

Penilaian Hasil Kegiatan Belajar Mahasiswa Menggunakan Metode *Cluster Non-Hierarki*

Haryoso Wicaksono

Program Studi Sistem Informasi
STMIK –IM, Jl. Jakarta 79 Bandung
email : yosnex@yahoo.com

ABSTRACT

Pengambilan keputusan dalam evaluasi pendidikan, salah satunya adalah penilaian atas kegiatan belajar mahasiswa dalam satu semester pada satu mata kuliah tertentu. Metode yang bisa diterapkan adalah Metode Data Mining, yaitu Clustering Non-Hierarki. Dipakai sebagai metode alternatif atas metode penilaian dengan pembobotan yang selama ini lazim digunakan. Dengan segala kelebihan dan kekurangannya, metode Cluster ini bisa dipakai sebagai metode alternatif penilaian atas kegiatan belajar mahasiswa dalam satu semester pada satu mata kuliah tertentu. Metode Cluster yang digunakan adalah Metode Non-Hierarki. Data yang diolah sebanyak 56 record dengan tujuan mengelompokkan isi dari variabel-variabel Hadir, Nilai Hadir, Nilai Tugas, Nilai UTS & Nilai UAS. Jumlah Cluster yang direncanakan adalah 5, mengacu kepada nilai-nilai huruf yang dipakai untuk penilaian. Yaitu, nilai huruf A, B, C, D & E. Hasil dari metode Cluster Non-Hierarki tadi dibandingkan dengan penentuan penilaian dengan metode pembobotan yang lazim dipakai selama ini. Hasil perbandingan didapatkan 73,21% didapatkan nilai huruf yang sama, antara metode Cluster Non-Hierarki dengan metode pembobotan yang lazim dipakai selama ini. Diperlukan lebih banyak record untuk meningkatkan kualitas penilaian menggunakan metode Cluster Non-Hierarki ini. Diperlukan kebijakan akademik untuk menerapkan metode penilaian menggunakan metode Cluster Non-Hierarki ini.

Kata Kunci : *Data Mining, Cluster, Non-Hierarki, Metode Penilaian*

1. Introduction

Pemanfaatan teknologi informasi telah memasyarakat secara luas. Aspek perkembangan teknologi elektronika & perangkat keras telah memudahkan masyarakat dalam memilih perangkat keras komputasi yang di inginkan sesuai kebutuhannya. Perkembangan perangkat lunak juga semakin meningkatkan kegunaan komputer sebagai tools wajib di masa sekarang. Salah satu organisasi yang memanfaatkan keberadaan teknologi informasi (didalamnya ada perangkat keras & perangkat lunak) adalah dunia pendidikan. Dunia pendidikan, khususnya pendidikan tinggi sangat perlu untuk mengoptimalkan segala kelebihan teknologi informasi. Dalam tulisan ini akan dicoba menerapkan salah satu aspek pengambilan keputusan dalam evaluasi pendidikan, yaitu bagaimana memberikan penilaian atas kegiatan belajar mahasiswa dalam satu semester pada satu mata kuliah tertentu. Penulis selama ini menggunakan teknologi informasi untuk memberikan penilaian atas kegiatan belajar mahasiswa.

Seperti yang kita maklumi bersama, bahwa penilaian atas kegiatan belajar mahasiswa dalam satu semester pada satu mata kuliah tertentu didasarkan atas ketentuan atau pedoman tertentu yang telah diketahui oleh Dosen & Mahasiswa yang bersangkutan. Metode penilaian yang akan Penulis sampaikan adalah terapan dari salah satu cabang Data Mining, yaitu Klasifikasi. Klasifikasi yang dimaksud adalah pengelompokkan atas komponen penilaian yaitu Tugas, Quiz, UTS, UAS & Kehadiran. Asumsi yang digunakan adalah meniadakan bobot tiap komponen, karena penilaian yg didasarkan atas klasifikasi ini masih belum diterapkan karena metode klasifikasi ini memerlukan perangkat lunak khusus, yaitu SPSS.

