

Volume 6 Nomor 3 September 2021

# INFORMASI INTERAKTIF

JURNAL INFORMATIKA DAN TEKNOLOGI INFORMASI

**PROGRAM STUDI INFORMATIKA – FAKULTAS TEKNIK -UNIVERSITAS JANABADRA**

RANCANG APLIKASI SISTEM INFORMASI BANK SOAL DAN ANALISIS BUTIR SOAL DI FAKULTAS KEDOKTERAN XYZ

**Dinar Mustofa, Anggit Wirasto, Abdul Ghofur**

PERBANDINGAN ALGORITMA RED, SFQ, DAN AQM PADA JARINGAN *ENTERPRISE* DENGAN *VMWARE ESXI* DAN *ROUTER OS*

**Azriel Christian Nurcahyo, Listra Firgia, Ag. Rudatyo Himamunanto**

IMPLEMENTASI METODE *K-NEAREST NEIGHBOR* BERBASIS *EUCLIDEAN DISTANCE* UNTUK KLASIFIKASI PENERIMAAN VAKSIN COVID-19

**Yumarlin MZ, Rizqi Mirza Fadilla, Indra Pratama**

ANALISIS DATA UNTUK PENGELOMPOKAN MAHASISWA DENGAN METODE K-MEAN (STUDI KASUS : INSTITUT SHANTI BHUANA)

**Santi Thomas, Noviyanti P**

JARINGAN SYARAF TIRUAN MENGGUNAKAN METODE *BACKPROPAGATION* DALAM PREDIKSI PERSEDIAAN BAHAN BAKU (STUDI KASUS : PT. BINTANG TOBA LESTARI)

**Niko Surya Atmaja, Deri Lianda**



INFORMASI  
INTERAKTIF

Vol. 6

No. 3

Hal. 96 - 133

Yogyakarta  
September  
2021

ISSN  
2527-5240

**DEWAN EDITORIAL**

- Penerbit** : Program Studi Informatika Fakultas Teknik Universitas Janabadra
- Ketua Penyunting  
(Editor in Chief)** : Fatsyahrina Fitriastuti, S.Si., M.T. (Universitas Janabadra)
- Penyunting (Editor)** : 1. Jemmy Edwin B, S.Kom., M.Eng. (Universitas Janabadra)  
2. Ryan Ari Setyawan, S.Kom., M.Eng. (Universitas Janabadra)  
3. Yumarlin MZ, S.Kom., M.Pd., M.Kom. (Universitas Janabadra)
- Alamat Redaksi** : Program Studi Informatika Fakultas Teknik  
Universitas Janabadra  
Jl. Tentara Rakyat Mataram No. 55-57  
Yogyakarta 55231  
Telp./Fax : (0274) 543676  
E-mail: [informasi.interaktif@janabadra.ac.id](mailto:informasi.interaktif@janabadra.ac.id)  
Website : <http://e-journal.janabadra.ac.id/>
- Frekuensi Terbit** : 3 kali setahun

**JURNAL INFORMASI INTERAKTIF** merupakan media komunikasi hasil penelitian, studi kasus, dan ulasan ilmiah bagi ilmuwan dan praktisi dibidang Informatika. Diterbitkan oleh Program Studi Informatika Fakultas Teknik Universitas Janabadra di Yogyakarta, tiga kali setahun pada bulan Januari, Mei dan September.

