

**IMPLEMENTASI ALGORITMA K-MEANS *CLUSTERING*
UNTUK MENGETAHUI BIDANG SKRIPSI MAHASISWA MULTIMEDIA
PENDIDIKAN TEKNIK INFORMATIKA DAN KOMPUTER
UNIVERSITAS NEGERI JAKARTA**

Widodo, Dina Wahyuni

Program Studi S1 Pendidikan Teknik Informatika dan Komputer
Universitas Negeri Jakarta
widodo@unj.ac.id ,dina_wahyuni163@yahoo.co.id

ABSTRAK

Penelitian ini bertujuan untuk mengetahui bidang skripsi mahasiswa peminatan multimedia PTIK UNJ angkatan 2010 yang dapat digunakan sebagai saran bagi mahasiswa yang belum mengajukan skripsi. Sedangkan bagi mahasiswa yang telah mengajukan skripsi dapat dijadikan sebagai perbandingan antara hasil program bantu dengan bidang skripsi yang diteliti. Penelitian dilakukan dengan menggunakan implementasi dari algoritma K-Means *clustering*, yaitu setiap data dikelompokkan berdasarkan jarak minimum terdekat dengan *centroid*. Data yang diolah merupakan rata-rata nilai mata kuliah Proyek Video Digital, Desain Grafis, dan Animasi Komputer untuk *cluster* 1 yaitu bidang video dan animasi. Untuk *cluster* 2 yaitu bidang aplikasi pengembangan perangkat lunak menggunakan rata-rata nilai mata kuliah Algoritma dan Pemrograman, Struktur Data, serta Pemrograman Web. Proses perhitungan menggunakan *software* MATLAB. Input data nilai berjumlah 53 mahasiswa dengan lima kali uji *centroid*. Hasil saran bidang skripsi untuk 19 mahasiswa yang belum mengajukan skripsi adalah 2 mahasiswa di bidang video dan animasi dan 17 mahasiswa di bidang aplikasi pengembangan perangkat lunak. Untuk 34 mahasiswa yang telah mengajukan skripsi, dilakukan perbandingan hasil bidang skripsi dengan menggunakan *sign test*. Berdasarkan hasil uji diketahui $\chi^2_{hitung} < \chi^2_{tabel}$ yaitu terima H_0 dengan kesimpulan tidak terdapat perbedaan antara bidang skripsi hasil perhitungan program bantu algoritma K-Means dengan bidang skripsi yang telah diajukan mahasiswa.

Kata kunci: algoritma K-Means *clustering*, bidang skripsi, mahasiswa

1. PENDAHULUAN

1.1. Latar Belakang

Mahasiswa peminatan multimedia PTIK UNJ angkatan 2010 saat ini sudah mulai mengajukan skripsi sebagai salah satu syarat dalam memperoleh gelar sarjana. Penyusunan skripsi dapat disesuaikan dengan bidang peminatan atau konsentrasi yang telah dipilih mahasiswa. Tahap pertama yang harus dilakukan adalah mencari masalah yang akan diangkat dalam penelitian. Masalah dapat dikembangkan dari bidang skripsi yang diminati mahasiswa. Bidang skripsi dapat diketahui dari pengelompokan data *clustering non hierarchical* yang berusaha mempartisi data yang ada ke

sejumlah nilai mata kuliah yang saling berkaitan satu sama lain. Dengan mengetahui bidang skripsi yang tepat, maka mahasiswa dapat mencari permasalahan seputar bidang skripsi tersebut. Bidang skripsi dalam peminatan multimedia dapat dikelompokkan pada pembuatan video dan animasi serta pembuatan aplikasi pengembangan perangkat lunak.

Penelitian ini didasarkan pada penggunaan metode algoritma K-Means. K-Means merupakan salah satu metode dalam bentuk satu atau lebih *cluster*/kelompok.[2] Dengan diterapkan

algoritma K-Means dalam proses *clustering* bidang skripsi, maka diharapkan dapat mengelompokkan sejumlah *cluster* yang tepat serta dapat mengetahui bidang skripsi mahasiswa multimedia PTIK UNJ angkatan 2010.

1.2. Pembatasan Masalah

Ruang lingkup permasalahan dalam penelitian ini dibatasi hanya pada pengembangan program bantu dengan algoritma K-Means *clustering* untuk mengetahui bidang skripsi mahasiswa multimedia PTIK UNJ angkatan 2010.

1.3. Perumusan Masalah

Bagaimana cara mengimplementasi algoritma K-Means *clustering* untuk mengetahui bidang skripsi mahasiswa multimedia PTIK UNJ angkatan 2010?

1.4. Tujuan Penelitian

- Mengembangkan program bantu untuk mengetahui bidang skripsi mahasiswa multimedia PTIK UNJ angkatan 2010 dengan algoritma K-Means *clustering* menggunakan *software* MATLAB.
- Memberikan informasi bagi mahasiswa yang belum mengajukan skripsi sebagai saran bidang skripsi yang sesuai. Sedangkan bagi mahasiswa yang telah mengajukan skripsi dapat dijadikan sebagai perbandingan antara bidang skripsi hasil program bantu dengan bidang skripsi yang sedang diteliti.