Metode klasifikasi untuk penilaian ini masih debatable, terkait prosesnya yang relatif agak rumit. Istilah klasifikasi dikenal juga dengan istilah cluster atau klaster.

Ada beberapa metode penilaian yang kita kenal. Dari Buku Panduan Sekolah Tinggi Manajemen Informatika & Komputer Indonesia Mandiri, 2006 halaman 9, sbb :

Tabel 1. Penilaian Nilai Akhir

Nilai Akhir (NA)	Nilai absolut	Predikat
A (=4)	85–100	Sangat baik
B (=3)	75–84	Baik
C (=2)	60–74	Cukup
D (=1)	55–59	Hampir cukup
E (=0)	≤54	Kurang/Gagal
T/K		Tidak lengkap/Kosong

Secara mendasar nilai akhir terdiri atas nilai huruf A, B, C, D, atau E. Di sini berarti ada 5 (lima) klasifikasi. Bila ada perbedaan jumlah klasifikasi maka pada pengolahan data nilai, pada proses klasifikasi tinggal menyesuaikan dengan berapa banyak jumlah tingkatan nilai (klasifikasi) yang ditentukan. Jadi klasifikasi nilai didasarkan atas ketentuan yang ada. Misal, bila penilaian di bawah ini maka ada 10 klasifikasi.

Tabel 2. Penilaian Nilai Akhir [1]

NA	NM	AM	Sebutan Mutu
≥ 85 <100	A	4,00	Sangat Cemerlang
≥ 80 < 84	A-	3,50	Hampir Cemerlang
≥ 75 <79	B+	3,25	Sangat Baik
≥ 70 < 74	B	3,00	Baik
≥ 65 <69	B-	2,75	Hampir Baik
≥ 60 < 64	C+	2,25	Lebih dari cukup
≥ 55 <59	C	2,00	Cukup
≥ 50 < 54	C-	1,75	Hampir Cukup
≥ 40 <49	D	1,00	Kurang
≥ 0 <39	E	0,00	Gagal

1.1. Konsep Data Mining

Ada beberapa ahli mengemukakan pengertian tentang Data Mining, sebagai berikut :

“Data mining is the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition technologies as well as statistical and mathematical techniques” [2]

“Data mining is the analysis of (often large) observational data sets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner” [3].

“The extraction of interesting (non-trivial, implicit, previously unknown and potentially useful) patterns or knowledge from huge amount of data” [4].

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan metoda statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [5].

Menurut Gartner Group, data mining adalah suatu proses menemukan hubungan yang berarti, pola dan kecenderungan dengan memeriksa dan sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti metoda statistik dan matematika [6]

1.2. Pengelompokan Data Mining

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas/fungsi yang dapat dilakukan , yaitu [6] :

1. Deskripsi : Terkadang peneliti dan analis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data.
2. Estimasi : Estimasi hampir sama dengan klasifikasi, kecuali variabel target. Estimasi lebih ke arah numerik dari pada ke arah kategori.
3. Prediksi : Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada dimasa datang.
4. Klasifikasi : Dalam klasifikasi terdapat target variabel kategori, sebagai contoh penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.
5. Pengklasteran : Pengklasteran merupakan pengelompokan record, pengamatan atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Klaster atau cluster adalah kumpulan record yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidak miripan dengan record-record dalam cluster.
6. Asosiasi : Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang biasa.

1.3. Data Preprocessing

Sebelum data diolah dengan *data mining*, data perlu melalui tahap *preprocessing*. Tahap ini berhubungan dengan pemilihan dan pemindahan data yang tidak berguna (*data cleaning*), penggabungan sumber-sumber data (*data integration*), transformasi data dalam bentuk yang dapat mempermudah proses (*data transformation*), menampilkan data dalam jumlah yang lebih mudah dibaca (*data reduction*). Semuanya berasal dari data mentah (data transaksi) dan hasilnya akan menjadi data yang nantinya siap untuk diolah dengan data mining [7].

