

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**PENDAFTARAN SISWA BARU BERDASARKAN WILAYAH ATAU  
ZONA MENGGUNAKAN METODE CLUSTERING K-MEANS****TUGAS AKHIR**

Disusun Sebagai Salah Satu Syarat  
Untuk Memperoleh Gelar Sarjana Teknik  
Pada Jurusan Teknik Informatika

Oleh:



**IKHSANUL WARDANA**  
**NIM. 11750115190**



**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU**

**PEKANBARU**

**2024**



## LEMBAR PERSETUJUAN

### PENDAFTARAN SISWA BARU BERDASARKAN WILAYAH ATAU ZONA MENGGUNAKAN METODE KLUSTERING K- MEANS

#### TUGAS AKHIR

Oleh

IKHSANUL WARDANA

NIM. 11750115190

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir  
di Pekanbaru, pada tanggal 1 Juli 2024

Pembimbing I,

Dr. Fitri Wulandari, S.Si., M.Kom

NIP. 197410162000032002

UIN SUSKA RIAU

#### Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



## LEMBAR PENGESAHAN

### PENDAFTARAN SISWA BARU BERDASARKAN WILAYAH ATAU ZONA MENGGUNAKAN METODE KLUSTERING K- MEANS

Oleh

IKHSANUL WARDANA

NIM. 11750115190

Telah dipertahankan di depan sidang dewan penguji  
sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik  
pada Universitas Islam Negeri Sultan Syarif Kasim Riau

Pekanbaru, 1 Juli 2024

Mengesahkan,

Ketua Jurusan,



Dekan  
FAKULTAS TEKNIK  
UNIVERSITAS ISLAM NEGERI  
SULTAN SYARIF KASIM RIAU  
Dr. Hartono, M.Pd  
NIP. 19640301 1992031 003



Iwan Iskandar, S.T., MT

NIP. 19821216 201503 1 003

#### DEWAN PENGUJI

Ketua : Liza Afriyanti, S.Kom., M.Kom.

Pembimbing I : Dr. Fitri Wulandari, S.Si., M.Kom.

Penguji I : Dr. Alwis Nazir., M.Kom.

Penguji II : Iis Afrianty, S.T., M.Sc.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada tulisan. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggunaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman dan tanggal pinjam.

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



## LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan didalam daftar pustaka.

Pekanbaru, 1 Juli 2024

Yang membuat pernyataan,

**IKHSANUL WARDANA**

**NIM. 117501151190**

UIN SUSKA RIAU

### Hak Cipta dilindungi Undang-Undang

1. Dilarang menjiplak sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



## LEMBAR PERSEMBAHAN



© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

### Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## ABSTRAK

Sistem zonasi merupakan kebijakan dari Kementerian Pendidikan dan Kebudayaan (Kemendikbud) untuk pemerataan akses dan kualitas pendidikan di Indonesia yang diterapkan oleh SMAN 2 Bangkinang. Namun, masalah ketidakseimbangan calon siswa antar wilayah/zona. Penelitian ini menggunakan metode k-means untuk mengelompokkan pendaftaran siswa baru di SMAN 2 Bangkinang. Data penerimaan siswa dari 2018-2023 ( $\pm 1000$  data) diolah dengan algoritma k-means menggunakan Rapidminer. penggunaan metode k-means dalam pendaftaran siswa baru berdasarkan wilayah atau zona dapat memberikan banyak manfaat, seperti mempercepat proses pendaftaran, mengurangi kesalahan dalam pengelompokan wilayah atau zona, dan meningkatkan efisiensi pengelolaan pendidikan di suatu daerah atau wilayah. Dengan adanya pengelompokan wilayah atau zona yang lebih terorganisir, maka dapat dilakukan pengalokasian sumber daya pendidikan secara lebih tepat sasaran

**Kata Kunci : Clustering, Algoritma K-Means, Sistem Zonasi**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## ABSTRACT

*The zoning system is a policy by the Ministry of Education and Culture (Kemendikbud) to ensure equitable access and quality of education in Indonesia, which is implemented by SMAN 2 Bangkinang. However, the issue of imbalance among prospective students across different areas/zones has arisen. This research uses the K-means method to cluster the registration of new students at SMAN 2 Bangkinang. Student admission data from 2018-2023 (approximately 1000 data points) were processed using the K-means algorithm in RapidMiner. The use of the K-means method in the registration of new students based on area or zone can provide many benefits, such as speeding up the registration process, reducing errors in area or zone clustering, and increasing the efficiency of educational management in a certain area or region. With more organized area or zone clustering, educational resources can be allocated more accurately..*

**Keywords:** *Clustering, K-Means Algorithm, Zoning System*

### Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.





