

**PENERAPAN ALGORITMA K-MEANS UNTUK MENDUKUNG
KEPUTUSAN DALAM PEMILIHAN TEMA TUGAS AKHIR
PADA PRODI SISTEM INFORMASI
UNIVERSITAS BUANA PERJUANGAN KARAWANG.**

Tukino
tukino@ubpkarawang.ac.id
Program Studi Sistem Informasi
Fakultas Teknologi dan Ilmu Komputer
Universitas Buana Perjuangan Karawang

Baenil Huda
Baenilhuda@ubpkarawang.ac.id
Program Studi Sistem Informasi Fakultas
Teknologi dan Ilmu Komputer Universitas
Buana Perjuangan Karawang

ABSTRAK

*Tema Tugas Akhir pada Prodi Sistem Informasi meliputi E-Business (E-BS), Sistem Pendukung Keputusan (SPK) dan Perancangan Sistem Informasi (PSI). Mahasiswa bebas untuk memilih satu tema dari 3 tema yang ditawarkan, tetapi pada kenyataannya mahasiswa mempunyai kesulitan dalam menentukan tema Tugas Akhir. Ketersediaan data Nilai mata kuliah belum dimanfaatkan secara optimal. Pada penelitian ini untuk mengotimalkan data nilai dengan menggunakan metode data mining dengan metode clustering dengan algoritma K-Means. Pada penelitian ini, data yang digunakan adalah data nilai mata kuliah semester satu sampai dengan semester 7. Atribut yang digunakan dalam daftar nilai mahasiswa adalah atribut nilai rata-rata mata kuliah yang menunjang tema E-BS, atribut nilai rata-rata mata kuliah yang menunjang tema SPK dan atribut nilai rata-rata mata kuliah yang menunjang tema PSI. Aplikasi yang digunakan untuk mengelompokkan daftar nilai dan menentukan nilai rata-rata mata kuliah penunjang tema tugas akhir dengan menggunakan Rapidminer dan MS-Excel, metode data mining menggunakan fase-fase dari CRISP. Hasil penelitian terdapat empat kelompok mahasiswa berdasarkan kemampuan akademisi sebagai berikut : (1) **Cluster 0** memiliki Nilai rata-rata MK penunjang tertinggi adalah EBS sebesar 2,698, (2) **Cluster 1** memiliki nilai rerata MK penunjang tertinggi adalah PSI sebesar 3,481, (3) **Cluster 2** memiliki Nilai rerata MK penunjang tertinggi adalah PSI sebesar 1,242, (4) **Cluster 3** memiliki nilai rerata MK penunjang tertinggi adalah PSI sebesar 2,629.*

Kata Kunci : Data Mining, Clustering, K-Means, Tema Tugas Akhir, Kemampuan Akademisi

ABSTRACT

The Final Project theme in the Information System Study Program consists of E-Business (E-BS), Decision Support System (DSS) and Information System Design (PSI). Students are free to choose one theme from the 3 themes offered, but in reality students have difficulty in determining the theme of the Final Project. Availability of data The value of courses has not been used optimally. In this study to optimize value data by using data mining methods with clustering methods with the K-Means algorithm. In this study, the data used are value data from semester one to semester 7, class of 2015 students. The attributes used in the list of student grades are the value attributes of the subjects that support the E-BS theme, the average value attribute courses that support the SPK theme and subject value average attributes that support the PSI theme. The application used to classify the list of values and determine the average value of the subjects supporting the final assignment theme using Rapidminer and MS-Excel, the data mining method uses the phases of CRISP. The results of the study are four groups of students based on academic abilities as follows: (1) Cluster 0 has the highest MK average value of EBS of 2.698, (2) Cluster 1 has the highest MK average value of support is PSI of 3.481, (3) Cluster 2 has the highest mean MK support is PSI of 1.242 (4) Cluster 3 has the highest MK average value of support is PSI of 2.629.

Key Word : Data Mining, Clustering, K-Means, Final Project Theme, Academic Ability

1. PENDAHULUAN

1.1 Latar Belakang

Tema Tugas Akhir program studi sistem informasi dikelompokkan menjadi tiga bidang yaitu tema E-Business, Tema Sistem Pendukung Keputusan dan Tema Perancangan Sistem Informasi[1]. Pengelompokan tema yang dilakukan oleh Prodi Sistem Informasi berdasarkan perkembangan kebutuhan aplikasi yang dibutuhkan oleh stakeholder atau kalangan dunia usaha dalam menjalankan bisnisnya. Kurikulum yang berlaku di prodi sistem informasi disusun mengarah pada bidang E-Business, Sistem Pendukung Keputusan, Perancangan Sistem Informasi. Pengelompokan tema Tugas Akhir untuk memudahkan mahasiswa dalam menentukan tema Tugas

Akhir dan mahasiswa dapat lebih fokus dalam satu bidang yang didasarkan kemampuannya akademiknya.