1.4. Analisis Cluster

Pada algoritma clustering, data akan dikelompokkan menjadi cluster-cluster berdasarkan kemiripan satu data dengan yang lain. Prinsip dari clustering adalah memaksimalkan kesamaan antar anggota satu cluster dan meminimumkan kesamaan antar anggota cluster yang berbeda [3].

Cluster atau kluster dapat diartikan kelompok, dengan demikian pada dasarnya analisis cluster akan menghasikan sejumlah kluster (kelompok). Analisis ini diawali dengan pemahaman bahwa sejumlah data tertentu sebenarnya mempunyai kemiripan diantara anggotanya, karena itu dimungkinkan untuk

mengelompokkan anggota-anggota yang mirip atau mempunyai karakteristik yang serupa tersebut dalam satu atau lebih dari satu kluster [8].

Seperti diketahui, analisis cluster akan membagi sejumlah data pada satu atau beberapa cluster tertentu. Pertanyaan yang timbul adalah “apa yang menjadi batas bahwa sejumlah data dapat disebut sebagai satu cluster?”. Secara logika, sebuah cluster yang baik adalah cluster yang mempunyai :

1. Homogenitas (kesamaan) yang tinggi antara anggota dalam satu cluster (with-in cluster). Sebagai contoh, cluster konsumen rumah yang peduli lingkungan tentu terdiri atas orang-orang yang mengutamakan kebersihan dan kenyamanan lingkungan rumahnya. Mereka yang mengutamakan harga rumah murah tentu tidak dapat digabungkan menjadi anggota cluster tersebut.
2. Heterogenitas (perbedaan) yang tinggi antara cluster yang satu dengan cluster lainnya (between cluster). Dalam contoh diatas, anggota dari cluster konsumen rumah yang peduli lingkungan tentu mempunyai pendapat yang jelas berbeda dengan anggota-anggota cluster konsumen rumah yang mementingkan harga murah.

Sehingga, cluster yang baik adalah cluster yang mempunyai anggota-anggota yang semirip mungkin satu dengan yang lain, namun sangat tidak mirip dengan anggota-anggota cluster yang lain. Disini, istilah “mirip” diartikan sebagai tingkat kesamaan karakteristik antara dua data [8].

1.5. Proses Dasar dari Analisis Cluster

Analisis clustering pada dasarnya adalah mencari dan mengelompokkan data yang mirip satu dengan yang lain, maka kriteria mirip (similarity) adalah dasar dari metode clustering. Proses pengolahan data sehingga sekumpulan data mentah dapat dikelompokkan menjadi satu atau beberapa cluster adalah sbb. :

1. Menetapkan ukuran jarak antar-data
Mengukur kesamaan antar-obyek (similarity), secara prinsip dasar cluster yang mengelompokkan obyek yang mempunyai kemiripan, maka proses pertama adalah mengukur seberapa jauh ada kesamaan antar-obyek. Ada 3 metode yang digunakan :
 - a. Mengukur korelasi antara sepasang obyek pada beberapa variabel. Jika beberapa data memang akan tergabung menjadi satu cluster, tentulah diantara data tersebut ada hubungan yang erat, atau disebut berkorelasi satu dengan yang lain. Metode ini mendasarkan beberapa korelasi antara data untuk mengetahui kemiripan data satu dengan yang lain.
 - b. Mengukur jarak (distance) antara dua obyek. Salah satunya adalah metode Euclidean Distance. Cara ini akan memasukkan sebuah data ke dalam cluster tertentu dengan mengukur jarak data tersebut dengan pusat cluster. Jika data ada dalam jarak yang masih ada dalam batas tertentu, data tersebut dapat dimasukkan pada cluster tersebut.
 - c. Mengukur asosiasi antar-obyek. Cara ini akan mengasosiasikan sebuah data dengan cluster tertentu. Dalam praktek, cara ini tidak sepopuler kedua cara sebelumnya. Dalam praktek, penggunaan metode Euclidean Distance adalah yang paling populer.
2. Melakukan proses standarisasi data jika diperlukan
Setelah cara mengukur jarak ditetapkan, yang perlu diperhatikan adalah apakah satuan data mempunyai perbedaan yang besar. Sebagai contoh, jika variabel Penghasilan mempunyai satuan juta (000.000), sedangkan usia seseorang hanya mempunyai satuan puluhan (00), maka perbedaan yang mencolok ini akan membuat perhitungan jarak (distance) menjadi tidak valid. Jika ada perbedaan yang cukup signifikan, maka data harus di standarisasi dengan mengubah

data yang ada ke Z-Score. Proses standarisasi menjadikan dua data dengan perbedaan satuan yang lebar akan otomatis menjadi menyempit.

