

***DATA MINING CLUSTERING DALAM
PENGELOMPOKAN HASIL TES
SISWA DENGAN MENGGUNAKAN
METODE K-MEANS***

SKRIPSI



**Oleh:
Kevin Antonius
140210009**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2018**

***DATA MINING CLUSTERING DALAM
PENGELOMPOKAN HASIL TES
SISWA DENGAN MENGGUNAKAN
METODE K-MEANS***

SKRIPSI

**Untuk memenuhi salah satu syarat
guna memperoleh gelar Sarjana**



Oleh:

**Kevin Antonius
140210009**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2018**

PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Skripsi ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (sarjana, dan/atau magister), baik di Universitas Putera Batam maupun di perguruan tinggi lain.
2. Skripsi ini adalah murni gagasan, rumusan, dan penelitian saya sendiri, tanpa bantuan pihak lain, kecuali arahan pembimbing.
3. Dalam skripsi ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
4. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di perguruan tinggi.

Batam, 6 Agustus 2018

Yang membuat pernyataan,

Kevin Antonius
NPM 140210009

***DATA MINING CLUSTERING DALAM PENGELOMPOKAN
HASIL TES SISWA DENGAN MENGGUNAKAN METODE
K-MEANS***

**Oleh:
Kevin Antonius
140210009**

**SKRIPSI
Untuk memenuhi salah satu syarat
guna memperoleh gelar Sarjana**

**Telah disetujui oleh Pembimbing pada tanggal
seperti tertera di bawah ini**

Batam, Agustus 2018

**Yulia, S.Kom., M.Kom.
Pembimbing**

ABSTRAK

Pendidikan merupakan suatu hal yang sangat penting bagi setiap individu, dikarenakan pendidikan mewujudkan suasana proses belajar mengajar agar peserta didik dapat mengembangkan potensi dirinya secara aktif supaya memiliki pengendalian diri, kecerdasan, kekuatan spiritual keagamaan, kepribadian, serta keterampilan yang diperlukan masyarakat. SMAK Yos Sudarso merupakan sekolah katolik pertama di Batam dan termasuk salah satu sekolah tertua di Batam. Data nilai siswa yang terkumpul cukup banyak, karena penambahan siswa dari tahun ke tahun lebih banyak, sehingga membutuhkan waktu yang cukup lama jika pihak sekolah mengelompokkan nilai siswa tersebut secara manual, tujuan pengelompokan nilai siswa adalah untuk memudahkan dalam pemantauan sehingga pihak sekolah dapat membimbing atau membantu siswa dalam meningkatkan nilai siswa tersebut, data yang diolah belum dimanfaatkan secara maksimal, serta pihak sekolah masih menggunakan cara manual untuk mengelompokkan data nilai siswa. Pengelompokan data nilai siswa yang berukuran besar ini tidak dapat dilakukan dengan mudah. *Data mining* merupakan suatu rangkaian proses untuk dimana *data mining* ini menggali nilai tambah dari sekumpulan data yang berupa pengetahuan yang selama ini tidak diketahui secara manual *K-Means clustering* merupakan salah satu metode yang banyak digunakan untuk melakukan *clustering* dalam *data mining*. *K-Means* mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien.

Kata Kunci: *Data Mining, Nilai, Clustering, Algoritma K-Means*

ABSTRACT

Education is a very important thing for every individual, because the education teaching and learning process so the student can develop their potentials actively in order to have self-control, intelligence, spiritual strength, personality, and skills needed by the community. SMAK Yos Sudarso is the first Catholic school in Batam and is one of the oldest schools in Batam. the grade of students has accumulated quite a lot, because the addition of students from year to year more, so it takes quite a long time if the school grouping students grades manually, the purpose of grouping students' grades is to facilitate the monitoring so that the school can guide or help students in increasing the grade of these students, the processed data has not been maximally utilized, and the school still uses manual methods to classify student value data. This large student grade data clustering can not be done easily. Data mining is a series of processes for which data mining explores the added value of a set of data in the form of knowledge that has not been known manually K-Means clustering is one method that is widely used to cluster in data mining. K-Means has the ability to group data in large numbers with relatively fast and efficient computing time.

Keywords: *Data Mining, Grade, Clustering, K-means algorithm*

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke Tuhan Yang Maha Esa yang telah melimpahkan segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Teknik Informatika Universitas Putera Batam.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati.

Dengan segala keterbatasan, penulis menyadari pula bahwa skripsi ini takkan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Ibu Dr. Nur Elfi Husda, S. Kom., M.SI. selaku Rektor Universitas Putera Batam.
2. Ketua Program Studi Universitas Putra Batam Bapak Andi Maslan, S.T., M.SI.
3. Ibu Yulia, S. Kom., M. Kom. selaku Pembimbing Skripsi pada Program Studi Teknik Informatika Universitas Putera Batam .
4. Dosen dan Staf Universitas Putera Batam.
5. Bapak Stanislaus K. Teluma selaku Kepala Sekolah SMAK Yos Sudarso
6. Orang tua penulis yang selalu mendoakan dan memberi semangat.
7. Teman-teman seangkatan yang berjuang bersama dan saling memberi semangat dan motivasi.

Semoga Tuhan Yang Maha Esa membalas kebaikan dan selalu mencurahkan hidayah serta taufik-Nya.