## DAFTAR ISI

	<i>halaman</i>
Rancang Aplikasi Sistem Informasi Bank Soal Dan Analisis Butir Soal Di Fakultas Kedokteran XYZ <b>Dinar Mustofa, Anggit Wirasto, Abdul Ghofur</b>	96 - 100
Perbandingan Algoritma RED, SFQ, Dan AQM Pada Jaringan <i>Enterprise</i> Dengan <i>Vmware ESXI</i> Dan <i>Router OS</i> <b>Azriel Christian Nurcahyo, Listra Firgia, Ag. Rudatyo Himamunanto</b>	101-110
Implementasi Metode <i>K-Nearest Neighbor</i> Berbasis <i>Euclidean Distance</i> Untuk Klasifikasi Penerimaan Vaksin Covid-19 <b>Yumarlin MZ, Rizqi Mirza Fadilla, Indra Pratama</b>	111-117
Analisis Data Untuk Pengelompokan Mahasiswa Dengan Metode K-MEAN (Studi Kasus : Institut Shanti Bhuana) <b>Santi Thomas, Noviyanti P</b>	118-123
Jaringan Syaraf Tiruan Menggunakan Metode <i>Backpropagation</i> Dalam Prediksi Persediaan Bahan Baku (Studi Kasus : PT. Bintang Toba Lestari) <b>Niko Surya Atmaja, Deri Lianda</b>	124-133

## **PENGANTAR REDAKSI**

Puji syukur kami panjatkan kehadiran Allah Tuhan Yang Maha Kuasa atas terbitnya JURNAL INFORMASI INTERAKTIF Volume 6, Nomor 2, Edisi September 2021. Pada edisi kali ini memuat 5 (lima) tulisan hasil penelitian dalam bidang informatika.

Harapan kami semoga naskah yang tersaji dalam JURNAL INFORMASI INTERAKTIF edisi September tahun 2021 dapat menambah pengetahuan dan wawasan di bidangnya masing-masing dan bagi penulis, jurnal ini diharapkan menjadi salah satu wadah untuk berbagi hasil-hasil penelitian yang telah dilakukan kepada seluruh akademisi maupun masyarakat pada umumnya.

Redaksi

## ANALISIS DATA UNTUK PENGELOMPOKAN MAHASISWA DENGAN METODE K-MEAN (STUDI KASUS : INSTITUT SHANTI BHUANA)

*Santi Thomas<sup>1</sup>, Noviyanti P<sup>2</sup>*

<sup>1,2</sup>Teknologi Informasi, Institut Shanti Bhuna

Email : <sup>1</sup>santi.thomas@shantibhuana.ac.id, <sup>2</sup>noviyanti@shantibhuana.ac.id

### ABSTRACT

*Student achievement index (GPA) has always been used as a benchmark to assess the quality of a college. Likewise with ISB (Institut Shanti Bhuna) which has been working for 5 years, also hopes to create qualified graduates. ISB is different from the other universities because in addition to intellectual achievement in ISB, it also has an ICP (Integrity-Credit-Point). These two values are used as benchmarks to see whether students have succeeded in becoming qualified graduates as written in the ISB's vision and mission. The purpose of this study was to classify students who have good grades in both GPA and ICP. The K-mean algorithm is chosen because it is easy to adapt and the most widely used in data clustering. The data set was taken as many as 97 students from 4 study programs with 3 clusters. The results showed that cluster 1 was 18 people, cluster 2 was 48 people and cluster 3 was 31 people. It is hoped that this research can help the leadership get useful information in making decisions both in academic development and morality as well as decisions in granting scholarships based on considerations of the value of the GPA and ICP.*

**Keyword:** K-mean, clustering, outstanding students, data analysis

### 1. PENDAHULUAN

ISB mempunyai ciri khas yang berbeda dengan perguruan tinggi lainnya karena ISB mempunyai visi misi menjadi Perguruan Tinggi yang berfokus pada cinta tanah air, integritas, profesionalitas, dan Budaya Amare. Didirikan pada tahun 2016 dengan nama Sekolah Tinggi Shanti Bhuna (STIM) dengan 2 program studi yaitu Program Studi Manajemen (MNJ) dan Program Studi Kewirausahaan (KWU). Kemudian pada tahun 2020, Sekolah Tinggi Shanti Bhuna berubah bentuk menjadi Institut Shanti Bhuna dengan penambahan 2 program studi, yaitu Program Studi Pendidikan Guru Sekolah Dasar (PGSD) dan Program Studi Teknologi Informasi (TI).

