambangan data dengan menggunakan algoritma Bayesian. Algoritma Bayesian ini termasuk di dalam penambangan data model klasifikasi. Data yang akan digunakan adalah data karyawan yang ada dalam perusahaan. Untuk klasifikasi Bayesian ini, atribut-atribut yang dapat digunakan adalah atribut yang bersifat independen. Pertama-tama sistem akan menghitung nilai probabilitas prior yang mungkin dan menentukan kelas yang akan digunakan. Kemudian nilai-nilai tersebut akan digunakan untuk mengklasifikasikan tuple yang ada ke dalam kelas-kelas tersebut. Perangkat lunak penambangan data ini diharapkan dapat membantu perusahaan untuk mengklasifikasikan kinerja/prestasi seorang karyawan dilihat dari nilai atribut yang dimilikinya.

1.3 Definisi, Akronim dan Singkatan

Daftar definisi akronim dan singkatan :

Keyword/Phrase	Definisi
SKPL	Merupakan spesifikasi kebutuhan dari perangkat lunak yang akan dikembangkan.
SKPL-DMBC-XX-(YY)	Kode yang merepresentasikan kebutuhan pada DMBC (Data Mining Bayesian Classification).

	XX merupakan kode fungsi perangkat lunak yang ada dan (YY) merupakan kode subfungsi yang ada dalam DMBC (bersifat optional).
DFD	Data Flow Diagram merupakan teknis grafis yang menggambarkan aliran informasi dan transformasi yang diaplikasikan saat data bergerak dari input menjadi output
ERD	Entity Relationship Diagram merupakan teknis grafis/diagram yang menggambarkan objek dan hubungan antar objek.
DMBC	Data Mining Bayesian Classification, yaitu perangkat lunak penambangan data dengan algoritma klasifikasi Bayesian yang digunakan untuk menemukan informasi yang berguna dari kumpulan data karyawan pada perusahaan.

1.4 Referensi

Referensi yang digunakan pada perangkat lunak tersebut adalah:

1. GLO1, *Spesifikasi Kebutuhan Perangkat Lunak*, Jurusan Teknik Informatika - UAJY
2. Pressman Roger S., *Rekayasa Perangkat Lunak*, McGraw-Hill Book Co., Andi Yogyakarta, 1997.

1.5 Deskripsi umum (Overview)

Secara umum dokumen SKPL tersebut terbagi atas 3 bagian utama. Bagian utama berisi penjelasan mengenai dokumen SKPL tersebut yang mencakup tujuan pembuatan SKPL, ruang lingkup masalah dalam pengembangan perangkat lunak tersebut, definisi, referensi dan deskripsi umum tentang dokumen SKPL ini.

Bagian kedua berisi penjelasan umum tentang perangkat lunak DMBC yang akan dikembangkan, mencakup perspektif produk yang akan dikembangkan, fungsi perangkat lunak, karakteristik

pengguna, batasan dalam penggunaan perangkat lunak dan asumsi yang dipakai dalam pengembangan perangkat lunak DMBC tersebut.

Bagian ketiga berisi penjelasan secara lebih rinci tentang kebutuhan perangkat lunak DMBC yang akan dikembangkan.

2 Deskripsi Kebutuhan

2.1 Perspektif produk

DMBC adalah perangkat lunak yang dikembangkan untuk membantu pengguna dalam melakukan penambahan terhadap data karyawan perusahaan yang dapat memungkinkan ditemukannya informasi yang berguna untuk pengembangan perusahaan selanjutnya, seperti faktor apa saja yang dapat mempengaruhi prestasi kerja karyawan dalam perusahaan.

Pengguna berinteraksi dengan sistem yang ditampilkan dalam bentuk form yang dapat diisi. Kemudian sistem akan melakukan penghitungan nilai-nilai probabilitas suatu tuple untuk diklasifikasikan ke dalam kelas-kelas tertentu, dimana hasil pengklasifikasian ini dapat digunakan untuk mengambil keputusan yang lebih lanjut. Secara garis besar, proses diawali dengan adanya proses login dari pengguna, dimana sistem ini hanya dapat diakses oleh pihak-pihak yang berwenang saja, karena yang dapat mengakses sistem ini hanyalah pihak-pihak yang memiliki `login_id` dan `password` yang valid saja.

Inputan data yang dimasukkan berupa `login_id` dan `password`. Jika `login_id` dan `password` yang diinputkan valid, maka pengguna berhak untuk mengakses basis data karyawan. Kemudian ia dapat menginputkan nilai-nilai atribut dari tuple yang akan diklasifikasikan. Nilai-nilai tersebut yang akan digunakan untuk mengklasifikasikan tuple tersebut ke dalam kelas-kelas yang ada. Kemudian sistem akan melakukan penghitungan nilai-nilai probabilitas yang ada dengan menggunakan algoritma klasifikasi *naive Bayesian*.

Program Studi Teknik Informatika	SKPL-DMBC	9/ 25
Dokumen ini dan informasi yang dimilikinya adalah milik Program Studi Teknik Informatika-UAJY dan bersifat rahasia. Dilarang untuk me-reproduksi dokumen ini tanpa diketahui oleh Program Studi Teknik Informatika		

2.2 Fungsi Produk

Fungsi produk perangkat lunak DMBC adalah sebagai berikut :

2.2.1 Fungsi *Connect* (**SKPL-DMBC-01**) adalah fungsi yang digunakan oleh pengguna untuk dapat terkoneksi dengan server database.

2.2.2 Fungsi *Select database & table* (**SKPL-DMBC-02**) adalah fungsi yang digunakan untuk memilih database dan tabel yang akan ditambang.

2.2.3 Fungsi *Input Attribute Values* (**SKPL-DMBC-03**) adalah fungsi yang digunakan untuk menginputkan nilai-nilai atribut dari tuple baru.

2.2.4 Fungsi *Select Attribute Fields* (**SKPL-DMBC-04**) adalah fungsi yang digunakan untuk memilih field atribut yang akan digunakan untuk mengklasifikasikan tuple baru.

2.2.5 Fungsi *Klasifikasi Bayes* (**SKPL-DMBC-05**) adalah fungsi yang digunakan untuk mengklasifikasikan tuple dengan nilai-nilai atribut yang telah diinputkan sebelumnya dan berdasarkan field atribut yang dipilih. Dalam fungsi ini terdapat proses penghitungan beberapa nilai probabilitas, diantaranya :

2.2.5.1 Penghitungan Nilai $P(C_j)$

$$P(C_j) = \frac{\text{banyaknya suatu kelas terjadi dalam basis data}}{\text{banyaknya tuple dalam basis data}}$$

2.2.5.2 Penghitungan Nilai $P(t_i|C_j)$

$$P(t_i | C_j) = \prod_{k=1}^p P(x_{ik} | C_j)$$

2.2.5.3 Penghitungan Nilai *Likelihood of being C_j*

Program Studi Teknik Informatika	SKPL-DMBC	10/ 25
Dokumen ini dan informasi yang dimilikinya adalah milik Program Studi Teknik Informatika-UAJY dan bersifat rahasia. Dilarang untuk me-reproduksi dokumen ini tanpa diketahui oleh Program Studi Teknik Informatika		

$$\text{Likelihood of being } (C_j) = P(t_i | C_j) * P(C_j)$$

2.2.5.4 Penghitungan Nilai $P(t)$

$$P(t) = \sum_{j=1}^n \text{Likelihood of being } (C_j)$$

2.2.5.5 Penghitungan Nilai $P(C_j | t_i)$

$$P(C_j | t_i) = \frac{\text{Likelihood of being } (C_j)}{P(t)}$$

2.2.6 Fungsi *View Report* (**SKPL-DMBC-06**) adalah suatu fungsi yang digunakan untuk melihat report hasil pengklasifikasian tuple.

2.2.7 Fungsi *Print Report* (**SKPL-DMBC-07**) adalah fungsi yang digunakan untuk melakukan pencetakan report melalui printer.

2.3 Karakteristik Pengguna

Karakteristik pengguna perangkat lunak DMBC tersebut adalah sebagai berikut :

1. Memahami pengoperasian komputer secara aktif.
2. Memahami sistem komputer tempat perangkat lunak dijalankan.
3. Mengerti proses pengelolaan data sistem.