Batam, 6 Agustus 2018

Kevin Antonius

DAFTAR ISI

	Halaman
HALAMAN SAMPUL DEPAN	i
HALAMAN JUDUL	ii
PERNYATAAN.....	iii
HALAMAN PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xi
DAFTAR RUMUS	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	
1.1 Latar Belakang Penelitian	1
1.2 Identifikasi Masalah	4
1.3 Pembatasan Masalah	4
1.4 Rumusan Masalah	5
1.5 Tujuan Penelitian.....	5
1.6 Manfaat Penelitian.....	6
BAB II KAJIAN PUSTAKA	
2.1 <i>Knowledge Discovery in Database (KDD)</i>	7
2.2 <i>Data Mining</i>	9
2.2.1 Manfaat <i>Data Mining</i>	11
2.2.2 Penerapan <i>Data Mining</i>	12
2.2.3 Kategori <i>Data Mining</i>	14
2.3 Metode <i>Data Mining</i>	15
2.3.1 Teknik <i>Clustering</i>	17
2.3.2 Algoritma <i>K-Means</i>	19
2.4 Software Pendukung.....	22
2.5 Penelitian Terdahulu.....	24

2.6	Kerangka Pemikiran	28
2.7	Hipotesis	28

BAB III METODE PENELITIAN

3.1	Desain Penelitian	29
3.2	Teknik Pengumpulan Data	31
3.3	Operasional Variabel	32
3.4	Metode Analisis dan Rancangan Sistem	48
3.5	Lokasi dan Jadwal Penelitian	49
3.5.1	Lokasi Penelitian	49
3.5.2	Jadwal Penelitian	50

BAB IV HASIL DAN PEMBAHASAN

4.1	Analisis Data	51
4.1.1	Seleksi Data	51
4.1.2	<i>Cleaning</i> Data	52
4.1.3	Perhitungan Algoritma	52
4.2	Hasil Pengujian	67

BAB V SIMPULAN DAN SARAN

5.1	Simpulan	75
5.2	Saran	75

DAFTAR PUSTAKA

RIWAYAT HIDUP

SURAT KETERANGAN PENELITIAN

SURAT BALASAN KETERANGAN PENELITIAN

LAMPIRAN

DAFTAR TABEL

	Halaman
Tabel 3.1 Data Nilai Siswa.....	33
Tabel 3.2 Data Nilai Siswa.....	46
Tabel 3.3 Jadwal Penelitian.....	50
Tabel 4.1 Perhitungan jarak data	55
Tabel 4.2 <i>Cluster</i> Baru Iterasi ke-1	58
Tabel 4.3 Hasil Perhitungan Jarak Dan Pengelompokan Data Iterasi ke-2.....	58
Tabel 4.4 <i>Cluster</i> Baru Iterasi ke-2	61
Tabel 4.5 Hasil Perhitungan Jarak Dan Pengelompokan Data Iterasi ke-3.....	62
Tabel 4.6 Hasil Pengelompokan Data dari Iterasi ke 1 Sampai Iterasi ke 3	63
Tabel 4.7 Hasil Pengelompokan Siswa Dengan Nilai Cukup	65
Tabel 4.8 Hasil Pengelompokan Siswa Dengan Nilai Tinggi.	65
Tabel 4.9 Lanjutan.....	66
Tabel 4.10 Hasil Pengelompokan Siswa Dengan Nilai Rendah.....	66

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Tahapan <i>Knowledge Discovery in Database</i> (KDD).....	9
Gambar 2.2 Bidang Ilmu <i>Data Mining</i>	10
Gambar 2.3 Metode <i>Data Mining</i>	15
Gambar 2.4 <i>RapidMiner</i>	23
Gambar 2.5 Kerangka Pemikiran	28
Gambar 3.1. Desain Penelitian	29
Gambar 3.2 Rancangan Sistem.....	48
Gambar 4.1 <i>New Process</i>	67
Gambar 4.2 <i>Operator Read Excel</i>	68
Gambar 4.3 <i>Dialog Import Configuration Wizard</i>	68
Gambar 4.4 <i>Data Import Wizard</i>	69
Gambar 4.5 <i>Operator K-Means</i>	70
Gambar 4.6 Pengisian Jumlah <i>k Cluster</i>	70
Gambar 4.7 Proses penyambungan garis <i>read excel</i> ke <i>clustering</i>	71
Gambar 4.8 Hasil <i>RapidMiner</i> Pengelompokan Nilai Siswa	72
Gambar 4.9 Hasil <i>RapidMiner</i>	72
Gambar 4.10 Tampilan Nilai Siswa pada <i>Cluster 0</i>	73
Gambar 4.11 Tampilan Nilai Siswa pada <i>Cluster 1</i>	73
Gambar 4.12 Tampilan Nilai siswa pada <i>Cluster 2</i>	74

DAFTAR RUMUS

	Halaman
Rumus 2.1 Nilai Rata-rata pada <i>Centroid</i>	22
Rumus 4.1 Rumus perhitungan sampel.....	51
Rumus 4.2 Rumus <i>Euclidian Distance</i>	53

DAFTAR LAMPIRAN

Lampiran 1. Bukti Foto Penelitian

Lampiran 2. Perhitungan jarak untuk iterasi 1

Lampiran 3. Perhitungan jarak untuk iterasi 2

Lampiran 4. Perhitungan jarak untuk iterasi 3

Lampiran 5. Data Penelitian