

ANALISA PENERAPAN DATA MINING PADA PENERIMAAN MAHASISWA POLITEKNIK NEGERI LHOKSEUMAWE MENGGUNAKAN ALGORITMA *K-MEANS*

Desita Ria Yusian TB

Jaringan Cerdas Multimedia, Jurusan Teknik Elektro
Fakultas Teknologi Industri, Institut Teknologi Sepuluh Nopember
desita@uui.ac.id

Abstrak

Pemanfaatan teknologi informasi dalam proses penerimaan mahasiswa baru memungkinkan terjadi akumulasi data dalam jumlah besar tiap tahunnya. Data mining merupakan penggalian informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data dalam jumlah besar yang diharapkan dapat mengatasi kondisi tersebut. Dengan memanfaatkan data pendaftaran dan data lulus tes diharapkan dapat menghasilkan informasi pengclusteran dan hubungan saling keterkaitan antara Jurusan dengan Jalur Masuk Mahasiswa. Algoritma yang digunakan dalam penelitian ini adalah algoritma *K-Means*, Algoritma ini disusun atas dasar ide yang sederhana. Informasi yang dihasilkan berupa grafik pengclusteran data mahasiswa, grafik tertinggi dengan persentase jalur masuk umptn 323,0 dan jalur undangan 105.0 terlihat pada jurusan Teknik Elektro.

Kata kunci : *Data Mining, K-Means, Jurusan dan Jalur Masuk, Cluster.*

BAB 1. Pendahuluan

Seleksi masuknya mahasiswa dalam sebuah perguruan tinggi umumnya dengan memberikan soal-soal test yang harus mereka kerjakan, untuk mengetahui kemampuan dan pengetahuan mereka.

Setelah mahasiswa mengalami proses belajar mengajar, maka akan dapat diketahui prestasi mereka setiap akhir semester. Hal ini akan terjadi secara berulang pada sebuah perguruan tinggi. Jumlah data yang banyak ini membuka peluang untuk dihasilkan informasi yang berguna bagi pihak perguruan tinggi. Penggalian informasi pada sebuah data yang berukuran besar (mempunyai jumlah record dan jumlah field yang cukup banyak) tidak dapat dilakukan dengan mudah.

Teknologi data mining merupakan salah satu alat bantu untuk penggalian data pada basis data berukuran besar dan dengan spesifikasi tingkat kerumitan yang telah banyak digunakan pada banyak domain aplikasi seperti perbankan maupun bidang telekomunikasi [1]. Salah satu organisasi yang memanfaatkan teknologi informasi berupa penerapan data mining yaitu Politeknik Negeri Lhokseumawe (PNL), sehingga akan terjadi akumulasi data tiap tahunnya untuk membantu mentransformasikan data menjadi suatu informasi yang berguna. *Data Mining* yang

dibangun menggunakan metode *K-Means* yang bertujuan untuk menggali dan menemukan pola-pola yang tersembunyi antara data masukan pada saat pendaftaran masuk dengan data setelah lulus tes masuk politeknik.

K-Means merupakan salah satu metode data clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster. Metode ini mempartisi data ke dalam cluster sehingga data yang memiliki karakteristik sama dikelompokkan ke dalam satu cluster yang sama, yang dicluster disini adalah jalur masuk mahasiswa berdasarkan jurusannya. Analisa pengklusteran dari hasil pengklusteran menggunakan *K-Means* sebagai model untuk menentukan jalur masuk mahasiswa berdasarkan jurusan diperlukan untuk dapat menghitung keakuratan dari hasil penerapan metode *K-Means Clustering* dengan menghitung nilai error presisi.

BAB 2. Dasar Teori

Penelitian mengenai Aplikasi data mining ini telah dilakukan pada tahun 2008 oleh Yugi Trianto Purba yaitu Penerapan Data Mining untuk Menemukan Pola antara Nilai

Ujian Saringan Masuk Terhadap Indeks Prestasi, informasi yang dihasilkan dapat digunakan sebagai dasar analisis dalam pengambilan keputusan dan memberikan nilai tambah bagi database yang telah dibangun. Menggunakan metode *regresi linier*. *Microsoft Visual Studio. NET 2005* sebagai aplikasi *interface system*.

Tahun 2010 oleh Nuqson Masykur Huda yaitu Aplikasi *Data Mining* untuk Menampilkan Informasi Tingkat Kelulusan Mahasiswa, informasi yang ditampilkan berupa nilai *support* dan *confidence* dari masing-masing kategori tingkat kelulusan. Menggunakan algoritma *Apriori*. *Microsoft Visual Studio. NET 2005* sebagai aplikasi *interface system*.

