

**MODEL ALGORITMA K-MEANS CLUSTERING UNTUK
MENENTUKAN STRATEGI PROMOSI UNIVERSITAS
ISLAM NAHDLATUL ULAMA JEPARA**



SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Strata 1 (S.1) Program Studi Teknik Informatika
Fakultas Sains dan Teknoogi
Universitas Islam Nahdlatul Ulama Jepara

Oleh :

Damas Fahmi Assena
NIM : 161240000500

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NAHDLATUL ULAMA (UNISNU)
JEPARA
2020**

PERSETUJUAN PEMBIMBING

Assalamu'alaikum Wr. Wb

Setelah kami meneliti dan mengadakan perbaikan seperlunya, Bersama ini saya kirim naskah skripsi Saudara :

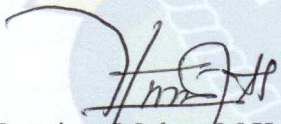
Nama : Damas Fahmi Assena
Nim : 161240000500
Program Studi : Teknik Informatika
Judul : Model Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Universitas Islam Nahdlatul Ulama Jepara

Skripsi ini telah disetujui pembimbing dan siap untuk dipertahankan dihadapan tim penguji program Sarjana Strata 1 (SI) Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama (Unisnu) Jepara.


Demikian harap menjadikan maklum.

Wassalamu'alaikum Wr. Wb.

Pembimbing I


Harminto Mulyo, M.Kom
NIDN. 0604028203

Jepara, 17 September 2020
Pembimbing II


Nur Aeni Widiastuti, S.Pd., M.Kom
NIDN. 0602078702

Mengetahui,
Ketua Program Studi Teknik Informatika


Akhmad Khanif Zyen, M.Kom
NIDN. 0621048602

PENGESAHAN SKRIPSI

Skripsi dengan judul “Model Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Universitas Islam Nahdlatul Ulama Jepara” karya :

Nama : Damas Fahmi Assena
Nim : 161240000500
Program Studi : Teknik Informatika

Telah diujikan dan dipertahankan dalam sidang oleh Dewan Penguji Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama Jepara dan dinyatakan lulus pada tanggal 25 September 2020.

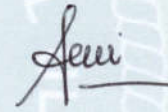
Selanjutnya dapat diterima sebagai syarat guna memperoleh gelar Sarjana Strata 1 (SI) Program Studi Teknik Informatika pada Fakultas Sains dan Teknologi UNISNU Jepara Tahun Akademik 2020/2021.

Jepara, 25 September 2020
Sekretaris Sidang

Ketua Sidang



Harminto Mulyo, M.Kom
NIDN. 0604028203



Nur Aeni Widiastuti, S.Pd., M.Kom
NIDN. 0602078702

Penguji I,




Teguh Tamrin, M.Kom
NIDN. 0620127603

Penguji II,



Buang Budi Wahono, S.Si., M.Kom
NIDN. 0603087802

Mengetahui,
Dekan Fakultas Sains Dan Teknologi
Unisnu Jepara


Ir. Gun Sudiryanto, M.M
NIDN. 0624056501

PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini :

Nama : Damas Fahmi Assena

Nim : 161240000500

Program Studi : Teknik Informatika

Saya menyatakan dengan penuh kejujuran dan tanggung jawab. Bahwa Skripsi yang saya susun sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata 1 (S1) Universitas Islam Nahdlatul Ulama (UNISNU) Jepara merupakan hasil karya saya sendiri dan belum pernah diajukan sebagai pemenuhan persyaratan untuk memperoleh gelar Sarjana dari Perguruan Tinggi lain.

Adapun bagian - bagian tertentu dalam penulisan Skripsi yang saya kutip dari karya orang lain telah dituliskan dengan jelas dengan norma, kaidah, dan etika penulisan ilmiah. Jepara, 17 September 2020.

