

Data Clustering Menggunakan Metodologi CRISP-DM Untuk Pengenalan Pola Proporsi Pelaksanaan Tridharma

Irwan Budiman^a, Toni Prahasto^b, dan Yuli Christyono^b

^aFakultas MIPA, Universitas Lambung Mangkurat
Banjarbaru, Kalimantan Selatan

^bMagister Sistem Informasi, Program Pascasarjana
Universitas Diponegoro, Semarang, Indonesia

Abstract

Quality of human resources faculty can be reflected from the implementation of productivity and quality Tridharma (education, research, community service and supporting field activities). Lecturer Workload and Evaluation of Higher Education Tridharma (BKD and the EPT-PT) aims to ensure the implementation of the faculty task runs according to the criteria set out in legislation. *Data clustering* Tridharma implementation is needed to get some knowledge of the pattern of Tridharma implementation at college. Clustering as a data mining technique should be scalable, reliable and meet an agreed standard. CRISP-DM is the standardization of data mining is used in this study. The results of data clustering found the pattern of proportion of Tridharma into 3 clusters representing patterns: professionals, managers and teachers.

Keywords : *Clustering, CRISP-DM, K-Means, Tridharma*

1. Pendahuluan

Data Mining sebagai salah satu cabang ilmu yang relatif baru mempunyai potensi pengembangan yang sangat besar dan diprediksi akan menjadi salah satu yang paling revolusioner pada dekade ini (Larose, 2005). *Data Mining* sendiri merupakan sebuah proses ekstraksi informasi untuk menemukan pola (*pattern recognition*) yang penting pada tumpukan data dalam *database* sehingga menjadi pengetahuan (*knowledge discovery*). Fungsi-fungsi dalam *data mining* antara lain: fungsi deskripsi, fungsi estimasi, fungsi Prediksi, fungsi Klasifikasi, fungsi *Clustering* dan fungsi asosiasi (Larose, 2006).

Clustering digunakan untuk pengelompokan data secara alamiah berdasarkan kemiripan pada objek data dan sebaliknya meminimalkan kemiripan terhadap kluster lain. *Clustering* adalah pengelompokan menggunakan teknik *unsupervised learning* dimana tidak diperlukan pelatihan pada metode tersebut atau dengan kata lain, tidak ada fase *learning* serta tidak menggunakan pelabelan pada setiap kelompok.

Proses *data mining* harus terukur, dapat dipercaya dan memenuhi suatu standar yang telah disepakati. CRISP-DM adalah standarisasi *data mining* yang disusun oleh tiga penggagas *data mining market*. Yaitu Daimler Chrysler (Daimler-Benz), SPSS (ISL), NCR. Kemudian dikembangkan pada berbagai *workshops* antara 1997-1999.

Kwalitas sumber daya manusia dosen dapat tercermin dari produktivitas dan kualitas pelaksanaan tridharma (pendidikan, penelitian, pengabdian pada masyarakat dan kegiatan pendukung lainnya). Direktorat Jenderal Pendidikan Tinggi (DIKTI) telah memberlakukan program

Beban Kerja Dosen dan Evaluasi Pelaksanaan Tridharma Perguruan Tinggi (BKD dan EPT-PT) mulai 2010 untuk menjamin pelaksanaan tugas dosen berjalan sesuai dengan kriteria yang telah ditetapkan dalam peraturan perundang-undangan.

Setiap dosen dapat berbeda pola proporsi pelaksanaan tridharma antara satu dengan yang lainnya. *Data clustering* pelaksanaan tridharma diperlukan agar mendapatkan suatu pengetahuan (*knowledge discovery*) tentang pola (*pattern recognition*) pelaksanaan tridharma pada perguruan tinggi. Hal ini diperlukan untuk mengontrol keseimbangan kegiatan tridharma dosen dan kemudian oleh lembaga dapat membuat kebijakan yang tepat dan terarah sesuai dengan kondisi dan kebutuhan perguruan tinggi dalam mencapai visi dan misinya.

Tujuan dari penelitian ini adalah menggali pengetahuan (*discovering knowledge*) tentang pola (*pattern*) proporsi pelaksanaan tridharma oleh dosen sehingga diketahui tingkat produktivitas, komposisi serta keseimbangan dalam pelaksanaan tridharma.