Tahun 2011 oleh Tacbir Hendro Pudjiantoro, Faiza Renaldi, dan Age Teogunadi yaitu Penerapan *Data Mining* untuk Menganalisa Kemungkinan Pengunduran Diri Calon Mahasiswa Baru, Dengan menggunakan data calon mahasiswa 2007-2009 yang digunakan sebagai *data sample* untuk pembentukan pohon keputusan, didapatkan hasil kecocokan prediksi dengan menggunakan 600 data calon mahasiswa baru 2010 yang di ujikan terdapat tingkat kecocokan hasil prediksi dengan hasil nyata sebesar 61,89%. Hasil uji tersebut menyatakan system cukup layak untuk digunakan dalam prediksi calon mahasiswa baru meskipun sistem belum menghasilkan tingkat akurasi yang maksimal. Menggunakan algoritma ID3 (*decision tree learning*).

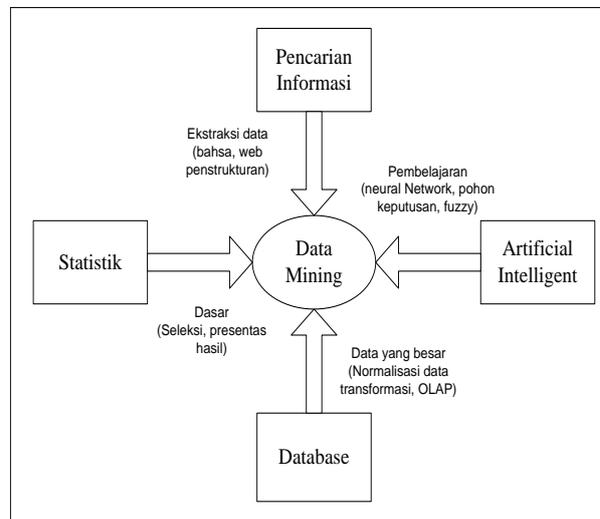
Dasar teori yang berhubungan dengan perancangan software dan pembuatan *Data Mining* Pada Penerimaan Mahasiswa Politeknik Negeri Lhokseumawe Menggunakan Algoritma *K-Means*, seperti penjelasan tentang kajian pustaka, penjelasan tentang *data mining*, penjelasan tentang tahap-tahap dalam proses KDD, penjelasan tentang pengelompokan *data mining*, penjelasan tentang algoritma *K-Means*, penjelasan tentang kelebihan dan kekurangan algoritma *K-Means*, penjelasan tentang data flow diagram, penjelasan tentang PHP.

Penjelasan teori-teori tersebut akan diuraikan pada subab-subab berikut ini.

Data mining

Data mining adalah suatu istilah yang digunakan untuk mengurangi penemuan pengetahuan di dalam database. *Data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakit dari berbagai database besar (Turban,dkk.2005).

Data mining bukanlah suatubidang yang sama sekali baru. Salah satu kesulitan untuk mendefinisikan *data mining* adalah kenyataan bahwa *data mining* mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang sudah mapan terlebih dulu. Gambar 1.1 menunjukkan bahwa *data mining* memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistik, database, dan juga *information retrieval* (Pramudiono,2006)



Gambar 2.1 Bidang Ilmu *Data Mining* (Sumber : Kusri dan Emha Taufik Luthfi, 2009)

Menurut Pramudiono (2006), *Data mining* adalah analisis otomatis dari data yang berjumlah besar atau kompleks dengan tujuan untuk menemukan pola atau kecenderungan yang penting yang biasanya tidak disadari keberadaannya.

Kemajuan yang luar biasa yang terus berlanjut dalam bidang *data mining* didorong oleh beberapa faktor, antara lain (Larose, 2005):

1. Pertumbuhan yang cepat dalam kumpulan data .
2. Penyimpanan data dalam *data warehouse*, sehingga seluruh perusahaan memiliki akses ke dalam database yang andal.
3. Adanya peningkatan akses data melalui navigasi web dan internet.
4. Tekanan kompetisi bisnis untuk meningkatkan penguasaan pasar dalam globalisasi ekonomi.
5. Perkembangan teknologi perangkat lunak untuk *data mining* (ketersediaan teknologi).
6. Perkembangan yang hebat dalam kemampuan komputasi dan pengembangan kapasitas media penyimpanan.

Dari definisi-definisi yang telah disampaikan, hal penting yang terkait dengan data mining adalah :

1. *Data mining* merupakan suatu proses otomatis terhadap data yang sudah ada.
2. Data yang akan diproses berupa data yang sangat besar.
3. Tujuan *data mining* adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat.

Pengklasteran

Peng-*klaster*-an adalah upaya menemukan sekelompok objek yang mewakili suatu karakter yang sama atau hamper sama (*similar*) antar satu objek dengan objek yang lainnya pada suatu kelompok dan memiliki perbedaan (*not similar*) dengan objek-objek pada kelompok lainnya. Tentunya persamaan dan perbedaan tersebut diperoleh berdasarkan informasi yang diberikan oleh obyek-obyek tersebut beserta hubungan (*relationship*) antar mereka. Dalam berbagai kesempatan, *klastering* juga sering disebut *unsupervised classification* yaitu, pengelompokan data

yang memiliki pengelompokan alami (seperti: jenis kelamin, cara berkembang biak, warna kulit, dan lain-lain).

