

**PREDIKSI DAN PEMETAAN DATA MAHASISWA FAKULTAS TEKNIK  
UNIVERSITAS NEGERI GORONTALO MENGGUNAKAN PENDEKATAN  
DATA MINING**

**PREDICTION AND DATA MAPPING of STUDENTS OF ENGINEERING FACULTY,  
UNIVERSITAS NEGERI GORONTALO USING DATA MINING**

Lillyan Hadjaratie

Jurusan Teknik Informatika Universitas Negeri Gorontalo

**ABSTRAK:** Penelitian “Prediksi dan Pemetaan Data Mahasiswa Fakultas Teknik Menggunakan Pendekatan Data Mining” bertujuan untuk membangkitkan informasi dan pengetahuan dari data akademik kemahasiswaan dengan: (1) Mengklasifikasi data mahasiswa aktif dan lulusan untuk memprediksi kategori IPK dan kategori Lama Studi; (2) Mengklaster data mahasiswa aktif dan lulusan berdasarkan kategori IPK dan Lama\_Studi. Metode klasifikasi yang digunakan untuk mengklasifikasi data mahasiswa aktif dan lulusan dalam penelitian ini adalah metode *Decision Tree*, *Artificial Neural Network* dan *K-Nearest Neighbour*, sedangkan metode yang digunakan untuk mengklaster data mahasiswa aktif dan lulusan adalah metode Hirarkis (*Average Between Linkage*) dan Non-Hirarkis (*K-Means*). Penelitian ini menghasilkan informasi akademik berupa hasil klasifikasi data mahasiswa aktif dan lulusan untuk memprediksi kategori IPK dan Lama\_Studi serta pemetaan data mahasiswa aktif dan lulusan berdasarkan sasaran mutu IPK, sebagai sebuah sistem peringatan dini (*early morning*) dan bahan pertimbangan dalam proses pengambilan kebijakan dan keputusan.

Kata Kunci : Data Mining, Klasifikasi, Klasterisasi, Data Mahasiswa, IPK, Lama Studi

**Abstract:** This research aims to generate information and knowledge of student academic data by: (1) Classify the data of current students and graduates to predict the category of GPA and period of Study, (2) cluster data of current students and graduates based on GPA and period of study category. This study uses classification method such as Decision Tree Method, Artificial Neural Networks and K - Nearest Neighbour to classifying the data of current students and graduates. Furthermore, to cluster the data of current students and graduates Hierarchical method (Average Linkage Between ) and Non - hierarchical (K - Means) are used. The result of this research is academic information such as classification data of current students and graduates to predict the GPA category and period of study, data mapping of student based on data quality of GPA, as a warning system (early morning ) and considerations in policy and decision making process .

Keywords: data mining, classification, clustering, current and graduate student, GPA, period of study

**PENDAHULUAN**

Program Studi merupakan garda terdepan dalam penyelenggaraan pendidikan dari sebuah Perguruan Tinggi, sehingga senantiasa melakukan evaluasi guna meningkatkan mutu dan efisiensi perguruan tinggi termasuk peningkatan kualitas lulusan. Fakultas Teknik Universitas Negeri Gorontalo merupakan salah satu fakultas di lingkungan Universitas Negeri Gorontalo yang memiliki jumlah mahasiswa yang cukup banyak. Hal ini terlihat dari peningkatan jumlah calon mahasiswa baru pada setiap tahun ajaran. Permasalahan yang sering terjadi adalah masih banyaknya jumlah mahasiswa yang lulus dengan lama studi melampaui waktu yang telah ditetapkan dengan perolehan Indeks Prestasi Kumulatif (IPK) yang relatif rendah yang dapat mempengaruhi mutu lulusan Perguruan Tinggi. Seiring dengan terus bertambahnya jumlah mahasiswa di Fakultas Teknik UNG maka jumlah data kemahasiswaan terus meningkat sehingga terjadi penumpukan data yang belum diolah dengan optimal untuk menggali informasi dan pengetahuan baru yang dapat digunakan sebagai bahan pertimbangan pimpinan dalam proses pengambilan kebijakan dan keputusan. Data ini juga sebagai peringatan dini (*early warning*) bagi mahasiswa tertentu yang berdasarkan hasil prediksi dinyatakan berpotensi lulus dengan melampaui ketentuan lama studi ataupun berpotensi lulus dengan perolehan IPK yang rendah.