2. KAJIAN TEORI

2.1. Pengertian Algoritma K-Means

K-Means merupakan salah satu metode data *clustering non hierarchical* yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster*/kelompok. Metode ini mempartisi data ke dalam *cluster*/kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama dan data yang

mempunyai karakteristik yang berbeda dikelompokkan ke dalam *cluster* lain. Tujuan dari data *clustering* ini adalah untuk meminimalkan *objective function* yang diset dalam proses *clustering*, yang pada umumnya berusaha meminimalkan variasi di dalam suatu *cluster* dan memaksimalkan variasi antar *cluster*. [2]

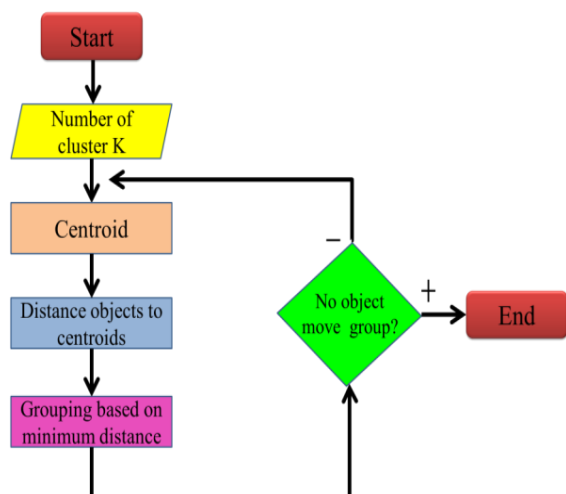
2.2. Tahapan Algoritma K-Means

Berikut ini adalah *statement* dari algoritma K-Means yaitu: [3]

- Pilih nilai k .
- Pilih objek k secara acak. Gunakan sebagai inisial *centroid* k awal.
- Menetapkan setiap objek ke *cluster* terdekat dengan pusat *centroid*.
- Menghitung ulang *centroid* dari *cluster* k .
- Ulangi langkah 3 dan 4 sampai *centroid* tidak lagi bergerak.

Cara kerja Algoritma K-Means (Gambar 1.), dilakukan sampai proses iterasi stabil atau nilai *centroid* tetap: [12]

- Menentukan data *centroid*.
- Menghitung jarak antara *centroid* dengan masing-masing data.
- Mengelompokkan data berdasarkan jarak minimum.



Gambar 1. Flowchart Algoritma K-Means

3. METODOLOGI PENELITIAN

3.1. Tempat dan Waktu Penelitian

Penelitian ini dilaksanakan di Program Studi Pendidikan Teknik Informatika dan Komputer, Jurusan Teknik Elektro, Fakultas Teknik, Universitas Negeri Jakarta. Waktu penelitian dilaksanakan pada perkuliahan semester genap periode 2013/2014.

3.2. Metode Penelitian

Metode penelitian yang digunakan adalah metode eksperimen dengan teknik pengumpulan data dokumentasi. Data penelitian berupa nilai beberapa mata kuliah serta judul skripsi mahasiswa yang telah diajukan, dikumpulkan melalui teknik dokumentasi.

3.3. Teknik Pengambilan Sampel

Teknik pengambilan sampel yang digunakan adalah sampel populasi. Populasi yang diteliti adalah mahasiswa aktif peminatan multimedia Program Studi S1 Pendidikan Teknik Informatika dan Komputer Universitas Negeri Jakarta angkatan 2010 periode 2013/2014.

3.4. Rancangan Penelitian

A. Analisis Kebutuhan

1. Kebutuhan input

- No.registrasi dan nama mahasiswa beserta nilai beberapa mata kuliah yang diperlukan berjenis file Microsoft Excel (.xls atau .xlsx).
- Jumlah *cluster* adalah dua, yaitu bidang skripsi video dan animasi serta bidang skripsi aplikasi pengembangan perangkat lunak.

2. Kebutuhan proses

- Setiap iterasi akan menghasilkan nilai *centroid* baru yang akan digunakan untuk input *centroid* baru pada iterasi selanjutnya.
- Proses *clustering* selesai apabila nilai *centroid* tetap atau tidak lagi berubah.

3. Kebutuhan output

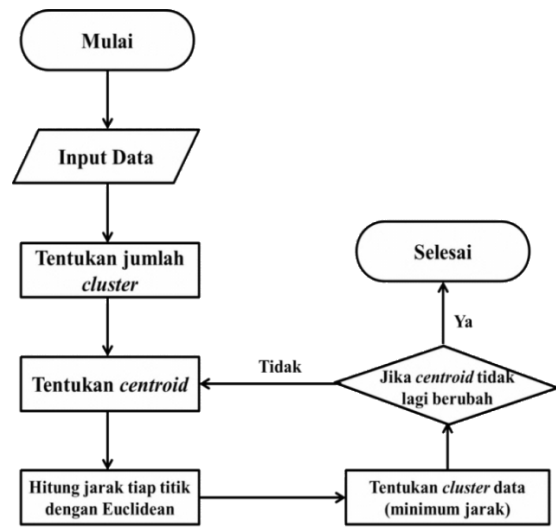
- Menghasilkan plot diagram pada bagian kiri menampilkan pergerakan *centroid* tiap iterasi dan pada bagian kanan

menampilkan titik *centroid* akhir setiap *cluster*.

