

IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA K-
MEANS CLUSTERING UNTUK MENGETAHUI POLA PEMILIHAN
PROGRAM STUDI MAHASISWA BARU IAIN SALATIGA



Disusun Oleh

Akhmad Choerudin Wakhid

12651098

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA

YOGYAKARTA

2017



PENGESAHAN TUGAS AKHIR

Nomor : B-2983/Un.02/DST/PP.00.9/11/2017

Tugas Akhir dengan judul : Implementasi Data Mining Menggunakan Algoritma K-Means Clustering Untuk Mengetahui Pola Pemilihan Program Studi Mahasiswa Baru IAIN Salatiga

yang dipersiapkan dan disusun oleh:

Nama : AKHMAD CHOERUDIN WAKHID
Nomor Induk Mahasiswa : 12651098
Telah diujikan pada : Kamis, 23 November 2017
Nilai ujian Tugas Akhir : A/B

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR

Ketua Sidang

Muhammad Didik Bohmad Wahyudi, S.T., MT.
NIP. 19760812 200901 1 015

Penguji I

Nurochman, S.Kom., M.Kom.
NIP. 19801223 200901 1 007

Penguji II

Sumarsono, S.T., M.Kom.
NIP. 19710209 200501 1 003

Yogyakarta, 23 November 2017
UIN Sunan Kalijaga
Fakultas Sains dan Teknologi
DEKAN



Dr. Murtono, M.Si.
NIP. 19691212 200003 1 001



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Permohonan

Lamp :

Kepada

Yth. Dekan Fakultas Sains dan Teknologi

UIN Sunan Kalijaga Yogyakarta.

di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Akhmad Choerudin Wakhid

NIM : 12651098

Judul Skripsi : Implementasi Data Mining Menggunakan Algoritma K-Means

Clustering Untuk Mengetahui Pola Pemilihan Program Studi Mahasiswa Baru Iain Salatiga

sudah dapat diajukan kembali kepada Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam bidang Teknik Informatika.

Dengan ini kami mengharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqsyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Yogyakarta, 14 November 2017

Pembimbing

M. Didik R Wahyudi, MT.

NIP. 19760812 200901 1 015

PERNYATAAN KEASLIAN SKRIPSI

Yang bertanda tangan di bawah ini :

Nama : Akhmad Choerudin Wakhid

Nim : 12651098

Program Studi : Teknik Informatika

Fakultas : Sains dan Teknologi

Menyatakan bahwa skripsi dengan judul "**Implementasi Data Mining Menggunakan Algoritma K-Means Clustering Untuk Mengetahui Pola Pemilihan Program Studi Mahasiswa Baru IAIN Salatiga**" tidak terdapat pada karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu Perguruan Tinggi dan sepengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Yogyakarta, 14 November 2017

Yang menyatakan



Akhmad Choerudin Wakhid

NIM 12651098

KATA PENGANTAR

Alhamdulillah rabbil'alamin, Puji dan syukur kehadiran Allah SWT yang telah memberikan rahmat serta hidayah-Nya kepada penulis sehingga penulis selalu dapat merasakan segala nikmatNya sehingga mampu menyelesaikan skripsi/tugas akhir yang berjudul “ Implementasi Data Mining Menggunakan Algoritma K-Means Clustering Untuk Mengetahui Pola Pemilihan Program Studi Mahasiswa Baru IAIN Salatiga “ Sholawat serta salam selalu tercurahkan kepada Nabi Muhammad SAW yang kita nanti nantikan syafaatnya di *yaumul qiyamah* kelak.

Skripsi/tugas akhir ini disusun untuk memenuhi salah satu syarat menyelesaikan studi Setara 1 dan untuk mendapatkan gelar Sarjana Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri (UIN) Sunan Kalijaga Yogyakarta.

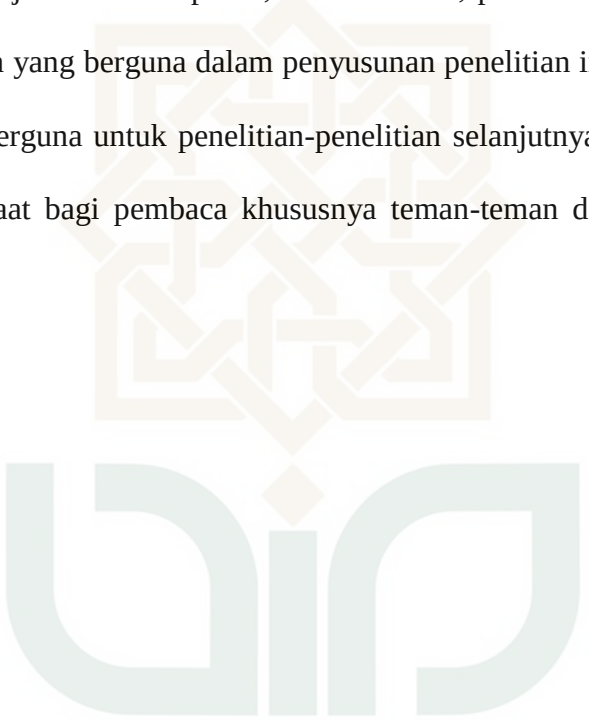
Dalam kesempatan ini penulis ingin mengucapkan terimakasih kepada banyak pihak yang telah memberikan bimbingan, arahan, dan bantuannya dalam penyusunan skripsi ini diantaranya kepada :

1. Bapak Prof. Drs. K.H. Yudian Wahyudi, M.A., Ph.D., selaku Rektor UIN Sunan Kalijaga Yogyakarta.
2. Bapak Dr. Murtono, M.Si, selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
3. Bapak Dr. Bambang Sugiantoro, S.Si., M.T. selaku Ketua Program Studi Teknik Informatika UIN Sunan Kalijaga Yogyakarta.

4. Bapak M. Didik R Wahyudi, S.T., MT. selaku Dosen Pembimbing skripsi yang telah menasehati, mengayomi, dan mengarahkan dengan segala kesabarannya.
5. Bapak Agus Mulyanto, S.Si., M.Kom., selaku Dosen Pembimbing Akademik selama masa perkuliahan yang telah memberikan pengarahan dan informasi seputar akademik.
6. Bapak dan Ibu Dosen Program Studi Teknik Informatika yang senantiasa memberikan bekal ilmu selama perkuliahan.
7. Bapak Syarifudin, S.Kom dan Ibu Nidaul Hasanah, S.T, yang telah memberikan izin penelitian sehingga penulis dapat memperoleh data-data yang dibutuhkan dalam tugas akhir ini.
8. Bapak Ahmari dan Ibu Asfiyatun tercinta, dan seluruh anggota keluarga besar yang tidak bisa penulis tulis satu per satu atas doa, perhatian, kasih sayang dan dukungan moril maupun materilnya.
9. Teman-teman Program Studi Teknik Informatika, khususnya angkatan 2012 (Katak'12) tercinta yang telah banyak memberi dukungan dan semangat yang sangat besar.
10. Saudari Nur Khasanah yang tak pernah lelah memberikan semangat dan dukungan yang luar biasa dalam proses penyelesaian skripsi ini.
11. Teman-teman KKN UIN Suka Angka 86 Kelompok 142, Serta keluarga disana. Terimakasih banyak atas pegalaman yang tidak ternilai harganya.
12. Teman-teman yang berjasa dalam penyusunan skripsi Akhmad Devianto, Nur Rohman, Rizky Terutama Krisna Adi Yogantara dan Andi Wijanarko yang bersedia saya repotkan berklali-kali untuk menyelesaikan skripsi ini.

13. Serta semua pihak yang tidak dapat disebutkan satu per satu, yang telah banyak memberikan dukungan, motivasi, inspirasi dan membantu dalam proses penyelesaian skripsi ini.

Penulis menyadari apa yang dilakukan dalam penyusunan laporan tugas akhir ini masih jauh dari sempurna, maka dari itu, penulis sangat mengharapkan kritik dan saran yang berguna dalam penyusunan penelitian ini di masa yang akan datang, serta berguna untuk penelitian-penelitian selanjutnya. Semoga skripsi ini dapat bermanfaat bagi pembaca khususnya teman-teman dan pihak-pihak yang bersangkutan.



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Yogyakarta, 14 November 2011

Penyusun

'Akhmad Choerudin Wakhid
NIM. 12651098

HALAMAN PERSEMBAHAN

Dengan mengucapkan segala rasa syukur penulis mempersembahkan tugas akhir ini untuk :

- ❖ Kedua Orang tuaku Ibu Asfiytun dan bapak Ahmari yang telah berjuang sejauh ini buatku, yang tak mengenal lelah untuk selalu mendoakan ku dan tetap menjadi motivasi terbesar dalam perjalanan hidupku. Kedua malaikat tanpa sayapku yang tak pernah bosan mendoakan dan menyayangiku, yang terus mendukungku sampai sejauh ini. Semoga Ayah dan ibuku panjang umur dan bisa melihatku menjadi anak yang membanggakan keluarga suatu hari nanti, amin.
- ❖ Keluarga beras Mbah Kalimah dan Mbah Rusmilah, Ibu Sriyah, Bapak Sugeng, Ibu Sopiya, Bapak Mulidi, Bapak Muhajir, Ibu Ning ,Bapak Ashuri, Ibu Marni, Ibu Siti, Bapak Sabar, Bapak Asrosi, Ibu Seni, Bapak Eswadi, Sri yamsih, Sriyono, Antok, Sahidun, Rowi, Anwar, Solikah, Eni Nurfuadah, Liha, Najid, Fadil, Zuhri, Isna, Sidiq, Eni, Fani, Fahri, yang selalu memberikan dukungan baik yang secara finansial maupun dengan segenap perhatian.
- ❖ Untuk kamu teman hidupku yang selalu menemani dalam setiap langkahku saudari Nur Khasanah yang tak pernah lelah untuk memberikan dukungan, motivasi, perhatian dan kasih sayang kepadaku serta selalu sabar menghadapi aku yang malas-malasan dalam menyusun skripsi, terima kasih sudah selalu mendengar segala keluh kesahku, terima

kasih sudah menemaniku dari awal pembuatan skripsi. Aku Sayang Kamu.

