

**LEMBAR  
HASIL PENILAIAN SEJAWAT SEBIDANG ATAU PEER REVIEW  
KARYA ILMIAH : JURNAL ILMIAH**

Judul Jurnal Ilmiah : Clustering Data UTS Mata Kuliah Data Mining Mahasiswa Menggunakan Algoritma K-Means

Penulis Jurnal Ilmiah : 1. Ginanjar Abdurrahman, S.Si, M.Pd

Identitas Jurnal Ilmiah :

- a. Nama Jurnal : Jurnal Sistem dan Teknologi Infomrasi Indonesia
- b. Nomor/Volume : 2/1
- c. Edisi/ISSN : Agustus 2016/ p.ISSN:2502-5724; e-ISSN:2541-5735
- d. Penerbit : Program Studi Teknik Informatika Muhammadiyah Universitas of Jember
- e. Jumlah Halaman : 68

Kategori Publikasi Makalah :

- Jurnal Ilmiah Internasional
- Jurnal Ilmiah Nasional Terakreditasi
- Jurnal Ilmiah Nasional Tidak Terakreditasi

Hasil Penilaian Peer Review :

Komponen yang Dinilai	Nilai Maksimal Jurnal Ilmiah			Nilai Akhir Yang Diperoleh
	Internasional <input type="checkbox"/>	Nasional Terakreditasi <input type="checkbox"/>	Nasional Tidak Terakreditasi <input checked="" type="checkbox"/>	
a. Kelengkapan unsur isi buku (10%)			7,5	0,75
b. Ruang lingkup dan kedalaman pembahasan (30%)			7,5	2,25
c. Kecukupan dan kemutakhiran data/informasi dan metodologi (30%)			7,5	2,25
d. Kelengkapan unsur dan kualitas penerbit (30%)			7,5	2,25
<b>Total = (100%)</b>				7,5

Jember, 30 September 2018

Reviewer 1

Wiwik Suharso, S.Kom, M.Kom  
NIP. 19760906 200501 1 003  
Unit kerja : FT UM Jember

**LEMBAR  
HASIL PENILAIAN SEJAWAT SEBIDANG ATAU PEER REVIEW  
KARYA ILMIAH : JURNAL ILMIAH**

Judul Jurnal Ilmiah : Clustering Data UTS Mata Kuliah Data Mining Mahasiswa Menggunakan Algoritma K-Means

Penulis Jurnal Ilmiah : 1. Ginanjar Abdurrahman, S.Si, M.Pd

Identitas Jurnal Ilmiah : a. Nama Jurnal : Jurnal Sistem dan Teknologi Infomasi Indonesia  
 b. Nomor/Volume : 2/1  
 c. Edisi/ISSN : Agustus 2016/ p.ISSN:2502-5724; e-ISSN:2541-5735  
 d. Penerbit : Program Studi Teknik Informatika Muhammadiyah Universitas of Jember  
 e. Jumlah Halaman : 68

Kategori Publikasi Makalah :  Jurnal Ilmiah Internasional  
 Jurnal Ilmiah Nasional Terakreditasi  
 Jurnal Ilmiah Nasional Tidak Terakreditasi

Hasil Penilaian Peer Review :

Komponen yang Dinilai	Nilai Maksimal Jurnal Ilmiah			Nilai Akhir Yang Diperoleh
	Internasional <input type="checkbox"/>	Nasional Terakreditasi <input type="checkbox"/>	Nasional Tidak Terakreditasi <input checked="" type="checkbox"/>	
a. Kelengkapan unsur isi buku (10%)			8	0,8
b. Ruang lingkup dan kedalaman pembahasan (30%)			7,5	2,25
c. Kecukupan dan kematangan data/informasi dan metodologi (30%)			7,5	2,25
d. Kelengkapan unsur dan kualitas penerbit (30%)			8	2,4
<b>Total = (100%)</b>				<b>7,7</b>

Jember, 30 September 2018

Reviewer 2

Agung Nilogiri, S.T, M.Kom  
 NIP. 19770330 200501 1 002  
 Unit kerja : FT UM Jember

# Clustering

*by* Ginanjar Abdurrahman

---

**Submission date:** 11-Aug-2018 09:22AM (UTC+0700)

**Submission ID:** 989082171

**File name:** Clustering\_UTS\_new.pdf (925.08K)

**Word count:** 2777

**Character count:** 16552

# **Clustering Data UTS Data Mining Menggunakan Algoritma K-means**

Ginanjar Abdurrahman<sup>1)</sup>

41

<sup>1)</sup>Jurusan Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember

Email: <sup>1)</sup> [abdurrahmanganjar@unmuhjember.ac.id](mailto:abdurrahmanganjar@unmuhjember.ac.id)