Penelitian ini mempunyai tujuan untuk membantu para pemimpin untuk mengambil keputusan dalam berbagai hal salah satunya memberikan beasiswa kepada mahasiswa yang kurang mampu karena syarat yang diajukan selain nilai IPK bagus, nilai dari ICP pun harus bagus. Nilai ICP meliputi iman, keterampilan, fisik, moral, kepribadian dan intelektualitas. Nilai ICP diberikan berdasarkan sikap mahasiswa selama perkuliahan maupun dalam hidup berasma. Salah satu kebijakan dari ISB

untuk mendidik mahasiswa dalam bersikap yaitu mereka pada tahun 1 dan 4 harus hidup dalam asrama yang sudah disediakan .

Metode pengelompokan mahasiswa untuk berbagai keperluan pada umumnya dilakukan dengan menggunakan metode K-mean. Salah satunya seperti yang dilakukan terhadap Universitas Cic Cirebon [1]. Penelitian ini mengelompokkan mahasiswa berdasarkan bidang keahlian dengan jumlah iterasi sebanyak 3 kali dan pengujian dilakukan dengan aplikasi Tanagra dengan nilai akurasi sebesar 80%.

Sementara penelitian terhadap nilai akademik mahasiswa di Universitas Muhammadiyah Magelang juga menunjukkan hasil yang sama [2]. Penelitian ini ingin memilih 5 mahasiswa yang paling berprestasi untuk mengikuti lomba. Dengan menggunakan metode K-mean dan software Weka Interface, penelitian ini berhasil melakukan clustering mahasiswa berdasarkan IPK dan beberapa atribut matakuliah sehingga 5 mahasiswa terpilih dari cluster 1.

Metode K-mean juga dapat dilakukan untuk menganalisa hasil akademik mahasiswa untuk merencanakan dan mengembangkan teknik pengajaran yang menyangkut interaksi antara guru dengan murid, lingkungan pembelajaran

interaksi antara guru dengan murid, lingkungan pembelajaran serta tugas-tugas dalam pembelajaran [3]. Penelitian ini menggunakan 5 *cluster* dan hasil terbaik ada pada *cluster* 2 dan 3. Penelitian ini juga menunjukkan bahwa nilai mahasiswi lebih baik daripada nilai mahasiwa. Hasil penelitian ini menjadi informasi penting bagi perguruan tinggi untuk merancang sistem akademik di masa depan.

Hal yang sama juga dilakukan dengan menganalisa jumlah data yang besar menjadi informasi yang mempunyai arti bagi pendukung keputusan [4]. Penelitian dengan menggunakan algoritma *K-mean* dan *software RapidMiner* ini mengelompokkan prestasi siswa tinggi, menengah dan cukup dengan atribut kegiatan ekstrakurikuler yang diikuti, nilai dari tugas, UTS, UAS, absensi dan nilai sikap.

## 2. LANDASAN TEORI

### 2.1 Analisis Data

Merupakan upaya atau cara untuk mengolah data menjadi informasi sehingga karakteristik data tersebut bisa dipahami dan bermanfaat untuk solusi permasalahan [5]. Analisis data dilakukan dengan 3 tahap, yaitu

- a. Reduksi data merupakan proses menfokus dan mengabstraksikan data mentah menjadi informasi yang bermakns.
- b. Sajian data adalah mengorganisir, menyajikan data dalam bentuk naratif, tabel, matrix atau bentuk lainnya.
- c. Menyimpulkan data adalah mengambil intisari dari sajian data yang telah terorganisir dalam bentuk pernyataan kalimat yang singkat dan padat tetapi mengandung pengertian luas.

### 2.2 Clustering

*Clustering* merupakan pengelompokan data yang didasarkan pada kesamaan antar objek, oleh karena itu *Clustering* digolongkan sebagai metode *unsupervised learning*. *Clustering* dapat dibagi menjadi dua, yaitu *hierarchical clustering* dan *non-hierarchical clustering*[6].