Menurut Kamus Besar Bahasa Indonesia (1996), *Klaster* adalah “kelompok” (p.35).

Berdasarkan konsep arsitektur, *klaster* merupakan “penataan beberapa rumah yang di kelompokkan kedalam ruang bersama untuk mendapatkan kepadatan yang tinggi pada suatu area, hingga lahan lainnya dapat dimanfaatkan untuk ruang terbuka” (kwanda, 2002,p.30).

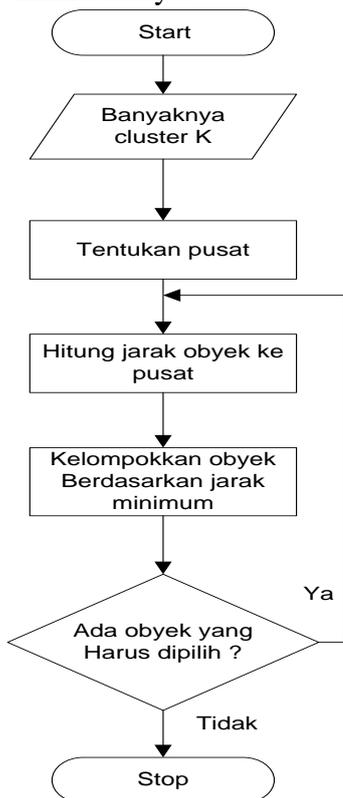
Klaster adalah suatu bentuk pola tertentu yang dibuat oleh manusia untuk tujuan atau maksud tertentu dan untuk membedakan klaster yang satu dengan klaster yang lain dalam sebuah proyek property yang berupa perumahan atau kawasan hunian (proyek dengan skala yang besar) pengembang akan memberikan tema atau identitas tertentu untuk klaster tersebut. Biasanya klaster (khususnya di Indonesia) hanya terdapat dikawasan hunian tertentu yang difokuskan untuk kalangan menengah ke atas.

Algoritma K-Means

Data Klastering merupakan salah satu metode *Data Mining* yang bersifat tanpa arahan (*unsupervised*). Ada dua jenis data klastering yang sering dipergunakan dalam proses pengelompokan data yaitu *hierarchical* (hirarki) data klastering dan *non-hierarchical* (non hirarki) data klastering. *K-Means* merupakan salah satu metode data klastering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih klaster/kelompok. Metode ini mempartisi data ke dalam klaster/kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu klaster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan dari data klastering ini adalah untuk meminimalisasikan *objective function* yang diset dalam proses klastering, yang pada umumnya berusaha meminimalisasikan variasi di dalam suatu klaster dan memaksimalkan variasi antar klaster.

Algoritma *K - Means* selanjutnya akan melakukan pengulangan langkah-langkah berikut sampai terjadi kestabilan (tidak ada obyek yang dapat dipindahkan):

1. menentukan koordinat titik tengah setiap kluster,
2. menentukan jarak setiap obyek terhadap koordinat titik tengah,
3. mengelompokkan obyek-obyek tersebut berdasarkan pada jarak minimumnya.



Gambar 1.2. Flow Algoritma K-Means

Kelebihan dan Kelemahan Algoritma K-Means

Algoritma *K - means* dinilai cukup efisien, yang ditunjukkan dengan kompleksitasnya $O(tkn)$, dengan catatan n adalah banyaknya obyek data, k adalah jumlah kluster yang dibentuk, dan t banyaknya iterasi. Biasanya, nilai k dan t jauh lebih kecil daripada nilai n . Selain itu, dalam iterasinya, algoritma ini akan berhenti dalam kondisi optimum lokal.

Hal yang dianggap sebagai kelemahan algoritma ini adalah adanya keharusan

menentukan banyaknya kluster yang akan dibentuk, hanya dapat digunakan dalam data yang *mean - nya* dapat ditentukan, dan tidak mampu menangani data yang mempunyai penyimpangan - penyimpangan (*noisy data* dan *outlier*). Berkhin menyebutkan beberapa kelemahan algoritma *K - means* adalah:

- (1) sangat bergantung pada pemilihan nilai awal centroid,
- (2) tidak jelas berapa banyak kluster k yang terbaik,
- (3) hanya bekerja pada atribut numerik.

Similarity dan Dissimilarity

Memperhatikan input dalam algoritma *K - Means*, dapat dikatakan bahwa algoritma ini hanya mengolah data kuantitatif. Hal tersebut juga diungkapkan oleh Berkhin, bahwa algoritma *K - means* hanya dapat mengolah atribut numerik.