## Abstrak:

Ilmu pengetahuan dan teknologi akan mempermudah pekerjaan. Namun, di sisi lain akan meningkatkan persaingan. Dalam menghadapi ketatnya persaingan, perlu dipersiapkan sumber daya manusia yang kompeten. Mahasiswa diharapkan siap secara akademik, berupa kesiapan <sup>32</sup>getahuan maupun keterampilan untuk menghadapi persaingan yang semakin ketat. Salah satu cara untuk melihat kompetensi mahasiswa adalah dengan melihat hasil belajar yang dapat direpresentasikan dengan nilai ujian yang ditempuh. Ujian tengah semester (UTS) merupakan salah satu bentuk ujian yang menjadi komponen penilaian. Dengan mengetahui nilai UTS, dosen mengetahui sebaran mahasiswa dalam hal kompetensi akademik. Untuk itulah, diperlukan pengelompokan (*clustering*) dengan algoritma *k-means* sebagai pertimbangan dosen dalam membentuk kelompok belajar mahasiswa, berdasarkan *cluster* nilai UTS.

Kata kunci: UTS, kompetensi, *clustering*, *k-means*.

## *Abstract*

*Science and technology will make task easier. But, in the other hand it will increase the competition. To face the competition, it need to be prepared the competent human resources. The students need to be prepared academically, that is the preparation of knowledge and skills to face the strictness of the competition itself. One way to see the competency of students is by seeing the study result of the students which can be represented by the examination grade of the student. Mid semester examination is one of the examination that become one of assesment. By knowing the mid semester grades, the lecturer can knowing the spread of students in academic competencies. Hence, it need the clustering methode with k-means algorithim as a review of lecturer for forming study group of students based on clustering mid semester examination grades.*

Keyword: *Mid semester examination, competencies, clustering, k-means*

## 1. PENDAHULUAN

Di satu sisi, ilmu perkembangan ilmu pengetahuan dan teknologi akan mempermudah pekerjaan manusia. Namun, di lain sisi hal ini akan meningkatkan persaingan baik itu persaingan lokal maupun global. Untuk menghadapi persaingan, perlu dipersiapkan sumber daya manusia, dalam hal ini mahasiswa yang kompeten, yakni dalam hal akademik berupa pengetahuan dan keterampilan. Salah satu cara untuk mengukur kompetensi <sup>21</sup>demik adalah dengan nilai mata kuliah. Permendikbud nomor 66 tahun 2013 tentang Standar Penilaian Pendidikan menyebutkan <sup>39</sup> bahwa penilaian adalah proses untuk mengukur pencapaian hasil belajar peserta didik, dalam hal ini adalah mahasiswa. Sedangkan untuk mengetahui tingkat kompetensi akademik, diperlukan suatu prosedur penilaian yakni tes/ujian (Gronlund, 1977:1). Ada beberapa komponen penilaian dalam menentukan nilai akhir mata kuliah di tingkat pendidikan tinggi, yang salah satunya adalah nilai Ujian Tengah Semester (UTS). Dengan mengetahui nilai UTS, dosen dapat mengetahui sebaran mahasiswa dalam hal kompetensi akademik.

Dengan pertimbangan tersebut, perlu diadakan pengelompokan mahasiswa berdasarkan nilai UTS agar dosen dapat menerapkan metode pembelajaran yang tepat dan menjadi bahan pertimbangan untuk pembentukan kelompok belajar sesuai dengan sebaran kemampuan akademik mahasiswa. Salah satu metode untuk pengelompokan adalah metode *clustering* dengan algoritma *k-means*. Algoritma *k-means* dipilih karena sederhana dan dimulai dengan deskripsi dasar algoritma (Tan, Steinbech, & Kumar, 2006:497). Selain itu, *k-means* juga dapat digunakan untuk data berukuran besar dan sangat efisien (Tan, Steinbech, & Kumar, 2006:510).