3. Melakukan proses clustering

Setelah data yang dianggap mempunyai satuan yang sangat berbeda diseragamkan & metode cluster ditentukan, misal. dipilih Euclidian, langkah selanjutnya adalah membuat cluster. Proses inti dari clustering adalah pengelompokan data, yang bisa dilakukan dengan 2 (dua) metode :

- a. Hierarchical Method. Metode ini memulai pengelompokan dengan dua atau lebih obyek yang mempunyai kesamaan paling dekat. Kemudian proses diteruskan ke obyek lain yang mempunyai kedekatan kedua. Demikian seterusnya sehingga cluster akan membentuk semacam pohon dimana ada hierarki (tingkatan) yang jelas antar-obyek. Dari yang paling mirip sampai paling tidak mirip. Semua obyek pada akhirnya hanya akan membentuk sebuah cluster. Dendogram biasanya digunakan untuk membantu memperjelas proses hierarki tersebut. Hasil keseluruhan dari algoritma hierarchical clustering secara grafik dapat digambarkan sebagai tree, yang disebut dengan dendogram. Tree ini secara grafik menggambarkan proses penggabungan dari cluster-cluster yang ada, sehingga menghasilkan cluster dengan level yang lebih tinggi. [7]
- b. Non-Hierarchical Method. Berbeda dengan metode hierarki, metode ini justru dimulai dengan menentukan terlebih dahulu jumlah cluster yang diinginkan (bisa 2, 3 atau lebih). Setelah jumlah cluster ditentukan, baru proses cluster dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan K-Means Cluster.

4. Melakukan penamaan cluster-cluster yang terbentuk

Setelah sejumlah cluster terbentuk, entah dengan metode hierarki atau non-hierarki, langkah selanjutnya adalah melakukan interpretasi terhadap cluster yang telah terbentuk, yaitu memberi nama spesifik untuk menggambarkan isi cluster tersebut. Misal., kelompok konsumen yang memperhatikan lingkungan sekitar sebelum membeli sebuah rumah bisa dinamai cluster Lingkungan.

5. Melakukan validasi dan profiling kluster

Kluster yang terbentuk kemudian diuji apakah hasil tersebut valid. Kemudian dilakukan proses profiling untuk menjelaskan karakteristik setiap cluster berdasar profil tertentu (seperti usia konsumen pembeli rumah, tingkat penghasilan, dsb.) [8].

1.6. Asumsi pada Analisis Cluster

Asumsi untuk analisis cluster terkait dengan sampel yang akan diolah.

1. Sampel yang diambil benar-benar bisa mewakili populasi yang ada. Memang tidak ada ketentuan jumlah sampel yang representatif, namun tetaplah diperlukan sejumlah sampel yang cukup besar agar proses clustering bisa dilakukan dengan benar
2. Multikolinieritas, yakni kemungkinan adanya korelasi antar-obyek. Sebaiknya tidak ada atau seandainya ada, besar multikolinieritas tersebut tidaklah tinggi (misal. diatas 0.5). Jika sampai terjadi multikolinieritas dianjurkan untuk menghilangkan salah satu variabel dari dua variabel yang mempunyai korelasi cukup besar [8].

2. Research Method

2.1 Persiapan Data

Data yang akan dijadikan sampel pengolahan data adalah data komponen nilai dari mata kuliah Aplikasi Komputer, dengan jumlah record 56. Data yang berisi 0 (nol) adalah data yang pada kasus ini dikarenakan tidak ada nilainya, atau ybs tidak mengikuti proses tsb.