BAB 3. Metode Penelitian

Dalam pembangunan Aplikasi *Data Mining* ini, sumber data diperoleh dari dua data yaitu data pendaftaran dan data lulus tes. Data mining membutuhkan data dalam jumlah besar, untuk itu diperlukan suatu data gabungan yang dapat menampung dan menyatukan dari kedua sumber data tersebut.

Alat dan Bahan

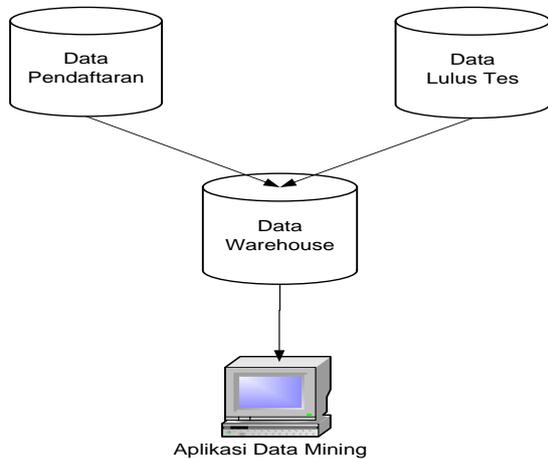
Kebutuhan alat dan bahan dalam penelitian ini meliputi perangkat keras (*hardware*) dan perangkat lunak (*software*). Perangkat keras yang digunakan meliputi spesifikasi:

1. Seperangkat komputer dengan *processor* intel Pentium 4
2. Memory 1 GB DDR3
3. Hardisk 320 GB

Data mining untuk penerimaan mahasiswa ini didukung kebutuhan perangkat lunak seperti:

1. Sistem Operasi *Microsoft Windows XP SP3*.
2. PHP 5.3.10
3. Appserver 2.6.0
4. Notepad++. 5.9.2

Data yang telah menyatu dapat diolah pada aplikasi *data mining*. Sumber data yang digunakan dalam penelitian ini yaitu data pendaftaran dan data lulus tes. Diambil seluruh jurusan yang ada di Politeknik Negeri Lhokseumawe.



Gambar 3.1 Aliran data dalam proses *data mining*

Data Pendaftaran

Data pendaftaran adalah data mahasiswa yang didata ketika mahasiswa mendaftar di Politeknik Negeri Lhokseumawe. Data yang dicatat adalah biodata pribadi mahasiswa. Atribut yang ada di data pendaftaran dapat dilihat dalam tabel 3.1

Tabel 3.1 Tabel Data Pendaftaran

Atribut	Keterangan
Nama Mahasiswa	Merupakan nama lengkap mahasiswa yang bersangkutan
Jenis Kelamin	Merupakan jenis kelamin mahasiswa yang bersangkutan
Tempat Lahir	Merupakan tempat lahir mahasiswa yang bersangkutan
Tanggal Lahir	Merupakan tanggal lahir mahasiswa yang bersangkutan

Data Lulus Tes

Data lulus tes adalah data mahasiswa yang telah dinyatakan lulus. Data yang dicatat adalah biodata mahasiswa dan kelengkapan

lulus. Atribut yang ada di data Lulus Tes dapat dilihat dalam tabel 3.2

Tabel 3.2 Tabel Data Lulus Tes

Atribut	Keterangan
NIM (Nomor Induk Mahasiswa)	Nomor Induk Mahasiswa atau yang disingkat dengan NIM adalah kode yang dimiliki mahasiswa sebagai number unik identitas di PNL
Nama Mahasiswa	Merupakan nama lengkap mahasiswa yang bersangkutan
Jurusan	Merupakan jurusan dari mahasiswa yang bersangkutan
Program Studi	Merupakan program studi dari mahasiswa yang bersangkutan
Tahun Masuk	Merupakan tahun masuk dari mahasiswa yang sudah di terima di politeknik negeri lhokseumawe
Nilai Masuk	Merupakan nilai masuk mahasiswa yang bersangkutan

Data Warehouse

Dalam aplikasi ini dibangun *data warehouse* yang digunakan untuk menampung data dari *database* induk mahasiswa dan *database* kelulusan mahasiswa. Tujuan utama pembangunan *data warehouse* adalah agar *database* sumber tidak terganggu bila terjadi *error*, selain itu *data warehouse* memudahkan dalam menyatukan data dari dua *database* sumber. Adapun atribut tabel yang terbentuk dapat dilihat pada tabel 3.3

