

# IMPLEMENTASI SISTEM PENENTUAN SKRIPSI MAHASISWA STIKOM JAMBI DENGAN EM- CLUSTERING DAN NAÏVE-BAYES

*Errissya Rasywir*

*Sekolah Tinggi Ilmu Komputer Dinamika Bangsa Jambi  
Jl.Jendral Sudirman Thehok-Jambi, Indonesia  
E-mail: [errissya.rasywir@gmail.com](mailto:errissya.rasywir@gmail.com)*

## Abstract

During this process of determining the title of students final assignment is done manually. That is, both lecturers who provide feedback or ideas derived from a variety of research papers. In the sense of the word, a process that has been running still traditional and manually without using a computerized system. Based on the results of previous studies, there is a correlation between students' scores on a particular subject with the title of students final assignment. Based on that, the researchers tried to build a system that is able to help students of Sekolah Tinggi Ilmu Komputer (STIKOM) to determine the final title to be taken. The system built in this study is a system that gives students the determination of final title is based on the value of the courses you have accomplished students. The results of this study are final title determination of student system using *EM Clustering* and Naïve Bayes which be able to give the example of the title of students final assignment .

Keywords: Naïve Bayes; *EM Clustering* algorithms; UML; Object Oriented Programming.

## Abstrak

Selama ini proses penentuan topik skripsi mahasiswa dilakukan secara manual. Yakni baik dosen yang memberi masukan atau ide yang diperoleh dari berbagai makalah penelitian. Dalam arti kata, proses yang telah berjalan masih tradisional dan manual tanpa menggunakan sistem yang terkomputerisasi. Berdasarkan hasil penelitian sebelumnya, terdapat korelasi antara nilai mahasiswa pada matakuliah tertentu dengan topik skripsi. Berdasarkan hal tersebut peneliti mencoba membangun sebuah sistem yang mampu membantu mahasiswa Sekolah Tinggi Ilmu Komputer (STIKOM) untuk menentukan topik skripsi yang akan diambil. Sistem yang dibangun dalam penelitian ini merupakan sistem yang memberikan penentuan topik skripsi mahasiswa berdasarkan nilai mata kuliah yang telah dicapai mahasiswa. Hasil penelitian ini berupa sistem penentuan topik skripsi mahasiswa menggunakan *EM Clustering* dengan Algoritma Naïve Bayes yang mampu memberi contoh referensi topik skripsi.

Kata kunci: Naïve Bayes; algoritma *EM Clustering* ; UML; Object Oriented Programming.

© 2017 Jurnal PROCESSOR.

---

## 1. Pendahuluan

Selama ini proses menentukan topik skripsi mahasiswa dilakukan secara manual. Yakni baik dosen yang memberi masukan atau ide yang diperoleh dari berbagai makalah penelitian. Berdasarkan hasil observasi penulis secara langsung pada STIKOM Dinamika Bangsa , proses yang telah berjalan masih tradisional dan manual tanpa menggunakan sistem yang terkomputerisasi. Pada Jurusan Teknik Informatika banyak sekali topik yang dipelajari seperti *software engineering, artificial intelligent, networking information system, mobile programming, business intelligent, troubleshooting, database* dan masih banyak topik lainnya. Topik tersebut masih merupakan suatu area yang masih luas ruang lingkupnya. Setiap topik

tersebut mempunyai sub area yang membawahi topik tersebut. Misalkan topik *software engineering* mempunyai area yang lebih kecil seperti *software design*, *software development lifecycle*, *software effort estimation*, *software testing* dan masih banyak yang lainnya.

Penelitian terdahulu telah menyatakan bahwa pencapaian nilai mahasiswa dapat digunakan dalam mendapatkan bobot kompetensi terhadap suatu bidang [1]. Berdasarkan hal tersebut, terdapat korelasi antara nilai mahasiswa pada matakuliah tertentu dengan dengan bidang skripsi yang diambil maka dapat diambil hipotesis sementara bahwa peminatan terhadap suatu topik skripsi dapat dilihat berdasarkan kompetensi nilai mata kuliah ilmu komputer yang pernah diambilnya. Berdasarkan hal tersebut peneliti mencoba membangun sebuah sistem yang mampu membantu mahasiswa Sekolah Tinggi Ilmu Komputer (STIKOM) untuk menentukan topik skripsi yang akan diambil. Sistem yang dibangun dalam penelitian ini merupakan sistem yang memberikan penentuan topik skripsi mahasiswa berdasarkan nilai mata kuliah yang telah dicapai mahasiswa.

Sistem pada penelitian ini menggunakan teknik *data mining* untuk mengelompokkan data dalam merekomendasikan topik skripsi kepada mahasiswa. Terdapat berbagai teknik *data mining* seperti *classification*, *association*, *clustering*, *estimation* dan *prediction* atau *forecasting* [6]. Pada penelitian ini akan digunakan *clustering* dalam mengelompokkan data mahasiswa kepada suatu topik skripsi tertentu berdasarkan nilai. Metode *clustering* digunakan karena teknik ini mampu membagi rangkaian data menjadi beberapa kelompok berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya [17].

*Clustering* terdiri dari jenis *hierarchical* dan *partitioning*. Metode *clustering* yang paling banyak digunakan adalah *Agglomerative*, *K-Means*, *Bisecting*, *K-Means*, *Buckshot* dan *EM Clustering* [7][8]. Dari berbagai jenis algoritma tersebut, algoritma *EM CLUSTERING* yang merupakan jenis *CLUSTERING* dengan perhitungan probabilistik memiliki performansi yang bagus dan mampu menghasilkan nilai prediksi yang baik [10][18]. Berdasarkan hal tersebut, maka sistem dalam penelitian ini akan menggunakan *EM Clustering* dalam merekomendasikan topik skripsi mahasiswa STIKOM Dinamika Bangsa Jambi. Selain itu, juga terdapat penelitian sejenis yang menggunakan algoritma berbasis probabilistik *naïve bayes* yang memiliki akurasi yang cukup baik [21]. Berdasarkan latar belakang di atas maka dapat dirumuskan masalah yang akan dibahas dalam penelitian ini adalah bagaimana membangun Sistem Penentuan topik skripsi Mahasiswa STIKOM Dinamika Bangsa Jambi dengan menggunakan algoritma *EM Clustering* dengan kombinasi perhitungan probabilistik menggunakan algoritma *Naïve Bayes*.

## 2. Tinjauan Pustaka/Penelitian Sebelumnya

Penelitian dengan judul Sistem Penentuan topik skripsi Mahasiswa STIKOM Dinamika Bangsa Jambi menggunakan *EM Clustering* dan *Naïve Bayes* ini mengacu pada berbagai penelitian yang telah melakukan teknik *data mining* yang menggunakan berbagai *Clustering* serta algoritma *Naïve bayes*. Antara lain, penelitian [3][6][7] yang bertopik penerapan *Clustering* untuk konsentrasi mahasiswa di STMIK AMIKOM Purwokerto dengan metode *K-Means Clustrering*. Penelitian tersebut menentukan konsentrasi mahasiswa didasarkan pada potensi akademik mahasiswa.