- Hasil *cluster* bidang skripsi dapat disimpan dalam format (.xls).

B. Perancangan atau Desain

Proses perancangan meliputi *flowchart* program bantu (Gambar 2.) dan antar muka (*user interface*) untuk memudahkan interaksi antara program bantu dengan *user*.



Gambar 2. Flowchart Program Bantu

C. Implementasi

- Jumlah *cluster* (k) yang dihasilkan yaitu 2 untuk bidang skripsi video dan animasi serta aplikasi pengembangan perangkat lunak.
- Pilih *centroid* awal dengan menggunakan *mean* dari data atau dengan input *centroid* secara acak atau random.
- Menghitung setiap objek ke *cluster* terdekat dengan *centroid* awal menggunakan rumus perhitungan jarak yaitu *euclidean distance*.
- Ulangi langkah 3 dan 4 sampai *centroid* tidak lagi berubah.

Program bantu dibuat pada software MATLAB versi 7.11.0.584 (R2010b) 32-bit

D. Pengujian Program Bantu (Testing)

Proses *testing* dalam penelitian meliputi input data nilai beberapa mata kuliah mahasiswa yang akan di *cluster*, serta dilakukan lima kali uji *clustering* dengan input *centroid* secara manual. Program bantu akan menghasilkan dua *cluster* bidang skripsi yang telah ditentukan.

3.5. Instrumen Penelitian

Berdasarkan teknik pengumpulan data dokumentasi, maka instrumen yang digunakan dalam penelitian ini adalah daftar atau tabel.

3.6. Teknik Analisis Data

A. Jenis Data

Jenis data yang digunakan adalah kuantitatif. Nilai mahasiswa dari beberapa mata kuliah termasuk jenis data kuantitatif karena berbentuk angka dan dapat dihitung.

B. Sumber Data

Sumber data yang digunakan adalah data primer. Data nilai mahasiswa diperoleh dari dosen pengampu setiap mata kuliah yang diteliti. Sedangkan data judul skripsi mahasiswa yang telah diajukan diperoleh dari pihak jurusan.

C. Teknik Analisis Data dengan Algoritma K-Means Clustering

Data nilai yang didapat digabungkan dengan mencari nilai rata-rata dari tiga mata kuliah pada setiap bidang skripsi, yaitu nilai rata-rata dari mata kuliah proyek video digital, desain grafis dan animasi komputer untuk variabel (i) serta mata kuliah algoritma&pemrograman, struktur data dan pemrograman web untuk variabel (j).

Nilai yang telah dicari rata-ratanya, selanjutnya dilakukan perhitungan jarak dengan *centroid* awal pada setiap *cluster*. Titik pusat *cluster* atau *centroid* dapat diinputkan secara manual oleh *user* (random) atau secara otomatis dengan menggunakan nilai mean data.

Rumus yang digunakan untuk melakukan perhitungan jarak adalah:

$$d_{(i,j)} = \sqrt{(x_{i1} - x_{j2})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{ip} - x_{jp})^2}$$

dimana (i) adalah nilai untuk bidang skripsi video animasi dan (j) adalah nilai untuk bidang skripsi aplikasi pengembang perangkat lunak.

Setelah *cluster* untuk setiap titik didapat, lakukan kembali perhitungan titik *centroid* dengan rumus perhitungan *centroid* baru:

$$C_{(i)} = \frac{i_1 + i_2 + i_{\dots} + i}{\sum i} C_{(j)} \\ = \frac{j_1 + j_2 + j_{\dots} + j_{\dots}}{\sum j}$$

Lakukan terus proses iterasi sampai titik *centroid* tetap yaitu titik *centroid* baru telah sama dengan titik *centroid* sebelumnya.

D. Teknik Analisis Data dengan Perbandingan Hasil Bidang Skripsi

Bidang skripsi mahasiswa yang dihasilkan dari implementasi algoritma K-Means *clustering* pada MATLAB akan dibandingkan dengan data yang didapat dari pihak jurusan mengenai judul skripsi yang telah diajukan oleh mahasiswa. Analisis perbandingan data akan dilakukan dengan menggunakan *sign test* atau uji tanda karena data yang akan dianalisis dinyatakan dalam bentuk tanda-tanda yaitu tanda positif (+) dan negatif (-).

Data akan dilihat berdasarkan terdapat perbedaan atau tidak antara hasil bidang skripsi. Apabila hasil *cluster* bidang skripsi sama dengan bidang skripsi yang telah diajukan maka diberi tanda positif (+). Sedangkan apabila hasil *cluster* bidang skripsi tidak sama atau berbeda dengan bidang skripsi yang telah diajukan maka diberi tanda negatif (-).