Tabel 3.3 Tabel Data Pendaftaran

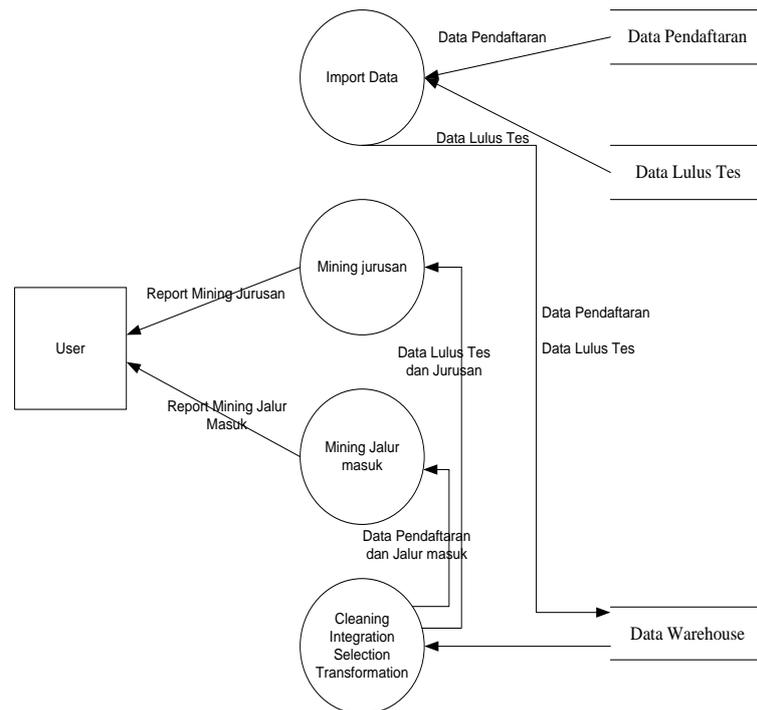
Atribut	Keterangan
NIM (Nomor Induk Mahasiswa)	Nomor Induk Mahasiswa atau yang disingkat dengan NIM adalah kode yang dimiliki mahasiswa sebagai number unik identitas di PNL

Nama Mahasiswa	Merupakan nama lengkap mahasiswa yang bersangkutan
Jurusan	Merupakan jurusan dari mahasiswa yang bersangkutan
Program Studi	Merupakan program studi dari mahasiswa yang bersangkutan
Tahun Masuk	Merupakan tahun masuk dari mahasiswa yang sudah di terima di politeknik negeri lhokseumawe
Nilai Masuk	Merupakan nilai masuk mahasiswa yang bersangkutan
Jenis Kelamin	Merupakan jenis kelamin mahasiswa yang bersangkutan
Tempat Lahir	Merupakan tempat lahir mahasiswa yang bersangkutan
Tanggal Lahir	Merupakan tanggal lahir mahasiswa yang bersangkutan

berupa Data pada data masuk dan data lulusan. Sedangkan untuk output yaitu cluster Jurusan dan Jalur Masuk. User disini tidak mengimputkan data tetapi hanya melihat informasi dari data yang sudah diimputkan terlebih dahulu.

DFD Level-1

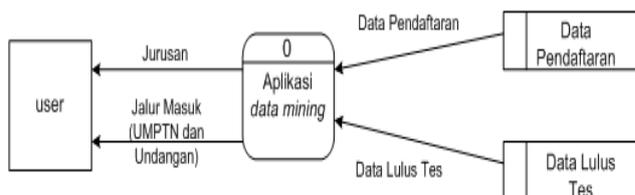
Pada gambar 3.2 merupakan DFD *level-1* dari *Data Mining* yang penjelasan proses-prosesnya lebih spesifik



Perancangan Sistem

Pada tahap awal melakukan perancangan DFD (*Data Flow Diagram*), perancangan DFD adalah representasi klaster dari sebuah system. DFD (*Data Flow Diagram*) untuk aplikasi data mining penerimaan mahasiswa dapat dilihat pada gambar 3.2 DFD Level-0 dan gambar 3.3 DFD Level-1 dibawah ini.

1. DFD Level-0



Gambar 3.2 DFD Level-0

Gambar 3.2 merupakan DFD level-0 pada aplikasi Penerapan *Data Mining* yang terdiri dari 2 input dan 2 output. Input disini

Berikut ini adalah penjelasan dari proses-proses yang terdapat pada DFD *level-1* :

1. *Import Data*
Proses *import* data adalah proses load data dari data pendaftaran dan data lulus tes ke data warehouse (gabungan). Semua akan di masukkan tanpa ada penyaringan.
2. *Cleaning, Integrasi, Selection, dan Transformasi*

- a. *Cleaning* data merupakan proses penghilangan *noise* dan data yang tidak konsisten atau data yang tidak relevan. Pada umumnya data yang diperoleh, baik dokumentasi suatu instansi pemerintahan maupun, *database* suatu perusahaan maupun hasil eksperimen, memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid

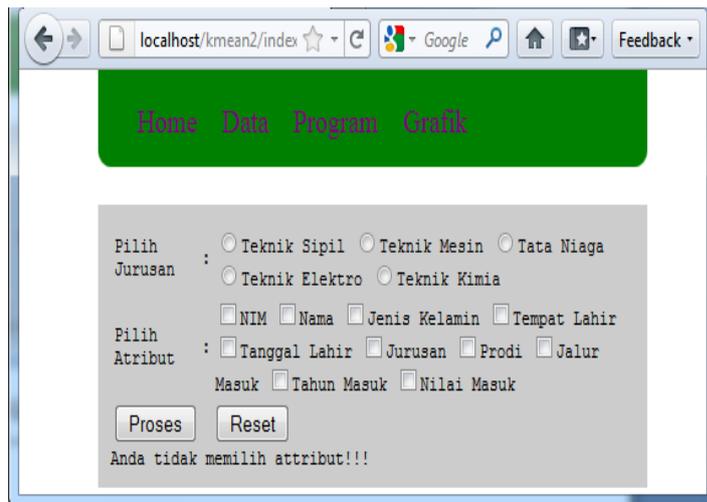
atau juga hanya sekedar salah ketik. Dalam tahap ini semua data yang akan digunakan baik dari data masuk maupun data lulusan dibersihkan dari *record* data yang tidak mempunyai atribut lengkap. Selain pembersihan *record* data yang tidak valid, juga dilakukan penghapusan atribut yang tidak dipakai.

- b. *Integrasi* data merupakan penggabungan data dari data pendaftaran dan data lulus tes kedalam data warehouse, *Integrasi* data dilakukan pada atribut yang mengidentifikasi entitas-entitas dengan satu atribut unik yaitu NIM
- c. *Selection* data adalah proses menyeleksi atribut apa yang akan diproses pada mining selanjutnya. Pada penelitian ini yang akan diproses adalah atribut Jurusan dan Jalur Masuk.
- d. *Transformasi* data merupakan proses mengubah data atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining*. Beberapa metode *data mining* membutuhkan format data yang khusus sebelum bisa diaplikasikan. Dalam Aplikasi *Data Mining* ini, data yang dirubah yaitu jalur masuk dan jurusan untuk melihat grafik dari jalur masuk UMPTN dan Undangan.

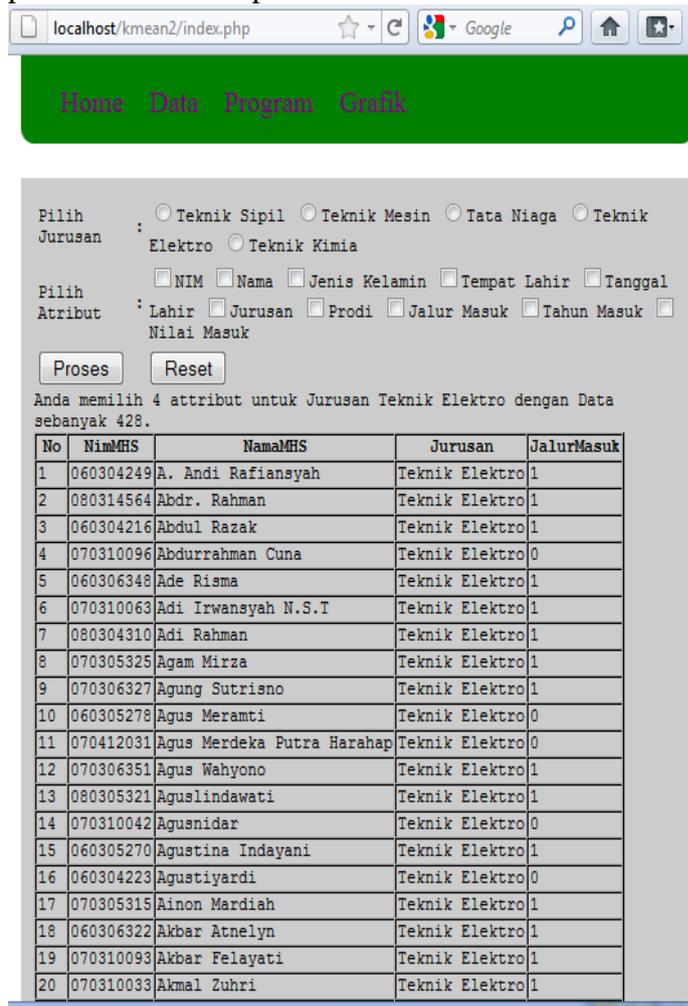
Bab 4. Implementasi dan Analisa

Secara umum aplikasi *data mining* ini berfungsi untuk membantu pihak Politeknik Negeri Lhokseumawe dalam penerimaan mahasiswa selanjutnya. Tampilan awal dari aplikasi ini adalah berupa tampilan halaman utama (tampilan link home), dilanjutkan dengan tampilan link data, link program, dan link grafik. Pada pengujian ini telah berhasil di uji dengan menggunakan PHP 5.3.10.

Hasil rancangan tampilan halaman home pada aplikasi data mining ini berisi perintah pemilihan atribut dan pilih jurusan. Pada gambar 4.1 dapat dilihat hasil tampilan halaman utama dari aplikasi data mining ini.