Keberhasilan mahasiswa setiap semester ditentukan oleh tingkat pencapaian Indeks Prestasi Semester (IPS) dan Indeks Prestasi Kumulatif (IPK) yang tertuang dalam Kartu Hasil Semester (KHS). Besarnya IPS dan IPK tergantung dari nilai per mata kuliah yang tercetak pada Kartu Hasil Semester (KHS). Kemampuan akademik mahasiswa terkait dengan pemilihan tema tugas akhir. IPS atau IPK yang diperoleh mahasiswa setiap semester selama tujuh semester dapat menggambarkan kemampuan mahasiswa dalam bidang tertentu. Permasalahan yang sering muncul bagi mahasiswa dalam menentukan tema tugas akhir adalah ketidakpercayaan diri mahasiswa atas kemampuan dirinya, tanpa ada perencanaan yang matang, tanpa adanya studi banding dengan tema tugas akhir yang telah ada, sehingga tema yang dipilih hanya berdasarkan saran dari temannya atau hanya sekedar melihat contoh yang sudah ada. Dampak dari pemilihan tema yang tidak terencana dan tanpa didasarkan pada kemampuan akademik mahasiswa, maka kualitas tugas akhir tidak sesuai dengan kompetensi prodi sistem informasi dan waktu yang dibutuhkan dalam menyelesaikan tugas akhir menjadi lama. Untuk mengatasi masalah tersebut diperlukan rekomendasi dalam menentukan tema tugas akhir yang didasarkan dengan kemampuan mahasiswa dalam bidang akademik.

Daftar nilai mahasiswa yang tersimpan dalam database yang ada server UBP Karawang yang setiap semester data nilai selalu bertambah sehingga didapatkan data nilai akademik yang sangat banyak jumlahnya. Penambahan data nilai setiap semester belum dimanfaatkan secara optimal, hanya sebatas untuk menampilkan Kartu Hasil Studi setiap semester dan transkrip akademik setiap mahasiswa yang sudah lulus. Penyimpanan data yang sangat besar, kalau dikelola dengan baik dapat menghasilkan informasi dan pengetahuan baru yang sangat berguna. Penggunaan data mining dalam mengolah data sangat tepat, karena data mining adalah proses dengan menggunakan teknik kecerdasan buatan, machine learning, matematika dan statistik dalam mengidentifikasi dan mengekstrak informasi dan Pengetahuan yang bermanfaat dalam berbagai database besar[7]. Diharapkan dengan penggunaan data mining dalam pengelolaan data nilai mahasiswa dapat memberikan pengetahuan dan informasi bagi Kaprodi, Dosen Pembimbing Akademik dan mahasiswa, yang selama ini data nilai mahasiswa kurang dimaksimalkan dalam pengelolaan dan penggunaannya. Menurut [6] menyatakan bahwa “clustering membagi dataset ke dalam kelompok yang berbeda, proses pengelompokan objek ke dalam kelompok tersebut bahwa objek dalam kelompok yang sama memiliki kemiripan dan objek dalam kelompok yang berbeda tidak memiliki kemiripan”.

Program Studi Sistem Informasi memiliki jumlah mahasiswa yang setiap tahunnya mengalami peningkatan kurang lebih 10%, sehingga setiap tahunnya ada proses penyusunan tugas akhir. Dalam penyusunan tugas akhir, program studi menentukan pilihan tema tugas akhir dengan memanfaatkan teknik data mining. Tujuan penelitian ini adalah bagaimana penerapan data mining dengan metode clustering menggunakan algoritma K-Means untuk mengelompokkan data nilai mahasiswa program studi sistem informasi berdasarkan kemampuan akademik, dan mahasiswa setiap semester yang dapat memberikan informasi dan pengetahuan dalam merekomendasikan mahasiswa dalam memilih tema tugas akhir. Sumber data nilai mahasiswa mulai dari semester 1 sampai dengan semester 7 dan atribut yang digunakan adalah NIM, IPK, rerata Nilai mata kuliah kelompok E-Business, rerata nilai mata kuliah kelompok Sistem pendukung keputusan, rerata nilai mata kuliah kelompok Perancangan Sistem Informasi.