- ❖ Bapak M. Didik R Wahyudi M.T, Terimakasih telah menjadi dosen pembimbing yang sangat memotivasi dan memberikan bimbingan serta saran dalam pembuatan aplikasi dan laporan tugas akhir ini sampai terselesaikan, Semoga bapak selalu dalam lindungan Allah SWT.
- ❖ Dosen dan keluarga besar Teknik Informatika, Pak Bambang Sugiantoro selaku ketua program studi yang selalu sedia dan terbuka menerima keluhan para mahasiswanya. Pak Didik yang selalu mengarahkan dan selalu peduli kepada anak bimbingnya, Ibu Ade, Pak Mustakim, Pak Agung, Ibu Uyun, Pak Rahmat, Pak Aulia dan Pak Agus Mulyanto yang selalu sabar memberikan ilmu-ilmunya. semoga Bapak dan Ibu dosen panjang umur, diberikan kesehatan, dan selalu bahagia sampai tua kelak, amin.
- ❖ Teman - teman seperjuangan dan keluarga besar Teknik Informatika Khusus 2012 (Katak'12) yang sangat saya sayangi dan saya banggakan Abdul Mukhrim, Afin Hilman, Ahmad Deviyanto, Ainul Yaqin Muntaqo, Andi Wijanarko, Asep Baeri, Bayu Resi Indrawan, Dwi Indah Permatasari, Dwi Kumala Mursid, Edita Rizki, Eri Kurniawan, Erin Septiana, Fahrieza Rahmadziba, Fajar Septian Nugraha, Firmansyah Gustap Hikmatyar, Gatot Cakra Sutradana, Ikhzan Zuhriyanto, Gumeta Sari Maharani, Hilyas Riza, Indra Faisol Alim, Irfan Andriyanto, Iwan Nasrullah, Juhdan, Kharizma Azdhi, Krisna Adi Yogantoro, Luqyana Khalda Kiki, Lusi Anggarini, M Berlian Rafsanjani, Malika Dhini,

Muhammad Edi Iswanto, Nanang Paminto Aji, Novie Astutie, Nur Kukuh Wicaksana, Nur Rohman, Nurul Zuni Rahmawati, Rahulloh Ayyatulloh Komaeni Noor Bintang, Ripah Gemah Nuripah, Rizky Fahrezi, Surahmat Laguni , Taufik Qurniawan, Tri Wiji Hastuti, Ulfa Septi Muslimah, Ulvi Ismaya dan Valdi Adrian Abrar. terimakasih atas semua kenangan indah yang telah kalian berikan. Semoga persahabatan ini tidak akan terpisahkan sampai hari akhir nanti.

- ❖ Sahabatku yang saya sayangi dan saya banggakan Miftahul Huda, Fariden, Mulyani, khoir, Muhroni, Mohari, Uzi, Sofyan BH, Eko, Ari, Iwan Nasrullah, Krisna Adi Yogantoro, Juhdan, Agung, Yunus, dani. Terimakasih atas pengorbanan dan petualangan yang sangat menyenangkan dan semoga persahabatan kita akan selalu terjaga sampai kita tua.
- ❖ Pihak - pihak yang selalu memberikan bantuannya, semangat, dan doanya baik secara langsung maupun tidak yang tidak dapat penulis sebutkan namanya satu per satu.

HALAMAN MOTTO

وَمَنْ يَتَّقِ اللَّهَ يَجْعَلْ لَهُ مِنْ أَمْرِهِ يُسْرًا

**Dan Barang Siapa Yang Bertakwa
Kepada Allah, Niscaya Allah
Menjadikan Baginya Kemudahan Dalam
Urusannya.**

~ (Q.S At-Talaq: 4) ~

**Try Not Become Man of Success, But
Rather Become a Man of Value.**

~ (Albert Einstein) ~

**Mandiri Dalam Bekerja Merdeka
Dalam Berkarya**

~ (Endang Soekamti) ~

DAFTAR ISI

Halaman Judul	i
Halaman Pengesahan	ii
Surat Persetujuan Skripsi	iii
Pernyataan Keaslian Skripsi	iv
Kata Pengantar	v
Halaman Persembahan	viii
Halaman Motto	xi
Dafta Isi	xii
Daftar Gambar	xvi
Daftar Tabel	xix
Daftar Rumus	xxi
Daftar Singkatan	xxii
Daftar Istilah	xiv
Intisari	xxv
Abstract	xxvi
BAB 1 PENDAHULUAN	1
1.1. Latar Belakang	1

1.2. Rumusan Masalah	2
1.3. Batasan Masalah	3
1.4. Tujuan Penelitian	4
1.5. Manfaat Penelitian	4
1.6. Keaslian Penelitian	5
1.7. Sistematika Penelitian	5
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI	7
2.1. Tinjauan Pustaka	7
2.2. Landasan Teori	11
2.2.1. Data Mining	11
2.2.2. Teknik Teknik Data Mining	12
2.2.3. Tahap Data Mining	18
2.2.4. Clustering	21
2.2.5. Algoritma K-Means	22
2.2.6. PHP (PHP: Hypertext Preprocessor).....	22
2.2.7. Basis Data dan DBMS	23
2.2.8. MySQL.....	23
2.2.9. Pengolahan Data Microsoft Excel	26
2.2.10. Data Flow Diagram (DFD)	29
2.2.11. Entity Relationship Diagram (ERD)	32
2.2.12. XAMPP	34

BAB III METODE PENELITIAN	36
3.1. Objek dan Lokasi Penelitian	36
3.2. Peralatan Peneitian	36
3.2.1. Perangkat Keras	36
3.2.2. Perangkat Lunak	37
3.3. Metode Penelitian	37
3.4. Tahap Tahap Penelitian	37
3.4.1. Studi Awal	38
3.4.2. Model Pengambilan Data	38
3.5. Pre-Processing Data	39
3.6. Proses K-Means Clustering	40
3.7. Hasil Output	44
 BAB IV HASIL DAN PEMBAHASAN	 45
4.1. Objek Penelitian	45
4.2. Lokasi Penelitian	45
4.3. Tahapan Proses Sistem	47
4.4. Pengambilan Data	47
4.5. Pre-Processing Data	50
4.5.1. Pembersihan Data (Data Cleaning).....	50
4.5.2. Integrasi Data (Data Integration)	52
4.5.3. Seleksi Data (Data Selection)	53
4.5.4. Transformasi Data (Data Transformation)	54

4.6. Penerapan Teknik Data Mining Menggunakan Algoritma	
K-Means Clustering Pada Ms.Word	60
4.7. Hasil Penerapan K-Means Menggunakan Microsoft Excel	71
4.8. Penerapan Algoritma K-Means Menggunakan PHP & MySQL. 97	
4.8.1. Perancangan Sistem	98
4.8.2. Implementasi Sistem Aplikasi K-Means Clustering	
Dengan Menggunakan PHP&MySQL	116
4.9. Tahapan Evaluasi	139
4.9.1. Tahapan Evaluasi Pola	139
4.9.2. Presentasi pengetahuan	141
BAB V Penutup	143
5.1. Kesimpulan	143
5.2. Saran	146
Daftar Pustaka	147
Lampiran	
Curriculum Vitae	

DAFTAR GAMBAR

Gambar 2.1 Classification	12
Gambar 2.2 Clustering	13
Gambar 2.3 Grafik 3-year price history	16
Gambar 2.4 Squence	17
Gambar 2.5 Tahap Data Mining (Han, 2006)	18
Gambar 2.6 Tampilan Microsoft Excel Fungsi Vlookup	28
Gambar 4.1 Tahapan Proses Sistem	47
Gambar 4.2 Data mahasiswa sebelum di olah	49
Gambar 4.3 Contoh Hasil setelah pembersihan data	51
Gambar 4.4 Pengambilan data dari nama sekolah	53
Gambar 4.5 Hasil Transformasi Data	60
Gambar 4.6 Flowchart Tahapan Proses K-Means Clustering	61
Gambar 4.7 Contoh Penggunaan Rumus K-Means di Ms.excel	65
Gambar 4.8 Hasil Proses Perhitungan K-Means Clustering	70
Gambar 4.9 Pola Prodi Berdasarkan Asal Sekolah Cluster 1	75
Gambar 4.10 Pola Prodi Berdasarkan Jurusan Sekolah cluster 1	76
Gambar 4.11 Pola Prodi Berdasarkan Asal sekolah pada cluster 2	80

Gambar 4.12 Pola Prodi Berdasarkan Jurusan Sekolah cluster 2	81
Gambar 4.13 Pola Prodi Berdasarkan Asal sekolah pada cluster 3	85
Gambar 4.14 Pola Prodi Berdasarkan Jurusan Sekolah cluster 3	86
Gambar 4.15 Pola Prodi Berdasarkan Asal sekolah pada cluster 4	90
Gambar 4.16 Pola Prodi Berdasarkan Jurusan Sekolah cluster 4	91
Gambar 4.17 Pola Prodi Berdasarkan Asal sekolah pada cluster 5	95
Gambar 4.18 Pola Prodi Berdasarkan Jurusan Sekolah cluster 5	96
Gambar 4.19 DFD Level 0 (Diagram Konteks)	99
Gambar 4.20 DFD Level 1	101
Gambar 4.21 Perancangan ERD Dalam Sistem	106
Gambar 4.22 Gambar Relasi Database (Sistem K-Means Clstering)	116
Gambar 4.23 Halaman Depan Aplikasi K-Menas Clustering	117
Gambar 4.24 Tampilan Menu About Pada Aplikasi K-Means Clustering	118
Gambar 4.25 Tampilan Menu K-Means Pada Aplikasi K-Means Clustering	119
Gambar 4.26 Tampilan Setelah Data Berhasil di Importkan	120
Gambar 4.27 Tampilan Halaman Transformasi data	121
Gambar 4.28 Pre-Processing Data dan Penentuan Titik Pusat Cluster	123
Gambar 4.29 Tampilan Fitur Manual	124