Gambar 4.1 Tampilan Halaman Utama Setelah jurusan dan atribut di pilih maka selanjutnya akan ditampilkan data yang dipilih tersebut beserta jumlahnya, dapat dilihat pada gambar 4.2. sebelum atribut di pilih maka tombol proses tidak akan aktif.



Gambar 4.2 Tampilan Halaman Utama

4.1 Tampilan Link Data

Pada halaman ini terdapat data yang nantinya akan di proses untuk menentukan cluster. Data tersebut adalah data hasil seleksi yaitu jurusan dan jalur masuk. Pada data jalur masuk ada inisial 0 dan 1, artinya untuk jalur masuk UMPTN (1) sedangkan jalur masuk undangan (0). Dikarenakan algoritma K-Menas hanya bekerja pada atribut numerik. Berikut hasil tampilan dari link data, dapat dilihat pada gambar 4.3

No	Nim MHS	Nama MHS	JK	Tmp Lahir	Tgl Lahir	Jurusan	Prodi	Jalur Masuk	Tahun Masuk
1	060304249	A. Andi Rafiansyah	L	Simpang Puet	1/4/1987	Teknik Elektro	Teknik Listrik	1	2006
2	080314564	Abdr. Rahman	L	Pangkalan Beranda	3/20/1985	Teknik Elektro	Teknik Telekomunikasi Konsentrasi TKJ	1	2008
3	060411008	Abdul Hadi	L	Desa Pusong	8/5/1987	Teknik Sipil	Reakayasa Bangunan Transportasi	1	2006
4	060302142	Abdul Halim	L	Kandang	1/23/1988	Teknik Kimia	Teknik Kimia (DIII)	1	2006
5	080301031	Abdul Hanif	L	Desa Sumbok	5/2/1990	Teknik Sipil	Teknik Sipil	0	2008
6	060301062	Abdul Khadir	L	Desa Geumata	1/28/1988	Teknik Sipil	Teknik Sipil (DIII)	1	2006
7	070302099	Abdul Manaf	L	Desa Keude	2/12/1988	Teknik Kimia	Teknik Kimia	1	2007
8	060301009	Abdul Mannan	L	Teupin Ara	1/20/1988	Teknik Sipil	Teknik Sipil (DIII)	0	2006
9	070302077	Abdul Munir	L	Blang Seunong	4/26/1988	Teknik Kimia	Teknik Kimia	1	2007
10	060302106	Abdul Rahman	L	Kandang	4/15/1988	Teknik Kimia	Teknik Kimia (DIII)	1	2006
11	060304216	Abdul Razak	L	Paya Punteuet	9/12/1987	Teknik Elektro	Teknik Listrik	1	2006
12	080303187	Abdullah M	L	Padang Meuri	2/16/1989	Teknik Mesin	Teknik Perawatan	1	2008
13	070303230	Abdullah Sani	L	Lhokseumawe	10/5/1988	Teknik Mesin	Teknik Mesin	1	2007
						Teknik			

Gambar 4.3 Tampilan Halaman Data

4.1.1 Tampilan Link Program

Pada halaman ini semua sintak-sintak program untuk menghasilkan cluster yang nantinya di cetak pada link Grafik. Tampilan link programnya dapat kita lihat pada gambar 4.4.

```

<?
// === Koneksi ke MySQL === //
include ("koneksi.php");

// === Tampilkan Error Code === //
error_reporting(E_ALL ^ E_NOTICE);

// === Ambil Nilai dari jurusan yang dipilih === //
$jurusan = $_POST['jurusan'];

// === Hitung Jumlah Data berdasarkan Jurusan yang dipilih (Default Semua Jurusan) === //
$query = "select count(*) as jumData from mahasiswa where Jurusan like '%$jurusan%'";
$hasil = mysql_query($query);
$data = mysql_fetch_array($hasil);
$jumData = $data['jumData'];

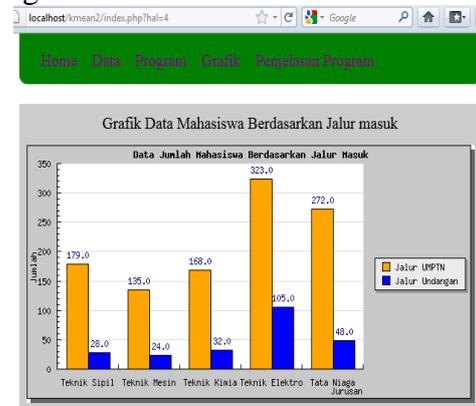
// === Ambil Nilai dari atribut yang dipilih === //
$atribut = $_POST['atribut'];

if(empty($atribut)) //Mengecek apakah atribut dipilih atau tidak
{
    echo "Anda tidak memilih atribut!!!";
}
else
{
    $N = count($atribut); //Hitung atribut yang dipilih
    if( $jurusan == "Sipil" || $jurusan == "Mesin" || $jurusan == "Elektro" || $jurusan == "Kimia")
    {
        echo "Anda memilih $N atribut untuk Jurusan Teknik $jurusan dengan Data sebanyak $jumData.
    <br/>"; //Menampilkan jumlah atribut yang dipilih
    }
    else if($jurusan == "Niaga")
    {
        echo "Anda memilih $N atribut untuk Jurusan Tata $jurusan dengan Data sebanyak $jumData.
    <br/>"; //Menampilkan jumlah atribut yang dipilih
    }
    echo "<table border='1' cellspacing='0'>";
    $join_attr = join(',', $atribut); //Menggabungkan atribut yang dipilih sebelum di query
    echo "<tr><td>";
    for($i=0;$i<$N;$i++)
    {
        echo "<td>".$atribut[$i];
    }
}
    
```