Pada Penelitian ini, data hasil UTS mata kuliah data mining akan dikelompokkan dengan

metode *clustering* menggunakan algoritma *k-means* dengan bantuan program aplikasi *weka 3.6*

20

## 2. TINJAUAN PUSTAKA

### 2.1 Data mining

Data mining merupakan proses dalam menemukan hubungan yang berarti, pola, dan tren dengan meriksa data berukuran besar dalam suatu penyimpanan dengan menggunakan teknologi pengenalan pola, misalnya statistik dan matematika (Gartner Group dalam Larose, 2005:2)

Data mining juga dapat artikan sebagai proses menemukan pola dalam data (Witten, Frank, & Hall, 2011:5) Sedangkan menurut Hand et al dalam (Larose, 2005:2) menyatakan bahwa data mining merupakan analisis data berukuran besar untuk menemukan hubungan dari suatu data agar dapat dipahami dan dapat digunakan.

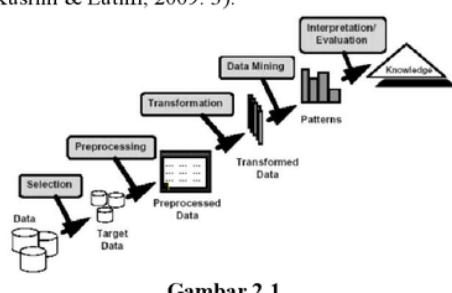
Menurut Pramudiono (Kusrini & Luthfi, 2009: 3) : "Data mining merupakan proses untuk menggali pengetahuan dari suatu himpunan data. Masih menurut Pramudiono (Kusrini & Luthfi, 2009:6) menyatakan bahwa data mining diidentikkan dengan *knowledge discovery in database* (KDD) yang mempunyai definisi proses penggalian suatu basis data yang besar untuk menemukan informasi.

Dari beberapa pengertian di atas, dapat disimpulkan bahwa data mining merupakan serangkaian proses dalam pencarian pola, hubungan, penggalian nilai tambah dari data dan informasi yang berukuran besar berupa pengetahuan dengan tujuan menemukan hubungan dan menyederhanakan data agar diperoleh informasi yang dapat dipahami dan bermanfaat dengan bantuan ilmu statistik dan matematika.

### 2.2 Tahapan Data Mining

38

Tahap-tahap data mining secara umum dapat dilihat pada gambar 2.1 (Fayyad dalam Kusrini & Luthfi, 2009: 3).



Gambar 2.1

### 2.3 Klasifikasi

Metode klasifikasi adalah metode yang melibatkan variabel kategori (Larose, 2005:14) yang artinya suatu objek dinyatakan ke salah satu kategori atau kategori yang lain. Metode dalam data mining yang merupakan klasifikasi

diantaranya adalah C.4.5, Nearest Neighbor, Naïve bayesian, back propagation, dan sebagainya.

### 2.4 Clustering

Metode *Clustering* merupakan metode pengelompokan data, observasi, atau kasus menjadi kelas objek-objek yang serupa. Metode *clustering* mencari segmen keseluruhan data menjadi sub grup-sub grup yang relatif homogen atau biasa disebut sebagai *cluster* (Larose, 2005: 16). Menurut Wu & Kumar (2009:33) algoritma *clustering* menempatkan data-data yang sama pada satu kelompok (*cluster*), sedangkan data yang tidak sama pada kelompok (*cluster*) yang lain. Contoh metode ini adalah algoritma *K-means*, algoritma *Agglomerative Hierarchical Clustering*, *Divisive Hierarchical Clustering*, dan sebagainya.

### 2.5 Asosiasi

Asosiasi sering disebut sebagai analisis keranjang belanja (*market basket analysis*) merupakan metode dalam *data mining* untuk menemukan aturan untuk mengukur hubungan antara dua atau lebih atribut. Aturan asosiasi biasanya dinyatakan dalam bentuk jika anteseden, maka konsekuensi, bersama dengan besarnya nilai *support* dan *confidence* yang berasosiasi dengan aturnannya (Larose, 2005:17). Contoh metode ini adalah algoritma *FP-Growth*, algoritma *Apriori*, dan sebagainya.

3

### 2.6 Algoritma K-means

*K-means clustering* merupakan metode *clustering* non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih *cluster/kelompok*. Data-data yang mempunyai kemiripan tinggi dikelompokkan dalam satu *cluster/kelompok* dan data yang memiliki kemiripan rendah dikelompokkan dengan *cluster/kelompok* yang lain (Agusta, 2007). Sedangkan menurut Wu & Kumar (2009:33) *K-means* adalah algoritma *clustering* yang mempartisi himpunan D menjadi *k* *cluster* data. Algoritma *K-means* mengcluster semua titik data pada D sedemikian sehingga titik data  $x_i$  menjadi satu-satunya *k* partisi. Dengan kata lain, satu titik data hanya masuk ke dalam satu *cluster*.