*Knowledge Discovery in Database* (KDD) adalah proses menentukan informasi yang berguna serta pola-pola yang ada dalam data. Informasi ini terkandung dalam basis data yang berukuran besar yang sebelumnya tidak diketahui dan potensial bermanfaat (Han & Kamber, 2006). Tahapan proses KDD terdiri dari: *Data Selection*, *Pre-Processing* dan *Cleaning Data*, *Transformation*, *Data Mining* dan *Interpretation / Evaluasi*. Fungsi-fungsi dalam data mining mengacu pada Larose (2005) yang membaginya ke dalam enam fungsi yaitu (Susanto & Suryadi, 2010), yaitu : Fungsi Deskripsi, Estimasi, Prediksi, Klasifikasi, Klusterisasi dan Asosiasi.

Teknik klasifikasi merupakan pendekatan untuk menjalankan fungsi klasifikasi dalam Data Mining yaitu untuk menggolongkan data. Teknik klasifikasi ini dapat pula digunakan untuk melakukan prediksi atas informasi yang belum diketahui sebelumnya. Beberapa algoritma yang dapat digunakan antara lain adalah algoritma Decision Tree C.45, *Artificial Neural Networks* (ANN), *K-Nearest Neighbor* (KNN), algoritma *Naive Bayes*, *Neural Network* serta algoritma lainnya. Beberapa penelitian yang menggunakan teknik data mining pada data set akademik dan kemahasiswaan telah banyak dilakukan, antara lain adalah penelitian yang dilakukan oleh Al-Radaideh, dkk (2006) yang menganalisa dan mengevaluasi data akademik dengan menggunakan metode pohon keputusan (*decision tree*); Pramudyo (2008) melakukan penelitian tentang klasifikasi mahasiswa baru berdasarkan prediksi Indeks Prestasi Semester dengan menggunakan metode *Case Base Reasoning* (CBR); Jananto (2010) melakukan penelitian untuk mengklasifikasi kinerja akademik mahasiswa dengan menggunakan algoritma *Supervised*

*Learning In Quest (SLIQ)*; Sunjana (2010) dalam penelitian Aplikasi Mining Data Mahasiswa dengan menggunakan metode klasifikasi *Decision Tree*.

Klasterisasi merupakan suatu teknik atau metode untuk mengelompokkan sejumlah besar data menjadi suatu bagian-bagian kecil data yang mempunyai atribut kemiripan dalam sifat, letak, ciri atau filter lain yang telah ditentukan. Teknik atau metode klasterisasi dapat diklompokkan menjadi dua kategori besar, yaitu : (1) Metode Hirarki (*Hierarchical Clustering*) dan (2) Metode Non-Hirarki/Partisi (*Partitional Hierarchical*). Dalam penelitiannya, Fahim dkk (2008), Gunawan (2009) dan Karlita (2011) menyatakan bahwa Algoritma *K-Modes* merupakan metode pengembangan dari *K-Means* yang mampu mengelompokkan data kategorikal dan menghasilkan klaster yang lebih stabil dengan waktu komputasi yang lebih singkat daripada metode *K-Means*. Martiana dkk (2009) dalam penelitiannya pernah mengkomparasikan penggunaan metode *Agglomerative Hierarchical (Single Linkage Hierarchical dan Centroid Linkage Hierarchical)* dengan metode *Partitional Clustering* yakni *K-Means*, untuk mengklaster bidang tugas akhir mahasiswa.

## METODE PENELITIAN

Penelitian ini dilakukan dalam beberapa langkah, yaitu : pengumpulan data, seleksi dan pembersihan data, transformasi data, klasifikasi data dan klasterisasi data.

## HASIL DAN PEMBAHASAN

### 1. Klasifikasi

Metode-metode yang digunakan dalam proses klasifikasi adalah metode *Decision Tree*, *Artificial Neural Network* dan *K-Nearest Neighbour*.

#### a. *Decision Tree*

Klasifikasi data mahasiswa aktif berdasarkan IPK dengan metode *Decision Tree* menghasilkan 7 klasifikasi ditunjukkan pada tabel 1.

Dari ketujuh klasifikasi yang terbentuk, dapat ditabulasi jumlah masing-masing mahasiswa yang memiliki IPK dengan kategori rendah, sedang dan tinggi dalam Tabel 2. Klasifikasi data mahasiswa lulusan berdasarkan IPK dengan metode *Decision Tree* menghasilkan 7 klasifikasi ditunjukkan pada table 3.