2. Kerangka Teori

2.1. Model Crisp-DM

Proses *data mining* berdasarkan CRISP-DM terdiri dari 6 fase yaotu: *Business Understanding, Data Understanding, Data preparation, Modeling, Evaluation, Deployment* (hrearer, 2000).

Business Understanding adalah pemahaman tentang substansi dari kegiatan *data mining* yang akan dilakukan, kebutuhan dari perspektif bisnis. Kegiatannya antara lain: menentukan sasaran atau tujuan bisnis, memahami situasi bisnis, menentukan tujuan *data mining* dan membuat perencanaan strategi serta jadwal penelitian.

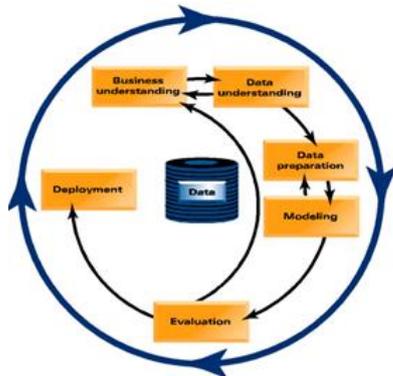
Data Understanding adalah fase mengumpulkan data awal, mempelajari data untuk bisa mengenal data yang akan dipakai, mengidentifikasi masalah yang berkaitan dengan kualitas data, mendeteksi subset yang menarik dari data untuk membuat hipotesa awal.

Data preparation sering disebut sebagai fase yang padat karya. Aktivitas yang dilakukan antara lain memilih *table* dan *field* yang akan ditransformasikan ke dalam *database* baru untuk bahan *data mining* (set data mentah).

Modeling adalah fase menentukan tehnik data mining yang digunakan, menentukan *tools data mining*, teknik *data mining*, algoritma *data mining*, menentukan parameter dengan nilai yang optimal.

Evaluation adalah fase interpretasi terhadap hasil data mining yang ditunjukkan dalam proses pemodelan pada fase sebelumnya. Evaluasi dilakukan secara mendalam dengan tujuan menyesuaikan model yang didapat agar sesuai dengan sasaran yang ingin dicapai dalam fase pertama.

Deployment atau penyebaran adalah fase penyusunan laporan atau presentasi dari pengetahuan yang didapat dari evaluasi pada proses *data mining* (Shearer, 2000).



Gambar 1. Model CRISP-DM

2.2. Clustering

Clustering adalah studi formal untuk partisi atau pengelompokan data dengan tidak menggunakan pelabelan kategori. *Clustering* bersifat *unsupervised learning* atau tidak mempunyai tahap pelatihan data, berbeda dengan klasifikasi. *Clustering* digunakan untuk pengelompokan data secara alamiah berdasarkan kemiripan pada objek data dan sebaliknya meminimalkan kemiripan terhadap kluster lain [4].

2.3. K-means

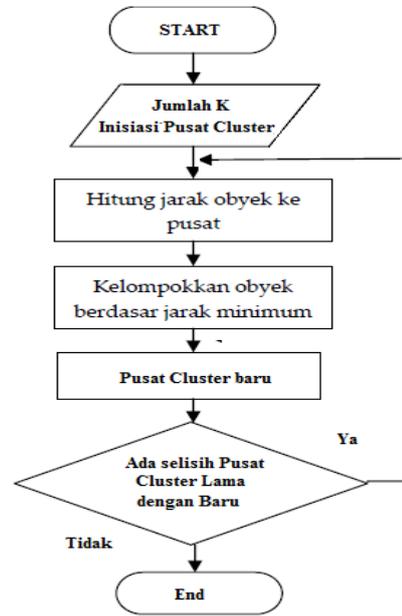
Konsep dasar dari algoritma K-means adalah pencarian pusat *cluster* (*centroid points*) secara iteratif. Pusat *cluster* ditetapkan berdasarkan jarak setiap data ke pusat *cluster*. Proses *clustering* dimulai dengan mengidentifikasi data yang akan dicluster, x_{ij} ($i=1, \dots, n; j=1, \dots, m$) dengan n adalah jumlah data yang akan dicluster dan m adalah jumlah variabel.