Pemilihan algoritma *K-Means* dalam penelitian ini berdasarkan karena algoritma *K-Means* sederhana dalam dijalankan dan diimplementasikan, juga relatif cepat, mudah dalam beradaptasi, dan umum penggunaannya [2]. Pada dasarnya algoritma *K-Means* dapat diterapkan untuk permasalahan dalam meringkas objek yang berjumlah besar, memahami perilaku konsumen, dan identifikasi peluang pada produk baru dipasaran. Sehingga lebih memudahkan dalam mendeskripsikan sifat ataupun karakteristik pada masing-masing kelompok [4].

1.2. Sitasi Literatur

Penerapan data mining untuk mengolah data menggunakan metode clustering dengan algoritma *K-means*. Proses pengelompokan sekumpulan obyek kedalam kelas-kelas obyek yang sama disebut clustering /pengelompokan [12]. Pengklasifikasian data dengan Algoritma pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah *k* cluster yang sudah ditetapkan di awal adalah algoritma *K-Mean*[2]. Algoritma *K-mean* mencari titik pusat (centroid) berdasarkan data yang memiliki cluster masing-masing. Langkah ini dilakukan secara berulang jika posisi centroid baru berubah dan akan mencapai kondisi konvergen ketika pengalokasian kembali titik data centroid tidak berubah.

Implementasi metode clustering dengan metode *K-Means* telah banyak digunakan oleh peneliti lain karena cukup bagus untuk menentukan cluster sebuah data. Contohnya peneliti [13]

penerapan metode clustering dengan algoritma *K-Means* untuk rekomendasi pemilihan jalur peminatan sesuai kemampuan pada Program Studi Teknik Informatika. Dari hasil peneliti tersebut didapat 5 kelompok mahasiswa berdasarkan tingkat kemampuan akademiknya, dimana setiap kelompok yang terbentuk dan direkomendasikan untuk memilih jalur peminatan sesuai dengan nilai akademiknya yang tinggi. Peneliti[14] Hasil penelitian didapatkan bahwa penyebaran burung banyak berasal dari daerah Jambi dengan tujuan daerah di pulau Jawa. Selain Jambi sebagai daerah asal penyebaran burung Jambi juga sebagai transit penyebaran burung yang akan dikirim ke pulau Jawa. Sehingga apabila terdapat kasus berjangkitnya wabah penyakit Flu Burung (Avian Influenza) dapat segera dilakukan tindakan antisipasi salah satunya dengan melakukan metode penyuluhan pencegahan dan penanggulangan flu burung pada daerah asal pengiriman burung dan daerah tujuan pengiriman burung tersebut.

2. METODOLOGI PENELITIAN

2.1. Metode Pengumpulan Data

Metode dalam Pengumpulan data pada penelitian ini sebagai berikut :

- A. Survey
Dilakukan untuk mengambil data nilai mahasiswa program studi sistem informasi angkatan 2015 dari tabel mahasiswa yang tersedia di server dengan mengajukan ijin ke bagian Pusat Data dan Informasi UBP Karawang
- B. Studi Pustaka
Dilakukan dengan mencari referensi yang terkait dengan tema penelitian melalui buku, jurnal, internet dan sebagainya
- C. Wawancara
Dilakukan dengan kaprodi untuk menanyakan tentang kurikulum dan tema tugas akhir yang berlaku di Prodi Sistem Informasi

2.2. Metode Analisis

Metode Analisis yang digunakan dalam penelitian ini adalah Cross-Industry Standart Process for Data Mining (CRISP-DM). Metode ini mampu menyediakan standar proses data mining sebagai strategi pemecahan masalah secara umum dari bisnis atau unit-unit penelitian [7]. Berikut ini merupakan langkah-langkah yang dilakukan berdasarkan dari fase CRISP-DM :