Tabel 3. Data Mentah untuk Analisis Cluster (Data diolah sendiri)

No	NIM	NAMA MAHASISWA	Hadir	Nilai Hadir	UAS	UTS	Tugas /Quiz #1	Tugas /Quiz #2
1	371461006	HERI SUPRIYADI	10	83	88	88	96	90
2	371463036	DIMAS PROGO PRAWIRA	11	91	0	90	0	0
3	371561001	AGUSTINUS DWIYATNA	12	99	97	95	88	90
4	371561002	GILANG ANDREINAPUTRI	12	99	96	91	91	90
5	371561003	OKI SRI HARTINI	12	99	96	86	91	90
6	371561004	ASEP DYANA	12	99	92	79	90	90
7	371561005	ROBBY PERMATA JIWA	10	83	82	42	96	90
54	371562042	PAXIA KHUSNUL KHOTIMAH	11	91	94	91	85	90
55	371562043	ERLINA SIMBOLON	9	74	88	79	80	90
56	371562044	ISNA KARISNA	12	99	85	75	95	90

2.2. Data View

Data disiapkan dan dibuka dengan software SPSS, bila awalnya data berbentuk data MS Excel, maka ada sedikit perubahan, khususnya nama variabel ketika dibuka pada SPSS, salah satunya adalah nama variabel yang tidak memperbolehkan penggunaan spasi antar kata pada penamaan variabel.

Tabel 4. Input Data di SPSS untuk Analisis Cluster (Data diolah sendiri)

	NIM	NAMAMAHASISWA	Hadir	NilaiHadir	UAS	UTS	TugasQuiz#1	TugasQuiz#2
1	371461006	HERI SUPRIYADI	10.0	82.5	88.0	88.0	96.0	90.0
2	371463036	DIMAS PROGO PRAWIRA	11.0	90.8	0	90.0	0	0
3	371561001	AGUSTINUS DWIYATNA	12.0	99.0	97.0	95.0	88.0	90.0
4	371561002	GILANG ANDREINAPUTRI	12.0	99.0	96.0	91.0	91.0	90.0
5	371561003	OKI SRI HARTINI	12.0	99.0	96.0	86.0	91.0	90.0
6	371561004	ASEP DYANA	12.0	99.0	92.0	79.0	90.0	90.0
7	371561005	ROBBY PERMATA JIWA	10.0	82.5	82.0	42.0	96.0	90.0
8	371561006	TUTI ALAWIYAH	12.0	99.0	94.0	77.0	91.0	90.0
9	371561008	LAIL ANGGRAINI	12.0	99.0	94.0	85.0	91.0	90.0
10	371561009	NADYA DAMAYANTI	8.0	66.0	85.0	90.0	96.0	90.0
11	371561011	ANDRI PURNAMA	12.0	99.0	85.0	71.0	90.0	90.0

2.3. Variable View

Struktur data bisa dilihat dengan mengaktifkan tab variable view, khususnya bila ingin melakukan pengeditan atas nama variable, type data atau lebar data. Berikut tampilan dari Variable View berupa list & tabel dibawah ini.

Variable: NIM	Type: Number	Width: 12	Dec: 0
Variable: NAMAMAHASISWA	Type: String	Width: 29	
Variable: Hadir	Type: Number	Width: 12	Dec: 1
Variable: NilaiHadir	Type: Number	Width: 12	Dec: 1
Variable: UAS	Type: Number	Width: 12	Dec: 1
Variable: UTS	Type: Number	Width: 12	Dec: 1
Variable: TugasQuiz#1	Type: Number	Width: 12	Dec: 1
Variable: TugasQuiz#2	Type: Number	Width: 12	Dec: 1

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	NIM	Numeric	12	0		None	None	7	Right	Scale	Input
2	NAMAMAHASISWA	String	60	0	NAMA MAHASISWA	None	None	19	Left	Nominal	Input
3	Hadir	Numeric	12	1		None	None	5	Right	Nominal	Input
4	NilaiHadir	Numeric	12	1	Nilai Hadir	None	None	8	Right	Scale	Input
5	UAS	Numeric	12	1		None	None	5	Right	Nominal	Input
6	UTS	Numeric	12	1		None	None	6	Right	Scale	Input
7	TugasQuiz#1	Numeric	12	1	Tugas/Quiz #1	None	None	9	Right	Nominal	Input
8	TugasQuiz#2	Numeric	12	1	Tugas/Quiz #2	None	None	10	Right	Nominal	Input