Langkah-langkah dari *K-means* yaitu: (Tan, Steinbech, & Kumar, 2006:497).

1. Tentukan *K* data sebagai *centroid*, *K* adalah jumlah *cluster* yang diinginkan (ditentukan oleh peneliti).
2. Tiap titik (data) kemudian dicari *centroid* terdekatnya
3. Setiap himpunan titik (data) yang menjadi *centroid* disebut *cluster*
4. Hitung kembali *centroid* dari setiap *cluster*
5. Ulangi langkah 1-4 sampai *centroid* tidak berubah

Metode *clustering* menggunakan Algoritma *K-means*, ukuran kedekatan data dihitung

menggunakan jarak *Euclidean*. Algoritma *K-means* bertujuan untuk meminimumkan jarak total *Euclidean* diantara setiap titik  $x_i$  dan *cluster* terdekat yakni  $c_j$  (Wu & Kumar, 2009:23). Jarak *Euclidean* ditentukan dengan menggunakan persamaan 2.1 berikut (Agusta, 2007):

$$d_{ij} = \sqrt{\sum_{k=1}^n \{x_{ik} - x_{jk}\}^2}$$

... (2.1)

Keterangan:

$d_{ij}$  = jarak antara data ke-i dan data ke-j  
 $n$  = dimensi data

$x_{ik}$  = koordinat data ke-i pada dimensi k  
 $x_{jk}$  = koordinat data ke-j pada dimensi k

Sedangkan menurut Hans & Kamber (2016) algoritma *k-means* mengelompokkan data ke dalam k-buah *cluster* yang telah ditentukan. Beberapa cara untuk menghitung jarak antara lain dengan menggunakan *Euclidean distance*, *Manhattan distance*, dan *Chebisev distance*.

#### 1. Euclidean Distance

Formula untuk menghitung jarak antar dengan *Euclidean Distance* untuk dua titik dalam satu, dua atau tiga dimensi secara berurutan ditunjukkan pada persamaan (2.2), (2.3), dan (2.4) berikut.

$$\sqrt{(x-y)^2} = |x-y| \quad ... (2.2)$$

$$d(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2} \quad ... (2.3)$$

$$d(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2} \quad ... (2.4)$$

#### 2. Manhattan Distance (Taxicab distance)

$$d_1(p,q) = \|p-q\|_1 = \sum_{i=1}^n |p_i - q_i| \quad ... (2.5)$$

#### 3. Chebisev Distance (Maximum Metric)

Untuk menentukan jarak dengan menggunakan *Chebisev Distance* dilakukan dengan cara mengambil nilai maksimum dari setiap koordinat dimensinya. Jika dinyatakan dalam persamaan matematika, maka *Chebisev Distance* dapat dilihat pada persamaan (2.6)

$$D_{cheb}(p,q) = \max(|p_i - q_i|) \quad ... (2.6)$$

Dalam penelitian ini, dipilih formula yang *Euclidean Distance* untuk menentukan jarak

#### 2.7 Ujian/Tes

Tes digunakan untuk mengukur tingkat keberhasilan prosedur pembelajaran (Gronlund,

1977:1). Ebel & Frisbie (1991:30) menyatakan bahwa tes memberikan informasi yang terbaik bagi guru dan siswa tentang keberhasilan setelah mengajar dan belajar. Ebel & Frisbie (1986: 20-21) menambahkan bahwa fungsi utama dari tes adalah untuk mengukur prestasi siswa dan kontribusinya terhadap evaluasi dan pencapaian proses pendidikan, tes juga membantu guru dalam menentukan nilai yang valid dan level yang reliabel bagi siswa.

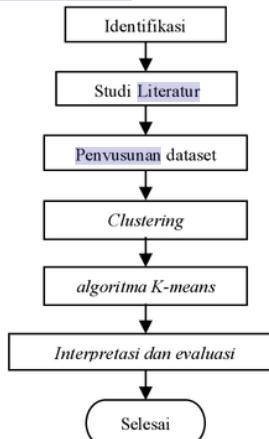
Nitko & Brookhart (2011:5) menyatakan bahwa tes merupakan instrumen atau prosedur sistematis untuk mengobservasi atau mendeskripsikan satu atau lebih karakteristik siswa dengan skala numerik atau skala pengklasifikasian.

