

***DATA MINING SELEKSI SISWA BERPRESTASI UNTUK
MENENTUKAN KELAS UNGGULAN MENGGUNAKAN
METODE K-MEANS CLUSTERING
(Studi Kasus di MTS Darul Fikri)***

SKRIPSI

Diajukan Sebagai Salah satu Syarat
Untuk Memperoleh Gelar Sarjana Jenjang Strata Satu (S1)
Pada Program Studi Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Ponorogo



**REYHAN PUTRA PRIMANDA
16532564**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH PONOROGO
2021**

HALAMAN PENGESAHAN

Nama : Reyhan Putra Primanda
NIM : 16532564
Program Studi : Teknik Informatika
Fakultas : Teknik
Judul Skripsi : **Data Mining seleksi siswa berprestasi untuk menentukan kelas unggulan menggunakan Metode K-Means Clustering (Studi kasus di MTs Darul Fikri)**

Isi dan Formatnya telah disetujui dan dinyatakan memenuhi syarat
Untuk melengkapi persyaratan guna memperoleh Gelar Sarjana
Pada Program Studi Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Ponorogo

Ponorogo, 26 Januari 2021

Menyetujui

Dosen Pembimbing 1



Dr. Aslan, S.Si, M.Cs
NIK. 19720324 20110113

Dosen Pembimbing 2



Dyah Mustikasari, ST, M.Eng
NIK. 19871007 201609 13

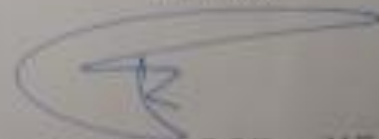
Mengetahui

Dekan Fakultas Teknik



Eby Kasmirwan, ST., MT
NIK. 19771026 200810 12

Ketua Program Studi Teknik
Informatika



Adi Fajaryanto C, S.Kom., M.Kom
NIK. 19840924 201309 13

PERNYATAAN ORISINILITAS SKRIPSI

Yang bertanda tangan dibawah ini :

Nama : Reyhan Putra Primanda
NIM : 16532564
Program Studi : Teknik Informatika

Dengan Ini menyatakan bahwa skripsi saya dengan judul "*Data mining seleksi siswa berprestasi untuk menentukan kelas unggulan menggunakan metode K-Means Clustering (Studi kasus di MTs Darul Fikri)*". Bahwa berdasarkan hasil penelusuran berbagai karya ilmiah, gagasan serta masalah ilmiah yang dirancang/teliti dalam Naskah Skripsi ini adalah asli dari pemikiran saya. Tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip dalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata dalam Naskah Skripsi ini dapat dibuktikan terdapat unsur unsur plagiatisme, saya bersedia ljasah saya dibatalkan, serta diproses sesuai dengan peraturan perundang undangan yang berlaku.

Demikian pernyataan ini dibuat dengan sesungguhnya dan dengan sebenar-benarnya.

Ponorogo, 26 Januari 2021

Mahasiswa,



Reyhan Putra Primanda

HALAMAN BERITA ACARA UJIAN

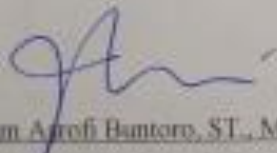
Nama : Reyhan Putra Primanda
NIM : 16532564
Program Studi : Teknik Informatika
Fakultas : Teknik
Judul Skripsi : *Data Mining* seleksi siswa berprestasi untuk menentukan kelas unggulan menggunakan *Metode K-Means Clustering* (Studi kasus di MTs Darul Fikri)

Telah diuji dan dipertahankan dihadapan
Dosen Penguji tugas akhir jenjang Strata (S1) pada

Hari : Senin
Tanggal : 25 Januari 2021
Nilai :

Dosen Penguji

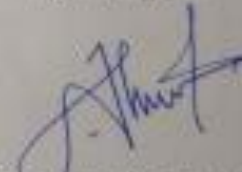
Dosen Penguji 1



(Gibulam Anrofi Hantoro, ST., M.Eng)

NIK.19870723 201603 13

Dosen Penguji 2



(Indah Pusp Astuti, S.Kom., M.Kom)