Pengujian dilakukan dengan menggunakan rumus Chi Kuadrat (χ^2) pada taraf signifikansi 5% (0,05). Adapun rumus yang digunakan untuk mencari nilai χ^2 hitung yaitu:

$$\chi^2 = \frac{[(n_1 - n_2) - 1]^2}{n_1 + n_2}$$

Dimana: n_1 = banyak data positif (+)

n_2 = banyak data negatif (-)

Hasil akhir dari perbandingan data bidang skripsi akan dilihat berdasarkan terdapat perbedaan atau tidak antara bidang skripsi yang dihasilkan dari perhitungan program bantu algoritma K-Means *clustering* dengan bidang skripsi yang telah diambil mahasiswa.

3.7. Hipotesis Statistik

Hipotesis statistik penelitian untuk uji bidang skripsi antara hasil perhitungan program bantu algoritma K-Means *clustering* dengan bidang skripsi yang telah diajukan beberapa mahasiswa adalah sebagai berikut:

H_0 : Tidak terdapat perbedaan antara bidang skripsi hasil perhitungan program bantu algoritma K-Means *clustering* dengan bidang skripsi yang telah diajukan mahasiswa.

H_1 : Terdapat perbedaan antara bidang skripsi hasil perhitungan program bantu algoritma K-Means *clustering* dengan bidang skripsi yang telah diajukan mahasiswa.

Dengan kriteria pengujian hipotesis yaitu:

Jika $\chi^2_{hitung} \geq \chi^2_{tabel}$, maka tolak H_0 .

Jika $\chi^2_{hitung} < \chi^2_{tabel}$, maka tolak H_1 .

4. HASIL PENELITIAN DAN PEMBAHASAN

4.1. Hasil Penelitian

A. Penyajian Data

Persiapan data dilakukan secara manual yaitu menggunakan Ms. Excel. Kolom ke-1 merupakan no.registrasi mahasiswa, kolom ke-2 nama mahasiswa, kolom ke-3 nilai 1 yaitu rata-rata nilai dari mata kuliah proyek video digital, desain grafis, dan animasi komputer, serta kolom ke-4 nilai 2 yaitu nilai rata-rata dari mata kuliah algoritma dan pemrograman, struktur data, serta pemrograman web.

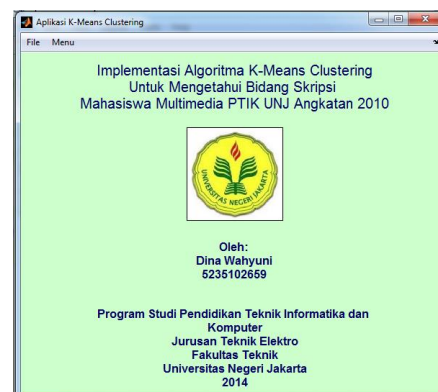
Hasil data yang disajikan setelah mengalami proses awal (*pre-processing*) berjumlah 53 *record* dari jumlah data awal 83 *record*.

B. Implementasi Program Bantu

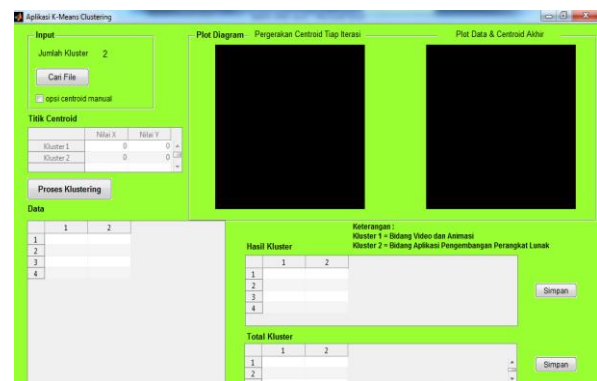
Menjalankan program bantu pada MATLAB dengan mengetikkan (>>AplikasiUtama) di *command window* (Gambar 3.). Buka program K-Means pada menu untuk menampilkan proses perhitungan dengan K-Means *clustering* (Gambar 4.).

Prosedur atau cara kerja yang dilakukan pada program bantu meliputi:

1. Jumlah *cluster* yang dihasilkan 2.
2. Masukkan data yang akan diproses dengan mengklik Cari File, dalam penelitian ini file yang digunakan adalah datanilai_mm.xlsx.
3. Tunggu sampai isi file tampil pada kolom Data (Gambar 5.).