Menurut Muijs & Reynolds (2005: 232) tes prestasi mengukur kinerja siswa dalam satu mata pelajaran dalam waktu tertentu. Berdasarkan Permendiknas nomor 66 tahun 2013 tentang Standar Penilaian Pendidikan disebutkan bahwa penilaian adalah proses pengumpulan dan pengolahan informasi untuk mengukur pencapaian hasil belajar siswa. Prestasi belajar siswa akan diketahui melalui penilaian yang dilakukan guru melalui suatu evaluasi belajar yang biasanya berupa tertulis.

Dari uraian di atas, maka dapat disimpulkan bahwa ujian/tes adalah alat untuk mengukur tingkat keberhasilan prosedur pembelajaran untuk memberikan informasi kepada dosen untuk mengobservasi atau mendeskripsikan satu atau lebih karakteristik siswa dengan menggunakan skala numerik.

### 3. METODE PENELITIAN

Langkah-langkah penelitian dapat dilihat pada Gambar 3.1 berikut.



Gambar 3.1 Langkah-Langkah Penelitian

#### 1. Identifikasi

Pada tahap ini, dilakukan identifikasi salah yang ada apa saja serta menawarkan solusi

yang dapat digunakan untuk menyelesaikan masalah tersebut.

## 2. Studi Literatur

Tahap ini merupakan tahap untuk mencari referensi, berupa jurnal penelitian, paper, buku-buku referensi, dan referensi lain terkait penelitian.

## 3. Penyusunan dataset

Penelitian ini melibatkan 103 data nilai UTS Mahasiswa semester IV Universitas Muhammadiyah Jember pada mata kuliah data mining.

## 4. Clustering

Pada tahap ini, metode *clustering* digunakan untuk mengelompokkan data menjadi tiga kelompok.

## 5. Algoritma K-Means

Pada tahap ini, algoritma *K-means* diterapkan untuk mengelompokkan dataset yang ada menjadi tiga cluster.

Contoh perhitungan k-means adalah sebagai berikut.

Perhaikan Tabel 3.1

data	Jml Transaksi	Tot. Belanja
1	4	2.255
2	2	1.788
3	2	1.255
4	4	3.322
5	4	3.012
6	3	1.261

Tabel 3.1

Langkah pertama, ditentukan terlebih dahulu jumlah cluster, misalnya  $k=3$  dan ditentukan *centroid* (pusat *cluster*) awal nya adalah  $C1=(2,1)$ ;  $C2=(4,3)$ ,  $C3=(3,2)$ .

Dengan menggunakan Tabel 3.1, hitung Jarak Euclidean setiap item (Jml Transaksi & Tot. Belanja) dengan centroid awal, maka diperoleh:

untuk data-1, diperoleh:

$$C1 = \sqrt{(4 - 2)^2 + (2.255 - 1)^2} = 2.36$$

$$C2 = \sqrt{(4 - 4)^2 + (2.255 - 3)^2} = 0.745$$

$$C3 = \sqrt{(4 - 3)^2 + (2.255 - 2)^2} = 1.03$$

dan seterusnya, sehingga diperoleh jarak antara data dengan centroid awal seperti pada Tabel 3.2

	C1	C2	C3
1	2.361149085	0.745	1.032000484
2	0.788	2.338577	1.022225024
3	0.255	2.654247	1.247006415
4	3.064585453	0.322	1.657613948
5	2.836925096	0.012	1.422724148
6	1.033499395	2.006021	0.739

Tabel 3.2

Dari Tabel 3.2 dipilih jarak Euclidean terkecil, sehingga diperoleh matriks penempatan cluster seperti tampak pada Tabel 3.3 berikut.

	7 C1	C2	C3
1	*		
2	*		
3	*		
4		*	
5		*	
6			*

Tabel 3.3

Ditentukan *centroid* baru yakni ( $C1=(C11,C12)$ ,  $C2=(C21,C22)$ ,  $C3=(C31,C32)$ ) dengan melihat matriks penempatan *cluster* dan menggunakan data pada Tabel 3.1.

$$C11=\text{Average } (2,2)=2$$

$$C12=\text{Average } (1.788,1.255)=1.5215$$

$$C21=\text{Average } (4,4,4)=4$$

$$C22=\text{Average } (2.255, 3.322, 3.012)=2.863$$

$$C31=\text{Average } (3)=3$$

$$C32=\text{Average } (1.261)=1.261$$

Sehingga *centroid* baru tersebut adalah:

$$C1=(C11,C12)=(2,1.5215)$$

$$C2=(C21,C22)=(4,2.863)$$

$$C3=(C31,C32)=(3,1.261)$$

Dengan menggunakan *centroid* baru, ditentukan kembali jarak setiap item terhadap *centroid* baru tersebut, sehingga diperoleh data jarak terhadap *centroid* baru seperti terlihat pada Tabel 3.4 berikut ini.

ini nilai awal maupun asumsi yang digunakan dalam analisis data adalah:

1. Jumlah *cluster* (*k*) ditentukan sebanyak 3 *cluster*.
2. *Cluster-cluster* yang terbentuk yakni *cluster* nilai rendah, *cluster* nilai sedang, dan *cluster* nilai tinggi.

