

**KLASTERISASI DATA PENERIMAAN MAHASISWA BARU UNTUK
MENENTUKAN TARGET PROMOSI UNIVERSITAS SRIWIJAYA
MENGUNAKAN ALGORITMA *K-MEANS CLUSTERING***

SKRIPSI

Sebagai salah satu syarat untuk menyelesaikan Studi

Di Program Studi Sistem Informasi S1



Oleh

NADYA ANGGRAINI

09031281823050

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS SRIWIJAYA

2022

LEMBAR PENGESAHAN

**KLASTERISASI DATA PENERIMAAN MAHASISWA BARU UNTUK
MENENTUKAN TARGET PROMOSI UNIVERSITAS SRIWIJAYA
MENGUNAKAN ALGORITMA *K-MEANS CLUSTERING***

SKRIPSI

Sebagai salah satu syarat untuk menyelesaikan studi

Di Program Studi Sistem Informasi S1

Oleh

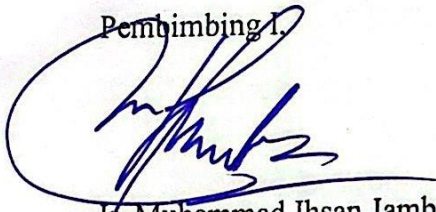
Nadya Anggraini

09031281823050

Palembang, Juni 2022

Pembimbing II,

Pembimbing I,



I. Muhammad Ihsan Jambak, M.Sc., M.M.
NIP. 196804052013081201



Ali Bardadi, S.SI., M.Kom.
NIP.198806292019031007

Mengetahui,

Ketua Jurusan Sistem Informasi,



Endang Lestari Ruskan, S.Kom., M.T.
NIP. 197811172006042001

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini :

Nama : Nadya Anggraini

NIM : 09031281823050

Program Studi : Sistem Informasi Bilingual

Judul Skripsi : Klasterisasi Data Penerimaan Mahasiswa Baru Untuk Menentukan Target Promosi Universitas Sriwijaya Menggunakan Algoritma *K-Means Clustering*

Hasil Pengecekan *Software iThenticate/Turnitin* : 18%

Menyatakan bahwa laporan skripsi saya merupakan hasil karya sendiri dan bukan penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan skripsi ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada unsur paksaan dari siapapun.



Palembang, Juni 2022



Nadya Anggraini
NIM. 09031281823050

HALAMAN PERSETUJUAN

Telah di uji dan lulus pada :

Hari : Kamis

Tanggal : 9 Juni 2022

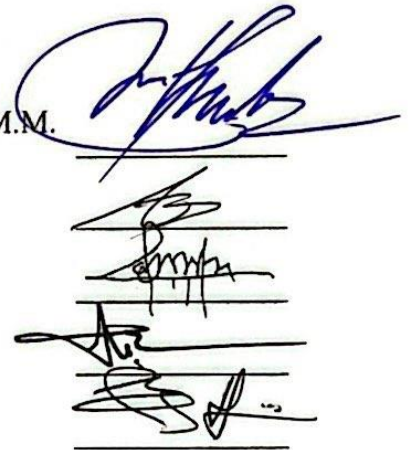
Nama : Nadya Anggraini

NIM : 09031281823050

Judul Tugas Akhir : Klasterisasi Data Penerimaan Mahasiswa Baru Untuk Menentukan Target Promosi Universitas Sriwijaya Menggunakan Algoritma K- Means Clustering

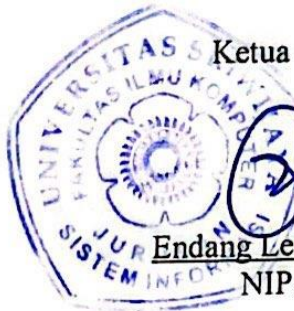
Tim Penguji :


1. Pembimbing I : Ir. Muhammad Ihsan Jambak, M.Sc.,M.M.
2. Pembimbing II : Ali Bardadi, M.Kom.
3. Ketua Penguji : Ermatita, M.Kom.
4. Penguji I : Fathoni, ST., MMSI.
5. Penguji II : Yadi Utama, M.Kom.



Mengetahui

Ketua Jurusan Sistem Informasi,




Endang Lestari Ruskan, S.Kom., M.T.
NIP. 197811172006042001

HALAMAN PERSEMBAHAN

“Hatiku tenang karena mengetahui bahwa
apa yang melewatkanmu tidak akan pernah menjadi takdirku,
dan apa yang ditakdirkan untukmu
tidak akan pernah melewatkanmu.”