Kemudian penelitian [14][15] dengan topik penerapan *data mining* untuk evaluasi kinerja akademik mahasiswa menggunakan algoritma *Naive Bayes*. Penelitian [12] dengan topik pengelompokan mahasiswa menggunakan algoritma *K-Means Clustering*. penelitian ini berfokus pada pengelompokan mahasiswa berdasarkan data akademik dengan menggunakan teknik *clustering*. Penelitian [7] melakukan penerapan algoritma klasifikasi *data mining* ID3 untuk menentukan penjurusan siswa SMA N 6 Semarang. Kemudian penelitian [2] melakukan penentuan jurusan sekolah menengah atas dengan algoritma *Fuzzy C-Means Clustering*. Berdasarkan hal tersebut, belum ada penelitian yang menggunakan jenis *EM Clustering*, sebuah teknik *clustering* yang mampu memberi prediksi baik [10][18]. Dengan demikian, penelitian ini mencoba membangun sistem dengan algoritma *EM Clustering* dengan perhitungan probabilistik dengan algoritma *Naïve Bayes*. Algoritma *Naïve Bayes* merupakan algoritma yang menggunakan konsep perhitungan probabilistik dengan akurasi yang tinggi [21] sehingga *EM Clustering* pada penelitian ini akan dilakukan menggunakan konsep algoritma *Naïve Bayes*.

2.1 EM Clustering

Algoritma EM Clustering merupakan algoritma clustering yang melakukan estimasi Maximum Likelihood dari parameter dalam sebuah model probabilistik. Algoritma ini akan melakukan clustering terhadap objek data termasuk data yang mengalami missing value. Algoritma EM termasuk algoritma clustering yang berbasiskan perhitungan probabilitas. Secara iteratif Algoritma EM Clustering melakukan 2 tahapan yaitu [10]:

a. Expectation Step

Pada tahap ini, akan menghitung probabilitas objek terhadap kluster.

b. Maximization Step

Pada tahap ini, akan dihitung ulang nilai parameter-parameter yang digunakan untuk memaksimalkan nilai probabilitas. Kedua tahap tersebut dilakukan secara berulang-ulang sampai mencapai nilai konvergen. Yang dimaksud dengan nilai konvergen ini adalah nilai threshold yang digunakan untuk menghentikan iterasi yang terus menerus dimana iterasi tersebut tidak akan berhenti jika nilai nya tak nol 0. Langkah-langkah algoritma EM Clustering dapat dilihat sebagai berikut:

1. Tentukan jumlah cluster (k) kemudian data-data yang ada di dalam cluster.
2. Tentukan mean ( $\mu$ ), covariance ( $\Sigma$ ) dan peluang cluster ( $\pi$ ) untuk setiap cluster dan hitung log-likelihood-nya ( $l^0$ )

$$\pi_k^0 = \frac{\text{jmlah\_point\_k}}{\text{Total\_point}} \tag{1}$$

$$\mu_k^0 = \frac{\sum_i^{n_c} x_i}{n_c} \tag{2}$$

$$\Sigma_k^0 = \frac{\sum (x_i - \text{mean})^2}{n_c - 1} \tag{3}$$

$$l^0 = \frac{1}{N} \sum_{n=1}^N \log(\sum_{i=1}^k \pi_k^0 \Phi(x_n | \mu_k^0, \Sigma_k^0)) \tag{4}$$

3. Lakukan tahap ekspektasi dengan menghitung  $\tau(x_{nk})$ :

$$\tau(x_{nk}) = \frac{\pi_k \Phi(x_n | \mu_k, \Sigma_k)}{\sum_{j=1}^k \pi_j \Phi(x_n | \mu_j, \Sigma_j)} \tag{5}$$

$$\Phi(x_n | \mu_k, \Sigma_k) = \frac{1}{\sqrt{2\pi \Sigma_k}} e^{-\frac{(x_n - \mu)^2}{2\Sigma}} \tag{6}$$

Bandingkan tiap nilai  $\tau(x_{nk})$ , jika  $\tau(x_{11}) > \tau(x_{12})$  maka  $x_1$  di letakan di cluster 1, Jika  $\tau(x_{21}) < \tau(x_{22})$  maka  $x_2$  di letakan di cluster 2, Jika  $\tau(x_{31}) = \tau(x_{32})$  maka  $x_3$  di letakan di cluster 1.

4. Tahap maximization

Hitung ulang nilai tiap parameter:

$$\pi_k^{new} = \frac{\text{jumlah\_objek\_k\_baru}}{N} \tag{7}$$

$$\mu_k^{new} = \frac{\sum_{n=1}^{N_c} \tau(x_{nk}) x_n}{N_k} \tag{8}$$

$$\Sigma_k^{new} = \frac{1}{N_k} \sum_{n=1}^{N_c} \tau(x_{nk}) (x_n - \mu_k^{new})^2 \tag{9}$$

$$N_k = \sum_{n=1}^{N_c} \tau(x_{nk}) \tag{10}$$

5. Evaluasi log-likelihood baru

$$l^{new} = \frac{1}{N} \sum_{n=1}^N \log(\sum_{i=1}^k \pi_k \Phi(x_n | \mu_k^{new}, \Sigma_k^{new})) \tag{11}$$

6. Masukkan nilai mean estimasi kedalam missing value sesuai klusternya, cek apakah konvergen. Jika  $|l^{new} - l^0| > \xi$  dengan  $\xi$  adalah nilai threshold, jika ya maka iterasi berhenti, sebaliknya kembali ke tahap 4.

Algoritma EM clustering merupakan sebuah algoritma berbasis probabilitas. Metode perhitungan probabilistik dalam algoritma ini dapat dilakukan dengan berbagai metode. Namun, dalam penelitian ini metode perhitungan probabilistik yang digunakan adalah algoritma naive bayes. Hal ini didasari pada

ulasan sebelumnya yang menjelaskan kelebihan *naïve bayes* yang mampu menghasilkan akurasi yang baik[21].

## 2.2 Algoritma Naives Bayes

Algoritma *Naive Bayes* berbasiskan perhitungan probabilistik dengan asumsi bahwa setiap fitur yang digunakan saling lepas. *Naive Bayes* merupakan metode klasifikasi teks yang paling populer digunakan. Algoritma ini memiliki kelebihan dari sisi kecepatan pembelajaran dan toleransinya terhadap nilai yang hilang dari fitur. Untuk menangani data numerik, algoritma ini menggunakan *probability density function*, artinya data dianggap mengikuti distribusi normal untuk kemudian dihitung nilai rata-rata dan simpangan bakunya [20].

Untuk merepresentasikan sebuah kelas, terdapat karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi yang berguna untuk menjelaskan bahwa peluang masuknya sampel karakteristik tertentu kedalam kelas *posterior*. Peluang munculnya suatu kelas (sebelum masuknya sampel tersebut, seringkali disebut *prior*), dikali dengan peluang kemunculan karakteristik sampel secara global disebut juga *evidence*. Nilai *evidence* selalu tetap untuk setiap kelas pada satu sampel. Nilai *posterior* tersebut dibandingkan dengan nilai *posterior* kelas lainnya untuk menentukan ke kelas apa suatu sampel [21].