*Tabel 1. Klasifikasi dan Karakteristik Data Mahasiswa Aktif berdasarkan IPK*

Klasifikasi	Karakteristik
Ke-1	Mahasiswa dengan Jurusan Sipil, Jenis Kelamin Laki-Laki, dan

---

	Asal Sekolah SMA
Ke-2	Mahasiswa dengan Jurusan Sipil, Jenis Kelamin Laki-Laki, dan Asal Sekolah SMK, Madrasah dan Penjenjangan
Ke-3	Mahasiswa dengan Jurusan Sipil, Jenis Kelamin Perempuan
Ke-4	Mahasiswa dengan Jurusan Informatika, Jenis Kelamin Laki-Laki
Ke-5	Mahasiswa dengan Jurusan Informatika, Jenis Kelamin Perempuan
Ke-6	Mahasiswa dengan Jurusan Elektro, Arsitektur, Industri
Ke-7	Mahasiswa dengan Jurusan Kriya

---

*Tabel 2. Klasifikasi dan Karakteristik Data Mahasiswa Aktif berdasarkan IPK*

Klasifikasi	IPK		
	Rendah	Sedang	Tinggi
Ke-1	88	17	2
Ke-2	129	5	0
Ke-3	60	31	0
Ke-4	205	<b>171</b>	9
Ke-5	70	133	8
Ke-6	<b>275</b>	105	5
Ke-7	30	42	<b>10</b>

---

Tabel 3. Klasifikasi dan Karakteristik Data Mahasiswa Aktif berdasarkan IPK

Klasifikasi	Karakteristik
Ke-1	Mahasiswa dengan Predikat Sangat Memuaskan, Jenjang D3
Ke-2	Mahasiswa dengan Predikat Sangat Memuaskan, Jenjang S1
Ke-3	Mahasiswa dengan Predikat Memuaskan dan Terpuji, Jenis Kelamin Laki-Laki, Jurusan Sipil, Informatika, Kriya
Ke-4	Mahasiswa dengan Predikat Memuaskan dan Terpuji, Jenis Kelamin Laki-Laki, Jurusan Elektro, Arsitektur, Industri
Ke-5	Mahasiswa dengan Predikat Memuaskan dan Terpuji, Jenis Kelamin Perempuan

---

Dari kelima klasifikasi yang terbentuk, dapat ditabulasi jumlah masing-masing mahasiswa yang memiliki IPK dengan kategori rendah, sedang dan tinggi dalam Tabel 4.

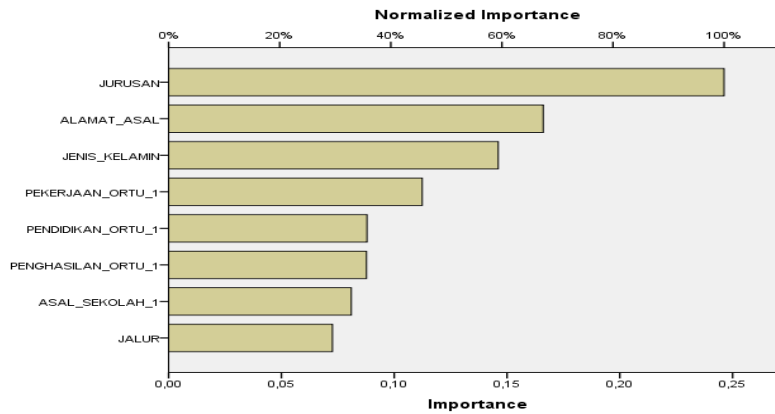
Tabel 4. Jumlah Mahasiswa Lulusan setiap Klasifikasi berdasarkan Kategori IPK

Klasifikasi	IPK		
	Rendah	Sedang	Tinggi
Ke-1	1	<b>993</b>	14
Ke-2	0	55	8
Ke-3	<b>130</b>	3	8
Ke-4	50	10	4
Ke-5	78	5	<b>19</b>