Jepara, 17 September 2020



Damas Fahmi Assena

NIM. 161240000500

PERSEMBAHAN

Dengan memanjatkan puji syukur kehadiran Allah Subhanahu Wa Ta'ala, karya ilmiah ini penulis persembahkan kepada:

1. Orangtua saya tercinta Bapak Asmudi Arva dan Ibu Kasniyati yang telah memberikan kasih sayang, perhatian, dan Do'a yang tak pernah putus.
2. Istri saya Rahma Nur Fadlila terima kasih atas kebersamaan dan dukungan untuk terus berjuang mengarungi kehidupan ini.

MOTTO

“Dan Dia memberinya rezeki dari arah yang tidak disangka-sangkanya. Dan barang siapa bertawakal kepada Allah, niscaya Allah akan mencukupkan (keperluannya)nya. Sesungguhnya Allah melaksanakan urusan-Nya. Sungguh, Allah telah mengadakan ketentuan bagi setiap sesuatu”.

(Qs. At-Talaq: 3)

“Dan apabila hamba-hamba-Ku bertanya kepadamu (muhammad) tentang Aku, maka sesungguhnya Aku adalah dekat. Aku mengabulkan permohonan orang yang berdo'a apabila ia memohon kepada-Ku. Hendaklah mereka itu memenuhi (perintah)-Ku dan beriman kepada-Ku, agar mereka memperoleh kebenaran”.

(Qs. Al-Baqarah: 186)

“Baginya (manusia) ada malaikat-malaikat yang selalu menjaganya bergiliran, dari depan dan belakangnya. Dan mereka menjaganya atas perintah Allah. Sesungguhnya Allah tidak akan mengubah keadaan suatu kaum sebelum mereka mengubah keadaan diri mereka sendiri. Dan apabila Allah menghendaki keburukan terhadap suatu kaum, maka tidak ada yang dapat menolaknya dan tidak ada pelindung bagi mereka selain Dia”.

(Qs. Ar-Ra'ad: 11)

“Dan milik Allah-lah kerajaan langit dan bumi; dan Allah maha kuasa atas segala sesuatu”.

(Qs. Ali-Imron: 189)

KATA PENGANTAR

Segala puji syukur kehadirat Allah SWT atas limpahan rahmat, hidayah serta karunia yang tak terhingga kepada kita semua. Sholawat serta salam senantiasa kita haturkan kepada junjungan Nabi Agung Muhammad SAW sebagai *uswatun hasanah* yang telah membawa dunia dari jaman *jahiliyyah* menuju jaman *islamiyyah*. Pada kesempatan ini, penulis mengucapkan terima kasih sebesar-besarnya kepada semua pihak yang telah membantu penulis dalam menyelesaikan skripsi dengan judul “Model Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Universitas Islam Nahdlatul Ulama Jepara”. Untuk itu penulis mengucapkan terima kasih kepada:

1. Rektor Universitas Islam Nahdlatul Ulama Jepara
2. Bapak Ir. Gun Sudiryanto, M.M. selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama Jepara
3. Bapak Akhmad Khanif Zyen, M.Kom selaku Kaprodi Teknik Informatika Universitas Islam Nahdlatul Ulama Jepara
4. Bapak Harminto Mulyo, M.Kom selaku Dosen pembimbing I yang selalu bijaksana memberikan bimbingan, nasehat serta waktunya selama penulisan skripsi ini
5. Ibu Nur Aeni Widiastuti, S.Pd., M.Kom selaku Dosen Pembimbing II yang telah memberikan perhatian, bimbingan dan kepercayaan bagi penulis.
6. Bapak dan Ibu Dosen Program Studi Teknik Informatika khususnya dan Bapak serta Ibu Dosen di lingkungan Fakultas Sains dan Teknologi pada umumnya yang telah memberikan ilmu dan wawasan kepada penulis
7. Bapak Asmudi, Ibu Kasniyati, Istri Rahma Nur Fadlila yang selalu memberikan kasih sayang, do’a yang tak pernah putus, dan dukungan yang tiada henti kepada penulis
8. Teman seperjuangan program studi Teknik Informatika angkatan tahun 2016
9. Teman-teman mahasiswa dilingkungan Program Studi Teknik Informatika, Fakultas Sains dan Teknologi dan Unisnu Jepara yang senantiasa menghibur dan memberikan nasehat kepada penulis sehingga terselesainya karya ilmiah ini