Algoritma K-means:

- Penetapan jumlah *cluster* (K), penetapan inisiasi pusat *cluster*
- Hitung jarak setiap data ke pusat *cluster*

- Kelompokkan data ke dalam *cluster* dengan jarak minimal
- Hitung pusat *cluster* baru berdasarkan rata-rata jarak terhadap pusat *cluster*
- Apakah ada selisih antara pusat *cluster* lama dengan pusat *cluster* baru? Jika ada, maka pusat *cluster* lama=pusat *cluster* baru.
- Ulangi langkah ii -iv hingga sudah tidak ada lagi selisih pada pusat *cluster*.

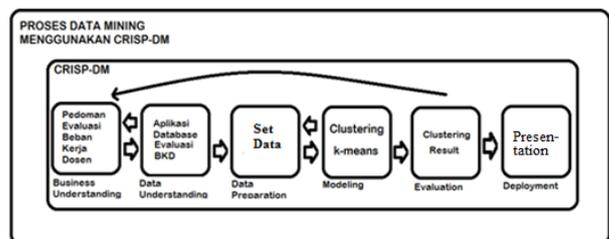
Algoritma:



Gambar 2. Flowchart algoritma K-means

3. Metodologi

CRISP-DM menguraikan langkah-langkah penambangan data secara *generic*. Meski menggunakan metodologi CRISP-DM namun pada pelaksanaannya penelitian ini menghasilkan satu model yang lebih spesifik (Gambar 3).



Gambar 3. Model pengenalan pola proporsi tridharma

Model pengenalan pola proporsi pelaksanaan tridharma dibagi menjadi 5 tahapan; tahap pemahaman bisnis (*business understanding*); tahap pemahaman data (*data understanding*); tahap persiapan data (*data preparation*); tahap pemodelan (*modelling*) dan tahap interpretasi hasil (*evaluation dan deployment*).

3.1. Tahap Pemahaman Bisnis

Pemahaman bisnis mengacu pada pedoman program Beban Kerja Dosen dan Evaluasi Pelaksanaan Tridharma yang diterbitkan Dikti pada tahun 2010. Pada tahap ini diperlukan pemahaman tentang latar belakang, tujuan, substansi dan semangat pada program tersebut

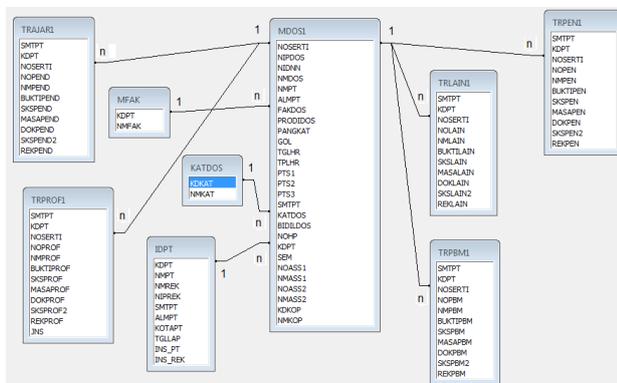
Pemahaman tujuan bisnis berdasarkan penjelasan pada pedoman program Beban Kerja Dosen antara lain untuk meningkatkan kinerja dosen dalam melaksanakan tridharma perguruan tinggi. Parameter suksesnya peningkatan kinerja dosen dalam pelaksanaan tridharma adalah dosen melaksanakan tridharma perguruan tinggi dengan beban kerja paling sedikit sepadan dengan 12 (dua belas) SKS dan paling banyak 16 (enam belas) SKS pada setiap semester dengan komposisi minimal 9 (sembilan) SKS untuk bidang pendidikan dan penelitian dan minimal 3 (tiga) SKS untuk bidang pengabdian masyarakat dan penunjang.

Tujuan data mining atau tujuan penelitian ini adalah menggali pengetahuan (*discovering knowledge*) tentang pola (*pattern*) proporsi pelaksanaan tridharma oleh dosen sehingga diketahui tingkat produktivitas, komposisi serta keseimbangan dalam pelaksanaan tridharma.

3.2. Tahap Pemahaman Data

Pemahaman data mengacu pada database BKD (Beban Kerja Dosen). Tahap memahami format data secara permukaan (format *form* dan *report*) dan secara lebih mendalam (bentuk fisik data).