2.4 Batasan-batasan

Batasan-batasan dalam pengembangan perangkat lunak DMBC tersebut adalah :

1. Kebijakan Umum

Berpedoman pada tujuan dari pengembangan perangkat lunak DMBC.

2. Keterbatasan perangkat keras

Program Studi Teknik Informatika	SKPL-DMBC	11/ 25
Dokumen ini dan informasi yang dimilikinya adalah milik Program Studi Teknik Informatika-UAJY dan bersifat rahasia. Dilarang untuk me-reproduksi dokumen ini tanpa diketahui oleh Program Studi Teknik Informatika		

Dapat diketahui kemudian setelah sistem ini berjalan (sesuai dengan kebutuhan).

3. Kebutuhan keandalan

Pengembangan perangkat lunak ini dibatasi pada kemudahan penggunaan dan kecepatan dalam proses pengolahannya.

3 Kebutuhan khusus

3.1 Kebutuhan antarmuka eksternal

Kebutuhan antar muka eksternal pada perangkat lunak DMBC meliputi kebutuhan antarmuka pemakai, antarmuka perangkat keras, dan antarmuka perangkat lunak.

3.1.1 Antarmuka pemakai

Pengguna berinteraksi dengan antarmuka yang ditampilkan dalam layar komputer dengan format form dengan pilihan fungsi untuk pengisian data dan tampilan informasi pada layar monitor.

3.1.2 Antarmuka perangkat keras

Antarmuka perangkat keras yang digunakan dalam perangkat lunak DMBC adalah:

1. PC
2. Keyboard dan mouse
3. Printer

3.1.3 Antarmuka perangkat lunak

Perangkat lunak yang dibutuhkan untuk mengoperasikan perangkat lunak DMBC adalah sebagai berikut :

1. Nama : Visual Basic .NET 2003
Sumber : Microsoft

Sebagai perangkat lunak yang digunakan untuk membuat perangkat lunak DMBC.

2. Nama : SQL Server 2000

Program Studi Teknik Informatika	SKPL-DMBC	12/ 25
Dokumen ini dan informasi yang dimilikinya adalah milik Program Studi Teknik Informatika-UAJY dan bersifat rahasia. Dilarang untuk me-reproduksi dokumen ini tanpa diketahui oleh Program Studi Teknik Informatika		

Sumber : Microsoft

Sebagai database yang dibutuhkan dalam mengoperasikan perangkat lunak DMBC.

3. Nama : Windows 2000/XP

Sumber : *Microsoft.*

Sebagai sistem operasi komputer.

3.2 Kebutuhan fungsional

3.2.1 Aliran informasi

3.2.1.1 Diagram Konteks DMBC

3.2.1.1.1 Entitas data

Entitas eksternal yang terlibat dalam pengembangan perangkat lunak DMBC tersebut dinyatakan dalam tabel berikut :

Nama	Kode
Pengguna	User

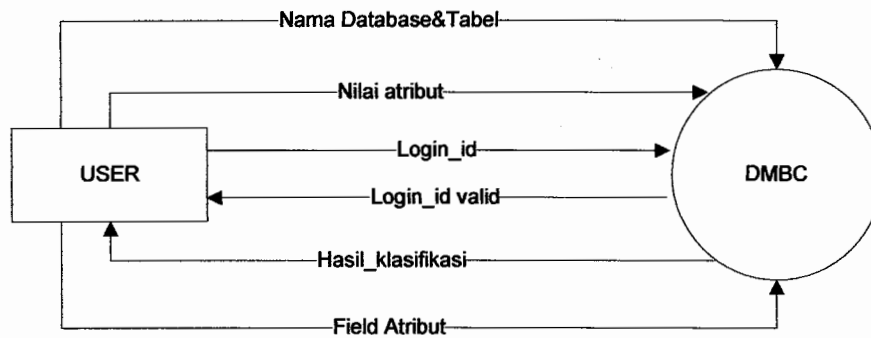
Seluruh entitas yang didefinisikan dalam tabel tersebut merupakan entitas yang terlibat dalam seluruh proses yang terjadi dalam perangkat lunak DMBC tersebut.

3.2.1.1.1.1 Proses

Proses yang terjadi dalam perangkat lunak DMBC tersebut adalah sistem menerima input data login berupa login_id dan password. Jika login_id dan password yang diinputkan adalah valid, maka sistem akan menampilkan form-form yang meminta inputan dari user. Input user inilah yang selanjutnya akan diproses menjadi informasi yang dikehendaki.

3.2.1.1.1.2 Topologi

Topologi dari proses perangkat lunak DMBC dapat dilihat pada Gambar 1 DFD Level 0.



DFD Level 0 DMBC

3.2.1.1.2 DFD Level 1 DMBC

3.2.1.1.2.1 Entitas Data

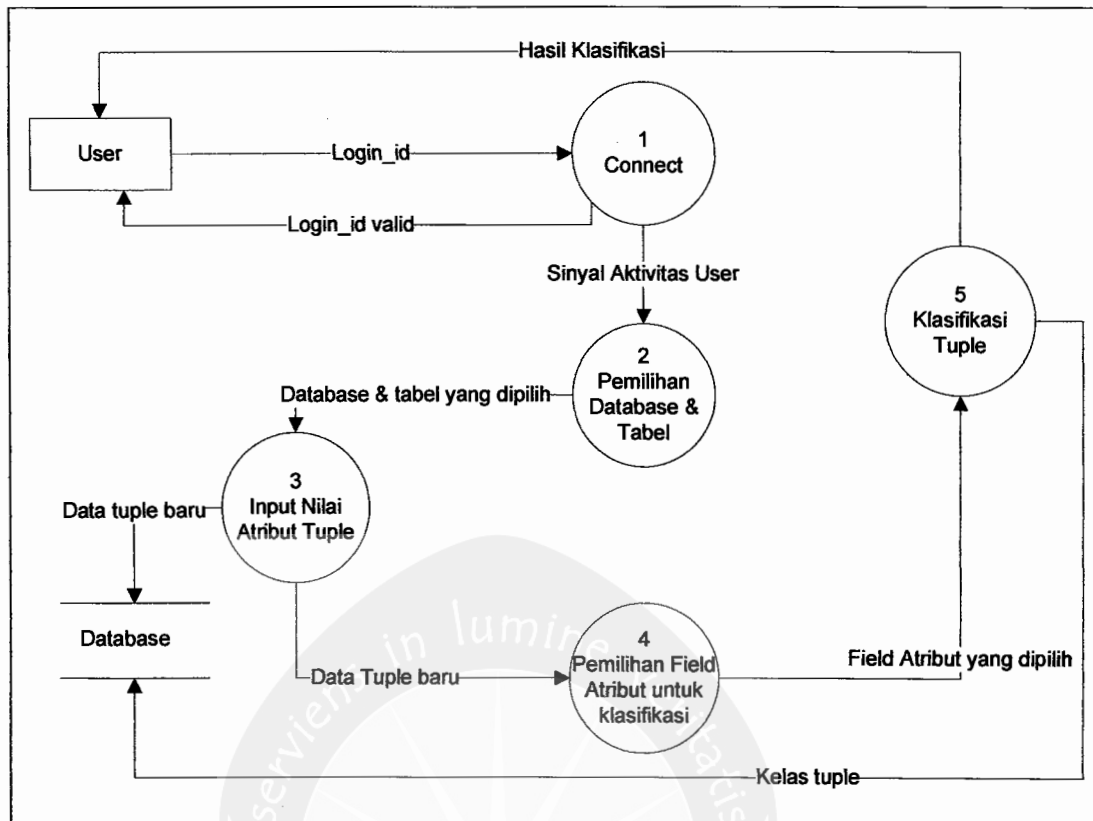
Entitas data eksternal sesuai dengan entitas data pada DFD Level 0.