**Hak Cipta Diindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dalam menyelesaikan Tugas Akhir ini.

Ibu Dr. Fitri Wulandari, S.Si, M.Kom. selaku pembimbing Tugas Akhir yang telah memberikan bimbingan, arahan, serta memberikan banyak kritik dan saran yang sangat membangun dalam penyusunan laporan Tugas Akhir ini.

Bapak Dr. Alwis Nazir, S.T., M.Kom. dan Ibu Iis Afrianty, S.T., M.Sc. selaku penguji Tugas Akhir yang telah memberikan masukan serta saran kepada penulis dalam penyusunan Tugas Akhir ini.

Ibu Fadhilah Syafria, S.T., M.Kom. selaku Koordinator Tugas Akhir Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.

Seluruh dosen di jurusan Teknik Informatika UIN Sultan Syarif Kasim Riau yang memberikan ilmu yang sangat bermanfaat serta wawasan yang luar biasa.

9. Kedua orang tua penulis tercinta, Ayahanda dan Ibunda dan adik penulis yang tersayang, serta Keluarga Besar yang tiada henti selalu memberikan dukungan, motivasi, semangat serta doa untuk kelancaran tahap demi tahap tugas akhir penulis.
10. Teruntuk teman-teman Fantastic<sup>17</sup> yang selalu memberikan kenangan indah didalam kelas, dan selalu memberikan semangat satu sama lain.
11. Serta kepada orang-orang yang terlibat dalam pembuatan laporan ini yang tidak bisa disebutkan satu persatu oleh penulis, semoga laporan ini bermanfaat bagi penulis dan orang lain.

Penulis berharap semoga laporan ini dapat bermanfaat bagi pembaca, serta dapat menjadi rujukan dalam melakukan penelitian. Terlebih penulis hanyalah manusia biasa yang tidak lepas dari kesalahan sekecil apapun. Akhir kata semoga laporan ini bermanfaat bagi pembaca dan terima kasih.

Pekanbaru, Juni 2024

Penulis

## DAFTAR ISI

<b>ABSTRAK .....</b>	<b>ii</b>
<b>ABSTRACT .....</b>	<b>iii</b>
<b>KATA PENGANTAR.....</b>	<b>iv</b>
<b>DAFTAR ISI.....</b>	<b>vi</b>
<b>DAFTAR GAMBAR.....</b>	<b>viii</b>
<b>DAFTAR TABEL.....</b>	<b>ix</b>
<b>BAB 1 PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
<b>BAB 2 KAJIAN PUSTAKA .....</b>	<b>6</b>
2.1 Kajian Metode .....	6
2.1.1 Data Mining .....	6
2.1.2 <i>Knowledge Discovery Database</i> (KDD).....	6
2.1.3 Tugas Utama <i>Data Mining</i> .....	7
2.1.4 <i>Clustering</i> .....	9
2.1.5 <i>K-Means Clustering</i> .....	10
2.2 Penelitian Terkait.....	12
<b>BAB 3 METODOLOGI PENELITIAN.....</b>	<b>17</b>
3.1 Mengidentifikasi Masalah .....	17
3.2 Studi Pustaka / Kajian Literatur .....	18
3.3 Pengumpulan Data.....	18
3.4 Analisa Tahapan KDD ( <i>Knowledge Discovery Database</i> ) .....	19
<b>BAB 4 PEMBAHASAN .....</b>	<b>22</b>
4.1 Data Penelitian .....	22
4.2 <i>Data Selection</i> .....	23
4.3 <i>Data Preprocessing / Cleaning</i> .....	24
4.4 <i>Data Transformation</i> .....	26

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



4.5	<i>Data Mining</i> .....	27
4.6	Pengujian <i>RapidMiner</i> .....	34
4.7	<i>Evaluation</i> .....	42
<b>BAB 5 PENUTUP</b> .....		<b>47</b>
5.1	Kesimpulan.....	47
5.2	Saran .....	47
<b>DAFTAR PUSTAKA</b> .....		<b>48</b>



**Hak Cipta Diilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



## DAFTAR GAMBAR

Gambar 1	Desain Penelitian.....	16
Gambar 2	Proses klustering pada tools RapidMiner.....	30
Gambar 3	Operator Read Excel .....	31
Gambar 4	Proses import data excel ke rapid .....	31
Gambar 5	Operator select atribut.....	32
Gambar 6	Proses pemilihan atribut.....	32
Gambar 7	Operator normalisasi .....	33
Gambar 8	Gambar Operator klustering.....	34
Gambar 9	Operator performance .....	35
Gambar 10	Nilai davies bouldin index .....	36
Gambar 11	cluster model .....	36
Gambar 12	Visualisation.....	37

© Hak Cipta Milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan pendidikan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.





Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

## BAB 1 PENDAHULUAN

### 1. Latar Belakang

“Mencerdaskan kehidupan bangsa” sebagaimana Alinea ke4 pembukaan UUD 1945 merupakan salah satu ide dasar dan alasan dalam membentuk pemerintahan negara Indonesia agar kita memiliki masyarakat terdidik dan cerdas. Ada juga pasal 31 UUD 1945 ayat 1 yang berbunyi “setiap warga negara berhak mendapat pendidikan”. Permasalahan muncul dalam bentuk adanya kesenjangan dari aspek mutu Pendidikan, kesempatan berkembang bagi peserta didik, dan adanya stigma “sekolah favorit”. Dengan diterapkannya sistem zonasi diharapkan dapat menanggulangi masalah-masalah tersebut.

Sistem zonasi merupakan salah satu kebijakan dari Kementerian Pendidikan dan Kebudayaan (Kemendikbud) untuk menghadirkan pemerataan akses pada layanan Pendidikan, serta pemerataan kualitas Pendidikan di Indonesia. Sistem zonasi sangat dipengaruhi dengan persebaran sekolah. Dengan terbatasnya jumlah sekolah di suatu zona menyebabkan kemungkinan anak yang tidak tertampung menjadi sangat besar terutama pada siswa yang rumahnya jauh dari zona sekolah. Pemerataan Pendidikan tidak dapat dicapai dengan kurang meratanya persebaran sekolah. Dibutuhkan analisis pemetaan zonasi untuk menilai baik tidaknya pncanangan sistem zonasi di Indonesia yang dalam studi kasus ini dalam lingkup Kampar(Firdaus 2017).

Pada perkembangan era digital saat ini, pendaftaran siswa baru tidak lagi terbatas pada pendaftaran secara langsung ke sekolah. Sekarang calon siswa baru dapat mendaftar melalui media online seperti website atau aplikasi yang disediakan oleh sekolah(Zulham and Hasugian 2022). Hal ini tentu memudahkan orang tua/wali siswa dalam melakukan pendaftaran siswa baru tanpa harus datang langsung ke sekolah.

SMAN 2 Bangkinang adalah sebuah Sekolah Menengah Atas Negeri yang terletak di Bangkinang, Kabupaten Kampar, Provinsi Riau, Indonesia. Sekolah ini merupakan salah satu sekolah negeri yang dioperasikan oleh pemerintah untuk memberikan pendidikan berkualitas kepada siswa di tingkat pendidikan menengah



#### Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

atas. Prosedur pendaftaran siswa baru di SMAN 2 Bangkinang menggunakan jalur zonasi. Proses pendaftaran siswa baru melalui jalur zonasi dibagi beberapa tahap yaitu, Pengumpulan Informasi, Verifikasi Kepesertaan Zonasi, Pengisian Formulir Pendaftaran, Persyaratan Dokumen, Penyerahan Dokumen, Seleksi dan Pengumuman, Registrasi dan Pembayaran. Prosedur pendaftaran siswa baru di SMAN 2 Bangkinang dapat berubah atau mengikuti peraturan yang ditetapkan oleh pemerintah daerah atau dinas pendidikan setempat.

Sangat penting bagi calon siswa dan orang tua untuk menghubungi sekolah langsung atau mengakses situs web resmi sekolah untuk mendapatkan informasi terbaru mengenai prosedur pendaftaran melalui jalur zonasi yang berlaku. Meskipun pendaftaran secara online telah menjadi alternatif yang populer, proses pendaftaran tetaplah membutuhkan waktu dan usaha. Selain itu, dengan banyaknya sekolah dan calon siswa yang ingin mendaftar, kadang kala terjadi ketidakseimbangan jumlah calon siswa baru di setiap wilayah/zona (Haris Ihsanil Huda and Utami Dewi 2018). Hal ini dapat menyebabkan beberapa wilayah/zona memiliki jumlah siswa baru yang berlebihan, sementara wilayah/zona lainnya kurang memiliki jumlah siswa baru (Yandri, Siboro, and Simanjuntak 2019). Pendaftaran siswa baru berdasarkan wilayah atau zona merupakan suatu proses yang penting dalam pengelolaan pendidikan di suatu daerah atau wilayah. Salah satu metode yang dapat digunakan dalam melakukan pengelompokan wilayah atau zona adalah metode k-means (Desi et al. 2022).

