

## **PRAKATA**

### **Bismillahirrahmanirrahim**

Alhamdulillahirabbil'alamin puji syukur kehadiran Allah SWT yang telah memberikan nikmat dan hidayah-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi dengan judul

**PENERAPAN METODE *CLUSTERING* DENGAN ALGORITMA *K-MEANS*  
PADA PENGELOMPOKAN DATA CALON MAHASISWA BARU DI  
UNIVERSITAS MUHAMMADIYAH YOGYAKARTA  
(STUDI KASUS: FAKULTAS KEDOKTERAN DAN ILMU KEPERAWATAN  
DAN FAKULTAS ILMU SOSIAL DAN ILMU POLITIK)**

Berbagai usaha dan upaya telah penulis lakukan untuk menyelesaikan penyusunan skripsi ini, karena keterbatasan kemampuan penulis, maka penulis meminta maaf apabila terdapat kesalahan dalam pemilihan kata, sistemika pembahasan dan penyusunan penulisan. Harapannya skripsi ini dapat memberikan manfaat bagi penulis khususnya dan pembaca pada umumnya.

Tentunya dalam proses penyusunan dan penulisan skripsi ini tidak terlepas dari bantuan, bimbingan serta dukungan dari berbagai pihak, penulis mengucapkan terima kasih diantaranya:

1. Bapak Helmi Zain Nuri, S.T., MT. selaku Ketua Program Studi Teknik Informatika Universitas Muhammadiyah Yogyakarta.

2. Bapak Ir. Eko Prasetyo, M.Eng. selaku dosen pembimbing I yang dengan sabar dalam membimbing, memberikan ilmu dan mengarahkan penulis sehingga selama melaksanakan penelitian hingga menyelesaikan penyusunan skripsi ini.
3. Bapak Asroni, S.T., M.Eng. selaku dosen pembimbing II yang telah memberikan ilmu dan pengetahuan, selalu menyemangati penulis dalam proses penyusunan skripsi.
4. Ibu Aprilia Kurnianti, S.T., M.Eng. selaku dosen penguji ujian pendadaran skripsi dengan senantiasa memberikan saran demi perkembangan ilmu pengetahuan serta perkembangan dunia pendidikan.
5. Kedua orangtua yang selalu mendoakan dan menyemangati penulis sejak awal perkuliahan hingga pada penyusunan skripsi.
6. Seluruh dosen dan staff program studi Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Yogyakarta.
7. Teman-teman angkatan 2012 program studi Teknik Informatika, khususnya kelas A yang mendorong penulis dalam penyelesaian penelitian skripsi ini.
8. Kakak – kakak dan adik – adik program studi Teknik Informatika.
9. Semua pihak yang telah membantu penulis dalam penyelesaian penelitian skripsi ini.

Penulis menyadari bahwa skripsi ini memiliki kekurangan, maka penulis mengharapkan kritik dan saran yang bersifat membangun sehingga menjadi masukan bagi penulis guna penyusunan karya-karya selanjutnya.

Yogyakarta, 11 Desember 2016

Penulis

## **PERNYATAAN**

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip dalam naskah ini dan disebutkan dalam daftar perpustakaan.

Yogyakarta, 11 Desember 2016

Hidayatul Fitri

## DAFTAR ISI

<b>HALAMAN PENGESAHAN I</b> .....	I
<b>HALAMAN PENGESAHAN II</b> .....	II
<b>PRAKATA</b> .....	III
<b>PERNYATAAN</b> .....	VI
<b>DAFTAR ISI</b> .....	VII
<b>DAFTAR GAMBAR</b> .....	IX
<b>DAFTAR TABEL</b> .....	XI
<b>DAFTAR LAMPIRAN</b> .....	XII
<b>INTISARI</b> .....	XIII
<b>ABSTRACT</b> .....	XIV
<b>BAB I PENDAHULUAN</b> .....	1
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah.....	3
1.3. Tujuan Penelitian .....	3
1.4. Batasan Masalah .....	3
1.5. Manfaat Penelitian.....	4
<b>BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI</b> .....	5
2.1. Tinjauan Pustaka.....	5
2.2. Landasan Teori.....	8
<b>BAB III METODOLOGI</b> .....	14
3.1 Prosedur Penelitian .....	14
3.1.1 Alur Penelitian.....	14
3.1.2 <i>Flowchart K-Means</i> .....	17
3.2 Alat dan Bahan.....	18
3.3 Lokasi Penelitian.....	19
3.4 Prosedur Validasi .....	19
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....	21
4.1 Pengumpulan Data.....	21

4.2 Penyeleksian Data.....	21
4.3 <i>Preprocessing</i> Data.....	23
4.4 Transformasi Data.....	24
4.5 Pengolahan Data.....	30
4.5.1 Pengujian dengan <i>Software WEKA</i> .....	32
4.5.2 Algoritma <i>K-Means</i> .....	56
<b>BAB V KESIMPULAN DAN SARAN</b> .....	<b>66</b>
5.1 Kesimpulan.....	66
5.2 Saran.....	67
<b>DAFTAR PUSTAKA</b> .....	<b>68</b>
<b>LAMPIRAN</b> .....	<b>70</b>