3.2.1.1.2.2 Proses

Proses yang terjadi dalam DFD Level 1 mencakup 5 bagian, yaitu :

1. **Connect**, adalah suatu proses untuk koneksi ke SQL Server.
2. **Pemilihan database dan tabel**, adalah suatu proses untuk memilih database yang ada pada server yang akan dipakai untuk melakukan penambahan data dan memilih tabel yang akan digunakan dalam database yang telah dipilih
3. **Pengisian nilai atribut**, adalah suatu proses untuk menginputkan nilai-nilai atribut tuple yang akan diklasifikasikan.
4. **Pemilihan field atribut**, adalah suatu proses memilih atribut apa saja yang akan digunakan untuk mengklasifikasikan suatu tuple yang baru.
5. **Pengklasifikasian tuple**, adalah suatu proses untuk melakukan klasifikasi terhadap suatu tuple berdasarkan nilai-nilai atribut yang diinputkan dengan menggunakan algoritma Bayesian.

3.2.1.1.2.3 Topologi



Gambar 2. DFD Level 1 DMBC

3.2.1.1.3 DFD Level 2 Proses Klasifikasi Tuple

3.2.1.1.3.1 Entitas Data

Entitas data yang terlibat dalam proses klasifikasi tuple tersebut adalah nilai atribut-atribut tuple yang ada dan kelas-kelas yang ada.

3.2.1.1.3.2 Proses

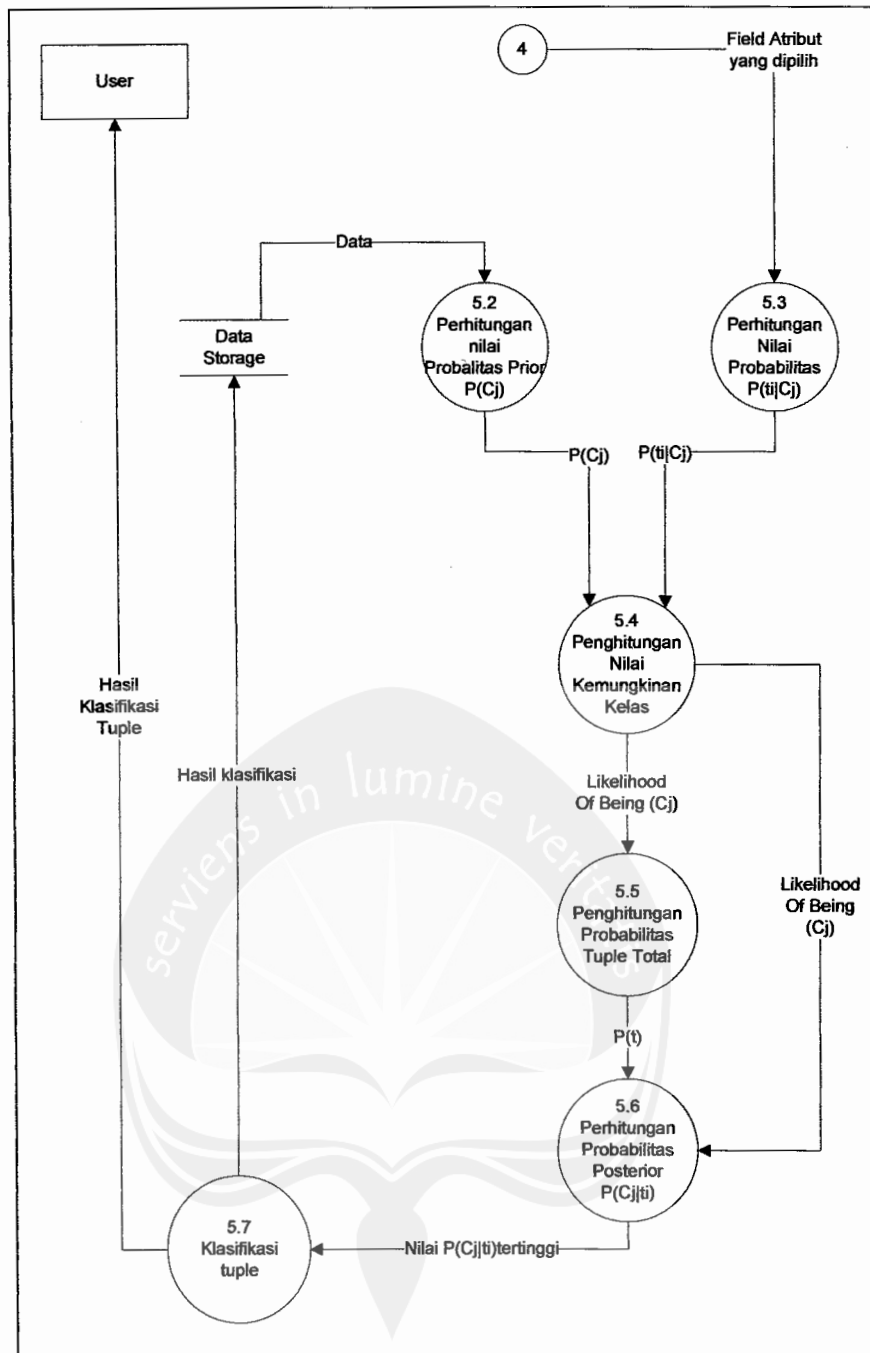
Proses yang terjadi dalam DFD Level 2 proses klasifikasi tuple dikelompokkan atas 5 bagian yaitu :

1. **Penghitungan probabilitas terjadinya kelas ($P(class)$)**, adalah suatu proses untuk menghitung nilai probabilitas terjadinya suatu keadaan kelas dalam basis data.
2. **Penghitungan Probabilitas Tuple per kelas ($P(tuple|class)$)**, adalah suatu proses untuk menghitung

probabilitas tuple berdasarkan nilai atribut yang bersangkutan.

3. **Penghitungan Kemungkinan Kelas (Likelihood of being class)**, adalah suatu proses untuk menghitung nilai kemungkinan suatu kelas yang merupakan hasil kali dari probabilitas tuple per kelas dengan probabilitas terjadinya kelas.
4. **Penghitungan Probabilitas tuple total ($P(t)$)**, adalah suatu proses untuk menghitung probabilitas total suatu tuple. Nilai ini akan digunakan sebagai perbandingan untuk setiap kelas yang ada.
5. **Penghitungan probabilitas aktual per kelas ($P(\text{short}\text{tuple})$)**, adalah suatu proses untuk menghitung probabilitas aktual bagi tuple untuk setiap kelas yang ada. Kelas dengan nilai probabilitas aktual tertinggi merupakan kelas dari tuple yang dimaksud.

3.2.1.1.3.3 Topologi



Gambar 3. DFD Level 2 Proses Klasifikasi Tuple

3.2.1.2 Deskripsi proses

3.2.1.2.1 Proses Connect

3.2.1.2.1.1 Entitas data masukan

Entitas data masukan dalam subproses connect ini tergantung pada tipe koneksi yang digunakan. Jika koneksi dengan Windows Authentication yang dipilih, maka tidak diperlukan input yang

berupa login_id dan password. Namun, jika koneksi dengan menggunakan SQL Server Authentication yang dipilih, maka diperlukan inputan login_id dan password.

3.2.1.2.1.2 Algoritma atau formula dari proses

Subproses connect tersebut akan menerima masukan berupa data login yang kemudian sistem akan mencocokkan data login tersebut, sehingga jika valid maka pengguna dapat melanjutkan ke proses berikutnya.

3.2.1.2.1.3 Entitas data terlibat

Entitas data yang terlibat dalam subproses connect adalah jenis koneksi yang dipilih, login_id dan password dari pengguna.

3.2.1.2.2 Proses Pemilihan Database dan tabel

3.2.1.2.2.1 Entitas data masukan

Entitas data masukan dalam subproses pemilihan database dan tabel adalah nama database yang ada pada server yang dipilih dan nama tabel yang ada pada database yang telah dipilih.

3.2.1.2.2.2 Algoritma atau formula dari proses

Subproses dalam pemilihan database adalah setelah user berhasil melakukan proses login, sistem akan menampilkan suatu form yang digunakan untuk menampilkan seluruh database yang ada pada server yang dipilih, sehingga user dapat memilih nama database yang akan digunakan dalam proses data mining. Setelah nama database dipilih, maka nama tabel yang ada dalam database yang dipilih juga akan ditampilkan, sehingga user dapat memilih nama tabel yang akan digunakan.