**b. Artificial Neural Network**

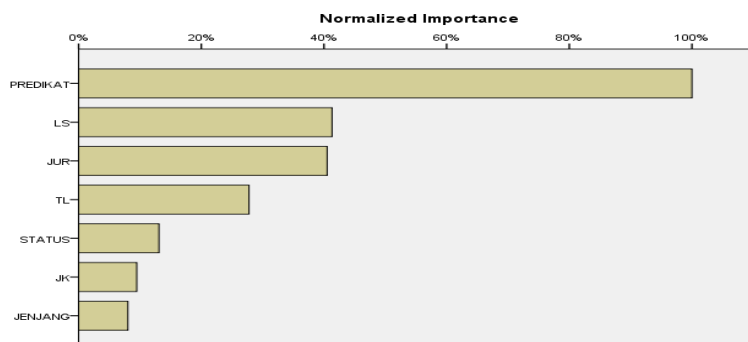
Klasifikasi dengan menggunakan metode Jaringan saraf Tiruan (*Artificial Neural Network*) untuk mengetahui variabel yang dianggap penting (*independent variabel importance*).

Klasifikasi data mahasiswa aktif berdasarkan IPK terlihat pada Gambar 1.



Gambar 1. Persentase Independent Variabel Importance berdasarkan IPK Mahasiswa Aktif

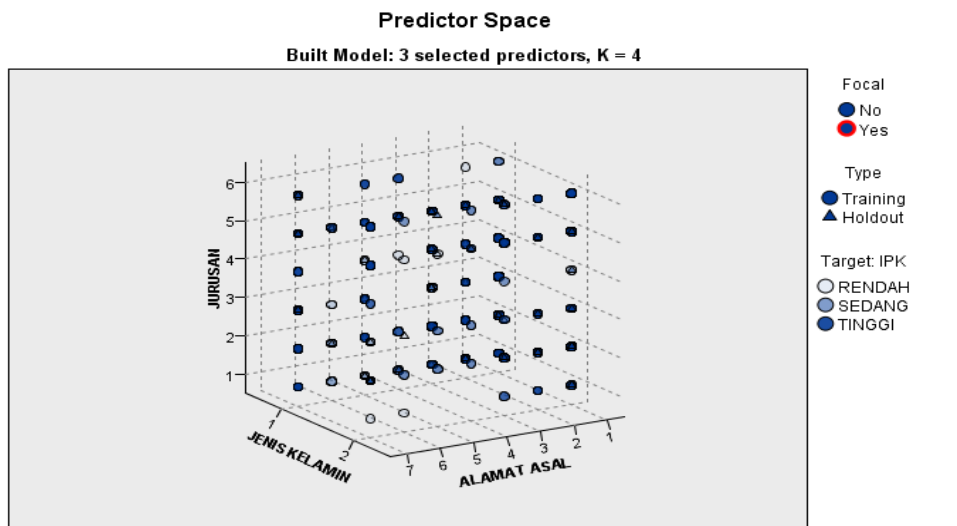
Klasifikasi data mahasiswa aktif berdasarkan IPK terlihat pada Gambar 2.



Gambar 2. Persentase Independent Variabel Importance berdasarkan IPK Mahasiswa Lulusan

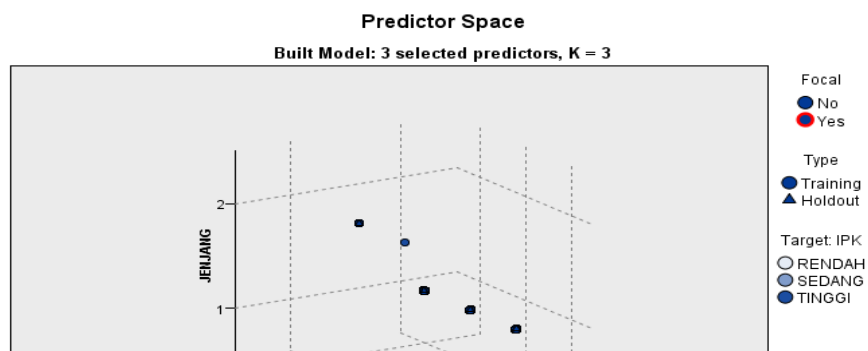
c. *K-Nearest Neighbour*

K-Nearest Neighbour (KNN) merupakan salah satu metode untuk mengklasifikasikan suatu data baru berdasarkan similaritas (kemiripan) dengan karakteristik pada data lama (*training set*) dan umumnya menggunakan metrik jarak (*euclidian*). Dalam penelitian ini, metode KNN digunakan untuk memprediksi kelas target (*variabel dependent*) IPK dari data mahasiswa aktif berdasarkan pola perolehan IPK dari data lulusan. Klasifikasi data mahasiswa aktif berdasarkan kategori IPK ditunjukkan oleh Gambar 3.