-Umar bin Khattab

SKRIPSI INI KUPERSEMBAHKAN UNTUK :

- ❖ **Allah SWT**
- ❖ **Diri sendiri, Nadya Anggraini**
- ❖ **Orang tua, Saudara, dan Keluarga**
- ❖ **Dosen Pembimbing**
- ❖ **Dosen Penguji**
- ❖ **Seluruh Dosen beserta Staff jurusan Sistem Informasi**
- ❖ **Teman-teman seperjuangan Sistem Informasi 2018**
- ❖ **Semua orang baik yang tidak bisa disebutkan satu-persatu**
- ❖ **Almamater yang kebanggakan, Universitas Sriwijaya.**

KLASTERISASI DATA PENERIMAAN MAHASISWA BARU UNTUK MENENTUKAN TARGET PROMOSI UNIVERSITAS SRIWIJAYA MENGUNAKAN ALGORITMA *K-MEANS CLUSTERING*

Oleh

Nadya Anggraini (09031281823050)

Sistem Informasi, Fakultas Ilmu Komputer, Universitas Sriwijaya

nadianggraini20@gmail.com

ABSTRAK

Saat ini, persaingan antar perguruan tinggi negeri maupun swasta dalam rangka penerimaan mahasiswa baru memang semakin pesat. Universitas Sriwijaya adalah salah satu perguruan tinggi terbesar di Sumatera Selatan yang secara terus-menerus melakukan peningkatan mutu kualitas terhadap sarana dan prasarana untuk menarik calon mahasiswa. Dikarenakan persaingan yang ketat, dibutuhkan strategi yang tepat bagi universitas untuk melakukan promosi. Dalam penelitian ini, menerapkan ilmu *data mining* dalam melakukan pengolahan data penerimaan mahasiswa baru Universitas Sriwijaya tahun 2019-2020 untuk mengelompokkan data sebaran asal sekolah mahasiswa. Dari pengolahan yang telah dilakukan melalui aplikasi *RapidMiner*, terbentuklah 3 *cluster* berdasarkan jumlah penerimaan mahasiswa baru. Dari analisis karakteristik *cluster* yang terbentuk, diketahui bahwa *cluster* yang paling efisien dan efektif untuk dijadikan target promosi adalah *cluster* 1 dengan jumlah anggota *cluster* sebanyak 60 sekolah. Hasil dari penelitian ini dapat dijadikan usulan untuk Tim PMB Universitas Sriwijaya dalam penentuan target promosi pada tahun-tahun berikutnya.

Kata Kunci: *Data Mining*, Klasterisasi, *K-Means*, Promosi, *RapidMiner*

CLASTERIZATION OF NEW STUDENT ADMISSION DATA TO DETERMINE THE SRIWIJAYA UNIVERSITY PROMOTIONAL TARGET USING K-MEANS CLUSTERING ALGORITHM

By

Nadya Anggraini (09031281823050)

Information System, Faculty of Computer Science, Sriwijaya University
nadianggraini20@gmail.com

ABSTRACT

Currently, competition between public and private universities in the context of new student admissions is indeed increasing rapidly. Sriwijaya University is one of the largest universities in South Sumatra which continuously improves the quality of facilities and infrastructure to attract prospective students. Due to intense competition, it takes the right strategy for universities to promote. In this study, applying data mining science in processing data on new student admissions at Sriwijaya University in 2019-2020 to classify data from the distribution of student schools. From the processing that has been done through the RapidMiner application, 3 clusters are formed based on the number of new student admissions. From the analysis of the characteristics of the cluster formed, it is known that the most efficient and effective cluster to be used as a promotion target is cluster 1 with a total of 60 cluster members. The results of this study can be proposed to the Sriwijaya University PMB Team in determining promotion targets in the following years.

Keywords: Data Mining, Clustering, K-Means, Promotion, RapidMiner

KATA PENGANTAR



Alhamdulillah rabbi'l'alamin segala puji dan syukur kehadiran Allah SWT atas berkat dan rahmat-Nya yang telah diberikan kepada Penulis sehingga dapat menyelesaikan Tugas Akhir ini dengan baik. Tugas Akhir ini disusun untuk memenuhi salah satu syarat guna menyelesaikan pendidikan program Strata-1 Program Studi Sistem Informasi pada Fakultas Ilmu Komputer di Universitas Sriwijaya.

Dalam menyelesaikan Tugas Akhir ini banyak pihak yang telah memberikan bantuan dan dukungan baik secara langsung maupun secara tidak langsung. Pada kesempatan ini, penulis ingin menyampaikan ucapan terima kasih kepada pihak-pihak yang telah membantu penulis dalam menyelesaikan Tugas Akhir ini, yaitu kepada:

1. Kedua Orang Tua serta keluarga penulis tercinta, yang telah memberikan doa restu serta dukungan yang sangat besar selama mengikuti dan melaksanakan perkuliahan di Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya hingga penulis dapat menyelesaikan Skripsi ini.
2. Bapak Jaidan Jauhari, S.Pd., M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Ibu Endang Lestari Ruskan, S.Kom M.T, selaku Ketua Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.

4. Bapak Ir. Muhammad Ihsan Jambak, M.Sc., M.M dan Bapak Ali Bardadi, S.SI., M.Kom sebagai pembimbing Tugas Akhir yang memberi motivasi dan ilmu dalam proses pengerjaan penelitian ini.
5. Bapak Mgs. Afriyan Firdaus, M.IT selaku dosen pembimbing akademik, yang telah membimbing dan memberikan motivasi penulis dalam proses perkuliahan.
6. Ibu Ermatita, M.Kom selaku ketua penguji serta Bapak Fathoni, ST., MMSI dan Bapak Yadi Utama, M.Kom selaku dosen penguji, yang telah memberikan masukan dan dorongan dalam proses pengerjaan Tugas Akhir.
7. Seluruh dosen Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.
8. Seluruh staf jurusan Sistem Informasi yang telah membantu dalam kelancaran proses administrasi dan akademik selama masa perkuliahan.
9. Para teman seperjuangan Rafli, Lulu, Fano, Ihsan, Said, Fiidha, Attiyah, Winnie, Nisa, Arin, Sultan, Ebi, Syeki, Luthfi, dan Rifqi yang telah membantu dalam menyelesaikan tugas akhir.
10. Teman-teman dari kelas SI BIL B 2018, kakak tingkat, adik tingkat, serta teman-teman lainnya yang telah mendengarkan keluh kesah penulis serta memberikan berbagai masukan selama menempuh Pendidikan di Fakultas Ilmu Komputer Universitas Sriwijaya.
11. BPH HIMSI Fasilkom UNSRI khususnya Dinas Bistra, BEM KM Fasilkom UNSRI Khususnya Bidang Relasi, dan AIESEC in UNSRI khususnya Wahaghari yang telah memberikan kesempatan penulis dalam berkarya serta