3.2.1.2.2.3 Entitas data terlibat

Entitas data yang terlibat dalam subproses pemilihan database adalah nama database yang dipilih dan nama tabel yang dipilih.

Program Studi Teknik Informatika	SKPL-DMBC	18/ 25
Dokumen ini dan informasi yang dimilikinya adalah milik Program Studi Teknik Informatika-UAJY dan bersifat rahasia. Dilarang untuk me-reproduksi dokumen ini tanpa diketahui oleh Program Studi Teknik Informatika		

3.2.1.2.3 Proses Input Nilai Atribut

3.2.1.2.3.1 Entitas data masukan

Entitas data masukan dalam subproses input nilai atribut adalah nilai masukan dari user untuk masing-masing field atribut tuple.

3.2.1.2.3.2 Algoritma atau formula dari proses

Subproses dalam input nilai atribut ini adalah setelah database dan nama tabel dipilih, sistem akan menampilkan suatu form yang digunakan untuk proses pemasukkan nilai atribut. User hanya dapat memilih nilai-nilai yang telah tersedia. Nilai harus diinputkan untuk setiap field atribut yang ada. Setelah semua field atribut terisi, sistem akan menyimpannya sebagai suatu record baru dalam database.

3.2.1.2.3.3 Entitas data terlibat

Entitas data yang terlibat dalam subproses input nilai atribut adalah nilai-nilai atribut yang diinputkan user.

3.2.1.2.4 Proses Pemilihan Field Atribut untuk Klasifikasi

3.2.1.2.4.1 Entitas data masukan

Entitas data masukan dalam subproses pemilihan field atribut adalah field-field tuple yang dipilih.

3.2.1.2.4.2 Algoritma atau formula dari proses

Subproses dalam pemilihan field atribut adalah sistem akan menampilkan field-field yang ada dalam suatu listbox. Kemudian user dapat memilih field-field yang ada sebagai field yang akan digunakan untuk mengklasifikasikan tuple.

3.2.1.2.4.3 Entitas data terlibat

Entitas data yang terlibat dalam subproses pemilihan field atribut adalah field-field yang dipilih dan nilai dari setiap field atribut yang dipilih.

3.2.1.2.5 Proses Penghitungan probabilitas terjadinya kelas

3.2.1.2.5.1 Entitas data masukan

Entitas data masukan dalam subproses penghitungan probabilitas terjadinya kelas ($P(C_j)$) atau probabilitas proir adalah banyaknya suatu kelas terjadi dalam basis data dan banyaknya tuple dalam basis data.

3.2.1.2.5.2 Algoritma atau formula dari proses

Subproses penghitungan probabilitas terjadinya kelas adalah memperoleh suatu nilai tertentu yang merupakan hasil pembagian dari banyaknya suatu keals terjadi dalam basis data dan banyaknya tuple dalam basis data.

$$P(C_j) = \frac{\text{banyaknya suatu kelas terjadi dalam basis data}}{\text{banyaknya tuple dalam basis data}}$$

3.2.1.2.5.3 Entitas data terlibat

Entitas data yang terlibat adalah banyaknya suatu kelas terjadi dalam basis data dan banyaknya tuple dalam basis data.

3.2.1.2.6 Proses Penghitungan Probabilitas Tuple per kelas

3.2.1.2.6.1 Entitas data masukan

Entitas data masukan dalam subproses penghitungan probabilitas tuple per kelas ($P(t_i|C_j)$) adalah nilai probabilitas terjadinya event per atribut untuk masing-masing kelas.

3.2.1.2.6.2 Algoritma atau formula dari proses

Subproses penghitungan probabilitas tuple per kelas adalah memperoleh suatu nilai tertentu yang merupakan hasil perkalian

dari nilai probabilitas terjadinya event untuk masing-masing kelas sesuai dengan masing-masing atribut ($P(x_{ik}|C_j)$).

$$P(t_i | C_j) = \prod_{k=1}^p P(x_{ik} | C_j)$$

3.2.1.2.6.3 Entitas data terlibat

Entitas data yang terlibat adalah nilai-nilai probabilitas terjadinya event per atribut yang ada untuk masing-masing kelas.

3.2.1.2.7 Proses Penghitungan Kemungkinan Kelas

3.2.1.2.7.1 Entitas data masukan

Entitas data masukan dalam subproses penghitungan kemungkinan kelas (*Likelihood of being* (C_j)) adalah nilai probabilitas terjadinya kelas ($P(C_j)$) dan nilai probabilitas tuple per kelas ($P(t_i|C_j)$).

3.2.1.2.7.2 Algoritma atau formula dari proses

Subproses penghitungan kemungkinan kelas adalah memperoleh suatu nilai tertentu yang merupakan hasil perkalian dari nilai probabilitas terjadinya kelas ($P(C_j)$) dengan nilai probabilitas tuple per kelas ($P(t_i|C_j)$).

$$\text{Likelihood of being } (C_j) = P(t_i | C_j) * P(C_j)$$

3.2.1.2.7.3 Entitas data terlibat

Entitas data yang terlibat adalah nilai probabilitas terjadinya kelas dengan nilai probabilitas tuple per kelas.

3.2.1.2.8 Penghitungan Probabilitas tuple total

3.2.1.2.8.1 Entitas data masukan

Entitas data masukan dalam subproses penghitungan probabilitas tuple total ($P(t)$) adalah nilai kemungkinan kelas (*Likelihood of being* (C_j)) yang ada.

3.2.1.2.8.2 Algoritma atau formula dari proses

Subproses penghitungan probabilitas tuple total adalah memperoleh suatu nilai tertentu yang merupakan hasil penjumlahan seluruh nilai kemungkinan kelas dari masing-masing kelas yang ada.

$$P(t) = \sum_{j=1}^n \text{Likelihood of being } (C_j)$$

3.2.1.2.8.3 Entitas data terlibat

Entitas data yang terlibat adalah nilai kemungkinan kelas yang ada.

3.2.1.2.9 Penghitungan probabilitas posterior

3.2.1.2.9.1 Entitas data masukan

Entitas data masukan dalam subproses penghitungan probabilitas posterior ($P(C_j|t_i)$) adalah nilai kemungkinan kelas (*Likelihood of being* (C_j)) dan nilai probabilitas tuple total ($P(t)$).

3.2.1.2.9.2 Algoritma atau formula dari proses

Subproses penghitungan probabilitas posterior adalah memperoleh suatu nilai tertentu yang merupakan hasil pembagian antara nilai kemungkinan kelas (*Likelihood of being* (C_j)) dengan nilai probabilitas tuple total ($P(t)$).

$$P(C_j|t_i) = \frac{\text{Likelihood of being } (C_j)}{P(t)}$$

3.2.1.2.9.3 Entitas data terlibat

Entitas data yang terlibat adalah nilai kemungkinan kelas dan nilai probabilitas tuple total.