10. Semua Pihak yang telah membantu sehingga skripsi ini dapat terselesaikan dengan baik

Semoga segala amal baik yang telah dilakukan mendapat ridho dan balasan pahala dari Allah SWT. Penulis menyadari bahwa dalam menyusun skripsi ini masih banyak sekali kekurangan dan masih jauh dari kata sempurna. Semoga skripsi ini bermanfaat bagi semua pihak. Amin.

Jepara, 17 September 2020

Penulis

Damas Fahmi Assena

NIM. 161240000500

ABSTRAK

Damas Fahmi Assena, 16124000500, Model Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Universitas Islam Nahdlatul Ulama Jepara. Harminto Mulyo, M.Kom, Nur Aeni Widiastuti, S.Pd., M.Kom.

Pada proses penerimaan mahasiswa baru Universitas Islam Nahdlatul Ulama Jepara akan menghasilkan data-data yang baru yang berlimpah berupa profil dari mahasiswa baru. Berdasarkan hal diatas informasi yang tersembunyi dapat diketahui dengan cara melakukan pengolahan terhadap data mahasiswa sehingga menghasilkan informasi yang bermanfaat bagi pihak universitas. Informasi yang didapatkan dapat membantu pihak universitas sebagai bahan pertimbangan dalam menentukan strategi promosi penerimaan mahasiswa baru di tahun depan. Penelitian ini bertujuan untuk mengelompokan data mahasiswa kedalam sebuah *cluster* dengan menggunakan metode *Data Mining* algoritma *K-Means Clustering*. Data mahasiswa dikelompokan berdasarkan kemiripan data tersebut sehingga data dengan karakteristik yang sama akan berada dalam satu *cluster*. Atribut yang digunakan adalah kecamatan asal, asal sekolah, dan IPK (*Indeks Prestasi Kumulatif*) mahasiswa selama 2 semester awal. *Cluster* yang terbentuk setelah proses *K-Means Clustering* adalah tiga *cluster* dengan *cluster* pertama berjumlah 709 data mahasiswa, *cluster* kedua 1116 data mahasiswa, dan *cluster* ketiga 2085 data mahasiswa. *Cluster* terbanyak berada pada *cluster* 3. Hasil dari penelitian ini digunakan sebagai salah satu dasar pengambilan keputusan untuk menentukan strategi promosi berdasarkan *cluster* yang terbentuk oleh pihak Universitas Islam Nahdlatul Ulama Jepara.

Kata Kunci : Kecamatan asal, asal sekolah, IPK (*Indeks Prestasi Kumulatif*), *K-Means Clustering*, *Data Mining*, Universitas Islam Nahdlatul Ulama Jepara

ABSTRACT

Damas Fahmi Assena, 161240000500, K-Means Clustering Algorithm to Determine the Promotion Strategy of the Jepara Nahdlatul Ulama Islamic University. Harminto Mulyo, M.Kom, Nur Aeni Widiastuti, S.Pd., M.Kom.

In the process of admitting new students at the Nahdlatul Ulama Islamic University Jepara, new data will be abundant in the form of profiles from new students. Based on the information above, hidden information can be found by processing student data so as to produce useful information for the university. The information obtained can help the university as a consideration in determining the promotion strategy for new student admissions next year. This study aims to classify student data into a cluster using the *Data Mining K-Means Clustering* algorithm. Student data is grouped based on the similarity of the data so that data with the same characteristics will be in one *cluster*. The attributes used are the sub-district of origin, the origin of the school, and the student's GPA (*Grade Point Average*) for the beginning of 2 semesters. The *clusters* formed after the *K-Means Clustering* process are three *clusters* with the first *cluster* totaling 709 student data, the second *cluster* 1116 student data, and the third *cluster* 2085 student data. The most *clusters* are in the third *cluster*. The results of this study are used as one of the basis for making decision to determine a promotion strategy based on the cluster formed by the Nahdlatul Ulama Islamic University Jepara.