turut andil dalam menjalankan berbagai tugas yang diberikan sehingga penulis dapat menerapkan tugas tersebut ke lingkungan yang lebih luas.

12. Semua orang yang tak tertuliskan dalam kata pengantar ini namun turut membantu dan melancarkan dalam proses untuk mencapai salah satu syarat gelar sarjana ini.

Penulis menyadari dalam penyusunan Tugas Akhir ini masih terdapat banyak kekurangan disebabkan keterbatasan pengetahuan dan pengalaman, oleh karena itu kritik dan saran yang membangun sangat diharapkan untuk kemajuan penelitian selanjutnya.

Akhir kata semoga Tugas Akhir ini dapat berguna dan bermanfaat bagi kita semua.

Palembang, Juni 2022



Nadya Anggraini
NIM. 09031281823050

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	ii
HALAMAN PERNYATAAN	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PERSEMBAHAN	v
ABSTRAK	vi
ABSTRACT	vii
KATA PENGANTAR	viii
DAFTAR ISI	xi
DAFTAR GAMBAR	xiv
DAFTAR TABEL	xv
BAB I PENDAHULUAN	1
1.1 Pendahuluan	1
1.2 Latar Belakang.....	1
1.3 Identifikasi Masalah.....	5
1.4 Rumusan Masalah.....	6
1.5 Tujuan Penelitian	6
1.6 Manfaat Penelitian	7
1.7 Batasan Masalah	7
1.8 Sistematika Penulisan	8
BAB II TINJAUAN PUSTAKA	9
2.1 Pendahuluan	9
2.2 Profil Universitas Sriwijaya	9
2.2.1 Sejarah	9
2.2.2 Lambang	12
2.2.3 Visi dan Misi.....	14
2.3 <i>Data Mining</i>	15
2.3.1 Pengertian <i>Data Mining</i>	15
2.3.2 Tahapan <i>Data Mining</i>	16
2.3.3 Metode Pengembangan <i>Data Mining</i> (CRISP-DM).....	18

2.4	Metode <i>Clustering</i>	21
2.4.1	Pengertian <i>Clustering</i>	21
2.4.2	Konsep <i>Clustering</i>	26
2.4.3	Tujuan Analisis <i>Cluster</i>	26
2.4.4	Manfaat <i>Cluster</i>	26
2.5	Algoritma <i>K-Means</i>	26
2.5.1	Pengertian Algoritma <i>K-Means</i>	25
2.5.2	Kelebihan dan Kekurangan Algoritma <i>K-Means</i>	26
2.5.3	Cara Kerja Algoritma <i>K-Means</i>	26
2.6	Algoritma <i>K-Medoids</i>	28
2.5.4	Pengertian Algoritma <i>K-Medoids</i>	28
2.5.5	Kelebihan dan Kekurangan Algoritma <i>K-Medoids</i>	29
2.5.6	Cara Kerja Algoritma <i>K-Medoids</i>	29
2.7	<i>Euclidean Distance</i>	30
2.8	<i>Davies-Bouldin Index</i> (DBI)	31
2.9	<i>RapidMiner</i>	32
2.10	Promosi	33
2.11	Penerimaan Mahasiswa Baru	33
2.12	Penelitian Terdahulu	34
	BAB III METODOLOGI PENELITIAN	37
3.1	Pendahuluan	37
3.2	Objek Penelitian	37
3.3	Data	37
3.3.1	Sumber Data	37
3.3.2	Jenis Data	38
3.4	Tahapan Penelitian	38
3.4.1	Pemahaman Bisnis dan Pengumpulan Data	39
3.4.2	Persiapan Data	39
3.4.3	Pemodelan <i>RapidMiner</i>	40
3.4.4	Perbandingan Algoritma dan Penentuan Parameter	40
3.4.5	Analisis Hasil <i>Clustering</i> dan Kesimpulan	41
3.5	Analisis Penyelesaian Masalah	41