4. Spesifikasi data DMBC

4.1 Konstruksi data DMBC

Tipe record dalam konstruksi data DMBC adalah :

- a. Data User (berisi informasi tentang user)
- b. Data karyawan (berisi informasi tentang Data karyawan)

5. Kamus data

5.1 Data login

5.1.1 Elemen Data User_Id

Representasi	Domain	Range	Format	Presisi	Struktur Data
Untuk Id dari User	Text	-	-	-	Char(5)

5.1.2 Elemen Data Password

Representasi	Domain	Range	Format	Presisi	Struktur Data
Untuk Password dari User	Text	-	xxxxxx	All numeric	Char(5)

5.2 Data Karyawan

5.3.1 Elemen Data No_induk

Representasi	Domain	Range	Format	Presisi	Struktur Data
Untuk Id dari karyawan	Nilai	0 sampai 2,147,483,647	-	-	Integer

5.3.2 Elemen Data Lama_kerja

Representasi	Domain	Range	Format	Presisi	Struktur Data
Untuk lama kerja	Text	< 5 thn, 5 - 15 thn, > 15	-	-	Char(8)

karyawan		thn			
----------	--	-----	--	--	--

5.3.3 Elemen Data Presensi

Representasi	Domain	Range	Format	Presisi	Struktur Data
Untuk presensi kehadiran kerja karyawan	Text	> 75%, 50% - 75%, < 50%	-	-	Char(8)

5.3.4 Elemen Data Produktivitas

Representasi	Domain	Range	Format	Presisi	Struktur Data
Untuk produktifitas karyawan	text	Bagus, kurang	-	-	Char(8)

5.3.5 Elemen Data Kemampuan kerja

Representasi	Domain	Range	Format	Presisi	Struktur Data
Untuk kemampuan kerja karyawan	text	Bagus, kurang	-	-	Char(8)

5.3.6 Elemen Data Tanggung jawab

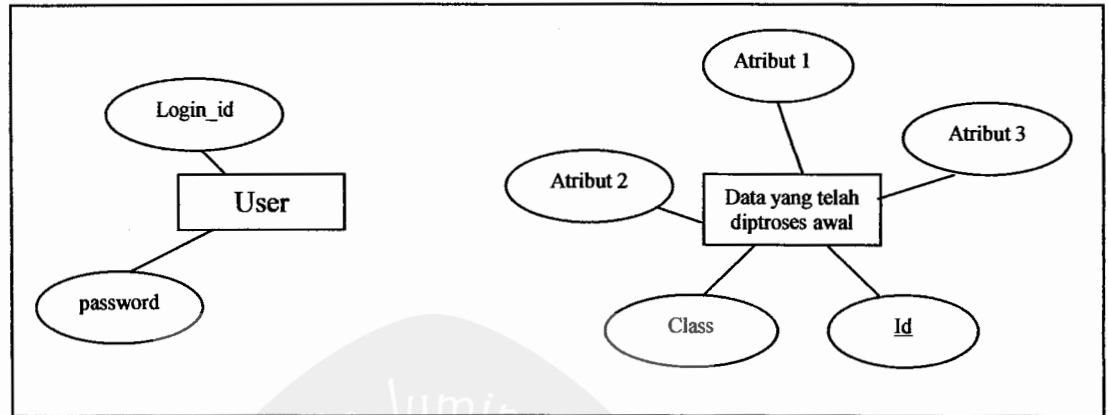
Representasi	Domain	Range	Format	Presisi	Struktur Data
Untuk tanggung jawab karyawan	text	Bagus, kurang	-	-	Char(8)

5.3.7 Elemen Data Class

Representasi	Domain	Range	Format	Presisi	Struktur Data

Untuk prestasi kerja karyawan	text	Baik, cukup, kurang	-	-	Char(8)
-------------------------------	------	---------------------	---	---	---------

6. Entity Relationship Diagram



Gambar 4. Entity Relationship Diagram

LAMPIRAN 2



DPPL

DESKRIPSI PERANCANGAN PERANGKAT LUNAK

Penambangan Data dengan Algoritma Bayesian

(DMBC - Data Mining Bayesian Classification)

Untuk :


PT. Iskandar Indah Printing Textile Surakarta

Dipersiapkan oleh:

Liliana Iskandar / 03493

Program Studi Teknik Informatika - Fakultas Teknologi Industri

Universitas Atma Jaya Yogyakarta

	Program Studi Teknik Informatika	Nomor Dokumen		Halaman
		DPPL-DMBC		1/16
		Revisi		Tgl : 3-12-2005

Program Studi Teknik Informatika	DPPL-DMBC	1 / 16
Dokumen ini dan informasi yang dimilikinya adalah milik Program Studi Teknik Informatika-UAJY dan bersifat rahasia. Dilarang untuk me-reproduksi dokumen ini tanpa diketahui oleh Program Studi Teknik Informatika		

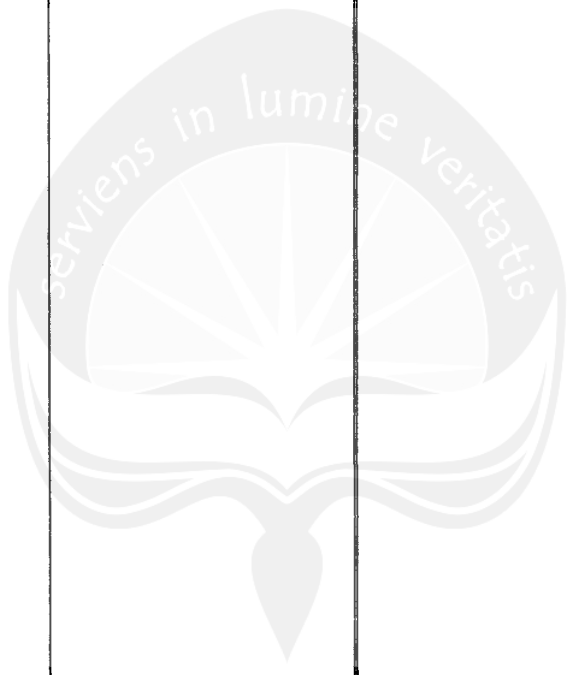
DAFTAR PERUBAHAN

Revisi	Deskripsi
A	
B	
C	
D	
E	
F	
G	

INDEX TGL	-	A	B	C	D	E	F	G
Ditulis oleh								
Diperik sa oleh								
Disetuj ui oleh								

Daftar Halaman Perubahan

Halaman	Revisi	Halaman	Revisi



Daftar Isi

1	Pendahuluan	6
1.1	Tujuan.....	6
1.2	Ruang Lingkup.....	6
1.3	Definisi dan Akronim.....	6
1.4	Referensi.....	7
2	Deskripsi Dekomposisi.....	7
2.1	Dekomposisi Data	7
2.1.1	<i>Deskripsi Entitas Data Login</i>	7
2.1.2	<i>Deskripsi Entitas Data Karyawan</i>	7
3	Dekomposisi Modul	8
3.1	Rancangan Arsitektur	8
4	Perancangan Antarmuka dan Fungsional	9
4.1	Antarmuka Halaman Utama.....	9
4.2	Antarmuka Connect to SQL Server.....	9
4.2.1	<i>Deskripsi Tombol OK</i>	10
4.2.2	<i>Deskripsi Tombol Cancel</i>	10
4.2.3	<i>Deskripsi Tombol Help</i>	10
4.3	Antarmuka Select Database dan Table.....	11
4.3.1	<i>Deskripsi Tombol New Tuple</i>	11
4.3.2	<i>Deskripsi Tombol Cancel</i>	12
4.4	Antarmuka Input Nilai Tuple Baru	12
4.4.1	<i>Deskripsi Tombol New Tuple</i>	12
4.4.2	<i>Deskripsi Tombol Add</i>	13
4.4.3	<i>Deskripsi Tombol Edit Tuple</i>	13
4.4.4	<i>Deskripsi Tombol Save</i>	13
4.4.5	<i>Deskripsi Tombol Cancel</i>	13
4.5	Antarmuka Select Attribute Fields	13
4.5.1	<i>Deskripsi Tombol ></i>	14
4.5.2	<i>Deskripsi Tombol >></i>	14
4.5.3	<i>Deskripsi Tombol <</i>	14
4.5.4	<i>Deskripsi Tombol <<</i>	15
4.5.5	<i>Deskripsi Tombol Start Classify</i>	15
4.5.6	<i>Deskripsi Tombol Cancel</i>	15
4.5.7	<i>Deskripsi Tombol View Report</i>	15
4.6	Antarmuka About DMBC	15
4.7	Antarmuka Credits DMBC.....	16

Daftar Gambar

<i>Gambar 1. Conceptual Data Model</i>	8
<i>Gambar 2. Rancangan Arsitektur DMBC</i>	8
<i>Gambar 3. Desain Antarmuka Menu Utama</i>	9
<i>Gambar 4. Desain Antarmuka Connect</i>	10
<i>Gambar 5. Desain Antarmuka Select Database & Table</i>	11
<i>Gambar 6. Desain Antarmuka Input Attribute Values</i>	12
<i>Gambar 7. Desain Antarmuka Select Attribute Fields</i>	14
<i>Gambar 8. Desain Antarmuka About DMBC</i>	16
<i>Gambar 9. Desain Antarmuka Credits DMBC</i>	16



<i>Tabel 1. Deskripsi Entitas Data Login</i>	7
<i>Tabel 2. Deskripsi Entitas Data Karyawan</i>	7

1 Pendahuluan

1.1 Tujuan

Dokumen Deskripsi Perancangan Perangkat Lunak (DPPL) bertujuan untuk mendefinisikan perancangan perangkat lunak yang akan dikembangkan. Dokumen DPPL tersebut digunakan oleh pengembang perangkat lunak sebagai acuan untuk implementasi pada tahap selanjutnya.