NIK.1986042420160913

Mengetahui

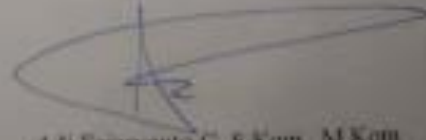
Dekan Fakultas Teknik



Edy Kurniawan, ST., MT

NIK.19771026 200810 12

Ketua Program Studi Teknik
Informatika




Adi Fajaryanto C, S.Kom., M.Kom

NIK.19840924 201309 13

**BERITA ACARA
BIMBINGAN SKRIPSI**

- 1 Nama : Reyhan Putra Primanda
 2 NIM : 16532564
 3 Program Studi : Teknik Informatika
 4 Fakultas : Teknik
 5 Judul Skripsi : *Data mining* seleksi siswa berprestasi untuk menentukan kelas unggulan menggunakan *Metode K-Means Clustering* (Studi kasus di MTs Darul Fikri)
 6 Dosen Pembimbing : Dr. Aslan, S.SI, M.Cs
 7 Konsultasi :
 8

NO	TANGGAL	URAIAN	TTD
1	19/5 2020	Penyusunan format skripsi	
2	26/5 2020	revisi Bab I	
3	1/6 2020	ACC Bab I, revisi Bab II	
4	3/7 2020	ACC Bab II, revisi Bab III	
5	4/7 2020	Penambahan Perhitungan.	
6	15/7 2020	ACC Bab III, revisi Bab IV	
7	25/7 2020	ACC Bab IV, revisi Bab V	
8	6/8 2020	ACC Bab V, ACC sidang	

- 9 Tgl Pengajuan :
 10 Tgl Pengesahan :

Ponorogo, 26 Januari 2021

Pembimbing I,



Dr. Aslan, S.SI, M.CS

NIK.19720324 201101 13

BERITA ACARA
BIMBINGAN SKRIPSI

- 1 Nama : Reyhan Putra Primanda
 2 NIM : 16532564
 3 Program Studi : Teknik Informatika
 4 Fakultas : Teknik
 5 Judul Skripsi : **Data Mining seleksi siswa berprestasi untuk menentukan kelas unggulan menggunakan Metode K-Means Clustering (Studi kasus di MTs Darul Fikri)**
 6 Dosen Pembimbing : Dyah Mustikasari, ST, M.Eng
 7 Konsultasi :
 8

NO	TANGGAL	URAIAN	TTD
1	22/5 2020	Penyusunan format skripsi	
2	28/5 2020	revisi Bab I	
3	17/6 2020	Acc Bab I, Revisi Bab II	
4	8/7 2020	Penambahan referensi	
5	15/7 2020	Acc Bab II, Acc Bab III	
6	28/7 2020	revisi Peruliisan Bab IV	
7	1/8 2020	Acc Bab III, revisi Bab I	
8	19/8 2020	Acc sidang	

- 9 Tgl Pengajuan :
 10 Tgl Pengesahan :

Ponorogo, 26 Januari 2021

Pembimbing II,



Dyah Mustikasari, ST, M.Eng

NIK.19871007 201609 13



UNIVERSITAS MUHAMMADIYAH PONOROGO
UNIT PELAKSANA TEKNIS PERPUSTAKAAN
Jalan Budi Utomo 10 Ponorogo 63471 Jawa Timur Indonesia
Telp (0352) 481124, 487662 Fax (0352) 461796, Website: library.umpo.ac.id
TERAKREDITASI A
(SK Nomor 00012/ LAP.PT/ L2017)

**SURAT KETERANGAN
HASIL PEMERIKSAAN PLAGIASI SKRIPSI MAHASISWA
UNIVERSITAS MUHAMMADIYAH PONOROGO**

Dengan ini kami nyatakan bahwa skripsi dengan rincian sebagai berikut:

Nama : Reyhan Putra Primanda

NIM : 16532564

Prodi : Teknik informatika

Judul : Data mining seleksi siswa berprestasi untuk menentukan kelas unggulan menggunakan metode k-means clustering di MTs Darul Fikri

Dosen pembimbing :

1. Dr. Aslan, S.SI, M.Cs

email : elangbijak4@gmail.com

2. Dyah Mustikasari, ST, M.Eng

email : dyah.mustikasari@gmail.com

Telah dilakukan check plagiasi di UPT. Perpustakaan Universitas Muhammadiyah Ponorogo dengan prosentase plagiasi sebesar 20 %

Demikian keterangan ini dibuat untuk digunakan sebagaimana mestinya.