3.5.1 Perhitungan Iterasi Pertama.....	44
3.5.2 Perhitungan Iterasi Kedua	47
3.5.3 Perhitungan Iterasi Ketiga	50
3.5.4 Perhitungan Iterasi Keempat.....	53
3.5.5 Perhitungan Iterasi Kelima	56
3.5.6 Perhitungan Iterasi Keenam.....	59
3.6 Pemodelan <i>RapidMiner</i>	64
3.6.1 Pengujian Perbandingan Algoritma <i>K-Means</i> dan <i>K-Medoids</i>	65
3.6.2 Pengujian Nilai K dan Parameter.....	66
3.7 Format Pengujian.....	68
BAB IV HASIL DAN PEMBAHASAN.....	69
4.1 Pendahuluan	69
4.2 Hasil Pengujian.....	69
4.2.1 Hasil Pengujian Perbandingan Algoritma.....	69
4.2.2 Hasil Pengujian Nilai K dan Parameter	70
4.3 Hasil <i>Clustering</i>	71
4.4 Pembahasan Hasil Pengujian.....	74
4.4.1 Pembahasan Hasil Pengujian Perbandingan Algoritma	74
4.4.2 Pembahasan Hasil Pengujian Nilai K dan Parameter	76
4.4.2.1 Uji Anova untuk Nilai K	77
4.4.2.2 Uji Anova untuk Parameter	78
4.4.2.3 Analisis Kesimpulan Pengujian Nilai K dan Parameter.....	80
4.5 Pembahasan Hasil <i>Clustering</i>	80
4.5.1 Uji Anova untuk <i>Centroid</i>	82
4.5.2 Analisis Hasil <i>Clustering</i>	84
4.5.3 Contoh Analisis Spesifik (Studi Kasus : Fasilkom).....	88
BAB V KESIMPULAN DAN SARAN	96
5.1 Pendahuluan	96
5.2 Kesimpulan.....	96
5.3 Saran	98
DAFTAR PUSTAKA	99
LAMPIRAN	103

DAFTAR GAMBAR

Gambar 2.1 Lambang Universitas Sriwijaya	11
Gambar 2.2 Tahapan <i>Data Mining</i>	16
Gambar 2.3 Tahapan Pemodelan <i>CRISP-DM</i>	18
Gambar 2.4 Proses <i>Clustering</i>	23
Gambar 2.5 <i>Flowchart</i> Algoritma <i>K-Means</i>	26
Gambar 2.6 <i>Flowchart</i> Algoritma <i>K-Medoids</i>	29
Gambar 2.7 Rumus <i>Euclidean Distance</i>	30
Gambar 2.8 Rumus <i>Davies Bouldin Index</i>	31
Gambar 3.1 Sumber Data (BAK INFO)	37
Gambar 3.2 Tahapan Penelitian	38
Gambar 3.3 <i>Import Data</i>	64
Gambar 3.4 <i>Input Data</i> ke Halaman Proses	64
Gambar 3.5 Pengujian Perbandingan Algoritma	65
Gambar 3.6 Pengaturan Parameter	65
Gambar 3.7 Hasil Pengujian Algoritma <i>K-Means</i>	66
Gambar 3.8 Hasil Pengujian Algoritma <i>K-Medoids</i>	66
Gambar 3.9 Pengujian Parameter	67
Gambar 3.10 Pengaturan Parameter	67
Gambar 4.1 <i>Example Set Clustering</i>	71
Gambar 4.2 Visualisasi <i>Cluster</i>	71
Gambar 4.3 Deskripsi Hasil <i>Clustering</i>	72
Gambar 4.4 Anggota <i>Cluster</i>	72
Gambar 4.5 Hasil Tabel <i>Centroid</i>	73
Gambar 4.6 Hasil Grafik	73
Gambar 4.7 Diagram <i>Scatter</i>	80
Gambar 4.8 Grafik Hasil <i>Cluster</i> (Fasilkom)	88

DAFTAR TABEL

Tabel 1.1 Penurunan Jumlah PMB 2019-2020.....	3
Tabel 3.1 Atribut Data yang Digunakan	39
Tabel 3.2 Data Penerimaan Mahasiswa Baru UNSRI	42
Tabel 3.3 <i>Centroid</i> Awal.....	43
Tabel 3.4 Perhitungan Iterasi Pertama	45
Tabel 3.5 Hasil Iterasi Pertama	46
Tabel 3.6 Jumlah <i>Cluster</i> Iterasi Pertama	46
Tabel 3.7 <i>Centroid</i> Kedua	47
Tabel 3.8 Perhitungan Iterasi Kedua.....	48
Tabel 3.9 Hasil Iterasi Kedua	49
Tabel 3.10 Jumlah <i>Cluster</i> Iterasi Kedua	49
Tabel 3.11 <i>Centroid</i> Ketiga	50
Tabel 3.12 Perhitungan Iterasi Ketiga	51
Tabel 3.13 Hasil Iterasi Ketiga	52
Tabel 3.14 Jumlah <i>Cluster</i> Iterasi Ketiga.....	52
Tabel 3.15 <i>Centroid</i> Keempat	53
Tabel 3.16 Perhitungan Iterasi Keempat	54
Tabel 3.17 Hasil Iterasi Keempat	55
Tabel 3.18 Jumlah <i>Cluster</i> Iterasi Keempat	55
Tabel 3.19 <i>Centroid</i> Kelima	56
Tabel 3.20 Perhitungan Iterasi Kelima	57
Tabel 3.21 Hasil Iterasi Kelima	58
Tabel 3.22 Jumlah <i>Cluster</i> Iterasi Kelima.....	58
Tabel 3.23 <i>Centroid</i> Keenam	59
Tabel 3.24 Perhitungan Iterasi Keenam	60
Tabel 3.25 Hasil Iterasi Keenam	61
Tabel 3.26 Jumlah <i>Cluster</i> Iterasi Keenam	61
Tabel 3.27 Hasil Akhir Perhitungan Iterasi 1-3.....	62
Tabel 3.28 Hasil Akhir Perhitungan Iterasi 4-6.....	63
Tabel 3.29 Kombinasi Parameter	68