1.2 Ruang Lingkup

Perangkat Lunak DMBC dikembangkan dengan tujuan untuk melakukan penambangan data terhadap data karyawan pada perusahaan PT. Iskandar Indah Printing Textile dengan menggunakan metode klasifikasi. Tujuan dari pengklasifikasian ini adalah untuk mengetahui apakah seorang karyawan termasuk dalam kelas prestasi yang baik, cukup, atau kurang berdasarkan nilai atribut yang dimiliki. Untuk mengklasifikasikan data karyawan tersebut, algoritma yang digunakan adalah algoritma *naïve Bayesian classification*.

1.3 Definisi dan Akronim

Daftar definisi akronim dan singkatan :

Keyword/Phrase	Definisi
DPPL	Deskripsi Perancangan Perangkat Lunak disebut juga Software Design Description (SDD) merupakan deskripsi dari perancangan produk / perangkat lunak yang akan dikembangkan.
DMBC	Data Mining Bayesian Classification, yaitu perangkat lunak penambangan data dengan algoritma klasifikasi Bayesian yang digunakan untuk menemukan informasi yang berguna dari kumpulan data karyawan pada perusahaan.

1.4 Referensi

Referensi yang digunakan pada perangkat lunak tersebut adalah:

1. GLO2, *Deskripsi Perancangan Perangkat Lunak*, Program Studi Teknik Informatika - UAJY
2. Pressman Roger S, *Rekayasa Perangkat Lunak*, McGraw-Hill Book Co., Andi Yogyakarta, 1997

2 Deskripsi Dekomposisi

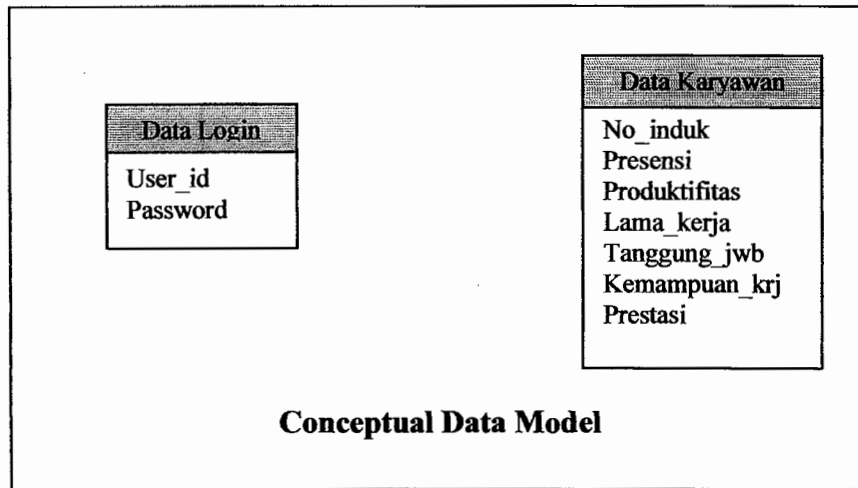
2.1 Dekomposisi Data

2.1.1 Deskripsi Entitas Data Login

Nama	Tipe	Panjang	Keterangan
User_id	Character	10	User_id pengguna, foreign key
Password	Character	10	Password pengguna, foreign key

2.1.2 Deskripsi Entitas Data Karyawan

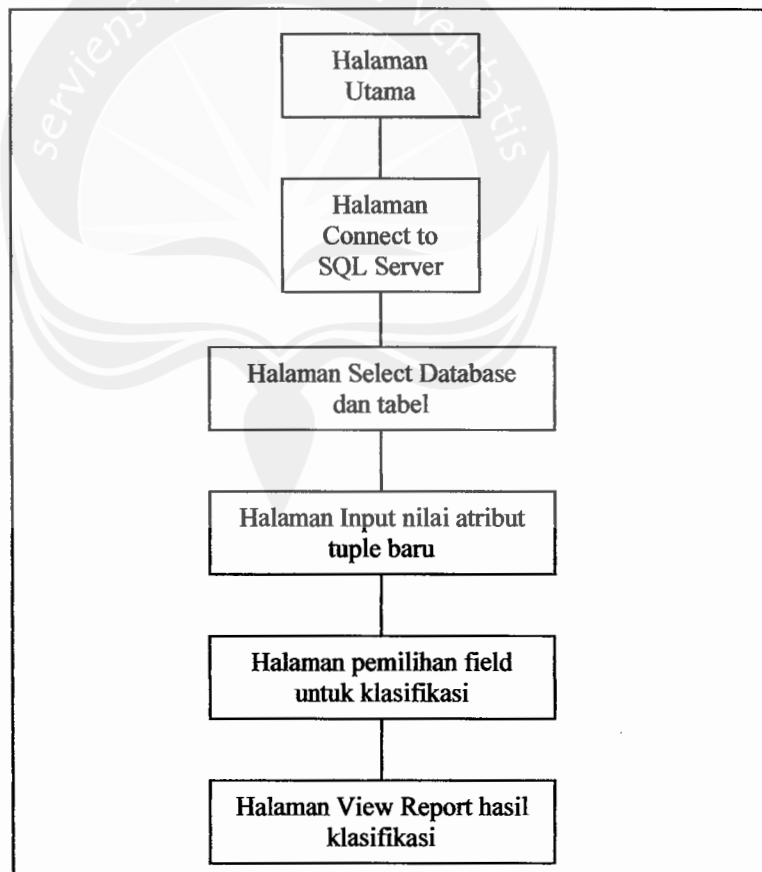
Nama	Tipe	Panjang	Keterangan
No_induk	Character	10	Id karyawan, primary key
Presensi	Character	10	Persentase kehadiran kerja karyawan
Lama_kerja	Integer	8	Lama kerja karyawan
Produktifitas	Character	10	Produktifitas kerja karyawan
Kemampuan_kerja	Character	10	Kemampuan kerja yang dimiliki karyawan
Tanggung_jawab	Character	10	Tanggung jawab karyawan dalam bekerja
Class	Character	10	Kelas prestasi kerja karyawan



Gambar 1. Conceptual Data Model

3 Dekomposisi Modul

3.1 Rancangan Arsitektur

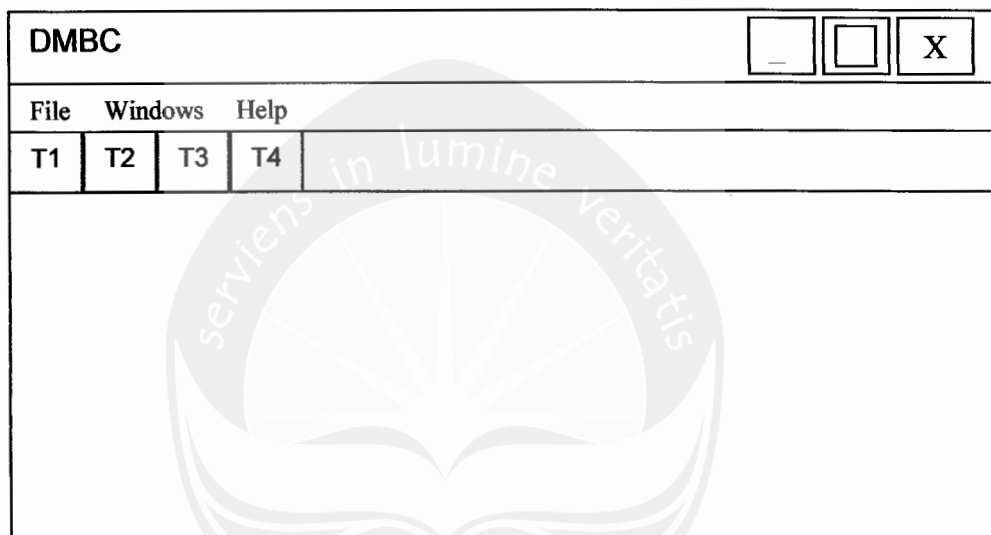


Gambar 2. Rancangan Arsitektur DMBC

4 Perancangan Antarmuka dan Fungsional

4.1 Antarmuka Halaman Utama

Merupakan halaman utama (MDIMain) yang merupakan form induk dari DMBC. Dalam form ini terdapat beberapa menu diantaranya : menu File (yang terdiri dari sub menu : connect, disconnect, dan exit) ; menu Windows (yang terdiri dari sub menu : Tile Horizontally, Tile Vertically, dan Cascade) ; menu Help (berisi Help tentang DMBC). Selain itu juga terdapat toolbar yang berisi icon-icon dari fungsi yang sering digunakan.