Ponorogo, 13 Januari 2021

Pemeriksa

(Mohamad Ulit Albab, SIP)

NIK.1989092720150322

Nb: Dosen pembimbing dimohon untuk mengecek kembali keaslian soft file karya ilmiah yang telah dicek di perpustakaan.



UNIVERSITAS MUHAMMADIYAH PONOROGO
UNIT PELAKSANA TEKNIS PERPUSTAKAAN
Jalan Budi Utomo No. 10 Ponorogo 63471 Jawa Timur Indonesia
Telp. (0352) 481124, Fax (0352) 461796, e-mail : lib@umpo.ac.id
website : www.library.umpo.ac.id

**SURAT KETERANGAN
HASIL PEMERIKSAAN ANTI PLAGIASI ARTIKEL ILMIAH MAHASISWA
UNIVERSITAS MUHAMMADIYAH PONOROGO**

Telah di periksa, artikel ilmiah dengan rincian sebagai berikut :

Nama : Reyhan Putra Primanda
NIM : 16532564
Fakultas / Prodi : Teknik informatika
Judul : Data mining seleksi siswa berprestasi untuk menentukan kelas unggulan menggunakan metode k-means clustering di MTs Darul Fikri

Dosen pembimbing :

1. Dr. Aslan, S.Si, M.Cs email : elangbijak4@gmail.com
2. Dyah Mustikasari, ST, M.Eng email : dyah.mustikasari@gmail.com

Tingkat kesamaan pada artikel sebesar 19%

Menggunakan aplikasi anti-plagiasi *Turnitin*. Demikian, atas perhatiannya di ucapkan terima kasih.

Ponorogo, 05 Februari 2021

Pemeriksa



[Mohamad Ulil Albab, SIP]
NIK.1989092720150322

Nb: Dosen pembimbing dimohon untuk mengecek kembali keaslian soft file karya ilmiah yang telah dicek di perpustakaan.

MOTTO

Memulai dengan Penuh Keyakinan , Menjalankan dengan Penuh Keikhlasan, Menyelesaikan dengan Penuh Kebahagiaan.



KATA PENGANTAR

Segala puji bagi Allah SWT yang telah memberikan rahmat dan karuniaNya kepada penulis, sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Shalawat dan salam senantiasa tercurah kepada Rasulullah SAW yang mengantarkan manusia dari zaman kegelapan ke zaman yang terang benderang ini. Penyusunan skripsi ini dimaksudkan untuk memenuhi gelar sarjana Program Studi Teknik Informatika di Universitas Muhammadiyah Ponorogo

Penulis menyadari bahwa penulisan ini tidak dapat terselesaikan tanpa dukungan dari berbagai pihak baik moril maupun materil. Oleh karena itu, penulis ingin menyampaikan ucapan terimakasih kepada semua pihak yang telah membantu dalam penyusunan skripsi ini terutama :

1. Kedua orang tua yang telah memberikan dukungan baik moril maupun materil serta doa kepada penulis.
2. Bapak Edy Kurniawan, ST., MT Selaku Dekan fakultas teknik Universitas Muhammadiyah Ponorogo.
3. Adi Fajaryanto C, S.Kom., M.Kom, Selaku Ketua Program Studi Teknik Informatika Universitas Muhammadiyah Ponorogo
4. Bapak Dr. Aslan, S.SI, M.Cs Selaku dosen pembimbing Skripsi I yang telah berkenan memberikan tambahan ilmu dan solusi pada setiap permasalahan atas kesulitan dalam penulisan skripsi ini
5. Ibu Dyah Mustikasari, ST, M.Eng, selaku dosen Pembimbing Skripsi II yang telah bersedia membimbing dan mengarahkan penulis selama menyusun skripsi dan memberikan banyak ilmu serta solusi pada setiap permasalahan atas kesulitan dalam penulisan skripsi ini.
6. Seluruh Bapak/Ibu Dosen Program Studi Teknik Informatika yang telah memberikan pengetahuan yang sangat bermanfaat selama masa perkuliahan.
7. Seluruh teman teman seangkatan, terutama kelas Teknik Informatika Kelas A Angkatan 2016 yang telah membantu dalam terselesaikanya skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna dikarenakan terbatasnya pengalaman dan pengetahuan yang dimiliki penulis. Oleh karena itu, penulis