1. Business understanding pada tahapan pertama penulis mencoba untuk memahami permasalahan yang ada dalam pemilihan tema tugas akhir pada program studi sistem informasi, sehingga dapat menentukan tujuan dan pola yang akan didapatkan dengan data mining.
2. Data understanding Pada tahap ini penulis melakukan pemahaman terhadap data yang dibutuhkan, untuk kemudian mengambil data yang relevan dan memiliki keterkaitan dengan tujuan penelitian. Adapun data yang digunakan yaitu data nilai mahasiswa.
3. Data preparation pada tahap ini penulis mengolah data yang didapat dengan melakukan pembersihan data atau data cleaning, melakukan integrasi data atau data integration, melakukan pemilihan data atau data selection, dan transformasi pada data atau data transformation. Tools yang digunakan dalam menunjang pengolahan datanya ialah Microsoft excel.
4. Modeling pada tahap ini penulis menentukan teknik data mining yang digunakan untuk mengolah data yang sudah disiapkan sebelumnya. Teknik yang dilakukan yaitu dengan clustering menggunakan algoritma *K-Means*.
5. Evaluation pada tahap evaluasi, akan diketahui apakah hasil daripada tahap pemodelan dapat menjawab tujuan yang telah ditetapkan pada fase pertama. Untuk itu akan dilakukan profilisasi pada setiap cluster yang telah terbentuk, untuk diketahui karakteristik pada kelompok tersebut. Disamping itu untuk diketahui kesesuaian dengan jalur pemilihan tema tugas akhir akan dilakukan analisis lebih lanjut untuk dihubungkan dengan atribut pemilihan tema tugas akhir, rerata nilai mata kuliah penunjang tema E-Bussiness, rerata nilai mata kuliah penunjang Sistem Pendukung Keputusan, rerata nilai mata kuliah penunjang Perancangan Sistem Informasi. Sehingga diharapkan mendapatkan informasi atau pola yang berguna sebagai rekomendasi pemilihan jalur pemilihan tema tugas akhir yang sesuai dengan kemampuan akademis mahasiswa.
6. Deployment Pada tahap ini penulis mencoba menyajikan hasil dari penerapan data mining metode clustering dengan algoritma *K-Means* yang telah dilakukan.

Dari 6 tahapan, masing-masing memiliki peran dan tugas untuk melakukan proses dan hasil dari proses menjadi sumber pada tahap berikutnya, seperti pada tahapan berikut ini :



Gambar 1 : Proses CRISP-DM

2.3. Algoritma K-Means

Metode ini digunakan dalam mengelompokkan data nilai mata kuliah dan IPK mahasiswa.

Langkah-langkah dari algoritma *K-Means* [2] adalah sebagai berikut :

1. Tentukan k sebagai jumlah *cluster* yang dibentuk
Dalam menentukan banyaknya *cluster* k dapat dilakukan dengan pertimbangan konseptual ataupun teoritis dan sesuai dengan tujuan dari pengelompokan.
2. Tentukan titik pusat (*centroid*) dari setiap *cluster*.
Untuk menentukan *centroid* awal setiap *cluster* dapat dilakukan secara random atau acak.
3. Alokasikan semua data ke centroid terdekat dengan matrik jarak yang sudah ditetapkan.
Rumus dalam menghitung jarak antara lain : jarak Euclidean, jarak Manhattan/Block City, dan jarak Minkowski. Dalam penelitian ini untuk menghitung jarak antara objek dengan centroid menggunakan Euclidean Distance dengan perhitungan seperti berikut :

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2; i = 1, 2, 3, \dots, n}$$

Dimana

x_i : objek x ke- i

y_i : objek y ke- i

n : banyaknya objek

4. Hitung kembali titik pusat (*centroid*) berdasarkan data mengikuti cluster masing-masing.
Untuk menghitung titik pusat (*centroid*) cluster ke- i berikutnya, menggunakan rumus sebagai berikut :

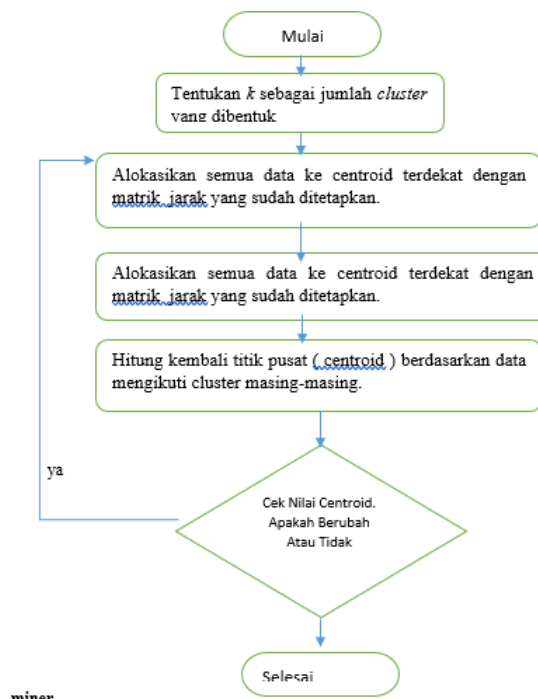
$$v = \sum_{i=1}^n x_i : i = 1, 2, 3, \dots, n$$

Dimana

v : centroid pada cluster

x_i : objek ke- i

n : banyaknya objek



5. Ulangi langkah 3 dan 4 jika posisi centroid baru berubah atau tidak sama.

Langkah-langkah dalam Algoritma *K*-Means memberikan kemudahan dalam melakukan pengelompokan dan data

2.4. Rapidminer

Pemilihan software rapidminer didasarkan pada fungsinya sebagai alat bantu dalam datamining dan knowledge Discovery dan merupakan software yang berlisensi gratis. Pada rapidminer terdapat 400 prosedur data mining, termasuk operator untuk masukan, output, data preprocessing dan visualisasi [5].