Gambar 3 Desain Antarmuka Halaman Utama

4.2 Antarmuka Connect to SQL Server

Merupakan form yang digunakan untuk melakukan koneksi ke database SQL Server. Prosesnya diawali dengan memilih server yang dituju dan memilih jenis koneksi yang digunakan. Jika jenis koneksi yang digunakan adalah Windows Authentication maka tidak diperlukan login_id dan password. Namun, jika jenis koneksi yang dipilih adalah SQL Server Authentication, maka diperlukan input user yang berupa login_id dan password.

Gambar 4 Desain Antarmuka Connect to SQL Server

4.2.1 Deskripsi Tombol OK

Merupakan tombol untuk mengeksekusi User ID dan password yang telah dimasukan sebelumnya untuk masuk ke dalam sistem. Sistem akan mengecek kebenarannya sesuai dengan jenis koneksi yang digunakan.

4.2.2 Deskripsi Tombol Cancel

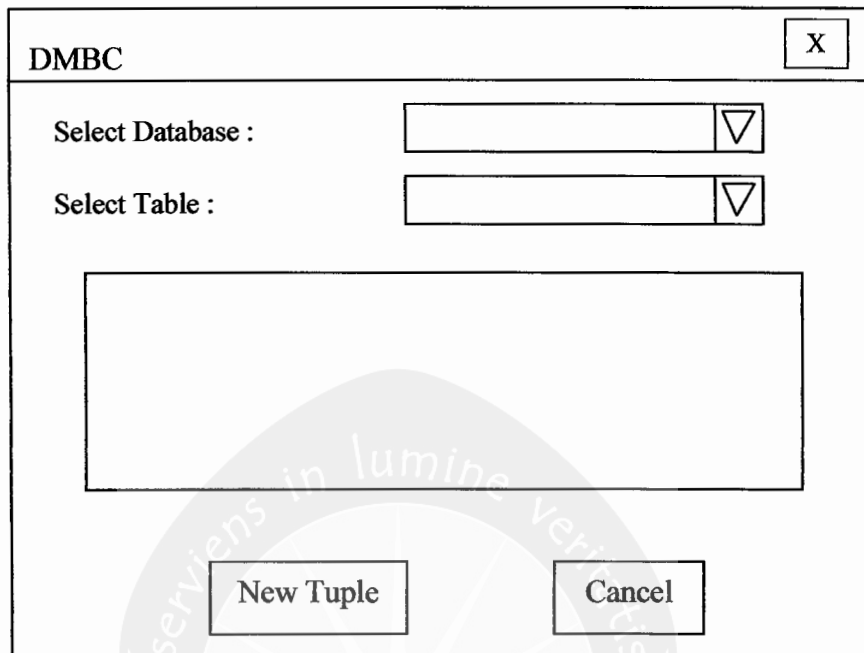
Merupakan tombol yang digunakan apabila user ingin membatalkan koneksi ke SQL Server.

4.2.3 Deskripsi Tombol Help

Merupakan tombol untuk menampilkan Help untuk koneksi ke database SQL Server 2000.

4.3 Antarmuka Select Database dan Table

Form ini digunakan untuk menampilkan nama database yang ada dalam server yang dipilih, sehingga user dapat memilih database mana yang akan digunakan. Dan setelah suatu database dipilih, maka akan ditampilkan juga daftar nama tabel yang ada dalam database terpilih :



Gambar 5 Desain Antarmuka Select Database & Table

Secara prosedural :

On_klik New Tuple

Show frmSelectDatabase

SQL : "SELECT name from dbo.sysdatabases"

"SELECT table_name from information_schema.tables where initial catalog = 'nama_database'"

4.3.1 Deskripsi Tombol New Tuple

Merupakan tombol yang jika diklik maka akan ditampilkan form New Tuple yang meminta user menginputkan nilai untuk masing-masing atribut tuple.

4.3.2 Deskripsi Tombol Cancel

Merupakan tombol yang digunakan apabila user ingin membatalkan koneksi ke database atau tabel yang dipilih.

4.4 Antarmuka Input Nilai Tuple Baru

Merupakan form untuk menginputkan data tuple baru. Pada form ini, user diminta untuk menginputkan nilai setiap atribut tuple baru dengan nilai-nilai yang telah tersedia.

The image shows a graphical user interface for adding or editing a tuple. The main window is titled "New Tuple" and contains two sub-sections: "New Tuple" and "Edit Tuple".

- New Tuple Section:** Includes an "Id" input field, a list of attributes (Atribut 1, Atribut 2, ..., Atribut n), and an "Add" button. Below this, a preview box shows the tuple structure: "Atribut 1 = Value 1", "Atribut 2 = Value 2", and "..."
- Edit Tuple Section:** Includes an "Id" dropdown menu, a list of attributes (Atribut 1, Atribut 2, ...), "Old Value" and "New Value" input fields, and a "Save" button.

At the bottom of the dialog are two buttons: "Classify" and "Cancel".

Gambar 6 Desain Antarmuka Input Attribute Value

4.4.1 Deskripsi Tombol New Tuple

Merupakan tombol yang jika diklik maka sistem mengenerate id untuk tuple yang baru. Kemudian user diminta untuk menginputkan nilai untuk masing-masing atribut yang ada. Secara prosedural :

On_klik New Tuple

Generate New_id

Input values for each Attribute

SQL : "INSERT INTO table_name (columns) VALUES (selected_values)"

4.4.2 Deskripsi Tombol Add

Tombol ini digunakan untuk memasukkan nama atribut yang dipilih beserta nilainya yang telah diinputkan oleh user ke dalam listbox di bawahnya dan sekaligus menyimpannya ke dalam database.

4.4.3 Deskripsi Tombol Edit Tuple

Merupakan tombol yang akan mengaktifkan kontrol untuk proses pengubahan data. User dapat memilih id data yang akan diubah pada combobox Id. Pada listbox akan ditampilkan seluruh atribut yang dimiliki oleh tuple dengan id seperti yang telah dipilih. Jika user memilih salah satu field pada listbox, sistem akan menampilkan nilai field yang dipilih tersebut pada textbox Old Value. Untuk mengubah data pada atribut yang dipilih, user dapat memilih nilai yang ada pada combobox New Value dan menyimpan nilai baru tersebut dengan menekan tombol Save.