mengharapkan segala bentuk saran serta masukan bahkan kritik yang membangun dari beberapa pihak. Semoga skripsi ini dapat bermanfaat bagi para pembaca dan semua pihak

Ponorogo, 26 Januari 2021

Mahasiswa,

Reyhan Putra Primanda

NIM. 16532564



DAFTAR ISI

HALAMAN SAMPUL DEPAN.....	i
HALAMAN PENGESAHAN	ii
PERNYATAAN ORISINILITAS SKRIPSI.....	iii
HALAMAN BERITA ACARA UJIAN	iv
BERITA ACARA BIMBINGAN SKRIPSI.....	v
MOTTO	ix
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	vivv
DAFTAR LAMPIRAN.....	xvv
BAB I.....	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian.....	2
1.4 Batasan Masalah.....	2
1.5. Manfaat Penelitian.....	3
BAB II.....	4
KAJIAN TEORI	4
2.1 Penelitian Terdahulu	4
2.2 Data Mining.....	6
2.3 Clustering	11
2.4 Alogaritma K-Means.....	11
2.5. Microsoft Excel 2016.....	16
2.6. Pemograman R dan Rstudio.....	16
BAB III	26
METODE PENELITIAN.....	26
3.1 Metode Pengumpulan Data	26
3.2 Analisis Data	27

3.3 Tahapan penelitian	28
BAB IV	33
HASIL DAN PEMBAHASAN.....	33
4.1 Gambaran Umum Lokasi Penelitian	33
4.1.1 Profil Instansi	33
4.1.2 Visi dan Misi instansi.....	33
4.2 Hasil Penelitian	33
4.3 Pembahasan.....	34
4.3.1 Seleksi Data	34
4.3.2 Preprocessing.....	35
4.4 Penerapan perhitungan K-means menggunakan Excel	36
4.4.1 Langkah pertama menentukan pusat cluster (k)	36
4.4.2 Menghitung jarak terdekat antara centroid dengan data	36
4.4.3 Pengelompokkan Data Nilai MTs Darul Fikri.....	38
4.4.4 Update Centroid.....	38
4.4.5 Menghitung kembali jarak dari setiap Cluster	39
4.4.6 Finish.....	39
4.5 Penerapan perhitungan K-means menggunakan Rstudio.....	40
4.5.1 Langkah Pertama	40
4.5.2 Langkah Ke-Dua.....	40
4.5.3 Langkah Ke-Tiga	41
4.5.4 Langkah Ke-Empat	42
4.5.5 Langkah Ke-Lima	44
4.5.6 Langkah Ke-Enam	44
4.5.7 Langkah Ke-Tujuh	45
4.6 Penerapan perhitungan <i>K-Means</i> menggunakan <i>Interfaces R Shiny</i>	46
4.6.1 Pengertian Packages Shiny	46
4.6.2 Aplikasi Antar Muka.....	46
4.7 Validasi Pengelompokan Kelas unggulan.....	48
4.7.1 Validasi Pengelompokan Menggunakan Teknik Manual	48
4.7.2 Validasi Menggunakan Metode K-Means Clustering.....	50

4.8	Pengujian Sistem.....	51
4.8.1	Pengujian.....	51
4.8.2	Rancangan Pengujian.....	52
4.8.3	Langkah Dan Hasil Pengujian.	52
4.8.4	Kesimpulan	53
BAB V		54
PENUTUP.....		54
5.1	Kesimpulan.....	54
5.2	Saran.....	54
DAFTAR PUSTAKA		56



DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	4
Tabel 3. 1 Data Atribut Siswa Kelas VII	27
Tabel 3. 2 Sebagian Data Seleksi MTs Darul Fikri	29
Tabel 3. 3 Pengelompokkan dari setiap Cluster.....	31
Tabel 3. 4 Flowchart Penelitian Secara Umum	32
Tabel 4. 1 Data Mentah Nilai Raport Kelas VII	34
Tabel 4. 2 Seleksi Data Kelas VII A	35
Tabel 4. 3 Seleksi Data Kelas VII B	35
Tabel 4. 4 Seleksi Data Kelas VII C	35
Tabel 4. 5 Seleksi Data Kelas VII D	35
Tabel 4. 6 Pengelompokkan Data	36
Tabel 4. 7 Pusat Cluster	36
Tabel 4. 8 Data hasil perhitungan jarak dari setiap Cluster	37
Tabel 4. 9 Pengelompokkan Data MTs Darul Fikri	38
Tabel 4. 10 Iterasi Pertama	38
Tabel 4. 11 Menghitung Rata-rata	39
Tabel 4. 12 Perhitungan Terakhir	39
Tabel 4. 13 Iterasi Terakhir	39
Tabel 4. 14 Tabel Data Uji.....	52
Tabel 4. 15 Langkah Dan Hasil Pengujian	52

DAFTAR GAMBAR

Gambar 2. 1 Teknik Gambar <i>Mining</i>	9
Gambar 2. 2 Tahapan <i>Data Mining</i>	10
Gambar 2. 3 Penentuan Nilai K	13
Gambar 2. 4 Penghitungan Jarak Suatu Kluster	14
Gambar 2. 5 Penghitungan Mean Untuk Centroid Baru	14
Gambar 2. 6 Penghitungan jarak suatu Kluster	15
Gambar 2. 7 Finish <i>Algoritma K-Means</i>	15
Gambar 2. 8 Data Penerapan <i>K-Means</i> Menggunakan <i>Rstudio</i>	18
Gambar 2. 9 Tampilan Data di <i>Rstudio</i>	19
Gambar 2. 10 Output Data Numerik	20
Gambar 2. 11 Grafik Claster K <i>Metode Elbow</i>	21
Gambar 2. 12 Grafik Claster K <i>Metode Silhouette</i>	22
Gambar 2. 13 Output <i>K-Means</i>	23
Gambar 2. 14 Visualisasi Cluster	23
Gambar 2. 15 Data yang telah ada Cluster	24
Gambar 4. 1 Tampilan data yang digunakan untuk penelitian	41
Gambar 4. 2 Tampilan Transformasi data menggunakan Scale (Sekala)	42
Gambar 4. 3 Tampilan hasil grafik Metode Elbow	43
Gambar 4. 4 Tampilan hasil grafik Metode Silhouette	43
Gambar 4. 5 Tampilan hasil K-Means	44
Gambar 4. 6 Hasil Visualisasi Cluster	45
Gambar 4. 7 Hasil Cluster berbentuk Tabel	46
Gambar 4. 8 Petunjuk aplikasi Rshiny dengan Algoritma K-Means	47
Gambar 4. 9 Input file yang akan di seleksi	47
Gambar 4. 10 Input file yang akan di seleksi	48
Gambar 4. 11 Hasil penyeleksian data siswa	48
Gambar 4. 12 Data validasi manual kelas A	49
Gambar 4. 13 Data validasi manual kelas B	49
Gambar 4. 14 Data validasi manual kelas C	49
Gambar 4. 15 Data validasi manual kelas D	50
Gambar 4. 16 Data validasi metode k-means clustering klaster 1	50
Gambar 4. 17 Data validasi metode k-means clustering klaster 2	51
Gambar 4. 18 Data validasi metode k-means clustering klaster 3	51

DAFTAR LAMPIRAN

Lampiran 1 Hasil Validasi Pengelompokan Menggunakan Teknik Manual	59
Lampiran 2 Hasil Validasi Pengelompokan Menggunakan Metode K-Means.....	61