Dari hasil pengolahan terhadap data nilai UTS mahasiswa dengan algoritma *k-means* untuk *clustering* menggunakan aplikasi *weka 3.6*, diperoleh hasil sebagai berikut:

1. Hasil *clustering* untuk kelas teknik informatika-A dari *output* aplikasi *weka 3.6* menunjukkan sebanyak 5 (10%) nilai UTS termasuk dalam *cluster* nilai rendah, sebanyak 17 (33%) nilai UTS mahasiswa termasuk dalam *cluster* nilai sedang, dan sebanyak 30 (58%) nilai UTS mahasiswa termasuk dalam *cluster* nilai tinggi. Selengkapnya, hasil *clustering* untuk nilai <sup>26</sup> UTS data mining kelas teknik informatika-A dapat dilihat pada Gambar 3.1.

	C1	C2	C3
1	2.130263423	0.608	1.409977305
2	0.2665	2.2706	1.130366755
3	0.2665	2.566255	1.000018
4	2.691059317	0.459	2.290790475
5	2.494311578	0.149	2.016432741
6	1.033373238	1.888493	0

Tabel 3.4

Dari Tabel 3.4 dipilih jarak Euclidean terkecil sehingga diperoleh matriks penempatan cluster seperti tampak pada Tabel 3.5.

	C1	C2	C3
1		*	
2	*		
3	*		
4		*	
5		*	
6			*

Tabel 3.5

Dengan membandingkan Tabel 3.3 dan Tabel 3.5. Ternyata tidak terjadi perpindahan *centroid*, sehingga iterasi *clustering* dihentikan.

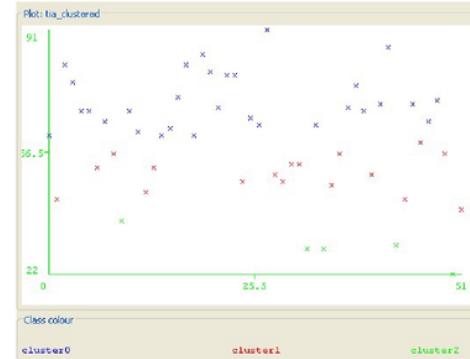
## 6. Interpretasi dan Evaluasi

Tahap interpretasi dan evaluasi <sup>2</sup> informasi yang dihasilkan dari proses data mining perlu disajikan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan.

## 4. HASIL DAN PEMBAHASAN

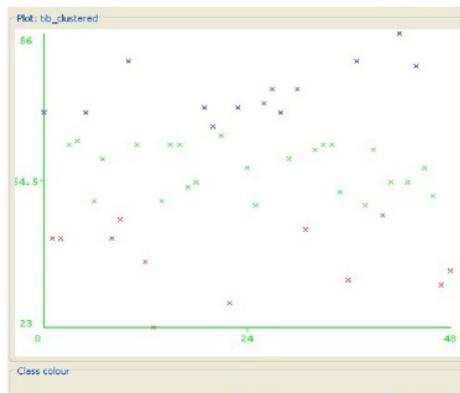
Penelitian ini menggunakan program aplikasi *weka 3.6* untuk mengelompokkan (*clustering*) nilai UTS mahasiswa dengan algoritma *clustering*, yakni algoritma *k-means*. *Clustering* dilakukan pada nilai Ujian Tengah Semester (UTS) mahasiswa Universitas Muhammadiyah Jember tahun akademik 20152 untuk mata kuliah *data mining* pada kelas teknik infomatika-A dan teknik informatika-B.

Data dari kelas teknik informatika-A dan kelas teknik informatika kelas-B, diproses secara terpisah dengan algoritma *k-means* menggunakan aplikasi *weka 3.6* untuk mendapatkan *clustering* mahasiswa berdasarkan nilai UTS. Pada pengujian



Gambar 3.1 Plot hasil clustering nilai UTS teknik informatika-A

2. Hasil *clustering* untuk kelas teknik informatika-B dari *output weka 3.6* menunjukkan sebanyak 12 (24%) nilai UTS termasuk dalam *cluster* nilai rendah, sebanyak 24 (49%) nilai UTS mahasiswa termasuk dalam *cluster* nilai sedang, dan sebanyak 13 (27%) nilai UTS termasuk dalam *cluster* nilai tinggi. Selengkapnya, hasil *clustering* untuk nilai <sup>2</sup> UTS data mining kelas teknik informatika-B dapat dilihat pada Gambar 3.2.