Gambar 3. Klasifikasi KNN Data Mahasiswa Aktif berdasarkan IPK

Klasifikasi data mahasiswa aktif berdasarkan kategori IPK ditunjukkan oleh Gambar 4.



Gambar 4. Klasifikasi KNN Data Mahasiswa Lulusan berdasarkan IPK

## 2. Klasterisasi

Tujuan dari analisis kluster adalah mengelompokkan obyek berdasarkan kesamaan karakteristik di antara obyek-obyek tersebut, tanpa memiliki kelas target seperti pada teknik klasifikasi. Metode pengelompokan yang umumnya digunakan pada analisis kluster adalah :

### a. Metode Hirarkis

Metode Hirarkis memulai pengelompokan dengan dua atau lebih obyek yang memiliki kesamaan paling dekat, kemudian diteruskan pada obyek yang lain dan seterusnya hingga cluster akan membentuk semacam “pohon” dimana terdapat tingkatan (hirarki) yang jelas antar obyek, dari yang paling mirip hingga yang paling tidak mirip. Alat bantu untuk memperjelas proses hirarki ini disebut dengan “Dendogram”.

Berdasarkan hasil *finalcluster* data mahasiswa aktif dengan jumlah  $k = 3$ , maka diperoleh karakteristik dari masing-masing *cluster* adalah seperti yang terlihat pada Tabel 5.

Tabel 5. Final Cluster Centers Mahasiswa Aktif

VARIABEL	Cluster		
	1	2	3
IPK	1	2	2
Jalur	3	3	3
Alamat_Asal	7	2	1
Jenis_Kelamin	1	2	1
Smean(Pendidikan_Ortu)	2	2	3
Smean(Pekerjaan_Ortu)	4	3	1
Smean(Penghasilan_Ortu)	2	2	3
Smean(Asal_Sekolah)	2	2	2

Dari Tabel 5 dapat didefinisikan bahwa :

- **Cluster-1**

Berisikan mahasiswa-mahasiswa dengan perolehan IPK kategori rendah, jalur masuk mandiri, alamat asal dari luar daerah Provinsi Gorontalo, jenis kelamin laki-laki, pendidikan

orang tua kategori menengah, pekerjaan orang tua Petani/Nelayan, penghasilan orang tua berkisar satu hingga tiga juta rupiah dan asal sekolah SMK.

- **Cluster-2**

Berisikan mahasiswa-mahasiswa dengan perolehan IPK kategori sedang, jalur masuk mandiri, alamat asal dari Kabupaten Gorontalo, jenis kelamin perempuan, pendidikan orang tua kategori menengah, pekerjaan orang tua TNI/POLRI, penghasilan orang tua berkisar satu hingga tiga juta rupiah dan asal sekolah SMK.

- **Cluster-3**

Berisikan mahasiswa-mahasiswa dengan perolehan IPK kategori rendah, jalur masuk mandiri, alamat asal dari Kota Gorontalo, jenis kelamin perempuan, pendidikan orang tua kategori tinggi, pekerjaan orang tua PNS, penghasilan orang tua berkisar tiga hingga lima juta rupiah dan asal sekolah SMK

Berdasarkan hasil *finalcluster* data mahasiswa lulusan dengan jumlah  $k = 3$ , maka diperoleh karakteristik dari masing-masing *cluster* adalah seperti yang terlihat pada Tabel 6.

*Tabel 6. Final Cluster Centers Mahasiswa Lulusan*

VARIABEL	Cluster		
	1	2	3
Jenis_Kelamin	2	2	1
Tempat_Lahir	2	1	7
Jenjang	1	1	1
Status	1	1	1
Predikat	2	2	1
IPK	2	2	1
Lama_Studi	1	2	2

Dari Tabel 6 dapat didefinisikan bahwa :

- **Cluster-1**, berisikan mahasiswa-mahasiswa dengan jenis kelamin perempuan, tempat lahir Kabupaten Gorontalo, Jenjang D3, Status Non-Kependidikan, IPK kategori sedang dan lama studi tepat waktu.
- **Cluster-2**, berisikan mahasiswa-mahasiswa dengan jenis kelamin perempuan, tempat lahir Kota Gorontalo, Jenjang D3, Status Non-Kependidikan, IPK kategori sedang dan lama studi melampaui batas waktu.