Klasifikasi Naive Bayes diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya. Persamaan dari teorema Bayes adalah sebagai berikut:

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (12)$$

Dengan:

X: Data dengan kelas yang belum diketahui;

H: Hipotesis data X merupakan suatu label kelas tertentu;

P(H|X): Probabilistik hipotesis H berdasarkan kondisi X (*posteriori probability*); P(H): Probabilistik hipotesis H (*prior probability*);

P(X|H): Probabilitas X berdasarkan kondisi pada hipotesis H;

P(X): Probabilistik X;

Untuk menjelaskan teorema *naive bayes*, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, teorema *bayes* tersebut akan disesuaikan sebagai berikut:

$$P(C|F1...Fn) = \frac{P(C)P(F1...Fn|C)}{P(f1...Fn)} \quad (13)$$

Dengan:

C: sebuah kelas; F1... Fn: karakteristik petunjuk.

$$Posterior = \frac{Prior \times likelihood}{evidence} \quad (14)$$

Penjabaran lebih lanjut sebagai berikut:

$$\begin{aligned} P(C|F1, \dots, Fn) &= P(C) P(F1, \dots, Fn | C) \\ &= P(C) P(F1|C) P(F2, \dots, Fn|C, F1) \\ &= P(C) P(F1|C) P(F2|C, F1) P(F3, \dots, Fn|C, F1, F2) \\ &= P(C) P(F1|C)P(F2|C, F1)P(F3|C, F1, F2), P(F4, \dots, Fn|C, F1, F2) \\ &= P(C)P(F1|C)P(F2|C, F1)P(F3|C, F1, F2), \dots, (Fn|C, F1, F2, \dots, Fn-1) \end{aligned} \quad (15)$$

## 2.3 Analisis dan Perancangan Berorientasi Objek

Analisis berorientasi objek merupakan pendekatan yang bertujuan untuk mengembangkan model yang menggambarkan bagaimana suatu perangkat lunak komputer bekerja untuk memenuhi seperangkat persyaratan yang ditetapkan pengguna. Analisa berorientasi objek seperti halnya metode analisis konvensional, membangun sebuah model analisis yang menggambarkan informasi, fungsi, dan perilaku objek [13].

Desain berorientasi objek dibagi menjadi dua kegiatan utama, yaitu desain sistem dan desain objek. Desain sistem menciptakan arsitektur produk, mendefinisikan serangkaian "layer" yang mencapai fungsi sistem tertentu dan mengidentifikasi kelas-kelas yang dienkapsulasi oleh subsistem yang berada di setiap layer. Desain sistem mempertimbangkan spesifikasi dari tiga komponen yakni *user interface*, fungsi manajemen data, dan tugas [13].

## 2.4 Pemodelan Sistem

### 2.4.1 Pemodelan Menggunakan UML (Unified Modelling Language)

Pemodelan sistem digunakan untuk menggambarkan kebutuhan sistem. Pemodelan dapat dilakukan dengan menggunakan pendekatan berorientasi objek, seperti *Unified Modelling Language* (UML). UML merupakan kesatuan dari pemodelan yang dikembangkan oleh Booch menjadi sangat terkenal dengan nama metode *Design Object Oriented*. Metode ini menjadikan proses analisis dan desain kedalam empat tahapan iteratif [11]:

1. Identifikasi kelas-kelas dan objek-objek.
2. Identifikasi semantik dari hubungan objek dan kelas tersebut.
3. Perincian *interface*.
4. Implementasi.

UML adalah salah satu model untuk merancang pengembangan *software* yang berbasis *object oriented*. UML juga memberikan standar penulisan sebuah sistem *blue print*, yang meliputi konsep bisnis proses, penulisan kelas-kelas dalam bahasa program yang spesifik, skema *database*, dan komponen-komponen yang diperlukan dalam sistem *software*. Selain itu dengan UML dapat dilakukan pendokumentasian dapat dilakukan seperti; *requirements*, arsitektur, *design*, *source code*, *project plan*, *tests*, dan *prototypes*.

Menurut Dennis [4][5] ada dua jenis diagram UML, yaitu :

1. Diagram Struktur (*structure diagram*)  
Diagram struktur untuk menggambarkan hubungan antar kelas. Diantaranya adalah *class diagram*, *object diagram*, *package diagram*, *deployment diagram*, *component diagram* dan *composite structure diagram*.
2. Diagram Tingkah Laku (*behavior diagram*)  
Diagram untuk menggambarkan interaksi antara orang-orang (*actor*) dan benda-benda yang ditunjuk sama. Diagram tingkah laku termasuk menggunakan *activity diagram*, *sequence diagram*, *communication diagram*, *interaction overview diagram*, *timing diagram*, *behavior state machine*, *protocol state machine* dan *use case diagram*.

## 2.5 Tool Pemrograman

### 2.5.1. PHP

PHP digunakan untuk membuat tampilan web menjadi lebih dinamis, dengan PHP bisa menampilkan atau menjalankan beberapa file dalam 1 file dengan cara di include atau require [9]. PHP (*Hypertext Preprocessor*) merupakan bahasa berbentuk skrip yang ditempatkan dalam server dan diproses di server, hasilnya yang dikirimkan ke klien, tempat pemakai menggunakan browser [1].

Kemudian sistem ini dibangun menggunakan Framework. Framework dapat langsung dipakai fungsinya oleh modul-modul atau fungsi yang akan kita kembangkan [16]. Framework yang digunakan adalah Code Igniter dan Bootstrap.

### 2.5.2 Code Igniter

Framework yang digunakan adalah Code Igniter. Framework ini adalah suatu kerangka kerja yang berupa sekumpulan folder yang memuat file-file php yang menyediakan class libraries, helpers, plugins dan lainnya. Framework menyediakan konfigurasi dan teknik coding tertentu.

### 2.5.3 Bootstrap

Bootstrap adalah sebuah *framework* CSS yang menyediakan kumpulan komponen-komponen antarmuka dasar pada web yang telah dirancang sedemikian rupa untuk digunakan bersama [16]. Bootstrap juga

menyediakan sarana untuk membangun layout halaman dengan mudah dan rapi, serta modifikasi pada tampilan dasar HTML untuk membuat seluruh halaman web yang dikembangkan senada dengan komponen-komponen lainnya.

#### 2.5.4. My SQL

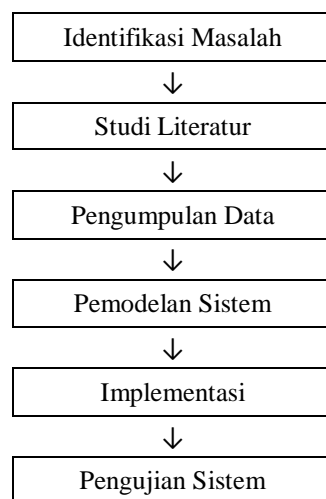
MySQL adalah sebuah aplikasi database guna menyimpan data-data yang akan disimpan. MySQL merupakan aplikasi database server. Pada MySQL, sebuah database mengandung satu atau sejumlah tabel. Tabel terdiri atas sejumlah baris, dan setiap baris mengandung satu atau beberapa kolom [1]

### 3. Metodologi

Metodologi penelitian menggambarkan tahapan proses, metode dan tools yang digunakan dalam melakukan penelitian agar penelitian dapat berjalan dengan baik dan tujuan yang telah ditetapkan dapat tercapai. Pada penelitian ini penulis menggunakan tahapan kegiatan penelitian yang berupa kerangka kerja penelitian yang penulis lakukan.