Keywords : Subdistrict of origin, school origin, GPA (*Grade Point Average*), *K-Means Clustering*, *Data Mining*, Nahdlatul Ulama Islamic University Jepara

DAFTAR ISI

SAMPUL HALAMAN	
PERSETUJUAN PEMBIMBING.....	i
PENGESAHAN SKRIPSI	ii
PERNYATAAN KEASLIAN.....	iii
PERSEMBAHAN	iv
MOTTO	v
KATA PENGANTAR	vi
ABSTRAK	viii
<i>ABSTRACT</i>	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN.....	xvi
BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Batasan Masalah.....	5
1.3 Rumusan Masalah	5
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian.....	6
1.6 Sistematika Penyusunan Skripsi.....	7
BAB 2 LANDASAN TEORI	8
2.1 Tinjauan Studi	8
2.2 Tinjauan Pustaka	9

2.2.1	Bussines Intelligence.....	9
2.2.2	Pengertian Data Mining	10
2.2.3	Tahapan Data Mining.....	12
2.2.4	Teknik-Teknik Data Mining	13
2.2.5	Metode Data Mining	16
2.2.6	Clustering	18
2.2.7	Algoritma K-Means Clustering.....	19
2.2.8	Transformasi Data.....	24
2.2.9	Pengolahan data	25
2.2.10	Pengujian Hasil Clustering K-Means.....	28
2.2.11	Metode Receiver Operating Characteristic (ROC)	29
2.3	Kerangka Pemikiran	31
BAB 3 METODE PENELITIAN		32
3.1	Desain Penelitian	32
3.2	Pengumpulan Data	32
3.3	Lokasi Penelitian	33
3.4	Pengolahan Data.....	33
3.4.1	Perangkat Keras	34
3.4.2	Perangkat Lunak.....	34
3.5	Tahapan Metode	35
3.5.1	Pre-processing Data	35
3.5.2	K-Means Clustering	36
3.6	Evaluasi	41
3.6.1	Pengujian Metode BCV dan WCV	42
3.6.2	Pengujian Metode ROC	43

BAB 4	HASIL PENELITIAN DAN PEMBAHASAN.....	45
4.1	Pre-Processing Data	45
4.1.1	Data Reduction.....	46
4.1.2	Data Cleaning.....	47
4.1.3	Data Transformation	48
4.1.4	Data Integration.....	52
4.2	Penerapan Algoritma K-Means Clustering	53
4.3	Penerapan Algoritma menggunakan Microsoft Excel 2019.....	54
4.4	Evaluasi	63
4.4.1	Pengujian Metode BCV dan WCV	63
4.4.2	Pengujian Metode ROC	64
4.5	Penerapan Algoritma menggunakan Rapidminer Studio	66
BAB 5	PENUTUP	73
5.1	Kesimpulan.....	73
5.2	Saran.....	74
	DAFTAR PUSTAKA	75
	LAMPIRAN.....	77