Gambar 4.30 Tampilan Fitur Otomatis	125
Gambar 4.31 Tampilan Hasil Proses 1 K-Means Clustering	126
Gambar 4.32 Tampilan Pusat Cluster Baru	127
Gambar 4.33 Tampilan Proses Selesai	128
Gambar 4.34 Hasil Proses <i>K-Means Cluster 1</i>	129
Gambar 4.35 Grafik Hasil Proses <i>K-Means Cluster 1</i>	130
Gambar 4.36 Hasil Proses <i>K-Means Cluster 2</i>	131
Gambar 4.37 Grafik Hasil Proses K-Means Cluster 2	132
Gambar 4.38 Hasil Proses K-Means Cluster 3	133
Gambar 4.39 Grafik Hasil Proses K-Means Cluster 3	134
Gambar 4.40 Hasil Proses K-Means Cluster 4	135
Gambar 4.41 Grafik Hasil Proses K-Means Cluster 4	146
Gambar 4.42 Hasil Proses K-Means Cluster 5	137
Gambar 4.43 Grafik Hasil Proses K-Means Cluster 5	138
Gambar 4.44 Hasil Pengujian Evaluasi	139
 Gambar 5.1 Prosentase Minat Mahasiswa Terhadap Program Studi Iain Salatiga	 145

DAFTAR TABEL

Tabel 2.1 Tabel Penelitian Terdahulu	8
Tabel 2.2 Tabel Simbol Data Flow Diagram (DFD)	31
Tabel 2.3 Simbol simbol Entity Relationship Diagram (ERD)	33
Tabel 4.1 Atribut Sebelum Di Olah	49
Tabel 4.2 Perbandingan data sebelum dan sesudah proses cleaning	52
Tabel 4.3 Inisialisasi Program Studi	55
Tabel 4.4 Tabel Inisialisasi Asal Sekolah	57
Tabel 4.5 Tabel Inisialisasi Jurusan Sekolah	58
Tabel 4.6 Nilai Centroid Pertama	63
Tabel 4.7 Tabel Titik Pusat Cluster Baru (Cluster 1)	67
Tabel 4.8 Tabel Titik Pusat Cluster Baru (Cluster 2)	68
Tabel 4.9 Tabel Titik Pusat Cluster Baru (Cluster 3)	68
Tabel 4.10 Tabel Titik Pusat Cluster Baru (Cluster 4)	68
Tabel 4.11 Tabel Titik Pusat Cluster Baru (Cluster 5)	68
Tabel 4.12 Verifikasi Centroid Lama Dengan Centroid Baru	70
Tabel 4.13 Verifikasi Jumlah Data Pada Setiap Cluster	71

Tabel 4.14 Hasil <i>Cluster</i> 1 Penerapan Algoritma K-Means Clustering	
(<i>Microsoft Excel</i>)	72
Tabel 4.15 Hasil Cluster 2 Penerapan Algoritma K-Means Clustering	
(Microsoft Excel)	77
Tabel 4.16 Hasil Cluster 3 Penerapan Algoritma K-Means Clustering	
(Microsoft Excel)	82
Tabel 4.17 Hasil Cluster 4 Penerapan Algoritma K-Means Clustering	
(Microsoft Excel)	87
Tabel 4.18 Hasil Cluster 5 Penerapan Algoritma K-Means Clustering	92
Tabel 4.19 Tabel Database Dataset	108
Tabel 4.20 Tabel Database Frek_Prodi	109
Tabel 4.21 Tabel Database Frek_Asko (Frekuensi Asal Sekolah)	110
Tabel 4.22 Tabel Database Frek_Jurusan	111
Tabel 4.23 Tabel Database Centroid_Lama	112
Tabel 4.24 Tabel Database Nomor	112
Tabel 4.25 Tabel Database Proses	114
Tabel 4.26 Tabel Database Grafik	115
Tabel 5.1 Jumlah Minat Manahiswa Tiap Cluster	145

DAFTAR RUMUS

Rumus 3.1 Rumus Mencari Jarak (<i>distance</i>) dari satu cluster.....	42
Rumus 4.1 Rumus <i>excel</i> mengambil 4 karakter dari sisi <i>cell</i> sebelah kiri	52
Rumus 4.2 Rumus Mencari Jarak Data ke Setiap <i>Cluster</i>	64
Rumus 4.3 Rumus Excel Mencari Jarak Data ke Setiap <i>Cluster</i>	66
Rumus 4.4 Rumus Excel Mengalokasikan Data kedalam cluster	66
Rumus 4.5 Rumus Excel Menentukan Titik Pusat Cluster Baru	66



DAFTAR SINGKATAN

Prodi	: Program Studi
STAIN	: Sekolah Tinggi Agama Islam Negeri
IAIN	: Institut Agama Islam Negeri
UIN	: Universitas Islam Negeri
IPK	: Indeks Prestasi Kumulatif
PHP	: PHP Hypertext Preprocessor
HTML	: Hyper Text Markup Language
URL	: Uniform Resource Locator
DNA	: Deoxyribose Nucleic Acid
SQL	: Struktur Query Language
DBMS	: Data Base Management System
DDL	: Data Definition Language
DML	: Data Manipulation Language
DCL	: Data Control Language
DFD	: Data Flow Diagram
ERD	: Entity Relationship diagram
HTTP	: Hypertext Transfer Protocol
GNL	: General Public License
NIM	: Nomor Induk Mahasiswa
BAKESBANGPOL	: Badan Kesatuan Bangsa dan Politik
DPMPTSP	: Dinas Pelayanan Modal dan Pelayanan Terpadu Satu Pintu
KKD	: Knowledge Discovery in Database
SMA	: Sekolah Menengah Atas
MA	: Madrasah Aliyah
SMK	: Sekolah Menengah Kejuruan
POND	: Pondok

SKB	: Sanggar Kegiatan Belajar
PS	: Perbankan Syariah
PAI	: Pendidikan Agama Islam
TBI	: Tadris Bahasa Inggris
PGMI	: Pendidikan Guru Madrasah Ibtidaiyah
ES	: Ekonomi Syariah
TM	: Tadris Matematika
KPI	: Komunikasi dan Penyiaran Islam
HES	: Hukum Ekonomi Syariah
TIPA	: Tadris Ilmu Pengetahuan Alam
SPI	: Sejarah Peradaban Islam
PBA	: Pendidikan Bahasa Arab
BSA	: Bahasa dan Sastra Arab
PIAUD	: Pendidikan Islam Anak Usia Dini
MD	: Manajemen Dakwah
IH	: Ilmu Hadis
IPA	: Ilmu Pengetahuan Alam
IPS	: Ilmu Pengetahuan Alam
ASKO	: Asal Sekolah
Iniprodi	: Inisial Prodi
Inijur	: Inisial Jurusan
Iniasko	: Inisial Asal Sekolah

DAFTAR ISTILAH

Data mining	: Serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data dengan melakukan penggalian pola-pola dari data dengan tujuan untuk memanipulasi data menjadi informasi yang lebih berharga
Clustering	: Proses untuk mengelompokkan suatu item berdasarkan kesamaan karakteristik.
Algoritma	: Susunan yang logis dan sistematis untuk memecahkan suatu masalah atau untuk mencapai tujuan tertentu
Cluster	: Pengelompokan / alokasi data yang saling berhubungan
Atribut	: kolom data
Centroid	: Titik pusat cluster
Database	: Sekumpulan file yang saling terkait dan membentuk suatu bangun data. Database minimal terdiri dari satu file yang cukup untuk dimanipulasi oleh komputer sedemikian rupa.
Field	: Bagian dari sebuah record, biasanya terdiri dari sebuah data dari informasi yang berelasi ke data lain dalam record tersebut
Kmeans	: Salah satu metode data non-hierarchical clustering yang dapat mengelompokkan data ke dalam beberapa cluster berdasarkan kemiripan dari data
Pemrograman	: Proses menulis, menguji dan memperbaiki (debug), dan memelihara kode yang membangun suatu program komputer.

IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING UNTUK MENGETAHUI POLA PEMILIHAN PROGRAM STUDI MAHASISWA BARU IAIN SALATIGA

Akhmad Choerudin Wakhid

NIM. 12651098

INTISARI

Data mining dapat digunakan untuk instansi, sekolah, perguruan tinggi dan lain sebagainya. Di perguruan tinggi dapat digunakan untuk mengolah data-data yang ada seperti data mahasiswa, karyawan, dosen dan lain sebagainya. Setiap tahun, perguruan tinggi akan membuka penerimaan mahasiswa baru. Oleh sebab itu, perguruan tinggi tersebut akan menghasilkan data-data mahasiswa baru yang harus disimpan seperti nama, alamat, asal sekolah, dan lain sebagainya. Tidak semua perguruan tinggi memanfaatkan banyaknya data mahasiswa tersebut selain hanya untuk administrasi.

Penelitian ini menggunakan penerapan data mining dengan metode *k-means clustering* agar dapat diketahui pola pemilihan program studi bagi mahasiswa baru di IAIN Salatiga dengan mengelompokkan data mahasiswa. Data mahasiswa dikelompokkan berdasarkan kemiripan data tersebut sehingga data dengan karakteristik yang sama akan berada dalam satu cluster. Atribut yang digunakan adalah prodi, asal sekolah, jurusan asal sekolah.

Cluster yang terbentuk setelah proses K-Means Clustering adalah tiga cluster dengan cluster pertama berjumlah 1492 data mahasiswa, cluster kedua 638 data mahasiswa, dan cluster ketiga 336 data mahasiswa. Dari hasil penelitian ini dapat disimpulkan bahwa pada cluster pertama merupakan yang paling tinggi peminatnya. Sehingga dapat terlihat minat mahasiswa baru dalam memilih program studi yang berada di IAIN Salatiga dengan melihat hasil dari tiga cluster yang telah melalui berbagai proses.