**Gambar 3.2 Plot hasil clustering nilai UTS teknik informatika-B**

8

## 5. KESIMPULAN DAN SARAN

### 5.1 Kesimpulan

Berdasarkan hasil dan pembahasan *output* dari aplikasi weka, maka dapat diambil beberapa kesimpulan sebagai berikut:

1. *Clustering* yang terbentuk adalah tiga *cluster* yakni, *cluster* nilai UTS tinggi, *cluster* nilai UTS sedang, dan *cluster* nilai UTS rendah.
2. Dari *clustering* yang terbentuk dari masing-masing kelas, dapat dijabarkan bahwa untuk kelas teknik informatika-A, sebanyak 5 (10%) nilai UTS termasuk dalam *cluster* nilai rendah, sebanyak 17 (33%) nilai UTS mahasiswa termasuk dalam *cluster* nilai sedang, dan sebanyak 30 (58%) nilai UTS mahasiswa termasuk dalam *cluster* nilai tinggi. Sedangkan untuk teknik informatika-B, sebanyak 12 (24%) nilai UTS termasuk dalam *cluster* nilai rendah, sebanyak 24 (49%) nilai UTS mahasiswa termasuk dalam *cluster* nilai sedang, dan sebanyak 13 (27%) nilai UTS termasuk dalam *cluster* nilai tinggi.

### 5.2 Saran

1. Perlu dilakukan perbandingan dengan algoritma *clustering* yang lain untuk data UTS sebagai perbandingan validitas *clustering* yang **25** tentukan.
2. Penelitian ini dapat dijadikan sebagai salah satu referensi untuk dosen pengampu dalam **22** ibentukan kelompok belajar mahasiswa.
3. Penelitian ini dapat dijadikan sebagai referensi untuk penelitian selanjutnya yang serumpun.

## DAFTAR PUSTAKA

- Agusta. (2007). *K-means-penerapan, permasalahan, dan metode terkait stmiK stikom bali, denpasar, bali.*
- Ebel, R.L., & Frisbie, D.A., (1986). *Essentials of education measurement (4<sup>th</sup> ed.).*

Engelwood Cliffs, New Jersey: Prentice Hall.

42  
..(1991). *Essentials of educational measurement (5<sup>th</sup> ed.).* Engelwood Cliffs, New Jersey: Prentice Hall.

Gronlund, N.E., (1977). *Constructing achievement test (2nd ed.).* Engelwood Cliffs, New Jersey: Prentice-Hall.

15  
Hans, J., & Kamber, M., (2006). *Data mining: concepts and techniques, 2<sup>nd</sup> ed.* Burlington: Morgan Kaufmann.

13  
Kemendikbud. (2006). *Peraturan menteri pendidikan dan kebudayaan republik indonesia nomor 66, tahun 2013, tentang standar penilaian pendidikan.*

6  
Husrini & Luthfi. (2009). *Algoritma data mining.* Yogyakarta: Andi.

Larose, D.T., (2005). *Discovering knowledge in data: an introduction to data mining.* New Jersey: John Wiley & Sons.

5  
Miller, M.D., Linn, R.L., & Gronlund, N.E., (2009). *Measurement and assessment in teaching (10<sup>th</sup> ed.).* Upper Saddle River, New Jersey: Pearson Education, Inc.

9  
Muijs, D., & Reynolds, D., (2005). *Effective teaching: evidence and practice(2<sup>nd</sup> ed).* London: SAGE Publications, Ltd.

33  
Nitko & Brookhart. (2011). *Educational Assessment of Students.* Boston: Allyn & Bacon.

16  
Tan, P.N., Steinbech, M., & Kumar, V., (2006). *Introduction to data mining.* Boston: Pearson Education, Ltd.

1  
Witten, I.H., Frank, E., Hall, M.A., (2011). *Data mining: practical machine learning tools and techniques third edition.* Burlington: Morgan Kaufmann.

Wu, X., & Kumar, V., (2009). *The top ten algorithms in data mining.* Boca Raton: CRC Press.