#### 3.1 Kerangka Kerja penelitian

Kerangka kerja penelitian menggambarkan tahapan proses yang dilakukan dalam penelitian agar penelitian dapat berjalan dengan baik dan tujuan yang telah ditetapkan dapat tercapai. Pada penelitian ini penulis menggunakan tahapan kerangka kerja penelitian sebagai berikut:



Gambar 1. Kerangka Kerja Penelitian

1. Identifikasi Masalah  
Pada tahap ini penulis melakukan observasi awal untuk mengidentifikasi dan merumuskan masalah dari penelitian yang dilakukan agar penelitian yang dilakukan dapat memberikan solusi atas permasalahan yang ada. Masalah yang akan dibahas adalah Implementasi Sistem Rekomendasi Topik Skripsi Mahasiswa STIKOM Dinamika Bangsa Jambi dengan menggunakan algoritma *EM Clustering* dan *Naïve bayes*.
2. Studi literatur  
Mempelajari dan memahami teori-teori yang menjadi pedoman dan referensi guna penyelesaian masalah yang dibahas dalam penelitian ini dan mempelajari penelitian yang relevan dengan masalah yang diteliti. Pada tahap ini peneliti mempelajari mengenai algoritma *EM-clustering*, dan *Naïve bayes* perancangan sistem dengan konsep orientasi objek serta pengujian sistem.
3. Pengumpulan data  
Mengumpulkan data-data yang berhubungan dengan organisasi yang penulis teliti. Proses ini dilakukan dengan mengumpulkan dokumen organisasi dalam hal ini adalah data nilai mahasiswa STIKOM Dinamika Bangsa Jambi.

#### 4. Pemodelan Sistem

Membuat rencana pemodelan sistem penentuan topik skripsi mahasiswa STIKOM Dinamika Bangsa Jambi dengan menggunakan algoritma *EM Clustering* dan *Naïve bayes*. Adapun metode pemodelan sistem yang digunakan adalah metode pendekatan berorientasi objek dengan menggunakan tools UML (*Unified Modeling Language*) dengan langkah-langkah sebagai berikut [11][13]:

a. Menentukan perencanaan awal

Pada tahap ini dibuat perencanaan mengenai kegiatan apa saja yang akan dilakukan beserta waktu yang dibutuhkan untuk masing-masing kegiatan.

b. Melakukan analisis proses

Melakukan analisis kebutuhan Sistem Penentuan topik skripsi Mahasiswa STIKOM Dinamika Bangsa Jambi untuk mendukung proses penilaian tersebut meliputi analisis kebutuhan fungsional dan analisis kebutuhan data.

c. Memodelkan sistem dengan menggunakan UML

Pada tahap ini dibuat pemodelan kebutuhan sistem dengan menggunakan diagram UML.

#### 5. Implementasi Sistem

Pada tahap ini dilakukan pembangunan sistem penentuan topik skripsi mahasiswa STIKOM Dinamika Bangsa Jambi menggunakan algoritma *EM Clustering* dan *Naïve Bayes* berbasis *Object Oriented Programming*. Proses ini meliputi perancangan *user interface*, *coding* dan *testing*. Untuk evaluasi output hasil *Clustering* yang berupa contoh topik skripsi yang sesuai dengan bidang yang cocok bagi mahasiswa tersebut. Sistem ini dibangun menggunakan PHP. PHP adalah sebuah bahasa *scripting* yang terpasang pada HTML.

#### 6. Pengujian Sistem

Setelah sistem dapat berjalan dan mampu melakukan rekomendasi topic skripsi, penelitian ini melakukan pengujian terhadap modul-modul sistem menggunakan *Blackbox Testing*. Pengujian ini akan menguji setiap modul, fungsi, menu serta prosedur system apakah sudah berjalan baik atau belum.

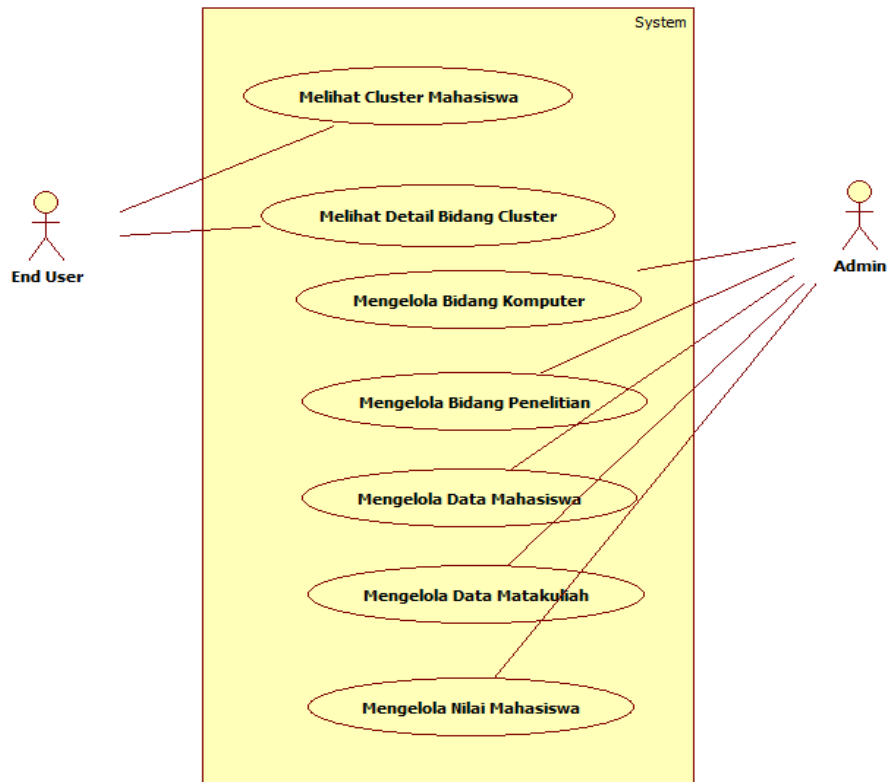
## 4. Hasil dan Pembahasan

### 4.1 Gambaran Umum Sistem Penentuan topik skripsi Mahasiswa

Sistem penentuan topik skripsi mahasiswa dengan menggunakan algoritma *EM Clustering* dan algoritma *naïve bayes* ini melakukan kluster mahasiswa teknik informatika menjadi kedalam beberapa kluster topik skripsi. Terdapat 18 area bidang ilmu computer berdasarkan Computer Science Curricula 2013 [19] yang di keluarkan oleh ACM dan IEEE. Kemudian data bidang skripsi tersebut disesuaikan dengan mata kuliah yang terkait dalam kurikulum tersebut. Selanjutnya, peneliti melakukan mencocokkan kode mata kuliah dalam kurikulum tersebut sesuai dengan kode mata kuliah yang diajarkan di STIKOM Dinamika Bangsa Jambi.