Kata Kunci : Program studi, Jurusan sekolah, Asal sekolah, Data mining, K-Means clustering, Cluster, Mahasiswa, IAIN Salatiga

**IMPLEMENTATION OF DATA MINING USING K-MEANS
CLUSTERING ALGORITHM TO KNOW THE SELECTION PATTERNS
OF NEW STUDENT PROGRAMS AT IAIN SALATIGA**

Akhmad Choerudin Wakhid

NIM. 12651098

ABSTRACT

Data mining can be used in agencies, schools, college and so on. In college can be used to process the existing data such as student data, employees, lecturers and others. Every year, colleges will open new admissions. Therefore, the college will produce new student data that must be stored like name, address, origin of school, and so forth. Not all colleges use all of the student data other than just for administration.

This study uses the application of data mining with k-means clustering method in order to know the pattern of study program selection for new students at IAIN Salatiga by grouping the student data. The student data are grouped based on similarity of data so that the data with the same characteristics will be in one cluster. The attributes used are program study, school, the major of origin of school.

Clusters formed after the K-Means Clustering process are three clusters with the first cluster amounting to 1492 student data, second cluster 638 student data, and third cluster 336 student data. From the results of this study can be concluded that the first cluster is the highest demand. So it can be seen the interest of new students in choosing courses that are in IAIN Salatiga by looking at the results of the three clusters that have been processed through various processes.

Keywords: Program of study, school, major of origin of school, Data mining, K-Means clustering, Student, IAIN Salatiga.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Data mining merupakan proses untuk mendapatkan informasi dari basis data yang besar dengan cara penggalian atau penyaringan data dengan memanfaatkan kumpulan data. Data mining dapat digunakan untuk instansi, sekolah, perguruan tinggi dan lain sebagainya. Di perguruan tinggi dapat digunakan untuk mengolah data-data yang ada seperti data mahasiswa, karyawan, dosen dan lain sebagainya. Setiap tahun, perguruan tinggi akan membuka penerimaan mahasiswa baru. Oleh sebab itu, perguruan tinggi tersebut akan menghasilkan data-data mahasiswa baru yang harus disimpan seperti nama, alamat, asal sekolah, dan lain sebagainya. Tidak semua perguruan tinggi memanfaatkan banyaknya data mahasiswa tersebut selain hanya untuk administrasi. Dengan menggunakan *data mining* menggunakan *algoritma k-means clustering* data-data yang begitu banyak tersebut dapat dijadikan sebagai hal yang positif bagi perguruan tinggi tersebut.

Apabila dilakukan pengolahan data pada data mahasiswa tersebut maka akan dapat diketahui berbagai informasi yang bermanfaat seperti pola pemilihan program studi berdasarkan asal sekolah, jurusan asal sekolah, alamat asal, dan lain sebagainya. Pengelolaan tersebut dapat dilakukan dengan menggunakan metode data mining dengan algoritma *k-means clustering* agar dapat diketahui pola pemilihan program studi bagi mahasiswa baru dan bagi perguruan tinggi.

Hal tersebut juga terjadi pada IAIN Salatiga. Data mahasiswa baru tersebut perlu dilakukan pengolahan data dengan hasil yang menjelaskan berbagai informasi. Ribuan mahasiswa baru yang masuk di IAIN Salatiga setiap tahunnya yang terbagi dalam 6 fakultas dan 20 program studi untuk program S1. Data yang telah ditampung pastinya banyak sekali. Jika data tersebut dimanfaatkan dengan penerapan data mining, maka akan dapat dilihat pola-pola pemilihan program studi berdasarkan asal sekolah, jurusan asal sekolah, alamat asal, pekerjaan orang tua, dan lain sebagainya.

Dari penelitian ini diharapkan dapat diketahui apakah mahasiswa tersebut benar-benar memilih program studi tersebut masih linear dengan bidang keilmuan yang dipelajari pada jenjang pendidikan sebelumnya, ataukah terjadi penyimpangan dari konsentrasi yang pernah ditempuh di SMA/SMK terdahulu terhadap pemilihan program studi yang ada di IAIN Salatiga.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas dapat dirumuskan permasalahan yang akan diselesaikan dalam penelitian ini yaitu :

1. Bagaimana cara mengetahui pola pemilihan program studi mahasiswa IAIN Salatiga dengan menerapkan metode data mining menggunakan algoritma k-means clustering?

2. Bagaimana pola minat mahasiswa terhadap program studi IAIN Salatiga berdasarkan jenjang pendidikan sebelumnya?
3. Bagaimana mengolah hasil pola minat mahasiswa IAIN Salatiga guna sebagai rujukan sasaran promosi ke depan?

1.3 Batasan Masalah

Berdasarkan identifikasi masalah yang telah diuraikan di atas, maka permasalahan dibatasi pada :

1. Penelitian dilakukan di Institut Agama Islam Negeri (IAIN) Salatiga.
2. Data yang digunakan adalah data mahasiswa S1 Institut Agama Islam Negeri (IAIN) Salatiga angkatan 2016.
3. Penelitian dilakukan dengan menggunakan metode Data Mining dengan Algoritma *K-means clustering*.
4. Pengolahan data menggunakan parameter atau atribut sebanyak 3 atribut yaitu prodi, asal sekolah dan jurusan asal sekolah.
5. pengolahan data menggunakan aplikasi *Microsoft Excel*.
6. Aplikasi yang dibangun menggunakan bahasa pemrograman *PHP* dan *MySQL*.
7. Aplikasi dengan menggunakan *PHP* dan *MySQL* bukan fokus utama melainkan hanya sebagai pembanding.
8. Dalam pengolahan data peneliti menggunakan cluster sebanyak 5 cluster.
9. Hasil penelitian berupa pola minat mahasiswa terhadap prodi di IAIN Salatiga yang terbagi dalam 5 cluster.

1.4 Tujuan Penelitian

Sesuai dengan masalah yang telah dirumuskan, maka tujuan dari penelitian ini sebagai berikut :

1. Mengetahui pola pemilihan program studi mahasiswa IAIN Salatiga menggunakan metode *Data Mining* dengan algoritma *K-Means Clustering* ke dalam aplikasi.
2. Mengetahui pola minat mahasiswa terhadap program studi IAIN Salatiga berdasarkan jenjang pendidikan sebelumnya.
3. Dengan menggunakan metode *Data Mining* dengan algoritma *K-Means Clustering* hasil pengolahan data dapat di jadikan sebagai rujukan agar promosi program studi IAIN salatiga dapat tepat sasaran.

1.5 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut :

1. Data mahasiswa yang berjumlah ribuan dapat cepat diproses dan di pisahan berdasarka kemiripan data dengan metode *data mining* menggunakan algoritma *k-means clustering*.
2. Hasil penelitian ini dapat memberikan gambaran dari pola pemilihan program studi berdasarkan asal sekolah dan jurusan sekolah pada jenjang pendidikan sebelumnya.

3. Bagi calon mahasiswa baru IAIN Salatiga, hasil dari penelitian ini dapat memberikan gambaran prodi apakah yang akan dipilih berdasarkan jenjang pendidikan dan jurusan sebelumnya.
4. Bagi IAIN Salatiga, dapat digunakan sebagai pertimbangan untuk memaksimalkan sosialisasi dan promosi program studi IAIN Salatiga dari hasil tingkat statistik mahasiswa dalam memilih program studi.

1.6 Keaslian penelitian

Penelitian tentang implementasi data mining menggunakan algoritma k-means clustering sudah pernah dilakukan sebelumnya, namun untuk penerapan dalam mengetahui pola pemilihan program studi mahasiswa baru IAIN Salatiga, sejauh pengetahuan penulis belum pernah dilakukan sebelumnya.

1.7 Sitematika Penelitian

Laporan penelitian tugas akhir ini disusun secara sistematis dibagi dalam 5 bab. Penyusunan laporan tugas akhir ini memiliki urutan yang di mulai dari BAB I sampai dengan BAB V Berikut ini gambaran sistematika penulisan pada masing-masing bab :

BAB I. PENDAHULUAN

Pada bab ini menerangkan tentang latar belakang, rumusan masalah, tujuan penelitian, batasan penelitian, keaslian penelitian dan sistematika penulisan

BAB II. TINJAUAN PUSTAKA DAN LANDASAN TEORI

Bagian ini berisikan tentang teori-teori penunjang penelitian. Terdiri atas teori data mining, algoritma k-means clustering, dan teori bahasa pemrograman yang nantinya digunakan untuk membuat pola pemilihan program studi.

BAB III. METODE PENELITIAN

Bagian ini berisikan tentang objek, rincian perangkat yang digunakan serta metode penelitian.

BAB IV. HASIL DAN PEMBAHASAN

Bagian ini memuat tentang hasil analisis penelitian dan pembahasan mengenai penerapan metode yang digunakan, hasil penelitian yang dicapai dari proses penyeleksian data sampai pengolahan data menggunakan cara manual dan dengan program aplikasi yang dibuat oleh peneliti dengan menggunakan bahasa pemrograman *PHP* dan *MySQL*.

BAB V. PENUTUP

Pada bagian ini berisikan kesimpulan dan saran-saran untuk perbaikan, sehingga dapat digunakan sebagai bahan pertimbangan bagi pihak-pihak yang berkepentingan serta kemungkinan perkembangan untuk penelitian selanjutnya.

BAB V

PENUTUP

5.1. KESIMPULAN

Kesimpulan yang dapat peneliti simpulkan dalam penelitian ini adalah sebagai berikut :

1. Penerapan data mining menggunakan algoritma K-Means Clustering untuk mengetahui pola pemilihan program studi IAIN Salatiga berdasarkan profil mahasiswa asal sekolah dan jurusan asal sekolah.
2. Dengan pengolahan data mining menggunakan algoritma k-means clustering untuk mengetahui pola pemilihan program studi IAIN Salatiga, dilakukan dengan menggunakan dua aplikasi, yang pertama dengan menggunakan aplikasi pengolah angka *Microsoft Excel* dan dengan menggunakan aplikasi yang peneliti rancang dengan menggunakan bahasa pemrograman *PHP* dan *Database MySQL*.
3. Hasil dari proses k-means clustering dengan menggunakan data mahasiswa IAIN Salatiga S1 angkatan 2016 terbagi ke dalam lima cluster, dimana cluster pertama berisikan tentang program studi yang paling diminati oleh mahasiswa baru IAIN Salatiga, cluster kedua berisikan prodi dengan peminat lebih sedikit dengan dari cluster pertama dan lebih banyak di bandingkan dengan cluster tiga, empat dan lima, pada cluster tiga berisikan program studi yang diminati, sedangkan pada cluster empat dan lima berisikan program studi yang kurang diminati oleh mahasiswa baru.