Tabel 5. Tampilan Input Data di SPSS untuk Analisis Cluster (Data diolah sendiri)

3. Result and Analysis

3.1. Data Preprocessing

Proses Cluster dimulai dengan menstandarisasi variable, dengan asumsi supaya menghindari kesenjangan antar data. Hasil proses standarisasi adalah :

Tabel 6. Tampilan Hasil proses standarisasi di SPSS (Data diolah sendiri)

	NIM	NAMAMAHASISWA	ZHadir	ZNilaiHadir	ZUAS	ZUTS	ZTugasQuiz#1	ZTugasQuiz#2
1	371461006	HERI SUPRIYADI	-.61870	-.61902	.19985	.46200	.61035	.23578
2	371463036	DIMAS PROGO PR...	-.12374	-.12106	-3.38144	.56814	-3.95964	-4.16548
3	371561001	AGUSTINUS DWIY...	.37122	.37090	.56612	.83349	.22952	.23578
4	371561002	GILANG ANDREINA...	.37122	.37090	.52542	.62121	.37233	.23578
5	371561003	OKI SRI HARTINI	.37122	.37090	.52542	.35586	.37233	.23578
6	371561004	ASEP DYANA	.37122	.37090	.36263	-.01564	.32473	.23578
7	371561005	ROBBY PERMATA ...	-.61870	-.61902	-.04433	-1.97925	.61035	.23578
8	371561006	TUTI ALAWIYAH	.37122	.37090	.44403	-.12178	.37233	.23578
9	371561008	LAIL ANGGRAINI	.37122	.37090	.44403	.30279	.37233	.23578
10	371561009	NADYA DAMAYANTI	-1.60863	-1.60893	.07776	.56814	.61035	.23578
11	371561011	ANDRI PURNAMA	.37122	.37090	.07776	-.44020	.32473	.23578
12	371561012	FEBI MUHAMAD F...	.37122	.37090	-3.38144	-.12178	.08671	.23578
13	371561013	NURYANA	.37122	.37090	-.81756	-.22792	.08671	.23578
14	371561014	RISCA ELIVIANA	.37122	.37090	.36263	-.20138	.08671	.23578
15	371561016	REGINA EFRILLIA	.37122	.37090	.11846	-.17485	.08671	.23578

3.2. Hasil Analisis Cluster

Hasil Cluster dinyatakan dalam angka 1 s/d 5 pada kolom QCL_1. Variabel QCL_1 berisi nomor cluster untuk setiap kasus, dan variabel QCL_2 yang berisi jarak antara data nilai tertentu dengan pusat cluster (cluster center).

Tabel 7. Tampilan Hasil Analisis Cluster di SPSS (Data diolah sendiri)

	NIM	NAMAMAHASISWA	QCL_1	QCL_2
1	371461006	HERI SUPRIYADI	5	1.33498
2	371463036	DIMAS PROGO PR...	2	.00000
3	371561001	AGUSTINUS DWIY...	5	.64921
4	371561002	GILANG ANDREINA...	5	.46246
5	371561003	OKI SRI HARTINI	5	.29970
6	371561004	ASEP DYANA	5	.31450
7	371561005	ROBBY PERMATA ...	4	1.38764
8	371561006	TUTI ALAWIYAH	5	.43845
9	371561008	LAIL ANGGRAINI	5	.22964
10	371561009	NADYA DAMAYANTI	4	1.90601
11	371561011	ANDRI PURNAMA	5	.74659
12	371561012	FEBI MUHAMAD F...	1	.00000
13	371561013	NURYANA	5	1.24142
14	371561014	RISCA ELIVIANA	5	.50816
15	371561016	REGINA EFRILLIA	5	.51617