3. HASIL DAN PEMBAHASAN

3.1. Business Understanding

Pada tahap ini peneliti menemui kaprodi Sistem Informasi untuk melakukan diskusi tentang tema tugas akhir yang dipilih oleh mahasiswa. Tema Tugas Akhir telah disahkan dan disampaikan ke mahasiswa melalui Panduan Penyusunan Tugas Akhir. Ada 3 tema tugas akhir yang dipilih oleh mahasiswa yaitu : E-Business (EBIS), Sistem Pendukung Keputusan (SPK) dan Perancangan Sistem Informasi (PSI) . Peneliti juga mengajukan ijin ke kepala Pusdatin untuk menggunakan data nilai mahasiswa prodi Sistem Informasi angkatan 2015.

3.2. Data Understanding

Pada tahap ini penulis mempelajari data yang dibutuhkan dalam penelitian, data yang dibutuhkan dalam penelitian ini adalah data kurikulum, data mahasiswa, data nilai mahasiswa per semester (KHS) dan transkrip nilai. Dari data yang tersedia, dilakukan pemilihan data atribut yang terkait langsung dengan pemilihan tema tugas akhir adalah Nomor Induk Mahasiswa (NIM), Indeks Prestasi Kumulatif (IPK), nilai rerata mata kuliah penunjang tema E-Business (R_EBIS), nilai rerata mata kuliah penunjang tema sistem pendukung keputusan (R_SPK), nilai rerata mata kuliah penunjang perancangan sistem informasi (R_PSI).

3.3. Data Preparation

3.3.1. Pembersihan Data dan Integrasi Data

Pada tahap ini dilakukan dengan menyeleksi data mahasiswa, data nilai, data kurikulum yang tidak memiliki arti (salah input data), data tidak relevan dengan tema penelitian, dan data tidak konsisten dengan data yang lainnya. Setelah data dibersihkan, maka dibuat dataset baru yang terdiri dari NIK, IPK, R_EBS, R_SPK dan R_PSI.

NO	NIM	R IBS	R SPK	R PSI	IPK
1	15416257201023	3.86	3.64	3.69	3.67
2	15416257201048	3.71	3.64	3.77	3.65
3	15416257201014	3.57	3.71	3.62	3.62
4	15416257201040	3.64	3.71	3.85	3.61
5	15416257201001	3.64	3.57	3.85	3.55
6	15416257201007	3.50	3.50	3.46	3.55
7	15416257201052	3.43	3.50	3.54	3.53
8	15416257201019	3.57	3.64	3.77	3.48
9	15416257201035	3.29	3.36	3.31	3.44
10	15416257201032	3.50	3.21	3.23	3.43
11	15416257201060	2.86	3.43	3.54	3.42
12	15416257201022	3.29	3.36	3.31	3.40
13	15416257201015	3.57	2.71	3.31	3.38
14	15416257201026	3.29	3.21	3.46	3.38
15	15416257201059	3.43	2.64	3.31	3.35
16	15416257201033	3.43	2.71	3.08	3.30
17	15416257201010	3.14	3.14	3.23	3.28
18	15416257201068	2.00	1.57	1.54	3.27
19	15416257201037	3.07	3.14	3.23	3.24
20	15416257201008	2.93	2.00	1.92	3.21
21	15416257201027	0.71	1.21	1.23	3.21
22	15416257201020	2.50	1.79	2.62	3.21
23	15416257201024	3.29	2.50	3.08	3.19
24	15416257201009	2.07	3.14	2.92	3.18
25	15416257201018	3.00	3.07	3.15	3.14
26	15416257201050	2.86	2.00	2.77	3.06
27	15416257201057	2.86	2.93	3.00	3.01
28	15416257201006	3.00	2.43	2.69	2.97
29	15416257201017	3.07	2.43	2.69	2.91
30	15416257201036	2.64	2.07	2.46	2.85
31	15416257201025	2.14	2.00	2.15	2.81
32	15416257201030	2.93	2.50	2.69	2.80
33	15416257201005	2.93	2.50	2.54	2.78
34	15416257201045	2.57	2.93	3.00	2.75
35	15416257201039	2.64	2.43	3.08	2.70
36	15416257201058	1.93	1.64	2.38	2.55
37	15416257201028	2.36	2.00	2.31	2.51
38	15416257201056	1.79	2.50	1.69	2.48
39	15416257201046	0.71	1.14	1.00	2.31
40	15416257201031	2.21	2.43	2.31	2.29
41	15416257201054	1.36	1.64	1.69	2.12
42	15416257201041	0.86	1.00	1.38	2.08
43	15416257201055	1.93	2.50	2.92	2.07
44	15416257201012	0.36	0.57	0.62	1.56
45	15416257201049	0.93	1.29	1.31	1.46
46	15416257201003	0.79	1.29	1.38	1.41
47	15416257201038	1.14	1.29	1.77	1.34