4. Pada aplikasi yang peneliti rancang terdapat fitur random dimana nilai titik pusat clusternya akan otomatis dan akan menjadikan nilai titik pusat cluster berbeda dengan nilai titik pusat cluster yang ada pada aplikasi *Microsoft excel*, hal tersebut akan mempengaruhi nilai-nilai pada proses k-means dan jumlah prosesnya, akan tetapi nilai yang dihasilkan pada akhir proses akan tetap sama.
5. Jika pada pusat nilai titik cluster antara Microsoft Excel dan dengan aplikasi yang peneliti rancang di inputkan nilai yang sama, maka hasilnya akan sama persis, jika ada perbedaan nilai hanya sedikit dikarenakan nilai di belakang koma yang berbeda, akan tetapi hasil pada proses k-means akan tetap sama.
6. Dari beberapa program studi yang ada di IAIN Salatiga, program studi yang paling diminati adalah program studi yang sudah dikelompokkan di *cluster* 1, dengan jumlah peminat yang berasal sekolah MA dengan jurusan asal sekolah IPA. Dan pada prodi yang cukup diminati dikelompokkan dalam *cluster* 2, sama dengan prodi *cluster* satu mahasiswa yang meminati prodi pada *cluster* dua berasal sekolah MA dengan jurusan asal sekolah IPA, dan pada prodi yang diminati dikelompokkan pada *cluster* 3, yang berasal sekolah MA dengan jurusan asal sekolah IPA. Begitu juga dengan *cluster* 3 dan 4 dimana peminatnya berasal sekolah MA dengan jurusan asal sekolah IPA, Berikut prosentase pola minat mahasiswa pada program studi IAIN Salatiga yang dikelompokkan dalam tiga *cluster* :

Tabel 5.1 tabel

CLUSTER 1	CLUSTER 2	CLUSTER 3	CLUSTER 4	CLUSTER 5
1111	668	351	207	129

**Gambar 5.1** prosentase minat mahasiswa terhadap program studi IAIN

Salatiga

Dari program studi yang sudah di kelompokkan menjadi lima cluster tersebut mahasiswa yang meminati berasal sekolah dan jurusan sekolah yang sama, yaitu berasal sekolah MA dengan jurusan IPA.

- Setelah dilakukan eksperimen perbandingan menggunakan 3 cluster, 4 cluster dan 5 cluster data yang diolah menunjukkan hasil yang sama yaitu berasalkan sekolah MA dengan jurusan asal sekolah IPA, jadi untuk

sosialisasi kedepan IAIN Salatiga sasaran promosi dapat dilakukan pada sekolah MA dengan jurusan asal sekolah IPA.

5.2. Saran

Dalam penelitian tentunya terdapat kekurangan, yang pasti dapat di perbaiki dalam penelitian selanjutnya. Bagi yang akan meneruskan penelitian ini atau melakukan penelitian baru dengan metode yang sama, penulis memberikan saran sebagai berikut:

1. Pada pengolahan data dengan menggunakan aplikasi yang peneliti rancang perlu pengembangan yang lebih baik, supaya pemrosesan lebih cepat ketika data yang di olah berjumlah banyak, dan perlu di tambahkan fitur lebih banyak dalam pengolahan data.
2. Untuk penelitian selanjutnya dapat menggunakan algoritma yang lain untuk melihat perbandingan dengan menggunakan k-means clustering dengan menggunakan metode lain.
3. Untuk objek penelitian bisa menggunakan dua tempat, kampus umum dan kampus yang berbasis agama, dengan begitu akan terlihat apakah ada kesamaan atau perbedaan minat mahasiswa berdasarkan jenjang pendidikan sebelum kuliah.

DAFTAR PUSTAKA

- Andi. 2007. *Aplikasi Manajemen Database Pendidikan Berbasis Web Dengan PHP Dan MySQL*, Yogyakarta:andi.
- Bahra, Al. 2004. *Konsep Sistem Basis Data Dan Implementasinya*, yogyakarta:Graha Ilmu.
- Dwiguna, Hanif. 2015. *Mengenal Relasi Tabel Database*, <http://www.belajarphp lengkap.com/2015/12/relasi-tabel-database.html>, (diakses pada tanggal 09 Oktober 2017).
- Febrian, Nopal. 2017. *Penerapan Data Mining Untuk Menentukan Pengelompokan Uang Kuliah Tunggal (UKT) UIN Sunan Kalijaga Menggunakan Algoritma K-Means Clustering*, Yogyakarta:UIN Sunan Kalijaga Yogyakarta.
- Fusna. 2014. *Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma Apriori Pada Data Transaksi Penjualan (studi Kasus: Pemella Supermarket)*, Yogyakarta:UIN Sunan Kalijaga Yogyakarta.
- IAIN Salatiga. 2017. Visi dan misi <http://iainsalatiga.ac.id/web/about/visi-dan-misi/>, (Diakses pada tanggl 19 maret 2017).
- Kusrini dan Luthfi, Emha Taufiq. 2009. *Algoritma Data Mining*, Yogyakarta:andi.
- Loonam, Brady. 2010. *Pengertian Entity Relationship Diagram (ERD)* <http://www.pendidikanku.org/2016/07/pengertian-entity-relationship-diagram.html>, (diakses pada tanggal 09 Oktober 2017).
- Pan-Ning-Tan dan Steinbach, Michael. 2006. *Introduction To Data Mining*, New York:Pearson

- Pramudiono, I. 2006. *Apa itu Data Mining?* Dalam <http://datamining.japati.net/cgi-bin/indodm.cgi>. Diakses tanggal 26 maret 2016.
- Pratama, Yoga. 2016. *Penerapan Data Mining Untuk Memprediksi Klasifikasi Jumlah Pembaca Sebuah Artikel Pada Situs Beranda.co.id menggunakan algoritma bayesian classification*, Yogyakarta:UIN Sunan Kalijaga Yogyakarta.
- Puspitosari, Heni A. 2011. *Pemrograman Web Database Dengan PHP Dan MySQL Tingkat Mahir*, Yogyakarta:Skripta Media Creative.
- Santoso, Singgih. 2003. *Microsoft Excel: Mengolah Data Secara Profesional*, Yogyakarta:Andi.
- Sianipar, Pandapotan. 2010. *Cara Mudah Menggunakan Microsoft Excel 2010*, Jakarta:Gramedia.
- Taufik, Teuku. 2016. *Pengertian dan Simbol DFD (Data Flow Diagram)*, <http://www.teukutaufig.com/2016/02/pengertian-dan-simbol-dfd-data-flow.html> (Diakses pada tanggal 09 Oktober 2017).
- Turban, E., dkk. 2005. *Decision Support Systems and Intelligent Systems*. Yogyakarta: Andi Offset.
- Waljiyanto. 2003. *Sistem Basis Data : Analisis Dan Pemodelan Data*, Yogyakarta:Graha Ilmu.
- Wijanarko, Andi. 2016. *Penerapan Data Mining Untuk Menentukan Strategi Promosi Universitas PGRI Yogyakarta Menggunakan Algoritma K-means Clustering*, Yogyakarta:UIN Sunan Kalijaga Yogyakarta.

LAMPIRAN

Lampiran source code inde.php

```
1 <!DOCTYPE html>
2 <!-- Halaman Awal aplikasi K-Means Clustering-->
3 <html>
4 <head>
5 <meta charset="UTF-8">
6 <title>Skripsi K-Means</title>
7 <link rel="stylesheet" href="css/style.css" type="text/css">
8 </head>
9 <body>
10 <div id="header">
11 <div id="logo">
12 
13 </div>
14 <div id="navigation">
15 <li class="selected">
16 <a href="index.php">HOME</a>
17 </li>
18 <li>
19 <a href="about.php">ABOUT</a>
20 </li>
21 <li>
22 <a href="kmeans.php">k-means</a>
23 </li>
24 </div>
25 </div>
26 </div>
27 <div id="contents">
28 <div class="body">
29 <div id="about">
30  <span class="info" /> </span>
31 <div class="details">
32 <h3>IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING UNTUK MENGETAHUI POLA PENILIHAN PROGRAM STUDI MAHASISWA BARU IAIN
33 SALATIGA</h3>
34 <p>
35 Ahmad Choerudin Wahid </p>
36 </div>
37 </div>
38 </div>
39 <div id="featured">
40 <div>
41  <a href="#"></a>
42 </div>
43 <div class="details">
44 <h3>Data Mining</h3>
45 </div>
46 </div>
47 </div>
```

Lampiran source code about.php

```
1 <!DOCTYPE html>
2 <!-- Halaman Awal aplikasi K-Means Clustering-->
3 <html>
4 <head>
5 <meta charset="UTF-8">
6 <title>about</title>
7 <link rel="stylesheet" href="css/style.css" type="text/css">
8 </head>
9 <body>
10 <div id="header">
11 <div id="logo">
12 
13 </div>
14 <div id="navigation">
15 <li>
16 <a href="index.php">Home</a>
17 </li>
18 <li class="selected">
19 <a href="about.php">About</a>
20 </li>
21 <li>
22 <a href="kmeans.php">k-means</a>
23 </li>
24 </div>
25 </div>
26 </div>
27 <div id="contents">
28 <div class="body" id="about">
29 <div id="about">
30 <h3>IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING UNTUK MENGETAHUI POLA PENILIHAN PROGRAM STUDI MAHASISWA BARU IAIN SALATIGA</h3>
31 <p>
32 Latar Belakang</p>
33 <p>
34 Data mining merupakan proses untuk mendapatkan informasi dari basis data yang besar dengan cara penggolongan atau penyaringan data dengan memanfaatkan kumpulan data. Data mining dapat di gunakan untuk paparan tinggi untuk mengolah data-data yang ada pada perguruan tinggi seperti data mahasiswa, karyawan, dosen dan lain sebagainya. dan pada perguruan tinggi akan membuka penerimaan mahasiswa baru pada tiap tahunnya dan akan menghasilkan data-data yang baru, hal tersebut akan terjadi pada tiap tahunnya dan akan menyipkan banyak data seperti nama, alamat, asal sekolah, dan lain sebagainya. Tidak semua perguruan tinggi memanfaatkan banyaknya data mahasiswa tersebut solusi hanya untuk administrasi. Dengan menggunakan data mining data data yang begitu banyak tersebut dapat di jadikan sebagai hal yang positive bagi perguruan tinggi tersebut.
35 </p>
36 <p>
37 Rumusan Masalah</p>
38 <p>
39 Berdasarkan latar belakang di atas dapat dirumuskan permasalahan yang akan diseleksi dalam penelitian ini adalah : </p>
40 <ol>
41 <li>1. Bagaimana menerapkan algoritma k-means clustering untuk pola
42 pemilihan program studi mahasiswa baru di IAIN Salatiga</li>
43 <li>2. Bagaimana pola minat mahasiswa masing-masing terhadap program
44 studi di IAIN Salatiga</li>
45 </ol>
46 </div>
47 </div>
```