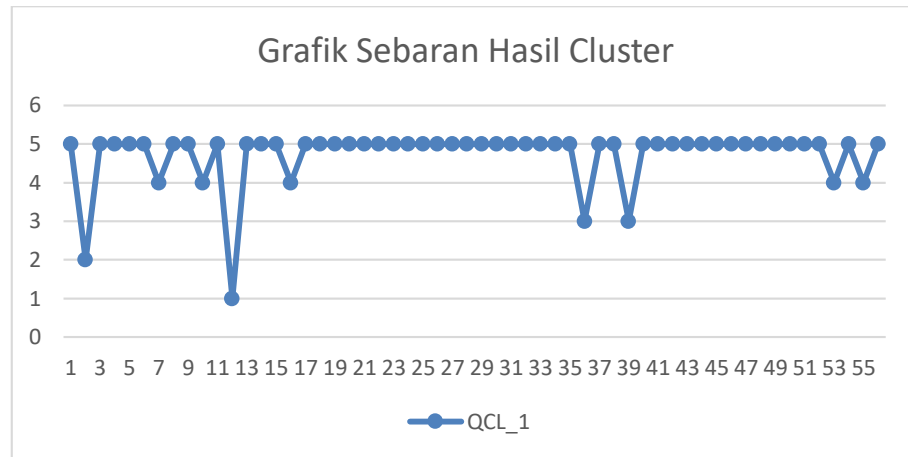
3.3. Melakukan Penamaan Cluster-cluster yang Terbentuk

Pada QCL_1 terdapat angka 1 s/d 5, hal tersebut dikarenakan pada saat awal telah ditentukan kriteria 5 cluster. Jumlah 5 (lima) menyakan kategori nilai akhir huruf, yaitu angka 5 identik dengan nilai huruf A, angka 4 identik dengan nilai huruf B, angka 3 identik dengan nilai huruf C, angka 2 identik dengan nilai huruf D, dan angka 1 identik dengan nilai huruf E. Jumlah masing-masing cluster adalah :

Tabel 8. Jumlah masing-masing Hasil Analisis Cluster di SPSS (Data diolah sendiri)

Cluster	1	1.000
	2	1.000
	3	2.000
	4	5.000
	5	47.000
Valid		56.000
Missing		.000

Hasil Cluster berupa nilai angka 1 s/d 5 digambarkan dalam grafik untuk menunjukkan sebaran datanya. Pada Gambar tersebut nampak nilai 5 atau identik dengan nilai A mendominasi sebaran data, atau banyaknya nilai 5 adalah 47 mahasiswa.



Gambar 1. Sebaran Hasil Cluster (Data diolah sendiri)

3.4. Melakukan Validasi dan Profiling Cluster

Validasi dilakukan dengan menghitung dengan metode yang lain. Dalam hal ini adalah menghitung dengan pembobotan tertentu. Bila Tugas1, Tugas2 & Hadir berbobot 10%, UTS berbobot 30% & UAS berbobot 40% maka nilai akhir bisa dihitung dibawah ini, sekaligus dibandingkan dengan metode Cluster. Ada hasil akhir nilai yang SAMA sebanyak 73,21%, dan yang berbeda 26,79%.

Tabel 9. Perbandingan Metode Penilaian (Data diolah sendiri)

No	NIM	NAMA MAHASISWA	Nilai Huruf Pembandingan	Hasil Cluster	Nilai Huruf dari Cluster	Uji kesamaan antara Pembandingan & Cluster
1	371461006	HERI SUPRIYADI	A	5	A	SAMA
2	371463036	DIMAS PROGO PRA	E	2	D	BERBEDA
3	371561001	AGUSTINUS DWIYA	A	5	A	SAMA
4	371561002	GILANG ANDREINA	A	5	A	SAMA
5	371561003	OKI SRI HARTINI	A	5	A	SAMA
6	371561004	ASEP DYANA	A	5	A	SAMA
7	371561005	ROBBY PERMATA	C	4	B	BERBEDA
8	371561006	TUTI ALAWIYAH	A	5	A	SAMA
9	371561008	LAIL ANGGRAINI	A	5	A	SAMA
10	371561009	NADYA DAMAYAN	A	4	B	BERBEDA
11	371561011	ANDRI PURNAMA	B	5	A	BERBEDA