Tahap mengumpulkan data kompilasi hasil laporan dari *database* Beban Kerja Dosen dan Evaluasi Pelaksanaan Tridharma Perguruan Tinggi. Tahap membiasakan diri dengan data-data yang sudah dikumpulkan dan berusaha menemukan wawasan awal mengenai informasi apa saja yang bisa didapatkan. Eksplorasi data pada aplikasi BKD dan menelusuran data hingga ke dalam *database*.



Gambar 4. Relasi antar tabel

3.3. Tahap Persiapan Data

Persiapan data merupakan tahap yang padat karya. Tahap membangun *database* baru sebagai set data akhir untuk pemodelan *data mining clustering*.

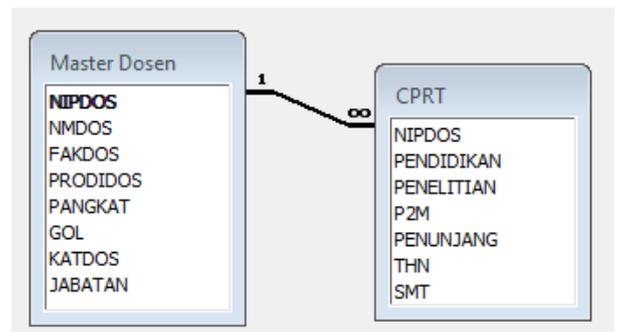
Tahap merancang format *set data* mentah pada *database* parsial yang digunakan sebagai gudang data yang digunakan untuk pemodelan *data mining*.

Tabel 1. Tabel rekapitulasi tridharma (CPRT)

No	Field Name	Type Data
1	NIPDos	Text
2	Pendidikan	Number
3	Penelitian	Number
4	P2M	Number
5	Penunjang	Number
6	THN	Text
7	SMT	Text

Tabel 2. Tabel master dosen

No	Nama Field	Type Data
1	NIPDos *	Text
2	NMDos	Text
3	FakDos	Text
4	ProdiDos	Text
5	Pangkat	Text
6	Gol	Text
7	KatDos	Text
8	Jabatan	Text



Gambar 5. Relasi antar tabel

Data yang dipilih pada setiap database Beban Kerja Dosen adalah data yang berhubungan dengan pelaksanaan tridharma diambil dari table Trajar1, Trpen1, Trpbm1, Trlain1 dan table Mdos1 sebagai data master dosen.

Setiap dosen dicari total SKS pelaksanaan setiap dharma-nya. Berikut ini adalah perintah memilih atribut dan menjumlahkan SKS setiap dharma dosen. Kemudian dibuat rekapitulasi jumlah dharma setiap dosen dalam table baru bernama CPRT.

```
SELECT MDOS1.NOSERTI,
Sum(TRAJAR1.SKSPEND2) AS SumOfSKSPEND2
FROM MDOS1 INNER JOIN TRAJAR1 ON
MDOS1.NOSERTI = TRAJAR1.NOSERTI GROUP
BY MDOS1.NOSERTI;
```

```
SELECT MDOS1.NOSERTI,
Sum(TRLAIN1.SKSLAIN2) AS SumOfSKSLAIN2
FROM MDOS1 INNER JOIN TRLAIN1 ON
MDOS1.NOSERTI = TRLAIN1.NOSERTI GROUP BY
MDOS1.NOSERTI;
```

```
SELECT MDOS1.NOSERTI,
Sum(TRPBM1.SKSPBM2) AS SumOfSKSPBM2
FROM MDOS1 INNER JOIN TRPBM1 ON
MDOS1.NOSERTI = TRPBM1.NOSERTI GROUP BY
MDOS1.NOSERTI;
```

```
SELECT MDOS1.NOSERTI, Sum(TRPEN1.SKSPEN2)
AS SumOfSKSPEN2
FROM MDOS1 INNER JOIN TRPEN1 ON
MDOS1.NOSERTI = TRPEN1.NOSERTI GROUP BY
MDOS1.NOSERTI;
```

```
SELECT MDOS1.NOSERTI, [SKS
Trajar1].SumOfSKSPEND2, [SKS
Trpen1].SumOfSKSPEN2, [SKS
Trpbm1].SumOfSKSPBM2, [SKS
Trlain1].SumOfSKSLAIN2, MDOS1.SMTPT,
MDOS1.SEM
FROM ((MDOS1 LEFT JOIN [SKS Trlain1] ON
MDOS1.NOSERTI = [SKS Trlain1].NOSERTI) LEFT
JOIN [SKS Trpbm1] ON MDOS1.NOSERTI = [SKS
Trpbm1].NOSERTI) LEFT JOIN [SKS Trpen1] ON
MDOS1.NOSERTI = [SKS Trpen1].NOSERTI) LEFT
JOIN [SKS Trajar1] ON MDOS1.NOSERTI = [SKS
Trajar1].NOSERTI;
```