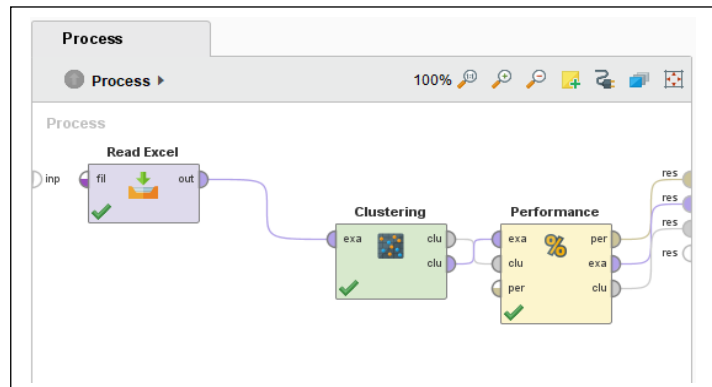
Gambar 2. Dataset Data Nilai Mahasiswa

3.3.2. Seleksi Data dan Transformasi Data

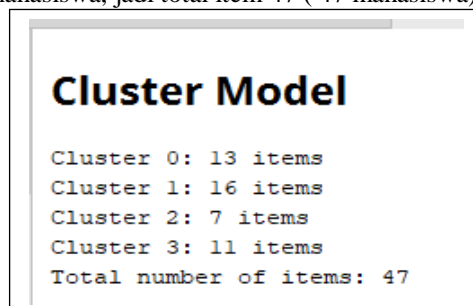
Data nilai mahasiswa semester 1 sampai dengan semester 7, dipilih dan dikelompokkan berdasarkan mata kuliah yang berkaitan dengan tema tugas akhir dan ditentukan nilai rerata per kelompok tema. Ada tiga tema tugas akhir yaitu E-Business, Sistem Pendukung Keputusan, Perancangan Sistem informasi. Hasil seleksi dan tranformasi data sebagai berikut :

3.4. Modeling

Dataset nilai mahasiswa yang telah diseleksi dan memilih atribut data yang diperlukan untuk diuji coba dengan menggunakan rapidminer dengan metode clustering dengan algoritma. Rancangan pada rapidminer dibutuhkan Read excel untuk menampung dataset nilai mahasiswa, Clustering untuk menentukan jumlah kelompok atau cluster dan performance untuk menentukan hasil Rancangan Rapidminer sebagai berikut :



Rancangan rapidminer dengan menggunakan dataset nilai mahasiswa, akan menghasilkan cluster model dengan 4 cluster, Cluster 0 terdapat 13 Item yang menunjukkan jumlah 13 mahasiswa, Cluster 1 terdapat 16 Item yang menunjukkan jumlah 16 mahasiswa, Cluster 2 terdapat 7 item yang menunjukkan jumlah 7 mahasiswa, Cluster 3 terdapat 11 item yang menunjukkan jumlah 11 mahasiswa, jadi total item 47 (47 mahasiswa)



Gambar 4 : Cluster Model

Dari cluster model terdiri dari empat cluster dan masing-masing cluster memiliki titik pusat.

Attribute	cluster_0	cluster_1	cluster_2	cluster_3
NO	7.500	21	33	43
NIM	15416257201...	15416257201...	15416257201...	15416257201...
R_EBS	2.796	2.830	2.253	2.397
R_SPK	2.694	2.610	2.240	2.484
R_PSI	2.797	2.805	2.510	2.624
IPK	3.001	3.126	2.641	2.867

3.5. Evaluation

Berdasarkan *dataset* yang telah di modelkan dalam rapidminer dengan menggunakan algoritma *K-Means* dan jumlah cluster yang ditentukan adalah 4 cluster. Pemodelan ini untuk menjawab tujuan yang telah ditetapkan pada fase pertama dengan melakukan profilisasi dan analisis. Hasil dataset cluster perlu dilakukan profilisasi sebagai langkah untuk mengidentifikasi dan menampilkan karakteristik yang diperoleh dari setiap cluster yang terbentuk :