Tabel 3.30 Format Pengujian Perbandingan Algoritma	68
Tabel 3.31 Format Pengujian Nilai K dan Parameter	68
Tabel 4.1 Hasil Pengujian Perbandingan Algoritma	69
Tabel 4.2 Hasil Pengujian Nilai K dan Parameter	70
Tabel 4.3 <i>Group Statics</i> Pengujian Algoritma	74
Tabel 4.4 Hasil Uji <i>Independent Sample T-Test</i>	75
Tabel 4.5 Hasil Uji Anova untuk Nilai K	77
Tabel 4.6 Hasil Uji Anova untuk Parameter	78
Tabel 4.7 Uji Homogenitas Parameter.....	79
Tabel 4.8 Himpunan Bagian Homogen Parameter.....	79
Tabel 4.9 Rata-Rata <i>Centroid</i>	81
Tabel 4.10 Hasil Uji Anova untuk <i>Centroid</i>	83
Tabel 4.11 Uji Homogenitas <i>Centroid</i>	83
Tabel 4.12 Himpunan Bagian Homogen <i>Centroid</i>	83
Tabel 4.13 Karakteristik <i>Cluster</i>	84
Tabel 4.14 Pengaruh <i>Cluster</i> terhadap Jumlah PMB.....	86
Tabel 4.15 Karakteristik <i>Cluster</i> Fasilkom	90
Tabel 4.16 Karakteristik <i>Cluster</i> Fasilkom	91
Tabel 4.17 Pengaruh <i>Cluster</i> terhadap Jumlah PMB Fasilkom	93
Tabel 4.18 <i>Cluster</i> 1 Fasilkom	95

BAB I

PENDAHULUAN

1.1 Pendahuluan

Pada bab ini menjelaskan secara detail tentang latar belakang masalah, identifikasi masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan dalam penelitian.

1.2 Latar Belakang

Perkembangan teknologi informasi sudah semakin pesat dan hampir mencakup segala bidang seperti ekonomi, industri, pendidikan, informasi dan komunikasi, promosi, dan bidang kehidupan lainnya sehingga menghasilkan ketersediaan data yang sangat besar yang mungkin menyimpan informasi penting dengan pengolahan tertentu. Penerapan teknologi informasi dalam dunia pendidikan juga dapat menghasilkan data yang berlimpah, mulai dari data proses pembelajaran maupun data pribadi pengajar dan pelajar. Pada suatu institusi perguruan tinggi, sejumlah data disimpan menjadi data historis yang jumlahnya akan bertambah secara terus menerus seperti data mahasiswa (Wijanarko, 2016).

Proses penerimaan mahasiswa baru dalam suatu perguruan tinggi menyebabkan penambahan akumulasi data tiap tahunnya dan menimbulkan tumpukan data yang berjumlah besar dalam basis data. Para mahasiswa yang lulus seleksi masuk suatu perguruan tinggi umumnya memberikan data profil kepada pihak kampus. Data yang disimpan akan terintegrasi ke masing-masing fakultas dan prodi terkait. Data tersebut kemudian akan digunakan untuk mahasiswa

mendapatkan akses proses pembelajaran selama masa kuliah di perguruan tinggi.

Salah satu perguruan tinggi yang ada di Indonesia adalah Universitas Sriwijaya, yang mana memiliki 2 lokasi kampus berbeda yaitu Kampus Indralaya dan Kampus Palembang. Universitas Sriwijaya merupakan perguruan tinggi negeri terbesar yang ada di Sumatera Selatan dengan memiliki 10 (sepuluh) fakultas yaitu Fakultas Ekonomi, Fakultas Hukum, Fakultas Kedokteran, Fakultas Teknik, Fakultas Pertanian, FKIP, FISIP, FMIPA, Fakultas Ilmu Komputer, dan Fakultas Kesehatan Masyarakat. Setiap fakultas memiliki beberapa program studi yang dapat menjadi pilihan bagi para mahasiswa baru untuk mengenyam pendidikan dan mengembangkan potensi serta menyalurkan minatnya.

Kegiatan penerimaan mahasiswa baru di Universitas Sriwijaya merupakan kegiatan rutin yang diselenggarakan secara tahunan untuk mendapatkan mahasiswa baru yang akan mengenyam pendidikan di tingkat perguruan tinggi. Universitas Sriwijaya terus-menerus melakukan peningkatan mutu kualitas baik dari segi fasilitas, tenaga pengajar, pegawai, administrasi, gedung, serta sarana dan prasarana lainnya untuk menarik calon mahasiswa. Namun, sangat disayangkan jika informasi mengenai masing-masing fakultas dan prodi yang ada tidak tersebar secara tepat sasaran. Apalagi jika melihat banyaknya bermunculan perguruan tinggi lainnya di sekitaran wilayah Sumatera Selatan maupun di Indonesia, menimbulkan persaingan yang sangat ketat bagi Universitas Sriwijaya dalam menjaring mahasiswa baru.