Tabel 10. Perbandingan Rekapitulasi Metode Penilaian (Data diolah sendiri)

Metode Cluster : Metode yg digunakan		Metode Pembeding : Sesuai dg Tabel 1.	
Nilai Huruf	Frekuensi	Nilai Huruf	Frekuensi
A	47	A	40
B	5	B	10
C	2	C	2
D	1	D	1
E	1	E	3
Jumlah	56	Jumlah	56

Perbandingan keduanya :

Nilai Huruf	Frekuensi	Prosentase
SAMA	41	73.21%
BERBEDA	15	26.79%
Jumlah	56	100.00%

4. Conclusion

Berdasarkan hasil analisis yang Penulis lakukan, dapat disimpulkan beberapa hal, antara lain :

1. Metode Cluster untuk menentukan penilaian hasil kegiatan belajar mahasiswa masih harus dikembangkan dengan melibatkan lebih banyak variabel yang diolah, disamping juga jumlah record yang dilibatkan dan pembobotan tiap variabel penilaian.
2. Dengan penyempurnaan baik dari sisi analisis, jumlah variabel & jumlah record, diharapkan metode penilaian hasil kegiatan belajar mahasiswa menggunakan analisis Cluster ini bisa dipertimbangkan sebagai metode penilaian yang terstandar. Melengkapi metode pembobotan yang selama ini digunakan.
3. Kelemahan analisis Cluster adalah digunakannya software SPSS, padahal SPSS bukan termasuk aplikasi yang lazim & hanya dipakai oleh kalangan tertentu saja. Berbeda dengan software MS Excel yang sudah sangat banyak digunakan masyarakat.

Untuk memperoleh pemahaman yang lebih baik tentang metoda Cluster ini beberapa hal berikut bisa disarankan, yaitu :

1. Pemahaman akan metoda Cluster, khususnya statistik multivariat sangat dianjurkan. Di samping konsep database dan data mining sebagai disiplin ilmu yang utama.
2. Analisis Cluster yang berbasis pada database akademik bisa lebih dikembangkan lebih lanjut dengan menggunakan variabel-variabel yang lain, yang diharapkan mampu memberikan sumbangan nilai pengetahuan untuk pengambilan keputusan di bidang akademik.
3. Perancangan database yang baik akan menjadikan kualitas database yang baik, sehingga kualitas database yang akan di analisis dengan data mining akan semakin baik. Sehingga nilai pengetahuan yang didapatkan akan berkualitas baik pula.

References

- [1] <http://hi.fisip.unand.ac.id/current-students/172-pengenalan-tentang-sistem-ujian-penilaian>, akses 1-Maret-2017
- [2] Athauda, Rukshan., Tissera, Menik., and Fernando, Chandrika., Data Mining Applications: Promise and Challenges. Publikasi pada Data Mining and Knowledge Discovery in Real Life Applications, Julio Ponce and Adem Karahoca, 2009.
- [3] Han, Jiawei. & Kamber, Micheline., Data Mining: Concepts and Techniques Second Edition, University of Illinois at Urbana-Champaign, 2006.

- [4] Han Jiawei and Micheline Kamber, Data Mining : Concepts and Techniques, Morgan Kaufmann, 2001
- [5] Turban, E., dkk , Decision Support System And Intelegent System, Andi , Yogyakarta, 2005
- [6] Larose, Daniel T., Discovering Knowledge ind Data : An Introduction to data mining, John Willey & Sons.inc, 2005
- [7] Budhi, Gregorius S., Handojo, Andreas, Wirawan, Christine Oktavina, Algoritma Generalized Sequential Pattern Untuk Menggali Data Sekuensial Sirkulasi Buku Pada Perpustakaan UK Petra, SNATI 2009, 2009
- [8] Santoso, Singgih., Statistik Multivariat, Konsep dan Aplikasi dengan SPSS, Elex Media Komputindo, 2010.