DAFTAR TABEL

Table 2.1 Data Mahasiswa	23
Table 2.2 Inisialisasi Data Wilayah Kota Asal	24
Table 2.3 Inisialisasi Data Jurusan	24
Table 2.4 Titik Pusat Awal Setiap <i>Cluster</i>	26
Table 2.5 Contoh Hasil Perhitungan Setiap Data ke Setiap <i>Cluster</i>	27
Table 2.6 Kriteria Pengukuran Rasio	29
Table 2.7 Standar Receiver Operating Characteristic (ROC)	30
Table 3.1 Data Mahasiswa 2016 sampai 2018	36
Table 3.2 Inisialisasi Data Kecamatan Asal	38
Table 3.3 Inisialisasi Data Asal Sekolah	38
Table 3.4 Titik Pusat Awal Setiap <i>Cluster</i>	39
Table 3.5 Contoh Hasil Perhitungan Setiap Data ke Setiap <i>Cluster</i>	40
Table 3.6 Nilai <i>Centroid</i> pada iterasi terakhir	42
Table 3.7 Nilai <i>Centroid</i> pada iterasi terakhir	44
Table 3.8 Standar Receiver Operating Characteristic (ROC)	44
Table 4.1 Inisialisasi Data Kecamatan Asal	50
Table 4.2 Inisialisasi Data Asal Sekolah	51
Table 4.3 Titik Pusat Awal Setiap <i>Cluster</i>	55
Table 4.4 Tabel Centroid Baru (Cluster 1)	58
Table 4.5 Tabel Centroid Baru (Cluster 2)	58
Table 4.6 Tabel Centroid Baru (Cluster 3)	58
Table 4.7 Hasil Cluster 1 Penerapan Algoritma K-Means Clustering (<i>Microsoft Excel</i>)	60
Table 4.8 Hasil Cluster 2 Penerapan Algoritma K-Means Clustering (<i>Microsoft Excel</i>)	61
Table 4.9 Hasil Cluster 3 Penerapan Algoritma K-Means Clustering (<i>Microsoft Excel</i>)	62
Table 4.10 Nilai <i>Centroid</i> pada iterasi terakhir	63
Table 4.11 Nilai <i>Centroid</i> pada iterasi terakhir	65

Table 4.12 Standar Receiver Operating Characteristic (ROC)	65
Table 4.13 Perbandingan Hasil Perhitungan.....	72

DAFTAR GAMBAR

Grafik 1.1 Data Mahasiswa 2016-2018	2
Gambar 2.1 Business Intelligence Environment [9]	10
Gambar 2.2 Bidang Ilmu Data Mining [15].....	11
Gambar 2.3 Tahapan Proses Knowledge Discovery in Database [14]	11
Gambar 2.4 Bidang Ilmu Data Mining [15].....	12
Gambar 2.5 Teknik Data Mining [16].....	14
Gambar 2.6 Clustering [17].....	15
Gambar 4.1 Data Mentah Mahasiswa UNISNU Jepara.....	45
Gambar 4.2 Tahapan <i>Preprocessing</i> Data	46
Gambar 4.3 Tahapan Data <i>Reduction</i>	47
Gambar 4.4 Proses Inisialisasi Data Kecamatan Asal di <i>Microsoft Excel</i>	49
Gambar 4.5 Proses Inisialisasi Data Asal Sekolah di <i>Microsoft Excel</i>	51
Gambar 4.6 Proses <i>Data integration</i> Mahasiswa UNISNU Jepara	53
Gambar 4.7 Menghitung Jarak Data Ke Setiap <i>Cluster</i> di Excel.....	56
Gambar 4.8 Mengalokasikan Data ke dalam Cluster di Excel.....	57
Gambar 4.9 Verifikasi Data Centroid (<i>Microsoft Excel</i>)	59
Gambar 4.10 Proses Impor Dataset kedalam RapidMiner.....	67
Gambar 4.11 Desain <i>Clustering</i> Menggunakan Algoritma <i>K-Means</i>	67
Gambar 4.12 <i>Result History</i> menggunakan Algoritma <i>K-Means Clustering</i>	68
Gambar 4.13 Centroid Distance <i>K-Means Clustering</i>	69
Gambar 4.14 Hasil <i>Cluster Model K-Means</i>	69
Gambar 4.15 Hasil Perhitungan Antara Jarak <i>cluster</i> dengan <i>centroid</i> menggunakan <i>Rapid Miner Studio</i>	70
Gambar 4.16 Hasil Set <i>Clustering (Diagram Scatter)</i>	71

DAFTAR LAMPIRAN

LAMPIRAN A Tabel IPK Semester 2 angkatan 2016 sampai 2018.....	77
LAMPIRAN B Tabel Dataset Mahasiswa	86
LAMPIRAN C Tabel Perhitungan Jarak Data ke Setiap <i>Cluster</i> (Ms. Excel)	96
LAMPIRAN D Tabel <i>Clustering</i> Data Mahasiswa (Ms. Excel)	105