### 4.2 Use Case Diagram

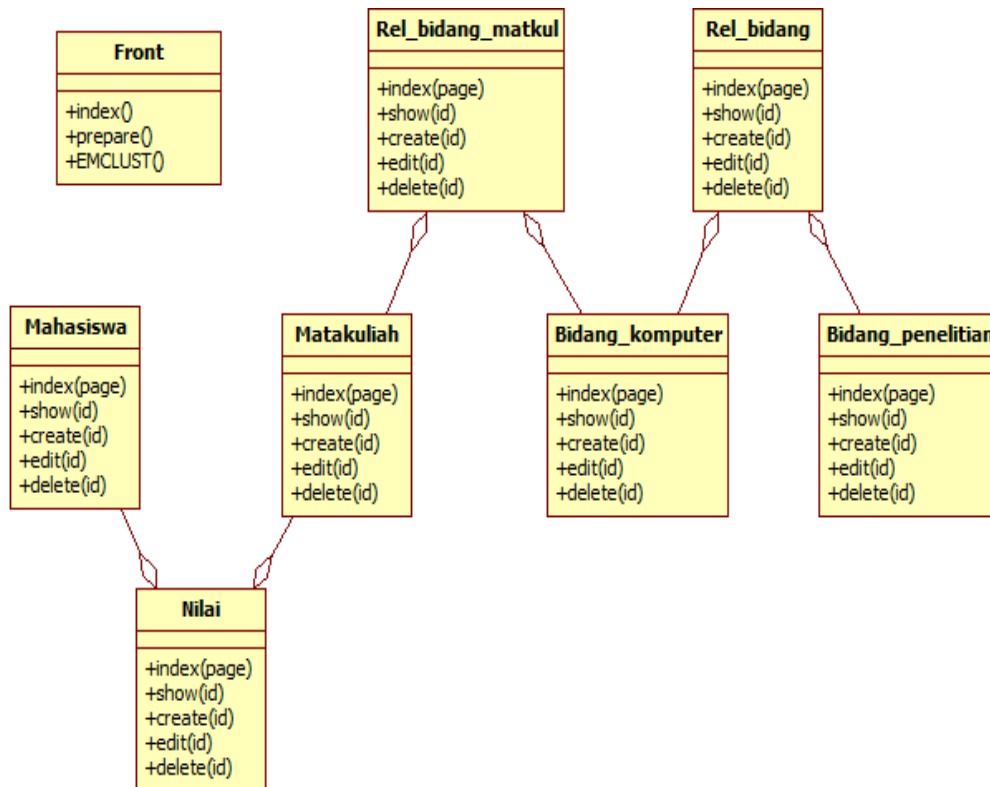
Usecase merupakan deskripsi dari fungsionalitas bagian Sistem Informasi yang akan dibuat [4][5]. Berikut adalah gambaran mengenai sistem:



Gambar 2. Use Case Diagram

### 4.3 Class Diagram

Pada gambar 3 berikut akan menjelaskan hubungan antar class atau objek yang ada dalam sistem penentuan topik skripsi mahasiswa dengan algoritma *EM Clustering* dan algoritma naïve bayes.



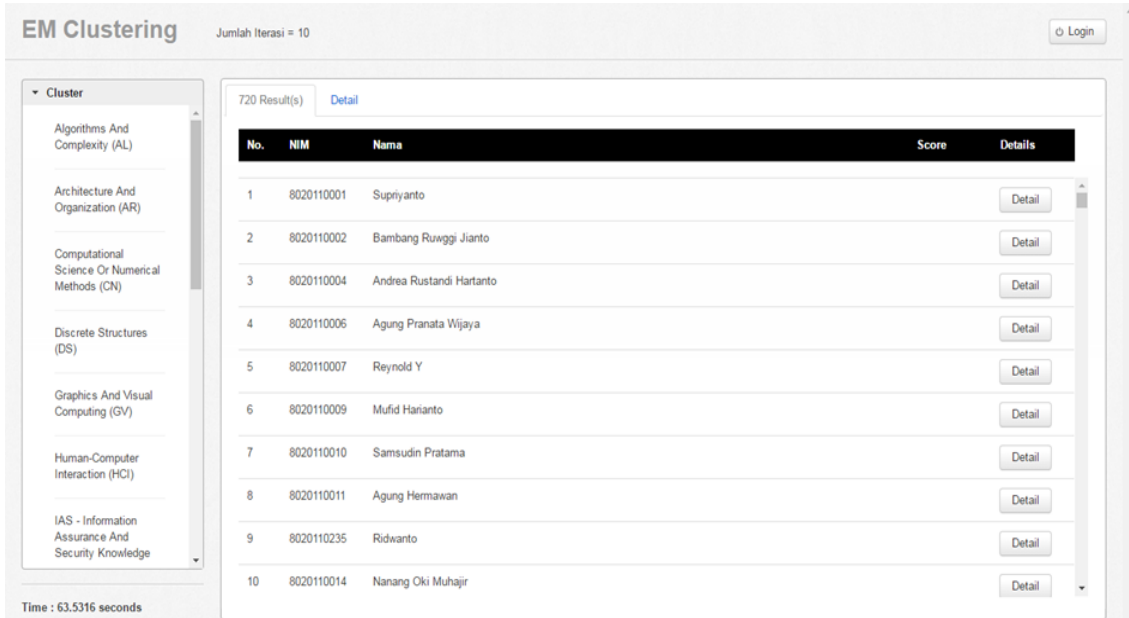
Gambar 3. Class Diagram Sistem Penentuan Topik Skripsi



#### 4.4 Hasil Implementasi

Pada bagian ini akan dijelaskan tentang implementasi dari sistem penentuan topik skripsi dengan algoritma *EM Clustering* dan algoritma naïve bayes. Implementasi pada setiap halaman sistem terdiri dari:

##### 4.4.1 Tampilan Awal Sistem



EM Clustering Jumlah Iterasi = 10

Cluster

- Algorithms And Complexity (AL)
- Architecture And Organization (AR)
- Computational Science Or Numerical Methods (CN)
- Discrete Structures (DS)
- Graphics And Visual Computing (GV)
- Human-Computer Interaction (HCI)
- IAS - Information Assurance And Security Knowledge