Dari data yang didapatkan dari laman BAK INFO, dapat dilihat bahwa terjadi penurunan yang cukup drastis terhadap jumlah mahasiswa baru yang diterima dari tahun 2019 ke tahun 2020. Penurunan jumlah penerimaan

mahasiswa baru tiap fakultas secara detail dapat dilihat pada tabel berikut:

Tabel 1.1 Penurunan Jumlah PMB 2019-2020

No	Nama Fakultas	Jumlah PMB Tahun 2019	Jumlah PMB Tahun 2020	Penurunan	Persentase Penurunan
1.	Fakultas Ekonomi	1172	463	709	60.49%
2.	Fakultas Hukum	690	648	42	6.09%
3.	Fakultas Kedokteran	1007	530	477	47.37%
4.	Fakultas Pertanian	1082	673	409	37.80%
5.	Fakultas Teknik	949	609	340	35.83%
6.	FASILKOM	737	664	73	9.91%
7.	FISIP	1038	981	57	5.49%
8.	FKIP	1741	860	881	50.60%
9.	FKM	412	394	18	4.37%
10.	FMIPA	744	660	84	11.29%
Grand Total		9572	6482	3090	32.28%

Saat ini, persaingan antar perguruan tinggi negeri maupun swasta dalam rangka penerimaan mahasiswa baru semakin pesat. Ada yang menghabiskan banyak biaya untuk media promosi dan ada juga yang memberikan berbagai penawaran seperti biaya perkuliahan yang terjangkau, menjanjikan kemudahan mendapat pekerjaan setelah lulus dan lain sebagainya (Budiman, 2019). Maka dari itu, perencanaan yang tepat sangat diperlukan untuk menentukan target promosi agar jumlah penerimaan mahasiswa baru di tiap tahun tercapai dan sesuai dengan ekspektasi supaya kelak menghasilkan SDM terbaik untuk mendukung perkembangan bangsa dengan lulusan mahasiswa yang bermutu.

Penerapan teknologi informasi pada bidang promosi dapat membantu pihak universitas mengelola data yang mungkin menyimpan informasi yang berharga dan bermanfaat sehingga memiliki nilai tambah tersendiri agar tujuan tercapai sesuai dengan yang diinginkan dalam menjaring calon mahasiswa terbaik (Heryati, 2020). Promosi merupakan alat dari kunci kesuksesan bagi suatu universitas dalam memperkenalkan lingkungan kampusnya untuk menarik

perhatian calon pendaftar. Promosi membutuhkan perencanaan yang tepat agar semua yang menjadi target promosi dari bisa terwujud, sehingga daya saing dapat meningkat juga disetiap daerah dengan adanya pemerataan informasi yang tepat sasaran (Silalahi, 2020).

Besarnya kumpulan data mahasiswa yang diperoleh sehingga menghasilkan data yang besar akan menjadi peluang ditemukannya suatu informasi dari beberapa kelompok berdasarkan karakteristik data di universitas, terbentuklah beberapa pengelompokan data mahasiswa sehingga dapat menentukan target promosi yang efektif dan efisien bagi universitas. Salah satu penyebab terjadinya penurunan jumlah mahasiswa pada perguruan tinggi adalah perencanaan strategi promosi yang kurang tepat akan berdampak pada fluktuasi jumlah mahasiswa baru (Heryati, 2020). Data yang besar tersebut dapat di dilakukan pengolahan menggunakan ilmu *data mining*. *Data Mining* dirancang untuk mengeksplorasi data untuk mencari pola yang konsisten dengan menerapkan pola yang terdeteksi ke subset baru (Handoko et al., 2020).

Pengelompokan dari data penerimaan mahasiswa baru tersebut dapat dikelompokkan menggunakan metode *K-Means Clustering*, pengelompokan data menjadi beberapa *cluster* yang di dasarkan kemiripan data tersebut, sedangkan yang mempunyai karakteristik yang tidak sama dikelompokan dalam *cluster* yang lain yang mempunyai kemiripan yang sama (Darmansah & Wardani, 2021). *K-Means Clustering* memiliki kelebihan yaitu dapat membuat kelompok data yang cukup besar dengan cepat dan efisiensi (Fajriah et al., 2019) sehingga dapat membantu pengambilan keputusan dalam menentukan target promosi yang efektif pada tahun berikutnya.

Untuk memenuhi ekspektasi jumlah penerimaan mahasiswa baru di tiap tahunnya, maka penelitian ini dilakukan untuk menganalisa dan mengolah data penerimaan mahasiswa baru tahun 2019-2020 di Universitas Sriwijaya sehingga mendapatkan hasil informasi yang dapat membantu dalam penentuan target promosi di tahun berikutnya. Atribut sebuah data yang akan dipakai dalam penelitian ini adalah NIM, asal sekolah, dan program studi para mahasiswa yang diharapkan dapat membantu serta memberi masukan kepada Tim PMB Universitas Sriwijaya dalam melakukan promosi ke berbagai sekolah di Indonesia khususnya di wilayah Sumatera Selatan. Pengolahan atribut data tersebut akan menggunakan *tools RapidMiner* sehingga menghasilkan beberapa *cluster* untuk penentuan target promosi di Universitas Sriwijaya. Dari permasalahan yang telah dirumuskan, maka judul yang ditetapkan pada penelitian ini adalah **“Klasterisasi Data Penerimaan Mahasiswa Baru untuk Menentukan Target Promosi Universitas Sriwijaya Menggunakan Algoritma *K-Means Clustering*”**.