Lampiran source code kmeans.php

```
kmeans.php
30 <div class="body" id="contact">
31 <div id="main">
32 <h3>SELAKSAN IMPORT FILE EXCEL</h3>
33 <php
34 //koneksi ke database, username password dan namadatabase menyesuaikan
35 mysql_connect('localhost', 'root', '');
36 mysql_select_db('skripsi2');
37
38
39 //memanggil file excel_reader
40 require "excel_reader.php";
41
42 //jika tombol import ditekan
43 if(isset($_POST['submit'])){
44
45     truncate="TRUNCATE TABLE dataset";
46     mysql_query($truncate);
47
48     $target = basename($_FILES['filepegawai1']['name']);
49     move_uploaded_file($_FILES['filepegawai1']['tmp_name'], $target);
50
51     $data = new Spreadsheet_Excel_Reader($_FILES['filepegawai1']['name'], false);
52
53     //menghitung jumlah baris file xls
54     $baris = $data->rowcount($sheet_index=0);
55
56     //import data excel mulai baris ke-2 (karena tabel xls ada header pada baris 1)
57     for ($i=2; $i<=$baris; $i++)
58     {
59         //membaca data (kolom ke-1 id terasbir)
60         $snim = $data->val($i, 1);
61         $prodi = $data->val($i, 2);
62         $asal_sekolah = $data->val($i, 3);
63         $jurusan_sekolah = $data->val($i, 4);
64
65         //setelah data dibaca, masukkan ke tabel program di
66         $query = "insert into dataset (idn_prodi, asal_sekolah, jurusan_sekolah) values ('$snim', '$prodi', '$asal_sekolah', '$jurusan_sekolah')";
67         $hasil = mysql_query($query); ini_set('max_execution_time', 30000);
68     }
69
70     if($hasil){
71         //jika import sukses
72
73         echo<script language="JAVASCRIPT">;
74         echo "Alert('Data berhasil Di Importkan')";
75         echo</SCRIPT>;
76     }
77 }
```

Lampiran source code frek.php

```
frek.php
18 <div id="logo">
19 <a href="index.php"></a>
20 </div>
21 <div id="navigation">
22 <div>
23 <a href="index.php">HOME</a>
24 </div>
25 <div>
26 <a href="about.php">ABOUT</a>
27 </div>
28 <div class="select">
29 <a href="contact.php">k-means</a>
30 </div>
31 </div>
32 </div>
33 <div id="contents">
34 <div>
35 <div class="body" id="contact">
36 <div id="main">
37 <h3>TAHAP TRANSFORMASI DATA</h3>
38 <a href="dataset.php?man-manual"><button> PROSES SELANJUTNYA </button></a>
39 </div>
40
41 <div>
42 <php $frek = frekuensi('frek_jurusan', 'jurusan_sekolah', 'injur'); $frek2 = frekuensi('frek_prodi', 'prodi', 'inprodi'); $freks =
43 frekuensi('frek_asking', 'asal_sekolah', 'inlasko'); ?>
44 </div>
45 <div>
46 <table border="1" width="500px" id="kiri"> <tr style="background-color : green"><th>frek_prodi, 'prodi', 'PRODI', 'inprodi'; ?> </tr>
47 </div>
48 <div>
49 <table border="1" width="500px" id="kanan"><th>frek_jurusan, 'jurusan_sekolah', 'JURUSAN SEKOLAH', 'injur'; ?> </table>
50 </div>
51 <div>
52 <table border="1" width="500px" id="kanan"><th>frek_asking, 'asal_sekolah', 'ASAL SEKOLAH', 'inlasko'; ?> </table>
53 </div>
54 </div>
55 <div id="footer">
56 <div class="navigation">
57 <a href="index.php">HOME</a>
58 </div>
59 <div>
60 <a href="about.php">ABOUT</a>
61 </div>
```


Lampiran source code dataset.php

[illegible]

Lampiran source code mencari inisial disetiap prodi dan jurusan asal

```
function vlookup($prodi,$table,$sub_tabel,$col){
```

Lampiran source code mencari nama id untuk tabel dataset

```
$res = mysql_query("SELECT * FROM $table");
$id = mysql_field_name($res,0);
```

Source code menghitung frekuensi prodi + inisial

```
$inisial=1;
    $sql = "SELECT $berdasarkan, COUNT($berdasarkan) as frekuensi FROM dataset GROUP BY $berdasarkan ORDER
BY frekuensi DESC , $berdasarkan ASC";
    $result = mysql_query($sql);
    while($row = mysql_fetch_array($result)){
```

```

        $frekuensi = $row['frekuensi'];
        $val      = $row[$berdasarkan];
        $otomatis = $frekuensi*$inisial;

        $q = "INSERT INTO $tabel ($iniapa,$berdasarkan,frekuensi,otomatis) VALUES
        ($inisial,$val,$frekuensi,$otomatis)";
        $r = mysql_query($q);

        $inisial++;
    }
}

```

Source code menampilkan berdasarkan frekuensi terbesar ke terkecil

```

function inisial($tabel,$atribut,$atr,$iniapa){

    echo "<tr>
        <th align='center'>NO</th>
        <th align='center'>$atr</th>
        <th align='center'>FREKUENSI</th>
        <th align='center'>INISIAL</th>
    </tr>";

    $no = 1;

    $sql = "SELECT * FROM $tabel ORDER BY frekuensi DESC , $iniapa ASC";

    $result = mysql_query($sql);
    while($row = mysql_fetch_array($result)){
        $frekuensi = $row['frekuensi'];
        $val      = $row[$atribut];
        $inisial  = $row[$iniapa];

        echo "<tr>
            <td align='center'>$no</td>
            <td>$val</td>
            <td align='center'>$frekuensi</td>
            <td align='center'>$inisial</td>
        </tr>";

        $no++;
    }
}

```

Lampiran source code menampilkan dataset yang akan diolah

```

function dataset(){

    echo "<tr>
        <th align='center' bgcolor='blue'>Nim</th>
        <th align='center' bgcolor='blue'>PRODI</th>
        <th align='center' bgcolor='blue'>INISIAL</th>
        <th align='center' bgcolor='blue'>ASAL SEKOLAH</th>
        <th align='center' bgcolor='blue'>INISIAL</th>
        <th align='center' bgcolor='blue'>JURUSAN</th>
        <th align='center' bgcolor='blue'>INISIAL</th>
    </tr>";
}

```

```

        $sno = 1;
        $inisial = 1;

        $sql = "SELECT * FROM dataset";

        $result = mysql_query($sql);
        while($row = mysql_fetch_array($result)){
            $nim = $row['NIM'];
            $prodi = $row['PRODI'];
            $asko = $row['ASAL_SEKOLAH'];
            $jurusan = $row['JURUSAN_SEKOLAH'];
            $siniprodi = vlookup($prodi,'frek_prodi','prodi',0);
            $siniasko = vlookup($asko,'frek_asko','asal_sekolah',0);
            $sinijur = vlookup($jurusan,'frek_jurusan','jurusan_sekolah',0);

            echo "<tr>
                <form method='POST' align='center'>
                    <td> $nim. </td>
                    <td> $prodi </td>
                    <td align='center'> $siniprodi </td>
                    <td> $asko </td>
                    <td align='center'> $siniasko </td>
                    <td> $jurusan </td>
                    <td align='center'> $sinijur </td>
                </form>
            </tr>";

            $sno++;
        }
    }
}

$inisial++;

```

Source code titik pusat cluster secara random

```

function random($kode){
    if($kode == '1'){
        $random = rand(1,5);
    }elseif($kode == '2'){
        $random = rand(1,2);
    }elseif($kode == '3'){
        $random = rand(1,3);
    }elseif($kode == '21'){
        $random = rand(5,10);
    }elseif($kode == '22'){
        $random = rand(2,3);
    }elseif($kode == '23'){
        $random = rand(3,6);
    }elseif($kode == '31'){
        $random = rand(9,11);
    }elseif($kode == '32'){
        $random = rand(3,4);
    }elseif($kode == '33'){
        $random = rand(6,7);
    }elseif($kode == '41'){
        $random = rand(12,15);
    }elseif($kode == '42'){
        $random = rand(4,5);
    }elseif($kode == '43'){
        $random = rand(8,12);
    }elseif($kode == '51'){
        $random = rand(16,20);
    }elseif($kode == '52'){
        $random = rand(5,5);
    }
}

```

```

}elseif($kode == '53'){
$random = rand(13,15);
}else{

}
return $random;
}

```

Lampiran source code memberi nomor proses

```

function nomorproses($proses){

    $q = "INSERT INTO nomor (nomor) VALUES ($proses)";
    $r = mysql_query($q);

    $sql = "SELECT MAX(id_nomor) as proses FROM nomor";
    $result = mysql_query($sql);
    $row = mysql_fetch_array($result);

    $p = $row['proses'];

    echo "$p";

}