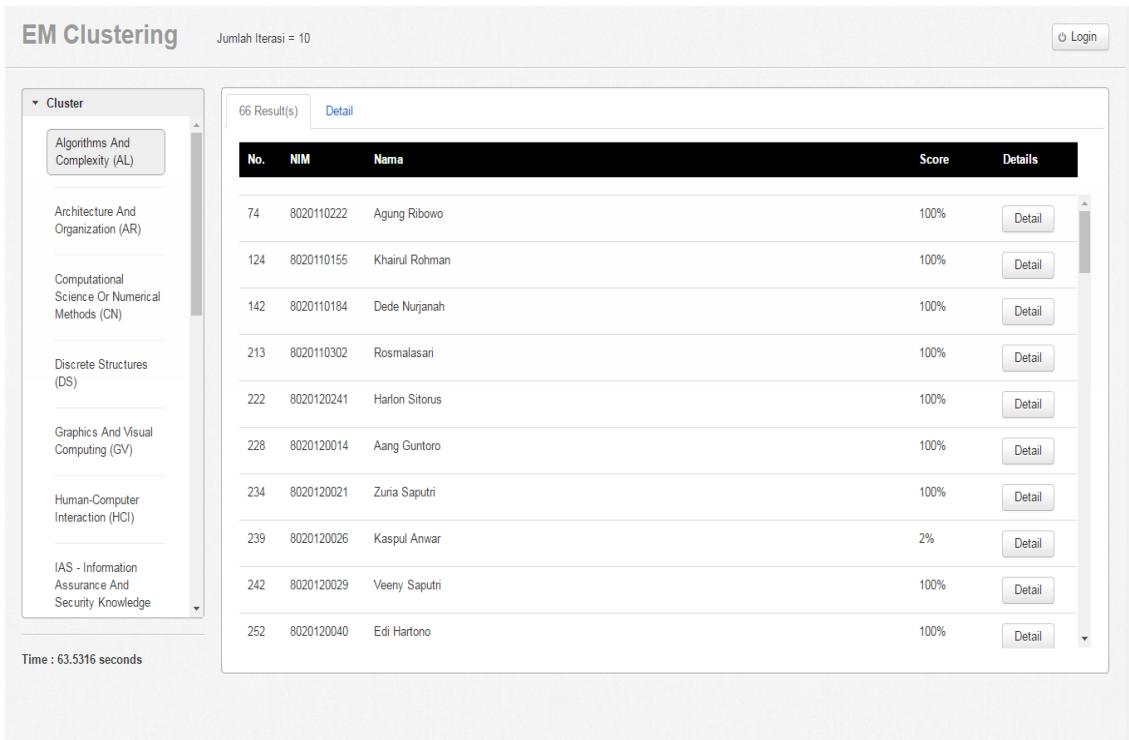
720 Result(s) Detail

No.	NIM	Nama	Score	Details
1	8020110001	Supriyanto		Detail
2	8020110002	Bambang Ruwggi Jianto		Detail
3	8020110004	Andrea Rustandi Hartanto		Detail
4	8020110006	Agung Pranata Wijaya		Detail
5	8020110007	Reynold Y		Detail
6	8020110009	Mufid Harianto		Detail
7	8020110010	Samsudin Pratama		Detail
8	8020110011	Agung Hermawan		Detail
9	8020110235	Ridwanto		Detail
10	8020110014	Nanang Oki Muhajir		Detail

Time : 63.5316 seconds

Gambar 4. Tampilan Awal Sistem

##### 4.4.2 Tampilan Sistem Dengan Pilihan Kluster



EM Clustering Jumlah Iterasi = 10

Cluster

- Algorithms And Complexity (AL)
- Architecture And Organization (AR)
- Computational Science Or Numerical Methods (CN)
- Discrete Structures (DS)
- Graphics And Visual Computing (GV)
- Human-Computer Interaction (HCI)
- IAS - Information Assurance And Security Knowledge

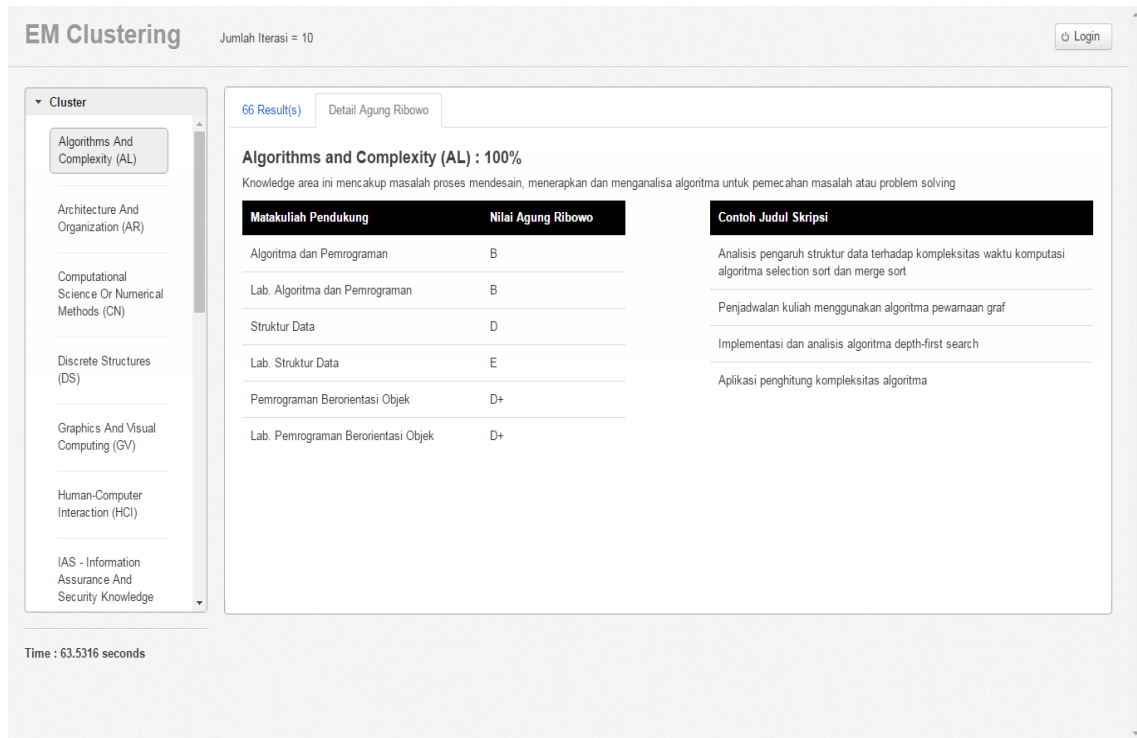
66 Result(s) Detail

No.	NIM	Nama	Score	Details
74	8020110222	Agung Ribowo	100%	Detail
124	8020110155	Khairul Rohman	100%	Detail
142	8020110184	Dede Nurjanah	100%	Detail
213	8020110302	Rosmalasari	100%	Detail
222	8020120241	Harlon Sitorus	100%	Detail
228	8020120014	Aang Guntoro	100%	Detail
234	8020120021	Zuria Saputri	100%	Detail
239	8020120026	Kaspul Anwar	2%	Detail
242	8020120029	Veeny Saputri	100%	Detail
252	8020120040	Edi Hartono	100%	Detail

Time : 63.5316 seconds

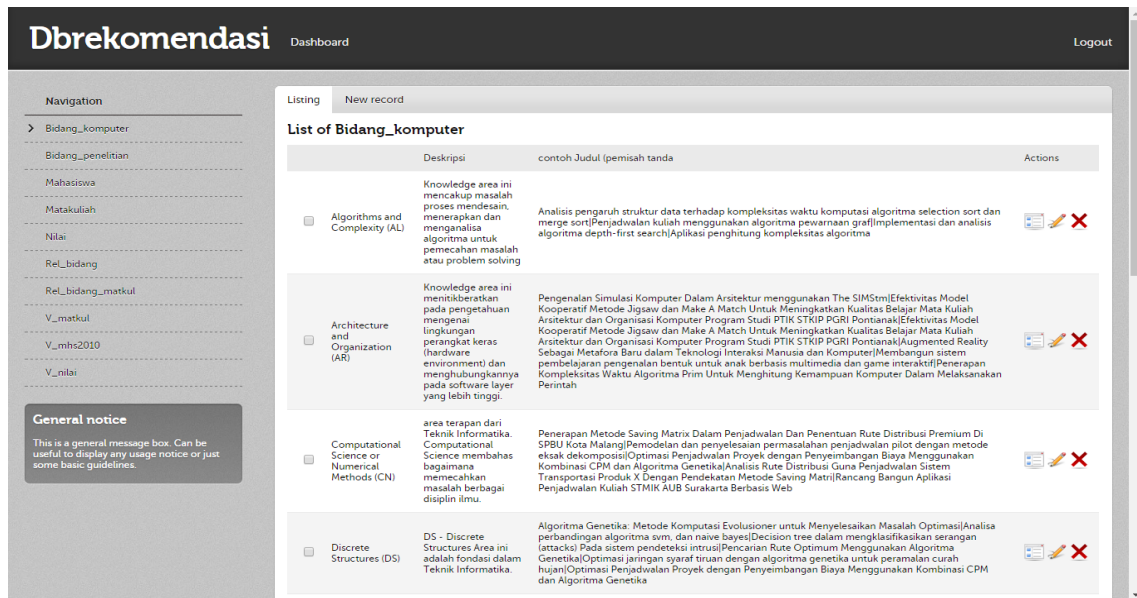
Gambar 5. Tampilan Sistem Dengan Pilihan Kluster