## DAFTAR GAMBAR

<b>Gambar 2. 1</b> Proses <i>data mining</i> (Fina Nasari, 2015). .....	10
<b>Gambar 3. 1</b> <i>Flowchart</i> alur penelitian.....	14
<b>Gambar 3. 2</b> <i>Flowchart</i> algoritma <i>k-means</i> .....	17
<b>Gambar 4. 1</b> Tampilan <i>view</i> dari <i>data warehouse</i> . .....	22
<b>Gambar 4. 2</b> <i>Filter</i> data dari <i>data warehouse</i> . .....	23
<b>Gambar 4. 3</b> Data penmaru FKIK format .CSV. ....	31
<b>Gambar 4. 4</b> Data penmaru FISIPOL format .CSV.....	32
<b>Gambar 4. 5</b> <i>WEKA explorer</i> . .....	33
<b>Gambar 4. 6</b> Data yang dianalisa berdasarkan atribut di <i>WEKA</i> pada FKIK. ....	34
<b>Gambar 4. 7</b> <i>Visualize</i> data berdasarkan <i>class</i> jurusan FKIK. ....	35
<b>Gambar 4. 8</b> <i>Visualize</i> dari 4 atribut menggunakan <i>WEKA</i> pada FKIK.....	36
<b>Gambar 4. 9</b> Atribut kota asal dari <i>visualize WEKA</i> pada FKIK.....	37
<b>Gambar 4. 10</b> Atribut jenis kelamin dari <i>visualize WEKA</i> pada FKIK. ....	38
<b>Gambar 4. 11</b> Atribut status lolos seleksi dari <i>visualize WEKA</i> pada FKIK. ....	39
<b>Gambar 4. 12</b> Atribut jalur masuk dari <i>visualize WEKA</i> pada FKIK. ....	40
<b>Gambar 4. 13</b> Tampilan <i>WEKA</i> metode <i>clustering</i> . .....	41
<b>Gambar 4. 14</b> <i>Cluster</i> data menggunakan algoritma <i>k-means</i> FKIK. ....	42
<b>Gambar 4. 15</b> <i>Cluster centroid</i> pada <i>WEKA</i> FKIK. ....	43
<b>Gambar 4. 16</b> Hasil <i>cluster centroids</i> dan <i>clustered Instances</i> dengan <i>WEKA</i> pada FKIK.....	43
<b>Gambar 4. 17</b> Grafik <i>clustering</i> FKIK.....	44

<b>Gambar 4. 18</b>	Data yang dianalisa berdasarkan atribut di <i>WEKA</i> pada FISIPOL. ....	46
<b>Gambar 4. 19</b>	<i>Visualize</i> data berdasarkan <i>class</i> jurusan FISIPOL. ....	47
<b>Gambar 4. 20</b>	Atribut kota asal dari <i>visualize WEKA</i> pada FISIPOL. ....	49
<b>Gambar 4. 21</b>	Atribut jenis kelamin dari <i>visualize WEKA</i> pada FISIPOL. ....	50
<b>Gambar 4. 22</b>	Atribut status lolos seleksi dari <i>visualize WEKA</i> pada FISIPOL. ....	51
<b>Gambar 4. 23</b>	Atribut jalur masuk dari <i>visualize WEKA</i> pada FISIPOL. ....	52
<b>Gambar 4. 24</b>	<i>Cluster</i> data menggunakan algoritma <i>k-means</i> FISIPOL. ....	53
<b>Gambar 4. 25</b>	Hasil <i>cluster centroids</i> dan <i>clustered Instances</i> dengan <i>WEKA</i> pada FISIPOL. ....	54
<b>Gambar 4. 26</b>	Grafik <i>clustering</i> FISIPOL. ....	55



## DAFTAR TABEL

<b>Tabel 4. 1</b> Data pada atribut kota asal pada FKIK. ....	24
<b>Tabel 4. 2</b> Data pada atribut kota asal pada FISIPOL. ....	26
<b>Tabel 4. 3</b> Data pada atribut status lolos seleksi pada FKIK.....	28
<b>Tabel 4. 4</b> Data pada atribut status lolos seleksi pada FISIPOL. ....	28
<b>Tabel 4. 5</b> Data pada atribut jalur masuk pada FKIK.....	29
<b>Tabel 4. 6</b> Data pada atribut jalur masuk pada FISIPOL. ....	29
<b>Tabel 4. 7</b> Data pada atribut jurusan pada FKIK.....	30
<b>Tabel 4. 8</b> Data pada atribut jurusan pada FISIPOL. ....	30
<b>Tabel 4. 9</b> Penentuan frekuensi dari data dalam menentukan jarak <i>centroid</i> pada FKIK.....	57
<b>Tabel 4. 10</b> Jarak antara <i>centroid</i> pada FKIK.....	58
<b>Tabel 4. 11</b> Pengelompokkan <i>centroid</i> pada FKIK.....	59
<b>Tabel 4. 12</b> Penentuan frekuensi dari data dalam menentukan jarak <i>centroid</i> pada FISIPOL. ....	62
<b>Tabel 4. 13</b> Jarak antara <i>centroid</i> pada FISIPOL. ....	62
<b>Tabel 4. 14</b> Pengelompokkan <i>centroid</i> pada FISIPOL.....	64

## DAFTAR LAMPIRAN

<b>Lampiran 1.</b> Data sumber dari <i>data warehouse Factual</i> Penmaru.....	70
<b>Lampiran 2.</b> Data Penmaru format CSV .....	72
<b>Lampiran 3.</b> Transformasi data .....	74
<b>Lampiran 4.</b> Inisiasi data nominal ke <i>numeric</i> .....	75
<b>Lampiran 5.</b> Jarak menggunakan rumus <i>euclidean distance</i> .....	76
<b>Lampiran 6.</b> Pengelompokkan data berdasarkan <i>centroid</i> terdekat .....	78
<b>Lampiran 7.</b> <i>Visualize</i> data pada <i>WEKA</i> .....	80
<b>Lampiran 8.</b> Grafik berdasarkan <i>cluster</i> .....	81