1. Cluster 0
Merupakan kelompok mahasiswa yang memiliki IPK sedang , nilai rata-rata MK penunjang tema E-BS sedang (2,698), nilai rerata MK penunjang tema SPK sedang(2,374) dan nilai rerata MK penunjang tema PSI sedang(2,580) .
2. Clsuter 1
Merupakan kelompok mahasiswa yang memiliki IPK tinggi, nilai rerata MK penunjang tema E- Business tinggi, nilai rerata MK penunjang SPK tinggi dan nilai rata-rata MK penunjang PSI tinggi.
3. Clsuter 2
Merupakan kelompok mahasiswa yang memiliki IPK kurang, nilai rerata MK penunjang tema E- BS kurang, nilai rerata MK penunjang tema SPK kurang dan nilai rata-rata MK penunjang tema PSI kurang.
4. Clsuter 3
Merupakan kelompok mahasiswa yang memiliki IPK sedang, nilai rerata MK penunjang tema E-BS sedang, nilai rerata MK penunjang tema SPK sedang dan nilai rerata MK penunjang tema PSI sedang.

3.6. Deployment

Hasil model algoritma *K-Means* terbentuk 4 cluster dan masing-masing cluster mempunyai karakteristik yang berbeda. Hasil cluster perlu dianalisis untuk mendapatkan informasi mengenai kesesuaian antara tema yang dipilih dengan kemampuan akademik setiap mahasiswa dengan menghubungkan antara atribut nilai rerata mata kuliah penunjang tema tugas akhir E-Business, atribut nilai rerata mata kuliah penunjang tema tugas akhir Sistem Pendukung Keputusan, atribut nilai rerata mata kuliah penunjang tema tugas akhir Perancangan Sistem Informasi. Dengan menghubungkan antara atribut tersebut diperoleh informasi dan pola karakteristik mahasiswa berdasarkan tingkat kemampuan akademiknya yang sesuai dengan pemilihan tema tugas akhir maka dapat dibuatkan acuan rekomendasi pemilihan tema tugas akhir sebagai berikut:

o	Karakteristik Kelompok Mahasiswa	Titik Pusat Cluster				Rekomendasi
		K	EBS	SPK	PSI	
	Data Mahasiswa yang memiliki: a. IPK :Sedang b. Nilai rerata MK penunjang tema E-Business: sedang c. Nilai rerata MK penunjang tema SPK d. Nilai rerata MK penunjang tema PSI:Sedang	3,001	2,698	2,374	2,580	Mahasiswa di rekomendasikan untuk Memilih tema E- Business
	Data Mahasiswa yang memiliki : a. IPK :Tinggi b. Nilai rerata MK penunjang E-Business tinggi c. Nilai rerata MK penunjang SPK:Sedang d. Nilai rerata MK penunjang tema PSI:Tinggi	3.126	698	362	481	Mahasiswa di rekomendasikan untuk Memilih tema Perancangan Sistem Informasi
	Data Mahasiswa yang memiliki : a. IPK :Kurang b. Nilai rerata MK penunjang E-Business:kurang c. Nilai rerata MK penunjang SPK:kurang d. Nilai rerata MK penunjang PSI:Kurang	461	786	112	242	Mahasiswa di rekomendasikan untuk Memilih tema Perancangan Sistem Informasi
	Data Mahasiswa yang memiliki : IPK :Sedang Nilai rerata MK penunjang E-Business:Sedang Nilai rerata MK penunjang	867	383	383	629	Mahasiswa di rekomendasikan untuk Memilih tema Perancangan Sistem Informasi

o	Karakteristik Kelompok Mahasiswa	Titik Pusat Cluster				Rekomendasi
		K	EBS	SPK	PSI	
	SPK: sedang Nilai rata-rata MK penunjang PSI: Sedang					

Gambar 5 : Rekomendasi pemilihan tema tugas akhir

R_EBS : Nilai Rata-rata MK Penunjang Tema E-Business

R_SPK : Nilai rata-rata MK Penunjang Tema SPK

R_PSI : Nilai Rata-rata MK Penunjang Tema PSI

Berdasarkan proses data mining metode clustering menggunakan algoritma *K-Means* untuk menentukan tema tugas akhir Program Studi Sistem Informasi angkatan 2015 memberikan suatu rekomendasi bagi pemangku kepentingan di UBP Karawang yaitu :

- Kaprodi menggunakan hasil datamining pengolahan data nilai mahasiswa sebagai alternative dalam memberikan rekomendasi bagi mahasiswa dalam memilih tema tugas akhir.
- Mahasiswa menggunakan hasil datamining sebagai acuan dan pertimbangan dalam menentukan tema tugas akhir berdasarkan kemampuan akademik yang dimiliki oleh setiap mahasiswa.