Hasil rekapitulasi SKS tridharma setiap semester pada Query Rekap Tridharma (Gambar atas) diintegrasikan pada Table CPRT.

```
INSERT INTO CPRT ( Nipdos, Pendidikan, Penelitian,
P2m, Penunjang, Thn, Smt )
SELECT [Rekap Tridharma].NOSERTI, [Rekap
Tridharma].SumOfSKSPEND2, [Rekap
Tridharma].SumOfSKSPEN2, [Rekap
Tridharma].SumOfSKSPBM2, [Rekap
Tridharma].SumOfSKSLAIN2, [Rekap
Tridharma].SMTPT, [Rekap Tridharma].SEM FROM
[Rekap Tridharma];
```

Hasil verifikasi pada tahap *data understanding* enunjukkan adanya beberapa data berkualitas buruk. Data tersebut perlu dibersihkan dengan cara diperbaiki, dirubah atau dihapus. Data yang dibersihkan antara lain:

- NoSerti (No Sertifikasi Dosen) yang berfungsi sebagai primary key diganti dengan NIPDos (Nip Dosen). Alasannya adalah karena Nip Dosen dimiliki oleh

semua dosen sedangkan Nomor Sertifikasi tidak. Ini semata untuk keseragaman.

- Nip Dosen dilakukan penyeragaman format penulisan agar menjadi konsisten.
- Data master dosen pada setiap semester selalu diisi, hal ini menjadi memiliki dualisme data master dosen sehingga dipilih salah satu saja.
- Pembersihan terhadap data yang tidak layak, seperti data tridharma yang null atau kosong dan jumlah SKS yang tidak wajar.
- Penyesuaian nama field sesuai rancangan pada set data .

Data yang akan ditambah adalah SKS rata-rata dari pelaksanaan setiap tridharma dosen. Table CPRT yang berisikan data pelaksanaan tridharma dihitung SKS rata-ratanya dengan menggunakan query “Rata2 Tridharma”. Adapun proses pembuatannya sebagai berikut:

```
SELECT CPRT.NIPDOS, Avg(CPRT.PENDIDIKAN)
AS Pendidikan, Avg(CPRT.PENELITIAN) AS
Penelitian, Avg(CPRT.P2M) AS Pengabdian,
Avg(CPRT.PENUNJANG) AS Penunjang
FROM CPRT GROUP BY CPRT.NIPDOS;
```

3.4. Tahap Pemodelan

Tahap pemodelan atau tahap data mining. Metode *clustering* dan algoritma K-means sangat tepat digunakan untuk mencapai tujuan awal penelitian ini yaitu menggali pengetahuan tentang pola pelaksanaan tridharma sekaligus untuk mengelompokan dosen berdasarkan pola proporsi pelaksanaan tridharma.

Langkah ke-1, penetapan jumlah cluster (*k*) dan inisiasi pusat cluster. Jumlah *cluster* (*k*) yang digunakan pada penelitian ini adalah 3 dengan pertimbangan bahwa *cluster* tersebut akan mewakili kelompok yang Bagus, Sedang dan Kurang. Meskipun metode *clustering* tidak ada pelabelan pada hasilnya, namun dengan menentukan jumlah $k=3$, diharapkan hasil cluster nanti mewakili kategori kelompok tersebut (bagus, sedang, kurang). Nama cluster ditentukan yaitu K1, K2 dan K3.

Kemudian menetapkan inisiasi nilai pusat cluster c_{kj} ($k=1, \dots, K; j=1, \dots, m$). Meskipun inisiasi nilai pusat cluster dapat bersifat acak karena sifat pengelompokan pada *clustering* yang alamiah, namun agar hasil lebih optimal maka inisiasi nilai pusat cluster sebaiknya mengacu pada kecenderungan data yang diperoleh pada tahap *data understanding*.