Tujuan dari *clustering* adalah agar objek di dalam grup menjadi serupa (atau terkait) dan berbeda dari (atau tidak terkait dengan) objek di grup lain. Semakin besar kesamaan (atau homogenitas) dalam suatu kelompok dan semakin besar perbedaan antar kelompok maka

semakin baik atau lebih jelas pengelompokannya [7]

*Clustering* merupakan jenis khusus dari klasifikasi [8] tapi ada perbedaan antara *clustering* dan klasifikasi seperti pada tabel 1 [9].

Tabel 1. Perbedaan antara *Clustering* dan Klasifikasi

	<b>Clustering</b>	<b>Klasifikasi</b>
Jenid data	Tidak berlabel	Berlabel
Metode	<i>Unsupervised Leraning</i>	<i>Supervised Leraning</i>
Tujuan	Membentuk grup berdasarkan pola kemiripan antar sampel	Membentuk class berdasarkan feature yangsedediakan oleh data set
Fungsi	Untuk memprediksi <i>class</i> dimasa depan	Untuk memahami data
Algoritama	<i>K-Mean, KNN, Naïve Bayes Classifier</i> , dsb	<i>Decision Tree, SVM, Neural Network</i> , dsb

### 2.3 K-mean

Algoritma *K-means* merupakan algoritma klusterisasi yang mengelompokkan data berdasarkan titik pusat *cluster* (centroid) terdekat dengan data. Tujuan dari *K-means* adalah pengelompokkan data dengan memaksimalkan kemiripan data dalam satu *cluster* dan meminimalkan kemiripan data antar *cluster*. Ukuran kemiripan yang digunakan dalam *cluster* adalah fungsi jarak. Sehingga pemaksimalan kemiripan data didapatkan berdasarkan jarak terpendek antara data terhadap titik centroid [2].

Secara sederhana, algoritma *K-mean* dapat dijabarkan sebagai berikut [7] :

Algoritma Dasar <i>K-mean</i>
1: Pilih titik K sebagai <i>centroid</i> awal
2: <b>Repeat</b>
3: Dari <i>cluster</i> K dengan menetapkan setiap titik ke centroid terdekat
4: Hitung ulang ke centroid untuk setiap <i>cluster</i>
5: <b>Until</b> centroid tidak berubah

Gambar 1. Algoritma *k-mean*

Langkah-langkah algoritma *K-means* adalah sebagai berikut [10] :

- a. Tentukan banyaknya *cluster* yang ingin dibentuk
- b. Pilih secara acak k buah data sebagai pusat *cluster*.
- c. Jarak antara data dan pusat *cluster* dihitung menggunakan Persamaan *Euclidian Distance* seperti berikut ini:

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

(1)

dimana:

- $d(x, y)$  = Jarak data ke x ke pusat *cluster* y
- $x$  = data pusat *cluster*
- $y$  = data pada atribut
- $i$  = setiap data
- $k$  = jumlah data
- $x_i$  = data pada pusat *cluster* ke i
- $y_i$  = data pada setiap data ke i

4. Dari hasil perhitungan jarak data dengan *centroid*, maka dikelompokan data berdasarkan pada jarak minimum.
5. Berdasarkan pengelompokan data tersebut, selanjutnya dicari *centroid* baru berdasarkan keanggotaan setiap *cluster* dengan menghitung rata-rata *cluster*. Untuk menentukan *centroid* baru dapat menggunakan rumus:

$$C(i) = \frac{x_1 + x_2 + \dots + x_n}{\sum x} \dots\dots\dots(2)$$

dimana :

- $x_1 \dots x_n$  = Nilai data record 1 sampai n
- $\sum x$  = Jumlah data record