Gambar 4.4 Tampilan Halaman Program

4.1.2 Tampilan Link Grafik

Pada halaman ini menampilkan grafik, dimana pada grafik dibawah ini diberitahukan atau diperlihatkan pengelompokan data jumlah mahasiswa berdasarkan jalur masuk menurut jurusan yang sipilih. Hasil akhir cluster dapat dilihat pada gambar 4.4 dibawah ini.



Gambar 4.5 Tampilan Halaman Hasil Cluster

Bab 5. Kesimpulan dan Saran

Kesimpulan

Dari hasil penelitian ini, diperoleh kesimpulan sebagai berikut :

1. Iputan data berupa data jalur masuk pada saat pendaftaran dengan data setelah lulus tes menjadi data yang diolah oleh sistem data mining menggunakan algoritma K-means clustering menghasilkan grafik tertinggi yaitu terlihat pada jurusan Teknik Elektro dengan persentase jalur masuk umptn 323,0 dan jalur undangan 105.0 mahasiswa.
2. Hasil cluster dari masing-masing jurusan untuk jalur umptn dan undangan 179.0 – 28.0 teknik sipil, 135.0 – 24.0 teknik mesin, 168.0 – 32.0 teknik kimia, 323 – 105.0 teknik elektro dan 272.0 – 48 tata niaga.

Saran

1. Pada pengembangan lebih lanjut dari system yang sudah dibuat tidak hanya menggunakan algoritma dipakai tetapi dapat menggunakan algoritma yang lain untuk menganalisa data. Misalnya *fuzzy c-means*, Dimana keanggotaannya tidak mutlak. Diharapkan dapat memberikan hasil yang lebih akurat.
2. Diharapkan aplikasi dapat dikembangkan untuk pengclusteran data menggunakan algoritma lain untuk memberikan tampilan berupa persentase dengan perhitungan yang lebih baik.

DAFTAR PUSTAKA

- Agusta, Yudi, PhD. (2007). *K-Means-Penerapan, Permasalahan dan Metode Terkait*.
- Berkhin, Pavel. *Survey on clustering data mining techniques*, http://www.ee.ucr.edu/~barth/EE242/clustering_survey.pdf
- Huda, Nuqson Masykur. (2010). *Aplikasi Data Mining untuk Menampilkan Informasi Tingkat Kelulusan Mahasiswa*.
- Han, Jiawei, and M. Kamber. (2006). *Data Mining: Concepts and Techniques*, Morgan Kaufman, USA.
- Kurniawati, Agelina Sartika. (2006). *Perancangan dan Pembuatan Aplikasi Cluster Analisis Terhadap Data Sirkulasi Buku di Perpustakaan UK*. Petra.
- Kadir, Abdul. (2008). *Dasar Pemrograman WEB Dinamis Menggunakan PHP*. Yogyakarta : C.V ANDI OFFSET (Penerbit ANDI).
- Kusri, dan Emha Taufik Luthfi. (2009). *Algoritma Data Mining*. Yogyakarta: C.V ANDI OFFSET.
- Teknomo, Kardi. (2006). Kardi. *Numerical Example of K-Means Clustering*, <http://people.revoledu.com/kardi/tutorial/kMean/NumericalExample.htm>
- Tacbir, Hendro Pudjiantoro. (2011). *Penerapan Data Mining untuk Menganalisa Kemungkinan Pengunduran Diri Calon Mahasiswa Baru*. Bali.
- Leman. (1995). *Pengantar Sistem Informasi Bisnis*. Jakarta: PT Elex Media Komputindo.
- William, Graham, *Data Mining Cluster*, http://datamining.anu.edu.au/student/math3346_2005/050809-maths3346-clusters-2x2.pdf
- Yoga, Pandika, dan Danny Prasetya Suryawan. (2008). *Identifikasi Pendekatan dalam Peng-Cluster-an Perumahan di Kawasan Wisata Bukit Tinggi*. Surabaya.
- Yugi, Trianto Purba. (2008). *Penerapan Data Mining untuk Menemukan Pola antara Nilai Ujian Saringan Masuk Terhadap Indeks Prestasi*. Bandung.