Input K dan Pusat Cluster				
Kn:	Kp1:	Kp2:	Kp3:	Kp4:
K1	5,7	4	1,3	2,4
K2	6,3	1,2	0,8	6,3
K3	10,4	1,5	1	2
*				

Gambar 6. Inisiasi pusat cluster

Langkah ke-2, menghitung jarak antara setiap data dengan setiap pusat cluster.

```
(Query: tahap1)
SELECT [Rata-Rata Tridharma].NIPDOS, [Rata-Rata
Tridharma].Pendidikan, [Rata-Rata
Tridharma].Penelitian, [Rata-Rata
Tridharma].Pengabdian, [Rata-Rata
Tridharma].Penunjang, Kluster.Kn, Kluster.Kp1,
Kluster.Kp2, Kluster.Kp3, Kluster.Kp4,
Sqr((([Pendidikan]-[kp1])*([Pendidikan]-
[kp1])+([Penelitian]-[kp2])*([Penelitian]-
[kp2])+([Pengabdian]-[kp3])*([Pengabdian]-
[kp3])+([Penunjang]-[kp4])*([Penunjang]-[kp4])) AS
bkp1
FROM Kluster, [Rata-Rata Tridharma]
ORDER BY [Rata-Rata Tridharma].NIPDOS,
Kluster.Kn;
```

Langkah ke-3, menentukan jarak terpendek setiap data dengan pusat cluster. dan menempatkan setiap data dalam sebuah cluster dengan jarak data tersebut ke pusat cluster bernilai paling kecil (dekat) jika dibandingkan dengan jarak ke pusat cluster lainnya.

```
(Query: tahap2)
SELECT tahap1.NIPDOS, tahap1.Pendidikan,
tahap1.Penelitian, tahap1.Pengabdian, tahap1.Penunjang,
Min(tahap1.bkp1) AS MinOfbcp1
FROM tahap1
GROUP BY tahap1.NIPDOS, tahap1.Pendidikan,
tahap1.Penelitian, tahap1.Pengabdian, tahap1.Penunjang;
```

```
(Query: tahap3)
SELECT tahap2.NIPDOS, tahap2.Pendidikan,
tahap2.Penelitian, tahap2.Pengabdian, tahap2.Penunjang,
tahap2.MinOfbcp1, tahap1.Kn
FROM tahap2 INNER JOIN tahap1 ON
(tahap2.MinOfbcp1 = tahap1.bkp1) AND
(tahap2.NIPDOS = tahap1.NIPDOS)
ORDER BY tahap2.NIPDOS;
```

Langkah ke-4, menghitung pusat cluster baru dan menyimpan hasilnya pada Table Kluster2.

```
(Query: tahap4)
SELECT tahap3.Kn, Avg(tahap3.Pendidikan) AS Kp1,
Avg(tahap3.Penelitian) AS Kp2,
Avg(tahap3.Pengabdian) AS Kp3,
Avg(tahap3.Penunjang) AS Kp4
FROM tahap3 GROUP BY tahap3.Kn;
```

```
(Query: tahap5)
INSERT INTO Kluster2 ( Kn, Kp1, Kp2, Kp3, Kp4)
SELECT tahap5.Kn, tahap5.Kp1, tahap5.Kp2,
tahap5.Kp3, tahap5.Kp4 FROM tahap5;
```

Pada iterasi ke-1 didapatkan pusat cluster baru seperti pada Gambar 7.

Kn	Kp1	Kp2	Kp3	Kp4
K1	6,139	3,646	1,392	2,395
K2	6,558	1,288	0,793	6,033
K3	10,549	1,031	0,677	1,450

Gambar 7. Pusat cluster baru pada iterasi ke-1

Langkah ke-5, memeriksa pusat *clustering* apakah sudah stabil (tidak ada perubahan) yaitu dengan membandingkan pusat cluster lama (Table Kluster) dengan yang baru (Table Kluster2) seperti pada Gambar 8.

```
SELECT Kluster.Kn, Kluster.Kp1 AS A1, Kluster.Kp2
AS B1, Kluster.Kp3 AS C1, Kluster.Kp4 AS D1,
Kluster2.Kp1 AS A2, Kluster2.Kp2 AS B2,
Kluster2.Kp3 AS C2, Kluster2.Kp4 AS D2, [A1]-[A2]
AS A, [B1]-[B2] AS B, [C1]-[C2] AS C, [D1]-[D2] AS
D FROM Kluster INNER JOIN Kluster2 ON Kluster.Kn
= Kluster2.Kn;
```

Pusat *clustering* terdapat selisih antara pusat cluster lama dengan pusat cluster baru dan ini menunjukkan bahwa data *clustering* belum stabil (ada perubahan).