6. Selanjutnya dilakukan langkah ke 3 kembali sampai beberapa iterasi hingga nilai *centroid* tidak berubah.

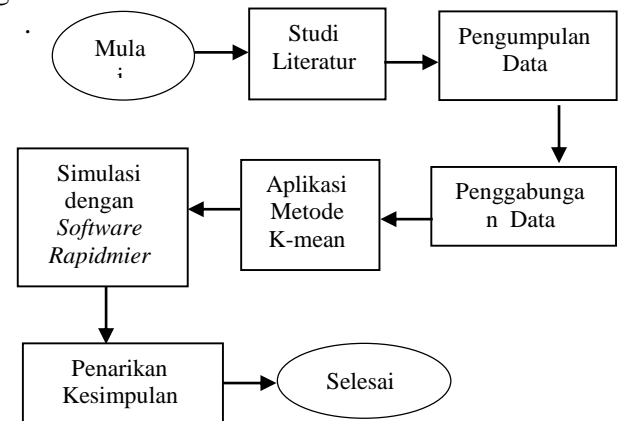
**2.4 RapidMiner**

*RapidMiner* merupakan perangkat lunak yang bersifat terbuka (*open source*). *RapidMiner* adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi. *RapidMiner* menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. *RapidMiner* memiliki kurang lebih 500 operator data mining, termasuk operator untuk input, output, data preprocessing

dan visualisasi. *RapidMiner* merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. *RapidMiner* ditulis dengan menggunakan bahasa Java sehingga dapat bekerja di semua sistem operasi[4].

**3. METODE PENELITIAN**

Adapun alur penelitian dapat dilihat dari gambar 2



Gambar 2. Alur Penelitian

Metode yang digunakan dalam penelitian ini adalah studi literatur, pengumpulan data, aplikasi metode *K-mean* dan simulasi dengan software *RapidMiner Studio*. Pengumpulan data dengan mengunduh nilai IPK dari *SIAKAD* dan ICP dari AIS. *SIAKAD* merupakan aplikasi berbasis web yang dirancang untuk menangani proses pengolahan data akademik dan data terkait lainnya sehingga seluruh proses kegiatan akademik dapat terkelola menjadi informasi yang bermanfaat dalam pengelolaan manajemen perguruan tinggi, pengambilan keputusan serta pelaporan di lingkungan perguruan tinggi. sedangkan AIS (*Academic Information System*) merupakan sistem penilaian yang dirancang sendiri oleh ISB untuk pengisian ICP. Oleh karena itu sebelum mengaplikasikan metode *K-mean*, kedua nilai ini perlu digabungkan terlebih dahulu. Dataset dalam penelitian ini menggunakan data dari mahasiswa angkatan 2020 karena 2 alasan sebagai berikut :

1. Berasal dari 4 prodi dengan jumlah 97 mahasiswa. Jika menggunakan data mahasiswa sebelum tahun 2020, jumlah dataset akan sedikit karena hanya dari 2 prodi.
2. Hanya bisa mengelompokkan berdasarkan 1 angkatan saja karena nilai ICP bertambah

sesuai dengan lamanya belajar. Nilai ICP mahasiswa tingkat 1 lebih rendah daripada nilai ICP mahasiswa tingkat 2 dan di atasnya.

Tabel 2 merupakan contoh data dari gabungan SIAKAD dan AIS yang sudah disederhanakan sesuai keperluan penelitian.

Tabel 2. Contoh Gabungan Data dari SIAKAD dan AIS

No	NIM	Nama	Prodi	IPK	ICP
1	2000103001	Anjas Saputra	PGSD	3.27	21
2	2000502004	Isori	KWU	3.05	41
3	2000501012	Dian Irmayanti	MNJ	3.2	25
4	2000104010	Petrus Sogian	TI	2.98	28

Untuk selanjutnya daftar mahasiswa tabel 2 ini yang akan dijadikan sampel dalam hasil dan pembahasan.

#### 4. HASIL DAN PEMBAHASAN

Dari langkah-langkah algoritma *K-mean* yang sudah dijelaskan diatas, maka langkah pertama yang harus dilakukan adalah menentukan jumlah *cluster*. Dalam penelitian ini menggunakan 3 *cluster* dengan 2 atribut.

Langkah ke 2 yaitu memilih acak 3 buah data sebagai sebagai centroid awal untuk iterasi 1. Adapun 3 centroid awal ini dapat dilihat dalam tabel 3.