1.3 Identifikasi Masalah

Berdasarkan uraian latar belakang dari penelitian ini maka dapat diidentifikasi permasalahan sebagai berikut :

- a. Banyaknya penumpukkan data mahasiswa dari tahun ke tahun sehingga membutuhkan pengolahan data yang tepat.
- b. Persaingan yang ketat antar perguruan tinggi lainnya dalam menjaring calon mahasiswa berkualitas sesuai dengan bidang potensinya.
- c. Penentuan target promosi ke sekolah hanya menggunakan perkiraan dan pengalaman dari tahun-tahun sebelumnya.

1.4 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, adapun rumusan masalah pada penelitian ini adalah “Bagaimana karakteristik sekolah asal calon mahasiswa baru yang tepat untuk dijadikan target promosi dengan menerapkan ilmu *data mining* menggunakan metode *clustering*?”.

1.5 Tujuan Penelitian

Dalam penelitian ini, tujuan yang ingin dicapai yaitu:

- a. Mengetahui algoritma *clustering* apa yang paling optimal dalam mengklasterisasi data penerimaan mahasiswa baru.
- b. Mengetahui daftar pengelompokan sekolah-sekolah asal calon mahasiswa baru Universitas Sriwijaya.
- c. Menentukan daftar sekolah untuk menjadi usulan target promosi yang efektif dan efisien.

1.6 Manfaat Penelitian

Adapun manfaat penelitian dari penelitian ini adalah:

- a. Dapat mengetahui perbandingan algoritma *clustering* yang paling optimal untuk mengelompokkan data.
- b. Dapat memperoleh kualitas output hasil klasterisasi data dari pengolahan melalui *Tools RapidMiner*.
- c. Dapat mengelompokkan data untuk dijadikan sebagai salah satu dasar pengambilan keputusan dalam menentukan target promosi yang tepat sasaran pada tahun-tahun berikutnya.

1.7 Batasan Masalah

Mengingat luasnya bahasan lingkup objek penelitian, maka diperlukannya batasan terhadap masalah ini agar memperoleh hasil yang baik dan optimal, maka rumusan masalah yang dibahas pada penelitian ini berfokus pada :

- a. Objek penelitian hanya dilakukan pada Data Penerimaan Mahasiswa Baru Universitas Sriwijaya tahun 2019-2020 dengan jalur seleksi SBMPTN, SNMPTN, USMPTN D3, dan USMPTN S1.
- b. Algoritma yang digunakan untuk proses *clustering* adalah algoritma K-Means dan algoritma K-Medoids dengan membandingkan antar kedua algoritma terlebih dahulu untuk mengetahui mana yang paling optimal.
- c. Mengadopsi konsep kerangka berpikir CRISP-DM (*Cross-Industry Standard Process for Data Mining*) sampai dengan tahapan *evaluation*.
- d. Pengolahan dan pengujian data menggunakan aplikasi *RapidMiner* dan SPSS untuk menemukan suatu pola atau informasi dalam mendukung penentuan target promosi Universitas Sriwijaya.

1.8 Sistematika Penulisan

Agar laporan penelitian ini mudah untuk dipahami, pemaparan materi dikelompokkan menjadi beberapa bab dengan sistematika penulisan sebagai berikut:

BAB I PENDAHULUAN

Membahas mengenai Latar Belakang, Permasalahan Penelitian yang terdiri dari Identifikasi Masalah, Ruang Lingkup Masalah, Rumusan Masalah, Tujuan dan Manfaat Penelitian, serta Sistematika Penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini membahas fundamental teori yang digunakan dalam penelitian, seperti definisi metode dan algoritma yang digunakan beserta tahapan kerja serta kajian literatur penelitian terdahulu yang relevan terhadap penelitian ini.

BAB III METODOLOGI PENELITIAN

Membahas mengenai metode pengumpulan data, tahapan penelitian, proses *clustering* dengan perhitungan manual dan pengelompokan hasil *clustering*.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini menyajikan hasil pengujian berdasarkan tahapan yang telah dilakukan. Analisis diberikan sebagai basis dari kesimpulan yang diambil dalam penelitian ini.

BAB V KESIMPULAN DAN SARAN

Pada bab ini berisi kesimpulan dari hasil penelitian yang telah dilakukan serta saran yang diharapkan dapat berguna pada penelitian berikutnya.