### 4.4.3 Tampilan Sistem Dengan Detail Kluster



Gambar 6. Tampilan Sistem Dengan Detail Kluster

### 4.4.4 Tampilan Back –End Tabel Bidang Komputer



Gambar 7. Tampilan Back –End Tabel Bidang Komputer

### 4.5 Pengujian Aplikasi

Pada tahap ini, penulis melakukan pengujian terhadap aplikasi untuk memeriksa dan mengurangi adanya kemungkinan terjadinya error yang akan berdampak besar pada hasil kelanjutan proses aplikasi nantinya. Aplikasi akan diuji dengan metode pengujian blackbox yang merupakan salah satu metode pengujian perangkat lunak. Metode *blackbox* merupakan sebuah metode yang digunakan untuk menemukan

kesalahandanmendemonstrasikan fungsional aplikasi saat dioperasikan, apakah input diterima dengan benar dan output yang dihasilkan telah sesuai dengan yang diharapkan. Adapun beberapa tahap pengujian yang telah penulis lakukan adalah sebagai berikut :

#### 4.5.1 Pengujian Modul Kluster Mahasiswa

Pada tahap ini, dilakukan pengujian pada modul mulai untuk mengetahui apakah proses Kluster Mahasiswa berjalan dengan baik atau tidak. Hasil pengujian pada modul mulai dapat dilihat pada Tabel berikut:

Tabel 1. *Pengujian Modul Kluster Mahasiswa*

Modul Uji	Tes	Input	Output	Status
Kluster Mahasiswa	- Buka halaman Awal - Klik Salah satu kluster	klik tombol Mulai	Tampilkan halaman Kluster Mahasiswa	Baik

#### 4.5.2 Pengujian Modul Detail Bidang Kluster

Pada tahap ini, dilakukan pengujian pada modul petunjuk untuk mengetahui apakah proses penampilan halaman Detail Bidang Kluster berjalan dengan baik atau tidak. Hasil pengujian pada modul petunjuk dapat dilihat pada Tabel berikut:

Tabel 2. *Pengujian Modul Detail Bidang Kluster*

Modul Uji	Tes	Input	Output	Status
Detail Bidang Kluster	- Buka halaman Awal - Klik Detail Bidang Kluster	Klik tombol Detail Bidang Kluster	Tampilkan halaman Detail Bidang Kluster	Baik

#### 4.5.3 Pengujian Modul Kelola Bidang Komputer

Pada tahap ini, dilakukan pengujian pada modul tentang untuk mengetahui apakah proses penampilan halaman Detail Bidang Kluster dengan baik atau tidak. Hasil pengujian pada modul tentang dapat dilihat pada Tabel berikut:

Tabel 3. *Pengujian Modul Kelola Bidang Komputer*

Modul Uji	Tes	Input	Output	Status
Kelola Bidang Komputer	- Buka halaman Awal - Klik Kelola Bidang Komputer	Klik tombol Kelola Bidang Komputer	Tampilkan halaman Kelola Bidang Komputer	Baik

#### 4.5.4 Pengujian Modul Kelola Bidang Penelitian

Pada tahap ini, dilakukan pengujian pada modul musik untuk mengetahui apakah proses Kelola Bidang Penelitian berjalan dengan baik atau tidak. Hasil pengujian pada modul tentang dapat dilihat pada Tabel berikut:

Tabel 4. *Pengujian Modul Kelola Bidang Penelitian*

Modul Uji	Tes	Input	Output	Status
Kelola Bidang Penelitian	- Buka halaman awal - Klik tombol Kelola Bidang Penelitian	Klik tombol Kelola Bidang Penelitian	Tampilkan halaman Kelola Bidang Penelitian	Baik

#### 4.5.5 Pengujian Modul Kelola Data Mahasiswa

Pada tahap ini, dilakukan pengujian pada modul materi untuk mengetahui apakah proses Kelola Data Mahasiswa berjalan dengan baik atau tidak. Hasil pengujian pada modul materi dapat dilihat pada Tabel berikut:

Tabel 5. *Pengujian Modul Kelola Data Mahasiswa*

Modul Uji	Tes	Input	Output	Status
Kelola Data Mahasiswa	- Buka halaman awal - Klik tombol Kelola Bidang Mahasiswa	Klik tombol Kelola Bidang Mahasiswa	Berhasil menampilkan halaman Kelola Bidang Mahasiswa	Baik

#### 4.5.6 Pengujian Modul Kelola Data Mata Kuliah

Pada tahap ini, dilakukan pengujian pada modul level 1 untuk mengetahui apakah proses menampilkan Data Mata Kuliah berjalan dengan baik atau tidak. Hasil pengujian pada modul level 1 dapat dilihat pada Tabel berikut:

Tabel 6. *Pengujian Modul Kelola Data Mata Kuliah*

Modul Uji	Tes	Input	Output	Status
Kelola Data Mata Kuliah	- Buka halaman awal - Klik tombol Kelola Data Mata Kuliah	Klik tombol Kelola Data Mata Kuliah	Menampilkan halaman Kelola Data Mata Kuliah	Baik

#### 4.5.7 Pengujian Modul Kelola Nilai Mahasiswa

Pada tahap ini, dilakukan pengujian pada modul level 2 untuk mengetahui apakah proses menampilkan halaman Kelola Nilai Mahasiswa berjalan dengan baik atau tidak. Hasil pengujian pada modul level 2 dapat dilihat pada Tabel 19:

Tabel 7. *Pengujian Modul Kelola Nilai Mahasiswa*

Modul Uji	Tes	Input	Output	Status
Kelola Data Nilai Mahasiswa	- Buka halaman awal - Klik tombol Kelola Nilai Mahasiswa	Klik tombol Kelola Nilai Mahasiswa	Berhasil menampilkan halaman Kelola Nilai Mahasiswa	Baik

#### 4.6 Analisis Kode Program

```

126
127     $iterasi = 0;
128     do{
129         //E Step
130         //classify use last model parameter
131         $newy = array();
132         $a=0;
133         $probcluster = array();
134
135         foreach ($nilaimhs as $nim => $mknilai) {
136             $prob = array();
137             $total = 0;
138             foreach ($model_par as $cluster => $mknilais) {
139                 $prob[$cluster] = 1;
140                 foreach ($mknilais as $key => $value) {
141                     $sh1 = ($mknilai[$key]==1?$value:1-$value);
142                     // $probcluster[$cluster][$nim]['mknilais'][$key] = $sh1;
143                     $prob[$cluster] *= $sh1;
144                 }
145                 if ($prob[$cluster]==0) {
146                     $prob[$cluster] = (1/pow(10, 308)) * $kelas["param"][$cluster];
147                 }
148                 else{
149                     $prob[$cluster] *= $kelas["param"][$cluster];
150                 }
151                 $probcluster[$cluster][$nim]['kelasparam'] = $kelas["param"][$cluster];
152                 $probcluster[$cluster][$nim]['hasil'] = $prob[$cluster];
153                 $total += $prob[$cluster];
154             }
155             foreach ($prob as $key => $value) {
156                 $newy[$nim][$key] = round($value/$total,2);