```

Source code proses k-means

```

Function
kmeans($prodi1,$asal_sekolah1,$jurusan1,$prodi2,$asal_sekolah2,$jurusan2,$prodi3,$asal_sekolah3,$jurusan3,$prodi4,$asal_sekolah4,$jurusan4,$prodi5,$asal_sekolah5,$jurusan5){

    $truncate ="TRUNCATE TABLE proses";
    mysql_query($truncate);

    $pro1 = number_format($prodi1,5); $asko1 = number_format($asal_sekolah1,5); $jur1 = number_format($jurusan1,5);
    $pro2 = number_format($prodi2,5); $asko2 = number_format($asal_sekolah2,5); $jur2 = number_format($jurusan2,5);
    $pro3 = number_format($prodi3,5); $asko3 = number_format($asal_sekolah3,5); $jur3 = number_format($jurusan3,5);
    $pro4 = number_format($prodi4,5); $asko4 = number_format($asal_sekolah4,5); $jur4 = number_format($jurusan4,5);
    $pro5 = number_format($prodi5,5); $asko5 = number_format($asal_sekolah5,5); $jur5 = number_format($jurusan5,5);

    echo "
    <tr>
      <th rowspan='2' align='center' bgcolor='94d03f'> NIM </th>
      <th rowspan='2' align='center' bgcolor='94d03f'> PRODI </th>
      <th rowspan='2' align='center' bgcolor='94d03f'> ASAL SEKOLAH </th>
      <th rowspan='2' align='center' bgcolor='94d03f'> JURUSAN </th>
      <th colspan='3' align='center' bgcolor='94d03f'>CENTROID C1</th>
      <th colspan='3' align='center' bgcolor='94d03f'>CENTROID C2</th>
      <th colspan='3' align='center' bgcolor='94d03f'>CENTROID C3</th>
      <th colspan='3' align='center' bgcolor='94d03f'>CENTROID C4</th>
      <th colspan='3' align='center' bgcolor='94d03f'>CENTROID C5</th>
      <th rowspan='2' align='center' bgcolor='94d03f'>C1</th>
      <th rowspan='2' align='center' bgcolor='94d03f'>C2</th>
      <th rowspan='2' align='center' bgcolor='94d03f'>C3</th>
      <th rowspan='2' align='center' bgcolor='94d03f'>C4</th>
      <th rowspan='2' align='center' bgcolor='94d03f'>C5</th>
    </tr>
    <tr>
      <th bgcolor='b3fb4d'>$pro1</th>
      <th bgcolor='b3fb4d'>$asko1</th>
      <th bgcolor='b3fb4d'>$jur1</th>
      <th bgcolor='b3fb4d'>$pro2</th>
      <th bgcolor='b3fb4d'>$asko2</th>
      <th bgcolor='b3fb4d'>$jur2</th>
      <th bgcolor='b3fb4d'>$pro3</th>

```

```

<th bgcolor='b3fb4d'>$asko3</th>
<th bgcolor='b3fb4d'>$jur3</th>
<th bgcolor='b3fb4d'>$pro4</th>
<th bgcolor='b3fb4d'>$asko4</th>
<th bgcolor='b3fb4d'>$jur4</th>
<th bgcolor='b3fb4d'>$pro5</th>
<th bgcolor='b3fb4d'>$asko5</th>
<th bgcolor='b3fb4d'>$jur5</th>
</tr>";

```

```
$no = 1;
```

```
$sql = "SELECT * FROM dataset";
```

```
$result = mysql_query($sql);
```

```
while($row = mysql_fetch_array($result)){
```

```
    $nim = $row['NIM'];
```

```
    $prodi = $row['PRODI'];
```

```
    $asko = $row['ASAL_SEKOLAH'];
```

```
    $jurusan = $row['JURUSAN_SEKOLAH'];
```

```
    $iniprodi = vlookup($prodi,'frek_prodi','prodi',0);
```

```
    $iniasko = vlookup($asko,'frek_asko','asal_sekolah',0);
```

```
    $inijur = vlookup($jurusan,'frek_jurusan','jurusan_sekolah',0);
```

```
    $jarak1 = distance($prodi1,$asko1,$jurusan1,$iniprodi,$iniasko,$inijur);
```

```
    $jarak2 = distance($prodi2,$asko2,$jurusan2,$iniprodi,$iniasko,$inijur);
```

```
    $jarak3 = distance($prodi3,$asko3,$jurusan3,$iniprodi,$iniasko,$inijur);
```

```
    $jarak4 = distance($prodi4,$asko4,$jurusan4,$iniprodi,$iniasko,$inijur);
```

```
    $jarak5 = distance($prodi5,$asko5,$jurusan5,$iniprodi,$iniasko,$inijur);
```

```
    $clustering = clustering($jarak1,$jarak2,$jarak3,$jarak4,$jarak5);
```

```
    $clustering2 = clustering2($jarak1,$jarak2,$jarak3,$jarak4,$jarak5);
```

```
    $clustering3 = clustering3($jarak1,$jarak2,$jarak3,$jarak4,$jarak5);
```

```
    $clustering4 = clustering4($jarak1,$jarak2,$jarak3,$jarak4,$jarak5);
```

```
    $clustering5 = clustering5($jarak1,$jarak2,$jarak3,$jarak4,$jarak5);
```

```
echo "<tr>
```

```
    <form method='POST' align='center'>
```

```
        <td align='center'>$nim </td>
```

```
        <td align='center'>$iniprodi</td>
```

```
        <td align='center'>$iniasko</td>
```

```
        <td align='center'>$inijur</td>
```

```
        <td align='center' colspan='3'>$jarak1</td>
```

```
        <td align='center' colspan='3'>$jarak2</td>
```

```
        <td align='center' colspan='3'>$jarak3</td>
```

```
        <td align='center' colspan='3'>$jarak4</td>
```

```
        <td align='center' colspan='3'>$jarak5</td>
```

```
        <td align='center'>$clustering</td>
```

```
        <td align='center'>$clustering2</td>
```

```
        <td align='center'>$clustering3</td>
```

```
        <td align='center'>$clustering4</td>
```

```
        <td align='center'>$clustering5</td>
```

```
    </form>
```

```
</tr>";
```

```
$q = "INSERT INTO proses
```

```
(iniprodi,iniasko,inijur,clustersatu,clusterdua,clustertiga,clusterempat,clusterlima) VALUES
```

```
('iniprodi','iniasko','inijur','$clustering','$clustering2','$clustering3','$clustering4','$clustering5')";
```

```
$r = "INSERT INTO centroid_lama (centroid_lama) VALUES
```

```
('prodi1'),('asal_sekolah1'),('jurusan1'),('prodi2'),('asal_sekolah2'),('jurusan2'),('prodi3'),('asal_sekolah3'),('jurusan3'),('prodi4'),('asal_sekolah4'),('jurusan4'),('prodi5'),('asal_sekolah5'),('jurusan5')";
```

```
$s = mysql_query($q);
```

```
$t = mysql_query($r);
```

```
$no++;
```

```
}
```

```
}
```

Source code menghitung jarak data ke cluster

```
function distance($p,$a,$j,$iniprodu,$iniasko,$inijur){  
  
    $jarak = SQRT( (pow(($iniprodu - $p),2)) + (pow(($iniasko - $a),2)) + (pow(($inijur - $j),2)) );  
    return number_format ($jarak,9);  
  
}
```

Source code mengelompokkan data kedalam cluster berdasarkan jarak terdekat data dengan pusat cluster (centroid)

```
function clustering($a,$b,$c,$d,$e){  
  
    IF ($a < $b AND $a < $c AND $a < $d AND $a < $e){  
        return "v";  
    }if ($b < $c AND $b < $a AND $b < $d AND $b < $e) {  
        return "";  
    } if ($c < $a AND $c < $b AND $c < $d AND $c < $e) {  
        return "";  
    } if ($d < $a AND $d < $b AND $d < $c AND $d < $e) {  
        return "";  
    } if ($e < $a AND $e < $b AND $e < $d AND $e < $d) {  
        return "";  
    } else {  
        return "fail";  
    }  
}  
function clustering2($a,$b,$c,$d,$e){  
  
    IF ($a < $b AND $a < $c AND $a < $d AND $a < $e){  
        return "";  
    }if ($b < $c AND $b < $a AND $b < $d AND $b < $e) {  
        return "v";  
    } if ($c < $a AND $c < $b AND $c < $d AND $c < $e) {  
        return "";  
    } if ($d < $a AND $d < $b AND $d < $c AND $d < $e) {  
        return "";  
    } if ($e < $a AND $e < $b AND $e < $d AND $e < $d) {  
        return "";  
    } else {  
        return "fail";  
    }  
}  
function clustering3($a,$b,$c,$d,$e){  
  
    IF ($a < $b AND $a < $c AND $a < $d AND $a < $e){  
        return "";  
    }if ($b < $c AND $b < $a AND $b < $d AND $b < $e) {  
        return "";  
    } if ($c < $a AND $c < $b AND $c < $d AND $c < $e) {  
        return "v";  
    } if ($d < $a AND $d < $b AND $d < $c AND $d < $e) {  
        return "";  
    } if ($e < $a AND $e < $b AND $e < $d AND $e < $d) {  
        return "";  
    } else {  
        return "fail";  
    }  
}
```

```

function clustering4($a,$b,$c,$d,$e){

    IF ($a < $b AND $a < $c AND $a < $d AND $a < $e){
        return "";
    }if ($b < $c AND $b < $a AND $b < $d AND $b < $e) {
        return "";
    } if ($c < $a AND $c < $b AND $c < $d AND $c < $e) {
        return "";
    } if ($d < $a AND $d < $b AND $d < $c AND $d < $e) {
        return "v";
    } if ($e < $a AND $e < $b AND $e < $d AND $e < $d) {
        return "";
    } else {
        return "fail";
    }
}

function clustering5($a,$b,$c,$d,$e){

    IF ($a < $b AND $a < $c AND $a < $d AND $a < $e){
        return "";
    }if ($b < $c AND $b < $a AND $b < $d AND $b < $e) {
        return "";
    } if ($c < $a AND $c < $b AND $c < $d AND $c < $e) {
        return "";
    } if ($d < $a AND $d < $b AND $d < $c AND $d < $e) {
        return "";
    } if ($e < $a AND $e < $b AND $e < $d AND $e < $d) {
        return "v";
    } else {
        return "fail";
    }
}