Tabel 3. Nilai centroid awal pada setiap *cluster*

Centroid	IPK	ICP
1	3.43	41
2	2.98	28
3	1.09	21

Langkah selanjutnya yaitu menghitung jarak antara data dengan pusat *cluster* dengan menggunakan Persamaan *Euclidian Distance* seperti pada rumus (1) diatas. Maka dilakukan proses perhitungan *Euclidian Distance* dengan contoh data untuk mahasiswa no 1 pada tabel 2 sebagai berikut:

$$d(\text{data1, centroid 1}) = \sqrt{(3.27 - 3.43)^2 + (21 - 41)^2} = 20$$

$$d(\text{data1, centroid 2}) = \sqrt{(3.27 - 2.98)^2 + (21 - 28)^2} = 7.01$$

$$d(\text{data1, centroid 3}) = \sqrt{(3.27 - 1.09)^2 + (21 - 21)^2} = 2.18$$

Hasil perhitungan jarak *Euclidian* pada iterasi 1 dapat dilihat pada tabel 4.

Tabel 4. Hasil jarak *Euclidian* untuk iterasi 1

No	IPK	ICP	Centroid d 1		Centroid d 2		Centroid d 3	
			3.4	4	2.9	2	1.0	2
1	3.27	21	3	1	8	8	9	1
2	3.2	41	0.38		13		20.10	
3	3.2	25	16		3.01		4.52	
4	2.98	28	13.01		0		7.25	

Langkah keempat adalah mengelompokkan data berdasarkan hasil perhitungan *Euclidian* yang diperoleh. Data dikelompokkan berdasarkan jarak minimum atau jarak terdekat dengan *cluster*nya. Dengan membandingkan nilai terkecil dari masing-masing *centroid*, maka dapat ditentukan data tersebut masuk ke dalam *cluster* keberapa. Hasil pengelompokan *cluster* dapat dilihat pada tabel 5.

Tabel 5. Pengelompokkan *cluster* pada iterasi 1

No	Cluster 1	Cluster 2	Cluster 3
1			2.18
2	0.38		
3		3.01	
4		0	

Dari tabel 5 diketahui bahwa mahasiswa 1 masuk dalam *cluster* 3, mahasiswa 2 masuk dalam *cluster* 1 sedangkan mahasiswa 3 dan 4 masuk dalam *cluster* 2 dstnya. Hasil pengelompokkan *cluster* iterasi 1 untuk 97 mahasiswa pada dapat dilihat pada tabel 6.

Tabel 6. Hasil pengelompokkan *cluster* pada iterasi 1 untuk keseluruhan data.

	x	y
Cluster 1	24	24
Cluster 2	42	42
Cluster 3	31	31
Jumlah	97	97



Langkah kelima adalah menentukan nilai *centroid* baru untuk iterasi selanjutnya. Penentuan nilai *centroid* baru dapat dilihat pada rumus (2) dimana untuk menentukan centroid 1 yang baru, semua nilai IPK dan ICP mahasiswa yang masuk dalam *cluster* 1 dijumlahkan dan dibagi dengan jumlah mahasiswa yang masuk dalam *cluster* 1. Hasil dari perhitungan centroid 1 adalah sebagai berikut:

$$\text{Centroid 1 baru untuk atribut x} = \frac{2.88+3.05+ 3.18...+3.2}{24} = 3.16$$

$$\text{Centroid 1 baru untuk atribut y} = \frac{35+41+45...+43}{24} = 40.83$$

Dengan melakukan cara perhitungan yang sama seperti diatas maka diperoleh nilai centroid 2 dan 3 seperti yang dapat dilihat pada tabel 7.

Tabel 7. Nilai Centroid baru

Centroid	X (IPK)	Y(ICP)
1	3.16	40.83
2	2.98	29.57
3	2.49	18.35

Setelah memperoleh nilai centroid baru, maka langkah selanjutnya mengulang lagi langkah ke 3 yaitu mencari jarak *Euclidian* , menentukan masuk *cluster* yang mana, mencari lagi *centroid* baru dan seterusnya sampai akhirnya nilai *centroid* tidak berubah. Tabel 8 menunjukkan perubahan jumlah *cluster* setiap iterasi sampai akhirnya iterasi terakhir yaitu iterasi 4 menunjukkan nilai *centroid* tidak berubah.