Kn	A1	B1	C1	D1	A2	B2	C2	D2	A	B	C	D
K1	5,700	4,000	1,300	2,400	6,569	3,356	1,422	2,249	-0,87	0,644	-0,12	0,151
K2	6,300	1,200	0,800	6,300	7,440	1,354	0,728	5,270	-1,14	-0,15	0,072	1,030
K3	10,400	1,500	1,000	2,000	11,464	0,640	0,462	0,840	-1,06	0,86	0,538	1,16

Gambar 8. Selisih pusat cluster pada iterasi ke-1

Langkah ke-6, menyalin isi Table Kluster2 (pusat cluster baru) ke table Kluster. Kemudian kembali ke langkah ke-2 hingga pusat cluster stabil atau tidak ada lagi selisih antara pusat cluster baru dengan pusat cluster lama.

```
INSERT INTO Kluster ( Kn, Kp1, Kp2, Kp3, Kp4 )
SELECT Kluster2.Kn, Kluster2.Kp1, Kluster2.Kp2,
Kluster2.Kp3, Kluster2.Kp4 FROM Kluster2;
```

Setelah melalui serangkaian proses iterasi, ditemukan pusat cluster stabil pada iterasi ke-12. Hasilnya seperti pada Gambar 9.

Kn	A1	B1	C1	D1	A2	B2	C2	D2	A	B	C	D
K1	6,569	3,356	1,422	2,249	6,569	3,356	1,422	2,249	0	0	0	0
K2	7,440	1,354	0,728	5,270	7,440	1,354	0,728	5,270	0	0	0	0
K3	11,464	0,640	0,462	0,840	11,464	0,640	0,462	0,840	0	0	0	0

Gambar 9. Selisih pusat cluster pada iterasi ke-12

Pada iterasi ke-12, pusat cluster sudah tidak ada selisih dan ini menunjukkan bahwa data *clustering* telah stabil, maka proses iterasi dihentikan.

3.4. Tahap Interpretasi Hasil

Evaluation merupakan tahap interpretasi terhadap hasil pemodelan *data mining* berdasarkan domain pengetahuan pada *business understanding*. Jika diaplikasikan dalam dunia nyata, tahap *evaluation* sebaiknya melibatkan pihak pengguna sistem atau yang kompeten misalnya pimpinan

fakultas, pimpinan perguruan tinggi, pimpinan kopertis bahkan pimpinan Dikti.

Deployment sebagai tahap pembuatan laporan atau presentasi tentang pengetahuan yang didapat dari hasil data mining berupa pengenalan pola proporsi pelaksanaan tridharma dosen dalam 3 kelompok (*cluster*).

4. Hasil dan Pembahasan

Hasil yang didapatkan dari aplikasi *data clustering* merupakan suatu pengetahuan tentang pola proporsi pelaksanaan tridharma dosen. Analisis menafsirkan pola tersebut berdasarkan domain pengetahuannya tentang sistem Beban Kerja Dosen Tahun 2010 pada tahap *business understanding*.

Cluster	Pendidikan	Penelitian	Pengabdian	Penunjang	Total
K1	6,976	3,389	1,421	2,238	14,024
K2	7,900	1,312	0,744	5,164	15,119
K3	11,354	0,514	0,417	0,699	12,983

Gambar 10. Pusat cluster final

Kelompok 1 (K1) memiliki jumlah SKS terendah dalam bidang pendidikan dan pengajaran dibanding 2 kelompok lain. Namun dosen pada kelompok ini menambah SKS bidang penelitian mereka hingga lebih dari 3 SKS. Dosen pada cluster ini dapat dikategorikan sebagai kelompok dosen ideal dalam proporsi pelaksanaan tridharma. Total SKS per semesternya sebesar 14,024 SKS dimana lebih dari batas minimum SKS persemester (12 SKS) dan tidak lebih dari batas maksimal (16 SKS). Komposisi antar bidang pun proporsional. Bidang pendidikan + penelitian lebih dari batas minimum 9 SKS yaitu 10,365 SKS dan bidang pengabdian + penunjang lebih dari 3 SKS yaitu 3,659 SKS.