- **Cluster-3**, berisikan mahasiswa-mahasiswa dengan jenis kelamin laki-laki, tempat lahir luar daerah Provinsi Gorontalo, Jenjang D3, Status Non-Kependidikan, IPK kategori rendah dan lama studi melampaui batas waktu

**b. Metode Non Hirarkis**

Metode Non-Hirarkis dimulai dengan menentukan terlebih dahulu jumlah cluster yang diinginkan (dua, tiga atau lainnya). Setelah jumlah cluster ditentukan, maka proses cluster dilakukan dengan tanpa mengikuti proses hirarki. Metode ini biasa disebut “*K-MeansCluster*”. Tabel 7 merupakan tabel yang berisikan anggota cluster dari data Mahasiswa Aktif Fakultas Teknik Universitas Negeri Gorontalo.

*Tabel 7. Anggota Cluster Data Mahasiswa Aktif*

Variabel	4 Clusters	3 Clusters	2 Clusters
Jenis Kelamin	1	1	1
Jurusan	2	2	1
Jalur	1	1	1
Alamat Asal	3	3	2
Smean(Pendidikan_Ortu)	1	1	1
Smean(Pekerjaan_Ortu)	4	2	1
Smean(Penghasilan_Ortu)	1	1	1
Smean(Asal_Sekolah)	1	1	1

Dari Tabel 7 di atas dapat dijabarkan bahwa :

- Apabila diinginkan dibentuk 4 cluster, maka :
  - ✓ Variabel-variabel yang digunakan dalam menentukan anggota dari Cluster 1 adalah Jenis\_Kelamin, Jalur\_Masuk, Pendidikan\_Ortu, Penghasilan\_Ortu dan Asal\_sekolah.
  - ✓ Variabel yang digunakan dalam menentukan anggota dari Cluster 2 adalah Jurusan.
  - ✓ Variabel yang digunakan dalam menentukan anggota dari Cluster 3 adalah Alamat\_Asal.
  - ✓ Variabel yang digunakan dalam menentukan anggota dari Cluster 4 adalah Pekerjaan\_Ortu.
- Apabila diinginkan dibentuk 3 cluster, maka :
  - ✓ Variabel-variabel yang digunakan dalam menentukan anggota dari Cluster 1 adalah Jenis\_Kelamin, Jalur\_Masuk, Pendidikan\_Ortu, Penghasilan\_Ortu dan Asal\_sekolah.
  - ✓ Variabel yang digunakan dalam menentukan anggota dari Cluster 2 adalah Pekerjaan\_Ortu dan Jurusan.
  - ✓ Variabel yang digunakan dalam menentukan anggota dari Cluster 3 adalah Alamat\_Asal.

- Apabila diinginkan dibentuk 2 cluster, maka :
  - ✓ Variabel-variabel yang digunakan dalam menentukan anggota dari Cluster 1 adalah Jenis\_Kelamin, Jurusan, Jalur\_Masuk, Pendidikan\_Ortu, Penghasilan\_Ortu, Pekerjaan\_Ortu dan Asal\_sekolah.
  - ✓ Variabel yang digunakan dalam menentukan anggota dari Cluster 2 adalah Alamat\_Asal.

Tabel 8 merupakan tabel yang berisikan anggota cluster dari data lulusan Mahasiswa Fakultas Teknik Universitas Negeri Gorontalo.

*Tabel 8. Anggota Cluster Data Mahasiswa Lulusan*

Variabel	4 Clusters	3 Clusters	2 Clusters
Jenis Kelamin	1	1	1
Tempat Lahir	2	2	2
Jurusan	3	3	1
Jenjang	4	1	1
Status	4	1	1
Predikat	1	1	1
IPK	1	1	1
Lama Studi	1	1	1

Dari Tabel 8 di atas dapat dijabarkan bahwa :

- Apabila diinginkan dibentuk 4 cluster, maka :
  - ✓ Variabel-variabel yang digunakan dalam menentukan anggota dari Cluster 1 adalah Jenis\_Kelamin, Predikat, IPK dan Lama Studi.
  - ✓ Variabel yang digunakan dalam menentukan anggota dari Cluster 2 adalah Tempat\_Lahir.
  - ✓ Variabel yang digunakan dalam menentukan anggota dari Cluster 3 adalah Jurusan.
  - ✓ Variabel yang digunakan dalam menentukan anggota dari Cluster 4 adalah Jenjang dan Status.
- Apabila diinginkan dibentuk 3 cluster, maka :
  - ✓ Variabel-variabel yang digunakan dalam menentukan anggota dari Cluster 1 adalah Jenis\_Kelamin, Jenjang, Status, Predikat, IPK dan Lama\_Studi.