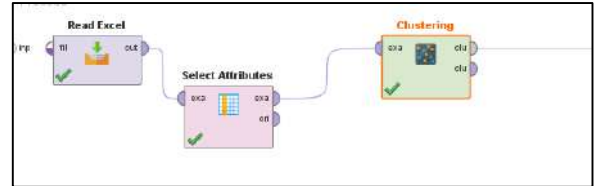
Tabel 8. Perubahan nilai *centroid* dan jumlah dalam *cluster* untuk setiap iterasi

Literasi		cluster 1	cluster 2	cluster 3	
1	Centroid	x	3.16	2.98	2.49
		y	40.83	29.57	18.35
	Jumlah (orang)	24	42	31	
2	Centroid	x	3.20	2.98	2.47
		y	42.00	29.91	18.17
	Jumlah (orang)	20	47	30	
3	Centroid	x	3.20	2.97	2.49
		y	42.00	30.04	18.35
	Jumlah (orang)	20	46	31	
4	Centroid	x	3.23	2.97	2.49
		y	42.67	30.29	18.35
	Jumlah (orang)	18	48	31	

Setelah melakukan perhitungan manual dengan menggunakan excel, maka perlu dilakukan pengujian dengan menggunakan

aplikasi pengujian data mining, yaitu *RapidMiner*. Tujuan dari penggunaan *software* ini adalah membandingkan hasil dengan perhitungan secara teoritis dengan hasil yang didapatkan dengan proses di *RapidMiner* ini.

Adapun tahapan yang dilakukan dengan *RapidMiner* ini adalah membaca data di excel, *select attribute* lalu *clustering* seperti pada gambar 3.



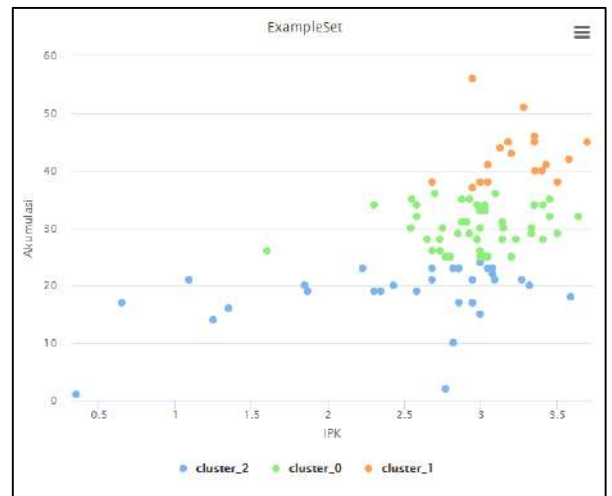
Gambar 3. Proses *clustering* di *RapidMiner*

Hasil yang didapat dapat dilihat dari gambar 4 berikut ini.

### Cluster Model

Cluster 0: 48 items  
 Cluster 1: 18 items  
 Cluster 2: 31 items  
 Total number of items: 97

Gambar 4. Hasil *cluster* dengan *software RapidMiner*



Gambar 5. Hasil *Clustering* dengan *Plot Scutter*

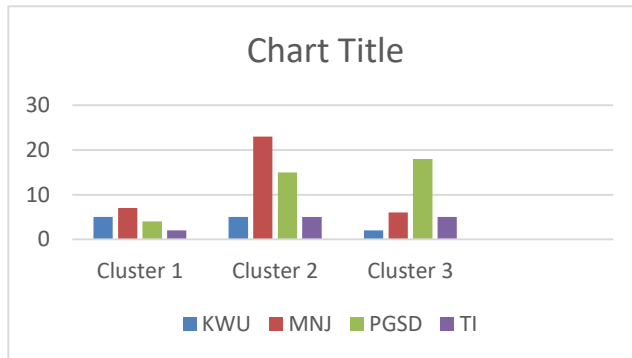
Sedangkan hasil *centroid* dapat dilihat pada gambar 6 berikut ini:

Attribute	cluster_0	cluster_1	cluster_2
IPK	2.971	3.230	2.488
Akumulasi	30.292	42.667	18.355

Gambar 6. Hasil *Centroid* dengan *software RapidMiner*

Attribute IPK dalam *RapidMiner* menjadi *attribute X* dalam perhitungan menggunakan Excel dan *Attribute* Akumulasi dalam *RapidMiner* menjadi *attribute Y* dalam perhitungan dengan Excel.