Gambar 3. Tampilan Awal Program Bantu

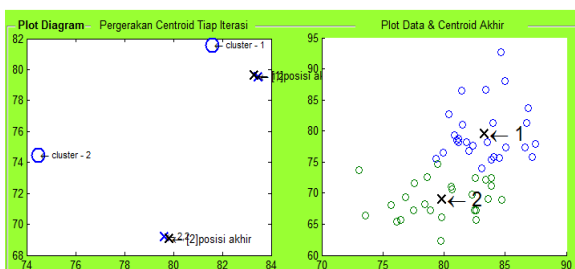


Gambar 4. Tampilan Program K-Means Clustering

4. Masukkan *centroid* awal secara acak dengan pilih opsi *centroid* manual. Jika tidak dipilih, maka program secara otomatis akan melakukan *clustering* dengan *centroid* awal dari nilai *mean* data.
5. Lakukan proses *clustering*, maka akan tampil hasil *clustering* bidang skripsi mahasiswa pada tabel Hasil Kluster. Program bantu menampilkan dua plot diagram hasil *cluster* (pada Gambar 6.), yaitu diagram pertama merupakan pergerakan *centroid* hasil iterasi pada tiap *cluster* dan diagram kedua merupakan plot tiap data beserta *centroid* akhirnya.
6. Hasil dari program bantu *clustering* bidang skripsi mahasiswa ini dapat disimpan dalam format .xls.

Data				
	No.Registrasi	Nama	Nilai 1	Nilai 2
1	5235100221	Immanuel G...	82.6000	67.2580
2	5235100222	Krisna Per...	83.5330	78.2160
3	5235100223	Hendro Sap...	87.5000	77.9360
4	5235100225	Adzy Soewa...	85.0330	77.3500
5	5235100227	Lenie Badhi...	81.4670	86.5580
6	5235100228	Nur Masitoh...	83.9670	81.3030
7	5235102652	Ayu Hardianti	84.0670	75.8190
8	5235102653	Arini Amalia	86.6000	77.3580
9	5235102655	Fajar Maula...	87.2330	75.8720
10	5235102656	Esti Putri Pr...	81.2000	78.7940
11	5235102657	Vannisa Irm...	79.9330	76.4930
12	5235102658	Melyanah	80.8670	79.3200
13	5235102659	Dina Wahyuni	81.5330	81.0640

Gambar 5. Data yang Akan di Cluster



Gambar 6. Diagram Hasil Cluster

Proses Clustering dengan Algoritma K-Means

Langkah-langkah algoritma K-Means pada proses *clustering* datanilai_mm.xlsx meliputi:

1. Menentukan *centroid* awal sebagai pusat tiap *cluster*. Dalam perhitungan ini menggunakan nilai *mean* dari data.
 $Centroid\ cluster\ 1 = (81,592 ; 81,592)$

2. Hitung jarak tiap data terhadap *centroid* awal pada tiap *cluster*. Hasil *cluster* pada tiap data akan dilihat berdasarkan kedekatan atau nilai jarak terkecil terhadap pusat *cluster*.
3. Setelah melakukan perhitungan jarak, hitung kembali *centroid* dengan data *cluster* yang baru yaitu merupakan hasil rata-rata dari semua nilai yang terdapat pada setiap *cluster*.
4. Ulangi langkah 1-3 hingga pusat *cluster* berada pada posisi tetap. Dalam program bantu ini proses *clustering* dikatakan selesai apabila titik *centroid* akhir telah sama dengan titik *centroid* pada iterasi sebelumnya.
5. Total *cluster* yang dihasilkan pada program bantu K-Means *clustering*:
 - a. Mahasiswa dengan bidang skripsi video dan animasi berjumlah 27 mahasiswa pada *cluster* 1 dengan nilai *centroid* akhir yaitu (83,305 ; 79,677).
 - b. Mahasiswa dengan bidang skripsi aplikasi pengembangan perangkat lunak berjumlah 26 mahasiswa pada *cluster* 2 dengan nilai *centroid* akhir yaitu (79,814 ; 69,092).

C. Uji Hasil Clustering dengan Titik Centroid yang Berbeda-Beda

1. Uji Pertama Menggunakan *Centroid* Nilai *Mean* Data (telah dilakukan pada proses *clustering* di atas)
2. Uji Kedua Menggunakan *Centroid* Nilai Tertinggi dan Terendah Data
 $Centroid\ cluster\ 1 = 87,500 ; 92,630$ (nilai tertinggi)
 $Centroid\ cluster\ 2 = 73,070 ; 63,350$ (nilai terendah)
Hasil uji kedua yaitu *cluster* 1 untuk bidang skripsi video dan animasi berjumlah 26 mahasiswa dengan *centroid* akhir (83,313 ; 79,898) serta *cluster* 2 untuk bidang skripsi aplikasi pengembangan perangkat lunak berjumlah 27 mahasiswa dengan *centroid* akhir (79,936 ; 69,272).