DAFTAR PUSTAKA

- Budiman, R. (2019). *Penerapan Data Mining Untuk Menentukan Lokasi Promosi Penerimaan Mahasiswa Baru Pada Univeristas Banten Jaya*. 8(5), 55.
- Darmansah, D., & Wardani, N. W. (2021). Analisis Pesebaran Penularan Virus Corona Di Provinsi Jawa Tengah Menggunakan Metode K-Means Clustering. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 8(1), 105–117.
- Erianto, D. (2020). *Sejarah Sistem Penerimaan Mahasiswa Baru di Indonesia*. <https://kompaspedia.kompas.id/baca/paparan-topik/sejarah-sistem-penerimaan-mahasiswa-baru-di-indonesia>
- Han, J., Kamber, M., & Pei, J. (2012). Data Mining Concepts and Techniuques. In *Data Mining*. <https://doi.org/10.1016/b978-0-12-381479-1.00001-0>
- Handoko, S., Fauziah, F., & Handayani, E. T. E. (2020). Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering. *Jurnal Ilmiah Teknologi Dan Rekayasa*, 25(1), 76–88.
- Hardiani, T., Sulisty, S., & Hartanto, R. (2015). Segmentasi Nasabah Tabungan Menggunakan Model RFM (Recency, Frequency, Monetary) dan K-Means Pada Lembaga Keuangan Mikro. *Seminar Nasional Teknologi Informasi Dan Komunikasi Terapan, November, 2015*.
- Haryati, S., Sudarsono, A., & Suryana, E. (2015). Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Universitas Dehasen Bengkulu). *Jurnal Media Infotama*,

11(2), 130–138.

- Jambak, M. I., Jambak, A. I. I., Febrianto, R. T., Saputra, D. M., & Jambak, M. I. (2021). Dimension reduction with extraction methods (principal component analysis - Self organizing map - Isometric mapping) in Indonesian language text documents clustering. *Advances in Intelligent Systems and Computing*, 1179 AISC, 1–9. https://doi.org/10.1007/978-3-030-49336-3_1
- Khakim, M. A., Rahmadhani, L., Purnomo, E. S. B., Idayani, R. W., & Rakhmawati, N. A. (2020). Analisa Jaringan Dokumentasi dan Informasi Hukum Kementerian Tentang Teknologi Informasi Menggunakan Metode K-Means Clustering. *Fountain of Informatics Journal*, 5(1), 27–34.
- Kotu, V., & Deshpande, B. (2015). *Predictive Analytics and Data Mining: Concepts and Practice with RapidMiner*.
- Manochandar, S., Punniyamoorthy, M., & Jeyachitra, R. K. (2020). Development of new seed with modified validity measures for k-means clustering. *Computers & Industrial Engineering*, 141, 106290. <https://doi.org/10.1016/J.CIE.2020.106290>
- Muni, A. (2020). Analisis Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Penjualan Sepeda Motor Studi Kasus PT. Alfa Scorpii. *JUTI UNISI*, 4(1), 1–8.
- Salshabila, R. (2021). *Pengujian Metode Neural Network Backpropagation Dan Linear Regression Dalam Memprediksi Target Penjualan Pupuk Urea PSO (Studi Kasus: PT. Pupuk Sriwidjaja Palembang)*. 3(2), 6.
- Saputra, N. E., Tania, K. D., & Heroza, R. I. (2016). Penerapan Knowledge Management System (KMS) Menggunakan Teknik Knowledge Data

- Discovery (KDD) Pada PT PLN (PERSERO) WS2JB Rayon Kayu Agung. *Paper Knowledge . Toward a Media History of Documents*, 8(2), 1038–1055.
- Sari, I., Maselena, A., Satria, F., & Muslihudin, M. (2018). Application Model of K-means Clustering: Insights Into Promotion Strategy of Vocational High School. *International Journal of Engineering & Technology*, 7(2.27), 182–187.
- Silalahi, N. (2020). Penentuan Strategi Promosi Universitas Budi Darma Menggunakan Algoritma K-Means Clustering. *TIN: Terapan Informatika Nusantara*, 1(1), 40–46.
- Sinaga, S. M., Hardinata, J. T., & Fauzan, M. (2021). Implementasi Data Mining Clustering Tingkat Kepuasan Konsumen Terhadap Pelayanan Go-Jek. *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer Dan Manajemen)*, 2(2), 118–124.
- Syafrinka, A. M. (2021). Penerapan Data Mining Untuk Clustering Data Penduduk Miskin Di Kabupaten Rejang Lebong Menggunakan Metode K-Means. In *Computer-Integrated Technologies: Education, Science, Production* (Issue 42). <https://doi.org/10.36910/6775-2524-0560-2021-42-26>
- Thomas, J. C. R., Peñas, M. S., & Mora, M. (2013). New Version of Davies-Bouldin Index for Clustering Validation Based on Cylindrical Distance. *Proceedings - International Conference of the Chilean Computer Science Society, SCCC*, 0(2), 49–53. <https://doi.org/10.1109/SCCC.2013.29>
- Trillo, C. (2021). *Promotion Strategy; How to Promote Your Business*. <https://www.evinex.com/promotion-strategy/>
- Verma, M., Srivastava, M., Chack, N., Diswar, A. K., & Gupta, N. (2012). A

comparative study of various clustering algorithms in data mining. *International Journal of Engineering Research and Applications (IJERA)*, 2(3), 1379–1384.

Wakhidah, N. (2014). Clustering Menggunakan K-Means Algorithm (K-Means Algorithm Clustering). *Fakultas Teknologi Informasi*, 21(1), 70–80.

Website Resmi Universitas Sriwijaya. (n.d.). <https://www.unsri.ac.id/>

Wurdianarto, S., Wurdianarto, S. R., Novianto, S., & Rosyidah, U. (2014). Perbandingan Euclidean Distance Dengan Canberra Distance Pada Face Recognition. *Techno.Com*, 13(1), 31–37. <https://publikasi.dinus.ac.id/index.php/technoc/article/view/539>