Penyebaran kinerja tridharma Kelompok 1 cukup merata. Dengan komposisi 50% bidang pendidikan, 24% bidang penelitian, 16% bidang penunjang dan 10% bidang pengabdian.

Kelompok 2 (K2) dengan Pola tridharma pada kelompok 2 dengan total SKS per semesternya sebesar 15,119 SKS juga tergolong proporsi ideal, dimana telah melebihi batas minimum (12 SKS) dan tidak lebih dari batas maksimal (16 SKS). Perbedaannya dengan kelompok 1 adalah bahwa komposisi antar bidang kurang proporsional. Bidang pendidikan dan bidang penunjang jauh lebih tinggi daripada bidang penelitian dan pengabdian pada masyarakat.

Penyebaran kinerja tridharma Kelompok 2 didominasi oleh bidang pendidikan sebesar 52% dan bidang penunjang 34%. Sisanya bidang penelitian 9% dan pengabdian 5%.

Faktornya antara lain bahwa dosen selain sebagai pendidik dan peneliti juga terkadang memiliki tugas tambahan, khususnya dosen yang diamanatkan menduduki jabatan tertentu di instansi perguruan tinggi. Dosen dengan tugas tambahan antara lain pimpinan perguruan tinggi, pimpinan fakultas, pimpinan program studi hingga pimpinan pada unit kerja.

Kelompok 3, pada kinerja tridharma didominasi oleh bidang pendidikan yang mencapai 11,354 SKS atau 88% dari seluruh kegiatan. Meskipun secara normatif jumlah total SKS lebih dari batas minimum 12 SKS dan kurang dari 16 SKS, namun komposisi ini tidak ideal karena penyebaran tridharma tidak merata.

Kelompok 3 merupakan gambaran dosen yang memang lebih senang mengajar dibanding meneliti. Hal yang bisa dipahami, karena mengajar memang bisa menghasilkan uang secara lebih nyata. Sedangkan meneliti atau pengabdian pada masyarakat memerlukan pengajuan yang prosesnya lama dan belum tentu disetujui. Selain itu meneliti dan melakukan pengabdian pada masyarakat perlu ketekunan, dana yang tidak sedikit, pengakuan yang masih kurang, dan banyak kekurangan lain.

5. Kesimpulan

Pola proporsi pelaksanaan tridharma berdasarkan hasil *data clustering* dan analisa menurut pemahaman bisnis dapat disimpulkan:

Kelompok 1 dengan komposisi tridharma 6,976 SKS bidang pendidikan, 3,389 SKS bidang penelitian, 1,421 SKS bidang pengabdian dan 2,238 SKS bidang penunjang, dapat dilabelkan sebagai kelompok "Dosen Profesional".

Kelompok 2 dengan komposisi tridharma 7,900 SKS bidang pendidikan, 1,312 SKS bidang penelitian, 0,744 SKS bidang pengabdian dan 5,164 SKS bidang penunjang, dapat dilabelkan sebagai kelompok "Dosen Manajer".

Kelompok 3 dengan komposisi tridharma 11,354 SKS bidang pendidikan, 0,514 SKS bidang penelitian, 0,417 SKS bidang pengabdian dan 0,699 SKS bidang penunjang, dapat dilabelkan sebagai kelompok "Dosen Pengajar".

Namun perlu diperhatikan bahwa hasil ini bukan merupakan gambaran umum dari kinerja dosen, karena database BKD yang digunakan sejak 2010 lalu baru berjalan 2 semester saat data ini diambil. Sehingga dapat dikatakan bahwa hasil penelitian ini merupakan gambaran sesaat tepatnya gambaran 2 periode awal program Beban Kerja Dosen dilaksanakan.

Daftar Pustaka

- Larose, D. T., 2005. *Discovering Knowledge in Data: An Introduction to Data Mining*. Hoboken. Wiley-Interscience, John Wiley & Sons, Inc.
- Larose, D.T., 2006. *Data Mining: Methods and Models*. Hoboken. Wiley-Interscience, John Wiley & Sons, Inc.
- Shearer, C., 2000. *The CRISP-DM Model: The New Blueprint for DataMining*. Journal of Data Warehousing, 5(4): 13-22.