3. Uji Ketiga Menggunakan *Centroid* Nilai Terendah dan Tertinggi Data
Centroid cluster 1 = 73,070 ; 63,350 (nilai terendah)
Centroid cluster 2 = 87,500 ; 92,630 (nilai tertinggi)
 Hasil uji ketiga yaitu *cluster 1* untuk bidang skripsi video dan animasi berjumlah 27 mahasiswa dengan *centroid* akhir (79,936 ; 69,272) serta *cluster 2* untuk bidang skripsi aplikasi pengembangan perangkat lunak berjumlah 26 mahasiswa dengan *centroid* akhir (83,313 ; 79,898).
4. Uji Keempat Menggunakan *Centroid* Nilai Acak (1)
Centroid cluster 1 = 79 ; 75
Centroid cluster 2 = 83 ; 74
 Hasil uji keempat yaitu *cluster 1* untuk bidang skripsi video dan animasi berjumlah 26 mahasiswa dengan *centroid* akhir (83,313 ; 79,898) serta *cluster 2* untuk bidang skripsi aplikasi pengembangan perangkat lunak berjumlah 27 mahasiswa dengan *centroid* akhir (79,936 ; 69,272).
5. Uji Kelima Menggunakan *Centroid* Nilai Acak (2)
Centroid cluster 1 = 65 ; 70
Centroid cluster 2 = 66 ; 73
 Hasil uji kelima, yaitu *cluster 1* untuk bidang skripsi video dan animasi berjumlah 26 mahasiswa dengan *centroid* akhir (79,814 ; 69,092) serta *cluster 2* untuk bidang skripsi aplikasi pengembangan perangkat lunak berjumlah 27 mahasiswa dengan *centroid* akhir (83,305 ; 79,677).

D. Analisis Hasil Uji *Clustering* dengan Data Sebenarnya

1. Perbandingan Hasil Uji Pertama dengan Data Sebenarnya
 Terdapat tanda (+) sebanyak 14 dan (-) sebanyak 20. Hasil uji tanda:

$$x^2 = \frac{[(14 - 20) - 1]^2}{14 + 20} = \frac{49}{34} = 1,441$$
 Nilai χ^2_{hitung} hasil perbandingan data pada uji pertama adalah 1,441.

2. Perbandingan Hasil Uji Kedua dengan Data Sebenarnya
 Terdapat tanda (+) sebanyak 15 dan (-) sebanyak 19. Hasil uji tanda:

$$x^2 = \frac{[(15 - 19) - 1]^2}{15 + 19} = \frac{25}{34} = 0,735$$
 Nilai χ^2_{hitung} hasil perbandingan data pada uji kedua adalah 0,735.
3. Perbandingan Hasil Uji Ketiga dengan Data Sebenarnya
 Terdapat tanda (+) sebanyak 19 dan (-) sebanyak 15. Hasil uji tanda:

$$x^2 = \frac{[(19 - 15) - 1]^2}{19 + 15} = \frac{9}{34} = 0,265$$
 Nilai χ^2_{hitung} hasil perbandingan data pada uji ketiga adalah 0,265.
4. Perbandingan Hasil Uji Keempat dengan Data Sebenarnya
 Terdapat tanda (+) sebanyak 15 dan (-) sebanyak 19. Hasil uji tanda:

$$x^2 = \frac{[(15 - 19) - 1]^2}{15 + 19} = \frac{25}{34} = 0,735$$
 Nilai χ^2_{hitung} hasil perbandingan data pada uji keempat adalah 0,735.
5. Perbandingan Hasil Uji Kelima dengan Data Sebenarnya
 Terdapat tanda (+) sebanyak 20 dan (-) sebanyak 14. Hasil uji:

$$x^2 = \frac{[(20 - 14) - 1]^2}{20 + 14} = \frac{25}{34} = 0,735$$
 Nilai χ^2_{hitung} hasil perbandingan data pada uji kelima adalah 0,735.

4.2. Pembahasan

A. Hasil Saran Bidang Skripsi Mahasiswa

Berdasarkan uji *clustering* yang telah dilakukan didapatkan hasil akhir bidang skripsi yang disimpulkan dengan banyaknya hasil *cluster* yang sama pada setiap uji pada Tabel 1. Adapun rincian dari setiap *cluster* adalah:

- a. Bidang skripsi video dan animasi berjumlah dua mahasiswa dengan nilai pusat *cluster* (83,305 ; 79,677) dan (83,313 ; 79,898).
- b. Bidang skripsi aplikasi pengembangan perangkat lunak berjumlah 17 mahasiswa dengan nilai pusat *cluster*

(79,814 ; 69,092) dan (79,936 ; 69,272).