4. HASIL DAN KESIMPULAN

4.1. Kesimpulan

Berdasarkan hasil penelitian penerapan data mining menggunakan metode clustering dengan algoritma *K-Means*, diperoleh 4 kelompok (cluster) mahasiswa berdasarkan kemampuan akademiknya dan dilakukan analisis hasil, maka dapat disimpulkan sebagai berikut

- Cluster 0
Nilai rerata MK penunjang tertinggi adalah EBS sebesar 2,698, maka mahasiswa yang terdapat dalam cluster ini direkomendasikan untuk memilih tema tugas akhir nya adalah E- Business.
- Cluster 1
Nilai rerata MK penunjang tertinggi adalah PSI sebar 3,481, maka mahasiswa yang terdapat dalam cluster ini direkomendasikan untuk memilih tema tugas akhir Perancangan Sistem Informasi
- Cluster 2
Nilai rerata MK penunjang tertinggi adalah PSI sebesar 1,242, maka mahasiswa yang terdapat dalam cluster ini direkomendasikan untuk memilih tema tugas akhir Perancangan Sistem Informasi
- Cluster 3
Nilai rerata MK penunjang tertinggi adalah PSI sebesar 2,629, maka mahasiswa yang terdapat dalam cluster ini direkomendasikan untuk memilih tema tugas akhir Perancangan Sistem Informasi

4.2. Saran

- Penelitian dapat dilanjutkan dengan mengetahui jumlah mahasiswa yanag mengambil tema tugas akhir sesuai dengan yang direkomendasikan oleh kaprodi.
- Penelitian dapat dilanjutkan dengan menentukan berapa lama mahasiswa mampu menyelesaikan tugas akhir.

DAFTAR PUSTAKA

- Universitas Buana Perjuangan Karawang , Buku Panduan Akademik 2017-2018.
- Eko Prasetyo, Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab. Yogyakarta: ANDI, 2014.
- Fajar Astuti Hermawati, Data Mining. Yogyakarta: CV ANDI OFFSET, 2013.
- Ediyanto, Muhlasah Novitasari Mara, and Neva Satyahadewi, "Pengklasifikasian Karakteristik Dengan Metode K-Means Cluster Analysis," Buletin Ilmiah Mat. Stat dan Terapannya (Bimaster), vol. 02, no. 2, pp. 133-136, 2013
- Lee Finn and Juan Santana, Data Mining : Meramalkan Bisnis Perusahaan. Jakarta: PT Elex Media Komputindo, 2010.
- Hasan Ziafat and Majid Shakeri, "Using Data Mining Techniques in Customer Segmentation," Journal of Engineering Research and Applications, vol. 4, no. 9, pp. 70-79, 2014. [Online]. www.ijera.com

- [7] Kusrini and Emha Taufiq, *Algoritma Data Mining*. Yogyakarta: CV ANDI OFFSET, 2009.
- [8] Totok Suprawoto, "Klasifikasi Data Mahasiswa Menggunakan Metode K-Means Untuk Menunjang Pemilihan Strategi Pemasaran," *Jurnal Informatika dan Komputer (JIKO)*, vol. 1, no. 1, pp. 12-18, Februari 2016.
- [9] Tamsir Hasudungan Sirait and Johan Oscar Ong, "Analisis Keberhasilan Mahasiswa Dengan Metode Clustering K-Means," *SNASTIA*, pp. 1-6, 2011.
- [10] Sofi Defiyanti and Mohamad Jajuli, "Integrasi Metode Klasifikasi Dan Clustering dalam Data Mining," *Konferensi Nasional Informatika (KNIF)*, pp. 39-44, 2015
- [11] Sumanto and Romi Satrio Wahono, "Penerapan Fuzzy C-Means Dalam Pemilihan Peminatan Tugas Akhir Mahasiswa," *Proceeding ISIT*, vol. 1, no. 1, pp. 1-8, 2011.
- [12] Han, Jiawei; & Kamber, Micheline. 2001. *Data Mining Concepts and Techniques Second Edition*. San Francisco: Morgan Kauffman.
- [13] Rizky Adrianto, and Amiq Fahmi, "Penerapan Metode Clustering Dengan Algoritma K-Means Untuk merekomendasikan Pemilihan jalur Peminatan Sesuai Kemampuan Pada Program Studi Teknik Informatika – S1 Universitas Dian Nuswantoro," *Journal Of Information System*, vol. 1, no. 2, (2016)
- [14] Sri Mulyati "Penerapan Data Mining Dengan Menggunakan Metode Clustering Untuk Pengelompokan Data Pengiriman Burung," *Prosiding Seminar Ilmiah Nasional Teknologi Komputer (SENATKOM 2015)* , vol. 1, Oktober 2015 ISSN : 2460-469