# Clustering

## ORIGINALITY REPORT



## PRIMARY SOURCES

- |   |   |            |
|---|---|------------|
| 1 | <b>Submitted to AUT University</b><br>Student Paper   | <b>1 %</b> |
| 2 | <b>mainaardi.blogspot.com</b><br>Internet Source  | <b>1 %</b> |
| 3 | <b>journals.ums.ac.id</b><br>Internet Source  | <b>1 %</b> |
| 4 | <b>munirmuhammad1991.blogspot.com</b><br>Internet Source  | <b>1 %</b> |
| 5 | <b>Submitted to Curtin University of Technology</b><br>Student Paper  | <b>1 %</b> |
| 6 | <b>Y A Gerhana, W B Zulfikar, A H Ramdani, M A Ramdhani. "Implementation of Nearest Neighbor using HSV to Identify Skin Disease", IOP Conference Series: Materials Science and Engineering, 2018</b><br>Publication | <b>1 %</b> |
| 7 | <b>www.golftracer.com</b><br>Internet Source  | <b>1 %</b> |
|   | <b>forester-untad.blogspot.com</b>  |            |

8

Internet Source

1 %

9

[www.oppekava.ee](http://www.oppekava.ee)

Internet Source

1 %

10

[makalahjurnalskripsi.com](http://makalahjurnalskripsi.com)

Internet Source

1 %

11

Submitted to Hellenic Open University

Student Paper

1 %

12

[rendymaulanaimron.blogspot.com](http://rendymaulanaimron.blogspot.com)

Internet Source

1 %

13

[mapendademak.org](http://mapendademak.org)

Internet Source

1 %

14

[web.mit.edu](http://web.mit.edu)

Internet Source

1 %

15

[www.irke.ir](http://www.irke.ir)

Internet Source

<1 %

16

Submitted to Colorado State University, Global Campus

Student Paper

<1 %

17

[jtiik.ub.ac.id](http://jtiik.ub.ac.id)

Internet Source

<1 %

18

Submitted to Syiah Kuala University

Student Paper

<1 %

Tania Dian Tri Utami, Agus Perdana Windarto,

- 19 Dedy Hartama, Solikhun Solikhun. "ANALISIS TINGKAT KEPUASAN PELANGGAN TERHADAP PENJUALAN AIR MINUM ISI ULANG DENGAN MENGGUNAKAN METODE ROUGH SET", Jurasik (Jurnal Riset Sistem Informasi dan Teknik Informatika), 2017  
Publication <1 %
- 
- 20 mantik.uinsby.ac.id <1 %  
Internet Source
- 
- 21 sitizujamilah.blogspot.com <1 %  
Internet Source
- 
- 22 repository.wima.ac.id <1 %  
Internet Source
- 
- 23 a-research.upi.edu <1 %  
Internet Source
- 
- 24 deddyztyawan.blogspot.com <1 %  
Internet Source
- 
- 25 eprints.undip.ac.id <1 %  
Internet Source
- 
- 26 repository.unpas.ac.id <1 %  
Internet Source
- 
- 27 www.gunadarma.ac.id <1 %  
Internet Source
- 
- 28 piru.alexandria.ucsb.edu <1 %  
Internet Source

---

29	arisdwisaputra.blogspot.com Internet Source	<1 %
30	ejournal-s1.undip.ac.id Internet Source	<1 %
31	www.kajianpustaka.com Internet Source	<1 %
32	www.eazy-solution.com Internet Source	<1 %
33	samples.jbpub.com Internet Source	<1 %
34	eprints.uns.ac.id Internet Source	<1 %
35	ardillaelfirasafitriipsauny.blogspot.com Internet Source	<1 %
36	irwan06nuklir.wordpress.com Internet Source	<1 %
37	wiki.answers.com Internet Source	<1 %
38	www.zahrasprei.com Internet Source	<1 %
39	raportterbaru.blogspot.com Internet Source	<1 %
40	KALANLAR, Bilge. "Sağlık Eğitiminde Staj	

---

Uygulaması ", Bayt Bilimsel Araştırmalar Basın  
Yayın ve Tanıtım, 2018.

Publication

<1 %

41

portalgaruda.ilkom.unsri.ac.id

Internet Source

<1 %

42

Tasdemir, Mehmet. "A comparison of multiple-choice tests and true-false tests used in evaluating student progress.(Repo", Journal of Instructional Psychology, Sept 2010 Issue

Publication

<1 %

43

Frank Y. Shih, Yi-Ta Wu. "Three-dimensional Euclidean distance transformation and its application to shortest path planning", Pattern Recognition, 2004

Publication

<1 %

Exclude quotes

Off

Exclude matches

Off

Exclude bibliography

Off