Tabel 1. Hasil Bidang Skripsi Mahasiswa Peminatan Multimedia

No.	No. Reg	Nama	Hasil Bidang Skripsi Tiap Uji <i>Clustering</i>					Hasil Saran Bidang Skripsi
			1	2	3	4	5	
1	5235107400	Khaulah Syahidah	1	1	2	1	2	1
2	5235108578	Hirzan Jundi Al Hak	1	1	2	1	2	1
3	5235100221	Immanuel Gesima R	2	2	1	2	1	2
4	5235102661	Riko Yulian	2	2	1	2	1	2
5	5235107369	Adi Putro Hastariyad	2	2	1	2	1	2
6	5235107376	Dwi Hartato	2	2	1	2	1	2
7	5235107380	M Immanudin AR	2	2	1	2	1	2
8	5235107388	Rizky Istiyono	2	2	1	2	1	2
9	5235107392	Mahfud Ramdhani A	2	2	1	2	1	2
10	5235107395	Bima Samudra	2	2	1	2	1	2
11	5235107398	Trisna Yudha	2	2	1	2	1	2
12	5235107401	Bagus Banyu Aji G	2	2	1	2	1	2
13	5235107415	Fauzia Resty M	2	2	1	2	1	2
14	5235107416	M Lomo Niroha	2	2	1	2	1	2
15	5235107418	Rina Rahmawaty	2	2	1	2	1	2
16	5235107421	Teguh Adin N	2	2	1	2	1	2
17	5235107423	M Fajar Siddiq	2	2	1	2	1	2
18	5235107428	Fahmi Yudaprawira	2	2	1	2	1	2
19	5235107429	Reza Syarif Alfath	2	2	1	2	1	2

Keterangan:

Cluster 1 adalah bidang skripsi video dan animasi

Cluster 2 adalah bidang skripsi aplikasi pengembangan perangkat lunak

B. Hasil Perbandingan Bidang Skripsi Mahasiswa

Bagi mahasiswa peminatan multimedia yang telah mengajukan judul skripsi, maka hasil *cluster* bidang skripsi yang didapat dari program bantu algoritma K-Means *clustering* dibandingkan dengan bidang skripsi yang telah diambil mahasiswa. Hasil rata-rata χ^2 hitung dari lima kali uji perbandingan adalah:

$$\chi^2_{hitung} = \frac{1,441 + 0,735 + 0,265 + 0,735 + 0,735}{5} = 0,782$$

Berdasarkan nilai $dk = 1$ pada taraf signifikansi 5%, maka nilai $\chi^2_{tabel} = 3,841$. Perbandingan antara χ^2_{hitung} dengan χ^2_{tabel} yaitu:

$\chi^2_{hitung} (0,782) < \chi^2_{tabel} (3,841)$, maka sesuai dengan kriteria pengujian hipotesis yaitu terima H_0 dengan kesimpulan tidak terdapat perbedaan antara bidang skripsi hasil perhitungan program bantu algoritma K-Means *clustering* dengan bidang skripsi yang telah diajukan mahasiswa.

5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan dapat disimpulkan:

1. Program bantu yang dikembangkan dari alur kerja implementasi algoritma K-Means dapat digunakan untuk

mengetahui pengelompokan bidang skripsi mahasiswa multimedia PTIK UNJ angkatan 2010 yaitu bidang video dan animasi serta bidang aplikasi pengembangan perangkat lunak.

2. Penentuan *centroid* awal pada proses *clustering* dapat mempengaruhi proses iterasi. Semakin dekat nilai *centroid* awal dengan nilai mean data, maka proses iterasi yang dihasilkan sedikit. Sebaliknya jika nilai *centroid* awal jauh dari nilai mean data, maka iterasi yang dihasilkan lebih banyak.
3. Saran bidang skripsi dari hasil program bantu algoritma K-Means clustering untuk 19 mahasiswa yang belum mengajukan skripsi adalah:
Cluster 1 = bidang skripsi video dan animasi berjumlah 2 mahasiswa dengan centroid akhir (83,305 ; 79,677) dan (83,313 ; 79,898).
Cluster 2 = bidang skripsi aplikasi pengembangan perangkat lunak berjumlah 17 mahasiswa dengan centroid akhir (79,814 ; 69,092) dan (79,936 ; 69,272).
4. Hasil dari uji sampel pada 34 mahasiswa yang telah mengajukan skripsi menunjukkan terdapat perbedaan antara bidang skripsi hasil perhitungan program bantu algoritma K-Means *clustering* dengan bidang skripsi yang telah diajukan mahasiswa. Hal ini dapat berarti bahwa proses *clustering* dengan algoritma K-Means dapat diterapkan untuk kasus seperti ini.
5. Implementasi dari algoritma K-Means *clustering* kurang cocok jika digunakan dalam pengelompokan yang berkaitan dengan kompetensi. Algoritma K-Means mengelompokan data berdasarkan rata-rata yang dimiliki pada setiap kelompok data, sedangkan pengelompokan yang berkaitan dengan kompetensi melihat data berdasarkan nilai tertinggi dan nilai terendah yang terdapat pada setiap kelompok tersebut.

5.2. Saran

Berdasarkan hasil penelitian yang telah disimpulkan, maka dapat dikemukakan beberapa saran untuk penelitian selanjutnya yaitu:

1. Pemberian kuesioner kepada mahasiswa mengenai bidang skripsi yang diminati, agar dapat digunakan sebagai perbandingan dengan prediksi bidang skripsi yang dihasilkan dari program bantu algoritma K-Means.
2. Penambahan *cluster* bidang skripsi peminatan multimedia menjadi 3 yaitu dalam bidang multimedia, bidang pendidikan, dan bidang aplikasi pengembangan perangkat lunak.
3. Disarankan agar dilakukan penelitian lebih lanjut dengan menggunakan metode atau algoritma lain dalam proses pengelompokan yang berkaitan dengan minat dan kompetensi mahasiswa.