4.4.4 Deskripsi Tombol Save

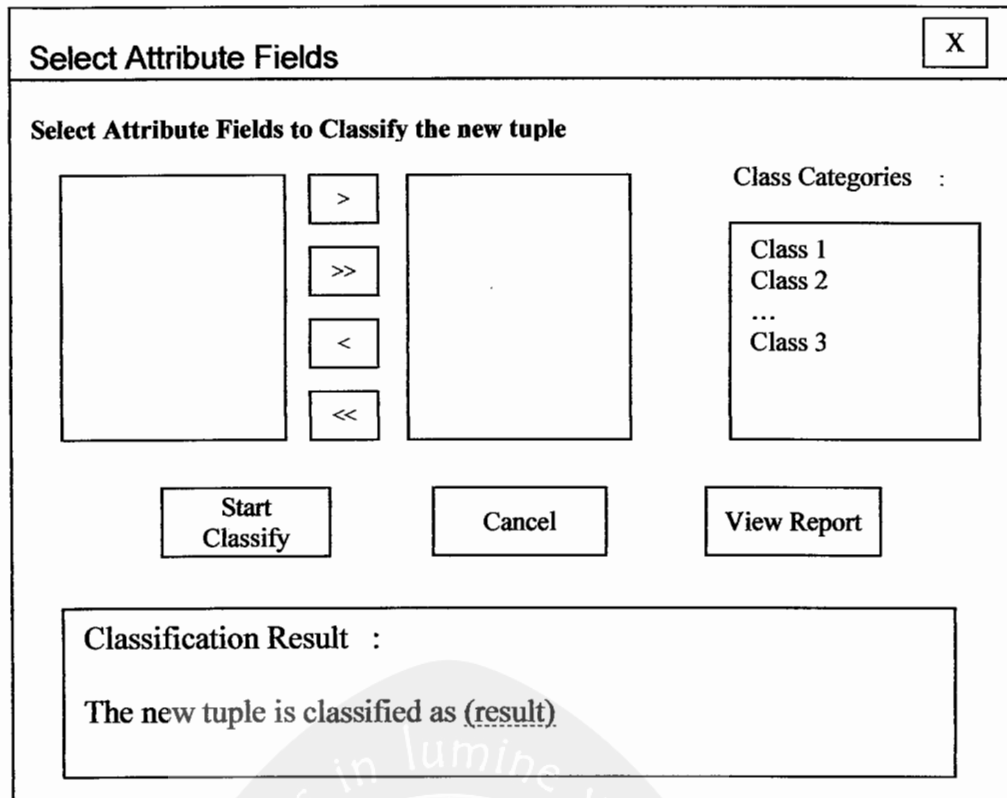
Merupakan tombol yang digunakan untuk menyimpan nilai baru yang telah ditetapkan user sebagai nilai dari field yang dipilih. Jika data telah tersimpan, maka sistem akan menampilkan pesan bahwa data telah ter-update.

4.4.5 Deskripsi Tombol Cancel

Merupakan tombol yang digunakan apabila user ingin membatalkan penginputan nilai tuple baru.

4.5 Antarmuka Select Attribute Fields

Merupakan halaman yang digunakan untuk meminta user memilih atribut apa saja yang akan digunakan untuk mengklasifikasikan tuple baru yang telah diinputkan nilai-nilainya sebelumnya.



Gambar 7 Desain Antarmuka Select Attribute Fields

4.5.1 Deskripsi Tombol >

Merupakan tombol yang digunakan untuk menambahkan atribut yang dipilih pada listbox sebelah kiri ke listbox sebelah kanan, yang berarti atribut tersebut dipilih untuk proses klasifikasi.

4.5.2 Deskripsi Tombol >>

Merupakan tombol yang digunakan untuk menambahkan seluruh atribut yang ada pada listbox sebelah kiri ke listbox sebelah kanan, yang berarti semua atribut tersebut dipilih untuk proses klasifikasi.

4.5.3 Deskripsi Tombol <

Merupakan tombol yang digunakan untuk memindahkan atribut yang dipilih pada listbox kanan ke listbox kiri.

4.5.4 Deskripsi Tombol <<

Merupakan tombol yang digunakan untuk memindahkan seluruh atribut yang pada listbox sebelah kiri ke listbox sebelah kanan.

4.5.5 Deskripsi Tombol Start Classify

Merupakan tombol yang digunakan untuk memulai proses pengklasifikasian tuple baru. Proses pengklasifikasian tuple dilakukan berdasarkan atribut yang telah dipilih oleh user, yaitu atribut-atribut yang terdapat pada listbox sebelah kanan. Setelah proses pengklasifikasian dilakukan, maka pada bagian Classification Result akan ditampilkan hasil klasifikasi tuple, yang berupa kelas dari tuple baru.

4.5.6 Deskripsi Tombol Cancel

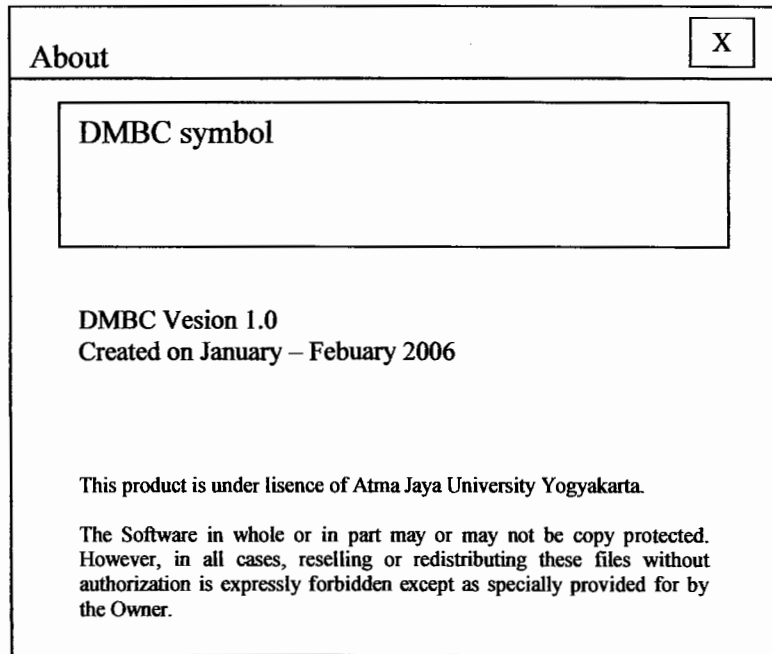
Merupakan tombol yang digunakan untuk membatalkan proses pemilihan atribut.

4.5.7 Deskripsi Tombol View Report

Merupakan tombol yang digunakan untuk menampilkan report. Dalam report ini, ditampilkan id tuple dan nilai probabilitas aktual dari masing-masing kemungkinan kelas tuple, atribut yang dipilih untuk pengklasifikasian tuple, dan hasil klasifikasi. Kelas dengan nilai probabilitas tertinggi merupakan kelas dari tuple yang diklasifikasikan.

4.6 Antarmuka About DMBC

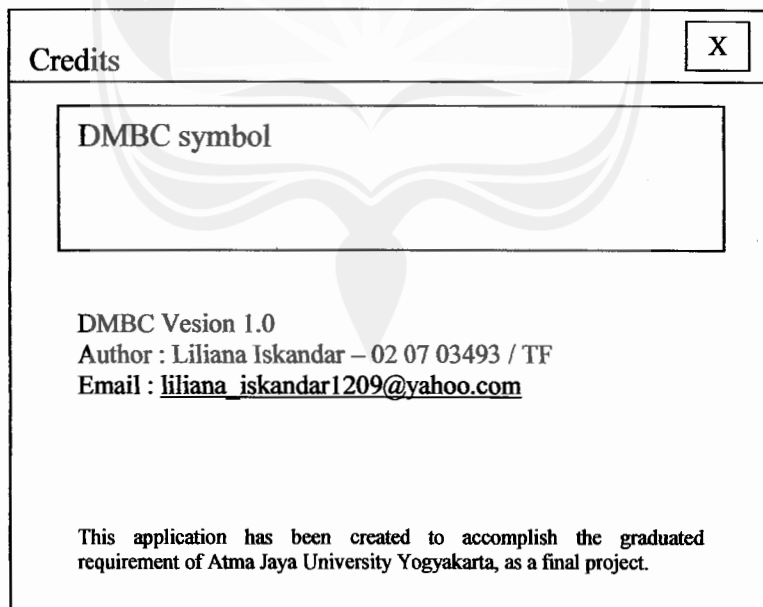
Merupakan halaman yang digunakan untuk memberikan informasi tentang sistem DMBC.



Gambar 8 Desain Antarmuka About DMBC

4.7 Antarmuka Credits DMBC

Merupakan halaman yang digunakan untuk memberikan informasi tentang programmer sistem DMBC.



Gambar 9 Desain Antarmuka Credits DMBC