Secara sederhana, hasil *clustering* dapat dilihat pada grafik 1.



Grafik1. Hasil *clustering* secara keseluruhan

## 5. KESIMPULAN

Hasil ini didapat dari perhitungan secara teoritis dengan menggunakan *excel* dan pengujian dengan menggunakan software *RapidMiner*. Dan nilai akurasi dari 2 perhitungan ini hampir mencapai 90%.

Jika dilihat dari grafiknya, *clustering* ini lebih mengarah kepada nilai ICP. *Cluster 1* adalah mahasiswa dengan nilai ICP 35 keatas, sedangkan *cluster 2* adalah mahasiswa dengan nilai ICP 25-35 dan *cluster 3* adalah mahasiswa dengan nilai ICP dibawah 25. Berbeda dengan IPK yang lebih stabil, ICP dapat meningkat sesuai dengan lamanya kuliah. Jika mahasiswa itu aktif dalam kegiatan kampus, integritasnya bagus maka nilai ICP itu dapat bertambah dengan cepat. Dengan demikian mahasiswa ini dapat disebut sebagai mahasiswa yang berintegritas, tidak hanya pintar dalam intelektual saja. Sebaliknya, nilai ICP itu juga dapat berkurang jika mahasiswa tersebut tidak menunjukkan integritasnya seperti yang dididik dalam ISB. Diharapkan penelitian ini dapat membantu pimpinan mengambil kebijakan untuk berbagai kepentingan mahasiswa dengan menimbang nilai IPK dan ICP mereka.

Untuk penelitian selanjutnya, akan mengklasifikasi dengan menggunakan *naïve bayes*.

## DAFTAR PUSTAKA

- [1] C. Nas, "Data Mining Pengelompokan Bidang Keahlian Mahasiswa Menggunakan Algoritma K-Means ( Studi Kasus : Universitas Cic Cirebon )," vol. 09, no. 1, pp. 1–14, 2020.
- [2] I. Method, K. C. Based, S. Value, W. Interface, C. Study, and I. U. M. M. Magelang, "Penerapan Metode K-Means Untuk *Clustering* Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang," vol. 18, no. 1, pp. 76–82, 2015.
- [3] D. Aggarwal and D. Sharma, "Application of *Clustering* for Student Result Analysis," no. July, 2019.
- [4] M. L. Sibuea and A. Safta, "Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering," *Jurteks*, vol. 4, no. 1, pp. 85–92, 2017, doi: 10.33330/jurteks.v4i1.28.
- [5] Ms. Prof. Dr. Suryana, "Metodologi Penelitian : Metodologi Penelitian Model Praktis Penelitian Kuantitatif dan Kualitatif," *Univ. Pendidik. Indones.*, pp. 1–243, 2012, doi: 10.1007/s13398-014-0173-7.2.
- [6] A. N. Khomarudin, "Teknik Data Mining : Algoritma K-Means *Clustering*," pp. 1–12, 2016.
- [7] C. Analysis and B. Concepts, "*Cluster Analysis* :"
- [8] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*. New Jersey: Prentice Hall, 2018.
- [9] B. Purnama, *Pengantar Machine Learning*. Bandung: Informatika Bandung, 2019.
- [10] W. Purba, S. Tamba, and J. Saragih, "The effect of mining data k-means *clustering* toward students profile model drop out potential," *J. Phys. Conf. Ser.*, vol. 1007, no. 1, 2018, doi: 10.1088/1742-6596/1007/1/012049.