6. DAFTAR PUSTAKA

- [1] Abudalfa, Shadi I dan Mikki, Mohammad. 2013. "K-Means Algorithm With A Novel Distance Measure", *Turkish Journal of Electrical Engineering & Computer Sciences* 21:1665-1684
- [2] Agusta, Yudi. 2007. "K-Means - Penerapan, Permasalahan dan Metode Terkait". *Jurnal Sistem dan Informatika* 3:47-60
- [3] Bramer, Max. 2007. *Principles of Data Mining*. London: Springer
- [4] Binanto, Iwan. 2010. *Multimedia Digital – Dasar Teori dan Pengembangannya*. Yogyakarta: Andi
- [5] Deskripsi Mata Kuliah. CD *Pedoman Akademik UNJ 2010/2011*. Fakultas Teknik, Jurusan Teknik Elektro, Program Studi Pendidikan Teknik Informatika dan Komputer (S1). File tik_isi
- [6] Design College, Yokohama. *Jurusan Video dan Animasi*.

- http://www.ydc.ac.jp/indonesia/course/v_video.html, [diakses 13 Juni 2014 pukul 13:40]
- [7] Fakultas Teknik, Jurusan Teknik Elektro. *SI Pend Teknik Informatika & Komputer*. <http://www.unj.ac.id/ft/jurusan/elektr o/content/s1-pend-teknik-informatika-komputer> [diakses 14 Maret 2014 pukul 00:39]
- [8] Gilat, Amos. 2011. *Matlab An Introduction with Applications Fourth Edition*. USA: John Wiley & Sons, Inc
- [9] Han, Jiawei; Kamber, Micheline; dan Pei, Jian. 2012. *Data Mining Concepts and Techniques, 3rd ed.* USA: Morgan Kaufmann Publishers
- [10] Hariyadi, Teguh dan Suprayogi. 2009. "Penerapan Algoritma K-Means Untuk Pengelompokan Data Nilai Siswa" https://www.academia.edu/5266790/Penerapan_Algoritma_K-Means_Untuk_Pengelompokan_Data_Nilai_Siswa [diakses 19 April 2014, pukul 20:30]
- [11] Hermawati, Fajar A. 2013. *Data Mining*. Yogyakarta: Andi Offset
- [12] Nango, Dwi N. 2012. *Penerapan Algoritma K-Means Untuk Clustering Data Anggaran Pendapatan Belanja Daerah di Kabupaten XYZ [Skripsi]*. Gorontalo: Jurusan Teknik Informatika, Program Studi Sistem Informasi, Fakultas Teknik, Universitas Negeri Gorontalo.
- [13] Narwati. 2010. "Pengelompokan Mahasiswa Menggunakan Algoritma K-Means". *Jurnal Dinamika Informatika* 2(2) <http://www.unisbank.ac.id/ojs/index.php/fti2/article/view/890/448> [diakses 20 Maret 2014 pukul 00:30]
- [14] Nurwakhida. *Clustering Menggunakan K-Means Algorithm (K-Means Algorithm Clustering)*. Fakultas Teknologi Informasi dan Komunikasi Universitas Semarang. <http://digilib.usm.ac.id/files/disk1/4/gdl-usm--nurwakhida-176-1-k-means-a.pdf> [diakses 12 Mei 2014 pukul 19:30]
- [15] Sugiyono. 2008. *Statistik Nonparametris Untuk Penelitian*. Bandung: Alfabeta
- [16] Suyanto, M. 2003. *Multimedia Alat Untuk Meningkatkan Keunggulan Bersaing*. Yogyakarta: Andi
- [17] Tan, P.N.; Steinbach, M.; Kumar, V. 2006. *Introduction to Data Mining*. USA: Pearson Education
- [18] Tim Penyusun. 2010. *Pedoman Akademik 2010/2011 Universitas Negeri Jakarta*. Jakarta: Kementerian Pendidikan Nasional Universitas Negeri Jakarta
- [19] Tim Penyusun. 2012. *Buku Pedoman Skripsi / Komprehensif / Karya Inovatif (SI)*. Jakarta: Gedung L Fakultas Teknik Kampus A Universitas Negeri Jakarta
- [20] Wahyudi, N.E.; Diartono, A.D.; dan Sulastri. 2010. "Analisa Profil Data Mahasiswa Baru Universitas Stikubank (UNISBANK) Semarang Tahun 2005-2010 Dengan Teknik Data Mining", *Jurnal Dinamika Informatika* 2(2) <http://www.unisbank.ac.id/ojs/index.php/fti2/article/view/895/454> [diakses 20 Maret 2014 pukul 00:40]
- [21] Wikipedia. *Pengembangan Perangkat Lunak*. http://id.wikipedia.org/wiki/Pengembangan_perangkat_lunak [diakses 13 Juni 2014, pukul 15:20]