```

Lampiran source code mencari cluster baru (rata-rata)

```

function rata_rata($colom,$cluster){

    $sql = "SELECT AVG($colom) as rata_rata FROM proses WHERE $cluster='v' ";
    $result = mysql_query($sql);
    $row = mysql_fetch_array($result);
    $val = $row['rata_rata'];
    return $val;

}

```

Lampiran source code mencari pusat (centroid) baru disetiap cluster

```

function newcluster(){

    //CLUSTER1
    $prodi1 = rata_rata('iniprodui','clustersatu'); $asko1 = rata_rata('iniasko','clustersatu'); $jur1 = rata_rata('inijur','clustersatu');
    //CLUSTER2
    $prodi2 = rata_rata('iniprodui','clusterdua'); $asko2 = rata_rata('iniasko','clusterdua'); $jur2 = rata_rata('inijur','clusterdua');
    //CLUSTER3
    $prodi3 = rata_rata('iniprodui','clustertiga'); $asko3 = rata_rata('iniasko','clustertiga'); $jur3 = rata_rata('inijur','clustertiga');
    //CLUSTER4
    $prodi4 = rata_rata('iniprodui','clusterempat'); $asko4 = rata_rata('iniasko','clusterempat'); $jur4 = rata_rata('inijur','clusterempat');
    //CLUSTER5
    $prodi5 = rata_rata('iniprodui','clusterlima'); $asko5 = rata_rata('iniasko','clusterlima'); $jur5 = rata_rata('inijur','clusterlima');
}

```



```

echo" <table border='1' width='210px'>
  <form name='newcluster' id='newcluster' action='proses.php' method='post' >
    <tr align='center'> <td colspan='2'>CLUSTER 1 </td></tr>
    <tr>
      <td> PRODI</td>
      <td> <input type='decimal' name='prodi1' value='$prodi1' hidden > $prodi1 </td>
    </tr>
    <tr>
      <td> ASAL SEKOLAH </td>
      <td> <input type='decimal' name='asko1' value='$asko1' hidden > $asko1 </td>
    </tr>
    <tr>
      <td> JURUSAN SEKOLAH</td>
      <td> <input type='decimal' name='jurusan1' value='$jur1' hidden > $jur1 </td>
    </tr>
    <br/>

    <tr align='center'> <td colspan='2'>CLUSTER 2 </td></tr>
    <tr>
      <td> PRODI</td>
      <td> <input type='decimal' name='prodi2' value='$prodi2' hidden > $prodi2 </td>
    </tr>
    <tr>
      <td> ASAL SEKOLAH </td>
      <td> <input type='decimal' name='asko2' value='$asko2' hidden > $asko2 </td>
    </tr>
    <tr>
      <td> JURUSAN SEKOLAH</td>
      <td> <input type='decimal' name='jurusan2' value='$jur2' hidden > $jur2 </td>
    </tr>
    <br/>

    <tr align='center'> <td colspan='2'>CLUSTER 3 </td></tr>
    <tr>
      <td> PRODI</td>
      <td> <input type='decimal' name='prodi3' value='$prodi3' hidden > $prodi3 </td>
    </tr>
    <tr>
      <td>ASAL SEKOLAH</td>
      <td> <input type='decimal' name='asko3' value='$asko3' hidden > $asko3 </td>
    </tr>
    <tr>
      <td> JURUSAN SEKOLAH</td>
      <td> <input type='decimal' name='jurusan3' value='$jur3' hidden > $jur3 </td>
    </tr>
    <br/>

    <tr align='center'> <td colspan='2'>CLUSTER 4 </td></tr>
    <tr>
      <td> PRODI</td>
      <td> <input type='decimal' name='prodi4' value='$prodi4' hidden > $prodi4 </td>
    </tr>
    <tr>
      <td>ASAL SEKOLAH</td>
      <td> <input type='decimal' name='asko4' value='$asko4' hidden > $asko4 </td>
    </tr>
    <tr>
      <td> JURUSAN SEKOLAH</td>
      <td> <input type='decimal' name='jurusan4' value='$jur4' hidden > $jur4 </td>
    </tr>
    <br/>

    <tr align='center'> <td colspan='2'>CLUSTER 5 </td></tr>
    <tr>

```

```

        <td> PRODI</td>
        <td> <input type='decimal' name='prodi5' value='$prodi5' hidden > $prodi5 </td>
    </tr>
    <tr>
        <td>ASAL SEKOLAH</td>
        <td> <input type='decimal' name='asko5' value='$asko5' hidden > $asko5 </td>
    </tr>
    <tr>
        <td> JURUSAN SEKOLAH</td>
        <td> <input type='decimal' name='jurusan5' value='$jur5' hidden > $jur5 </td>
    </tr>
    <br/>
</table> ";

    $cek =
verifikasi($prodi1,$asko1,$jur1,$prodi2,$asko2,$jur2,$prodi3,$asko3,$jur3,$prodi4,$asko4,$jur4,$prodi5,$asko5,$jur5);
    if ($cek == 'Lanjut') {
        echo " <br />
        <td colspan='4'> <input type='submit' name='submit' value='Next Proses' /> </td>
        </form>";
    } elseif ($cek == 'stop') {
        echo " <br />
        <td colspan='4'> <h2> <strong> PROSES SELESAI !! </strong> </h2> </td>
        <h1><span class='text1'>hasil proses k-means</span> <br/> <input type='submit' name='hasil'
value='HASIL' /></h1>";
    }
}

```

Lampiran source code memverifikasi apakah centroid baru yg didapatkan sama atau beda dengan centroid yang lama

```

function verifikasi($a,$b,$c,$a2,$b2,$c2,$a3,$b3,$c3,$a4,$b4,$c4,$a5,$b5,$c5) {

    $resultsatu = mysql_query("SELECT * FROM centroid_lama");
    $num_rowsatu = mysql_num_rows($resultsatu);
    $pertama = $num_rowsatu;

    $q = "INSERT IGNORE INTO centroid_lama (centroid_lama) VALUES
($a),($b),($c),($a2),($b2),($c2),($a3),($b3),($c3),($a4),($b4),($c4),($a5),($b5),($c5)";
    $cek = mysql_query($q);

    $resultkedua = mysql_query("SELECT * FROM centroid_lama");
    $num_rowskedua = mysql_num_rows($resultkedua);
    $kedua = $num_rowskedua;

    if ($pertama < $kedua) {
        return "Lanjut";
    }else{
        return "stop";
    }
}

} function verifikasi($a,$b,$c,$a2,$b2,$c2,$a3,$b3,$c3,$a4,$b4,$c4,$a5,$b5,$c5) {

    $resultsatu = mysql_query("SELECT * FROM centroid_lama");
    $num_rowsatu = mysql_num_rows($resultsatu);
    $pertama = $num_rowsatu;

    $q = "INSERT IGNORE INTO centroid_lama (centroid_lama) VALUES
($a),($b),($c),($a2),($b2),($c2),($a3),($b3),($c3),($a4),($b4),($c4),($a5),($b5),($c5)";
    $cek = mysql_query($q);

```

```

$resultkedua = mysql_query("SELECT * FROM centroid_lama");
$num_rowskedua = mysql_num_rows($resultkedua);
$kedua = $num_rowskedua;

if ($pertama < $kedua) {
    return "Lanjut";
}else{
    return "stop";
}
}

```

Lampiran source code hasil dari proses k_means

```

function hasil($sub_tabel,$cluster,$tabel_vlookup,$iniapa,$cluster_x){
$sql2 = "SELECT $sub_tabel, COUNT($sub_tabel) as jumlah FROM proses WHERE $cluster = 'v' GROUP BY $sub_tabel
ORDER BY jumlah DESC ";
$result2 = mysql_query($sql2);
while($row2 = mysql_fetch_array($result2)){

    $jumlah = $row2['jumlah'];
    $inisial = $row2[$sub_tabel];
    $val = vlookup($inisial,$tabel_vlookup,$iniapa,1);
    $sql3 = mysql_query("INSERT INTO grafik ( nama , jumlah, cluster) VALUES ('$val','$jumlah','$cluster_x');");
    echo "<tr>
        <form>
            <td> $val </td>
            <td align='center'> $jumlah </td>
        </form>
    </tr>";
}
}

```

Lampiran source code jumlah data mahasiswa setiap hasil cluster

```

function recordcluster($cluster){

$result = mysql_query("SELECT * FROM proses WHERE $cluster = 'v' ");
$num_rows = mysql_num_rows($result);
// Display the results
echo $num_rows;

}

```

Lampiran source code menghitung rata-rata setiap hasil cluster

```

function rata_rata_Asko($cluster){

$sql = "SELECT AVG(asko) as rata_rata FROM proses WHERE $cluster='v' ";
$result = mysql_query($sql);
$row = mysql_fetch_array($result);

$val = $row['rata_rata'];
echo $val;

}

```

CURICULUM VITAE

Identitas diri

Nama Lengkap : Akhmad Choerudin Wakhid
Tempat, Tanggal Lahir : Magelang, 15 April 1994
Kewarga Negara : Indonesia
Agama : Islam
Jenis Kelamin : Laki-laki
Email : Akhmadchoerudin4@gmail.com
Kontak : 0816696094



Riwayat Pendidikan

2000 - 2006 : MI AL-Ma'arif Kebumen, Banyubiru, Semarang
2006 - 2009 : MTs Negeri Ngablak, Magelang
2009 - 2012 : SMK Muhammadiyah 5 Karanganyar
2012 – 2017 : S1 Teknik Informatika UIN Sunan Kalijaga Yogyakarta

Pengalaman Organisasi

- ❖ Karisma (Keluarga Besar Mahasiswa Magelang)
- ❖ UKM (Unit Kegiatan Mahasiswa) Taekwondo
- ❖ HIMA (Himpunan Mahasiswa) Teknik Informatika