- ✓ Variabel yang digunakan dalam menentukan anggota dari Cluster 2 adalah Tempat\_Lahir.
- ✓ Variabel yang digunakan dalam menentukan anggota dari Cluster 3 adalah Jurusan.
- Apabila diinginkan dibentuk 2 cluster, maka :
  - ✓ Variabel-variabel yang digunakan dalam menentukan anggota dari Cluster 1 adalah Jenis\_Kelamin, Jurusan, Jenjang, Status, Predikat, IPK dan Lama\_Studi.
  - ✓ Variabel yang digunakan dalam menentukan anggota dari Cluster 2 adalah Tempat\_Lahir.

## **SIMPULAN**

Hasil klasifikasi data mahasiswa aktif dan lulusan yang diperoleh dengan menggunakan 3 metode, yaitu *Decision Tree*, *Artificial Neural Network* dan *K-Nearest Neighbour*, dapat digunakan untuk mengetahui faktor-faktor atau karakteristik penting dari data akademik dan non akademik mahasiswa yang dapat membangkitkan pola-pola untuk memprediksi kemajuan studi mahasiswa berupa kategori IPK, sehingga dapat dijadikan sebagai sebuah sistem peringatan dini (*early warning*) bagi mahasiswa yang masuk ke dalam klasifikasi mahasiswa yang berpotensi untuk lulus dengan IPK kategori rendah yang bisa menurunkan mutu lulusan sebuah Perguruan Tinggi.

## **DAFTAR PUSTAKA**

- Al-Radaideh, Q.A., Al-Shawakfa, E.M dan Al-Najjar, M.I. 2006. *Mining Student Data Using Decision Tree. International Arab Conference on Informational Technology (ACIT)*.
- Fahim, A.M. Saake, G. Salem, A.M. Torkey, F.A & Ramadan, M.A. 2008. *K-Means for Spherical Cluter with Large Variance in Size*. Proceedings of International Conference on Mathematical, Computational and Statistical Sience and Engineering.
- Gunawan. 2009. Klasterisasi Data Kategorikal Menggunakan Algoritma K-Modes. Jurnal Muria Sains. Universitas Muria Kudus.
- Han & Kamber. 2006. *Data Mining Concepts and Techniques* Edisi ke-2. San Fransisco : Morgan Kaufmann publisher.
- Jananto A. 2010. Penggunaan Algoritma SLIQ untuk Pengklasifikasian Kinerja Akademik Mahasiswa. Jurnal Teknologi Informasi DINAMIK Vol XV, No.1 : 66-72
- Karlita, T. 2011. Klasterisasi Data Kategorikal dengan Menggunakan Algoritma *Modes Linkage*. Proseding Industrial Electronics Seminar. Surabaya.
- Larose. 2005. *Discovering Knowledge in Data*. Canada : Wiley-Interscience.
- Martiana, E. Mutbada'i, N.R. Purnomo, E. 2009. Penggunaan Metode Pengklasteran Untuk Menentukan Bidang Tugas Akhir Mahasiswa Teknik Informatika PENS Berdasarkan Nilai. Proseding Industrial Electronics Seminar. Surabaya.

- Pramudyo, A. S. 2008. *Case Base Reasoning* untuk Klasifikasi Mahasiswa Baru berdasarkan prediksi Indeks Prestasi Semester I (studi kasus Program Studi Teknik Informatika Universitas Bina Darma Palembang). Tesis. Universitas Gajah Mada. Yogyakarta.
- Sunjana. 2010. Aplikasi Mining Data Mahasiswa dengan Metode Klasifikasi Decision Tree. Proseding Seminar Nasional Aplikasi Teknologi Informasi. Yogyakarta.
- Susanto & Suryadi. 2010. Pengantar Data Mining Menggali Pengetahuan dari Bongkahan Data.CV Andi Offset.Yogyakarta