```

Gambar 8. Kode Program EM Clustering

Berikut merupakan kode program berupa fungsi EMCLUST dengan parameter maxiterasi. Kode ini merupakan kode utama dari proses *EM Clustering* dan *Algoritma Naïve Bayes*. Ada 2 proses utama pada algoritma ini, yaitu proses *expectation* dan proses *Maximization*. Awalnya dilakukan proses *Expectation* seperti pada baris kode 129-158, kemudian dilakukan pengecekan apakah sudah convergence (tidak ada lagi perubahan pada nilai Y), jika sudah maka iterasi di stop, jika tidak maka dilanjutkan proses *Maximization*. Hal ini terus dilakukan berulang-ulang sehingga dicapai Convergence atau sudah mencapai Iterasi maksimum.

## 5. Kesimpulan

Pada bagian ini terdapat kesimpulan dan saran dari penelitian ini:

### 5.1 Simpulan

Adapun kesimpulan dari penelitian ini adalah Sistem Penentuan topik skripsi Mahasiswa STIKOM Dinamika Bangsa Jambi dengan menggunakan algoritma *EM Clustering* dan Algoritma Naïve Bayes dapat digunakan untuk merekomendasikan topik skripsi pada mahasiswa stikom. Algoritma *EM Clustering* dengan *Algoritma Naïve Bayes* dapat melakukan *Clustering* mahasiswa berdasarkan topik yang cocok. Sistem ini dapat digunakan untuk membantu mahasiswa STIKOM Dinamika Bangsa Jambi yang akan mengambil skripsi dengan memberikan sebuah rekomendasi mengenai topik ilmu komputer mana yang sesuai dengan kompetensi mereka serta membantu STIKOM Dinamika Bangsa Jambi dalam memberikan bimbingan dan arahan kepada mahasiswa dalam hal memilih topik skripsi yang sesuai dengan kompetensi mahasiswa. Output dari sistem ini sendiri berupa contoh referensi topik skripsi.

### 5.2 Saran

Perlu pengembangan implementasi dari penelitian ini dengan metode yang teknik pembelajaran mesin atau non pembelajaran mesin. Diharapkan penelitian ini dapat digunakan untuk rekomendasi dalam bidang lainnya serta perlu melakukan evaluasi terhadap sistem yang telah dibuat.

## 6. Daftar Rujukan

- [1] Sugiyanto, Suprapedi dan Himawan ., Heribertus, 2009, Penentuan Kompetensi Mahasiswa Berdasarkan Prestasi Akademik, Sertifikasi Kompetensi, Minat, Dan Kegiatan Pendukung. *Jurnal Teknologi Informasi*, Volume 5 Nomor 2, Oktober 2009, ISSN 1414-9999.
- [2] Bahar, 2011, Penentuan Jurusan Sekolah Menengah Atas dengan Algoritma Fuzzy C-Means. TESIS. Universitas Dian Nuswantoro
- [3] Barkah, Azhari Shouni dan Sinatrya, Elya., 2015, Penerapan Data Mining untuk Kosentrasi Mahasiswa Di STMIK AMIKOM Purwokerto
- [4] Dennis, Alan; Wixom, Haley Barbara;& Tegarden, David., 2005, *Systems Analysis and Design with UML Version 2.0 : An Object-Oriented Approach*. Second Edition. United States of America : John Wiley & Sons, Inc.
- [5] Dennis, Alan; Wixom, Haley Barbara;& Tegarden, David., 2010, *System Analysis and Design with UML: An Object-Oriented Approach*. 3rd Edition. United States of America : John Wiley & Sons, Inc.
- [6] Hermawati, Fajar Astuti, 2013, *Data Mining*. Yogyakarta: Andi.
- [7] Kristanto, Obbie. 2014. Penerapan Algoritma Klasifikasi Data Mining Id3 Untuk Menentukan Penjurusan Siswa SMAN 6 Semarang. SKRIPSI. Universitas Dian Nuswantoro.
- [8] Law rencus Sirait, 2013, Analisis Algoritma EM (Em) Dalam Penanganan Missing Value, Volume: 113040228, Nomor , 005.1.
- [9] Madcoms, 2013, Mahir dalam 7 Hari Adobe Dreamweaver CS6 dengan Pemrograman PHP & MySQL. Yogyakarta: Andi .
- [10] G. McLachlan and T. Krishnan, 1997, *The EM Algorithm and Extensions*. New York: Wiley Series in Probability and Statistics.
- [11] Munawar, 2005, *Pemodelan Visual dengan UML*. Edisi Pertama. Yogyakarta: Penerbit Graha Ilmu.
- [12] Narwati, 2010, Pengelompokan Mahasiswa Menggunakan Algoritma K-Means. *Jurnal Dinamika Informatika* Vol 2 No 2 tahun 2010 ISSN: 2085-3343
- [13] Pressman, Roger S, 2001, *Software Engineering : A Practitioner's Approach*. Fifth Edition. New York : McGraw-Hill.
- [14] Pudjo, Widodo, Prabowo., Trias, Handayanto, Rahmadya., Herlawati, 2013, *Penerapan Data Mining dengan Matlab*. Bandung: Rekayasa
- [15] Ridwan, Mujid. Hadi, Suyono. dan Sarosa, M, 2013, Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naïve Bayes Classifier. *Jurnal EECCIS* Vol.7, No. 1, Juni 2013 ISSN: 1978-3345
- [16] Septian, Gungun, 2011, *Trik Pintar Menguasai Codeigniter*. Jakarta: Gramedia
- [17] Widodo, PB. Handayanto, RT. dan Herlawati, 2013, *Penerapan Data Mining Dengan Matlab*. Bandung: Rekayasa Sain
- [18] Elya Sinatrya, 2016, *Jurnal aplikasi untuk menentukan matakuliah konsentrasi mahasiswa STMIK Amikom Purwokerto*.
- [19] The Joint Task Force on Computing Curricula: Association for Computing Machinery and IEEE Computer Society. *Computer Science Curricula*. 2013. Ironman Draft (Version 0.8).
- [20] Saraswati, N. W. (2013): *Naïve Bayes Classifier Dan Support Vector Machines Untuk Sentiment Analysis*. Seminar Nasional Sistem Informasi Indonesia.
- [21] Samodra, J., Sumpeno, S., & Hariadi, M. (2009): *Klasifikasi Teks Dokumen Bahasa Indonesia dengan menggunakan Naive Bayes*. Seminar Nasional Electrical, Informatics and It's Education.
- [22] Abdul Kadir., 2003, *Dasar Pemrograman WEB Dinamis Menggunakan PHP*. Yogyakarta : Penerbit Andi Yogyakarta