Metode k-means merupakan salah satu metode clustering yang paling banyak digunakan dalam pengolahan data. Metode ini bertujuan untuk mempartisi data menjadi beberapa kelompok atau cluster, di mana setiap data dikelompokkan ke dalam cluster yang memiliki karakteristik yang sama atau sejenis (Desi et al. 2022).

Pengumpulan data penerimaan siswa baru berdasarkan faktor wilayah atau zona. Wilayah adalah sebuah zona yang berkelanjutan dan berada di antara level lokal dan nasional (Yandri, Siboro, and Simanjuntak 2019). Wilayah juga tergantung pada tujuan analisis atau kebijakan pengembangan wilayah yang akan diumumkan. Dalam melakukan pengelompokan wilayah atau zona menggunakan metode k-means, terlebih dahulu perlu dilakukan pemilihan jumlah cluster atau

#### Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

kelompok yang belum diketahui. Setelah itu, dilakukan pengelompokan wilayah atau zona berdasarkan jarak antara wilayah atau zona tersebut dengan pusat cluster yang belum ditentukan sebelumnya. Pusat cluster merupakan titik rata-rata dari semua data pada cluster yang bersangkutan. Setelah proses pengelompokan selesai dilakukan, maka wilayah atau zona dapat dikelompokkan ke dalam kelompok-kelompok yang homogen atau sejenis, sehingga memudahkan dalam proses pendaftaran siswa baru berdasarkan wilayah atau zona yang telah dibuat. Implementasinya, penggunaan metode k-means dalam pendaftaran siswa baru berdasarkan wilayah atau zona dapat memberikan banyak manfaat, seperti mempercepat proses pendaftaran, mengurangi kesalahan dalam pengelompokan wilayah atau zona, dan meningkatkan efisiensi pengelolaan pendidikan di suatu daerah atau wilayah.

K-Means merupakan salah satu metode data clustering non-hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster/kelompok (Yandri, Siboro, and Simanjuntak 2019). Teknik clustering dapat digunakan dalam berbagai bidang seperti dalam ilmu sosial, bisnis, kedokteran, dan lain sebagainya. Algoritma k-means bekerja dengan cara menginisialisasi titik pusat kelompok secara acak dan mengelompokkan data berdasarkan jarak terdekat dari titik pusat. Kemudian, algoritma akan menghitung kembali titik pusat dari masing-masing kelompok dan mengulang proses pengelompokan hingga titik pusat tidak berubah atau mencapai konvergensi.

Berdasarkan permasalahan di atas pada penelitian ini akan menggunakan algoritma k-means untuk proses penerimaan siswa/i baru pada suatu sekolah di Kabupaten Kampar yang nantinya akan menghasilkan data siswa/i yang sangat banyak dan menjadi faktor untuk menentukan pengelompokan siswa/i diterima sekolah sesuai dengan wilayah atau zona dari pendaftar tersebut. Implementasi algoritma k-means pada data ini menggunakan *tools Rapidminer* sebagai alat bantu untuk menemukan nilai yang akurat. Hasil yang akan diharapkan berupa tingginya tingkat keberhasilan siswa/i dan rendahnya tingkat kegagalannya dalam penentuan kelulusan diterima pada suatu sekolah. Clustering juga memegang peran sangat penting dalam eksplorasi data ilmu pengetahuan, pengaksesan informasi dan analisis web (Hidayat 2022).



Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## 1.1 Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah dari Tugas Akhir ini adalah: Bagaimana mengelompokkan wilayah/zona pendaftaran siswa baru menggunakan metode k-means?

## 1.2 Batasan Masalah

Adapun batasan masalah dalam Tugas Akhir ini adalah sebagai berikut:

1. Implementasi metode k-means pada pendaftaran siswa baru berdasarkan wilayah/zona.
2. Atribut dan variable dalam penelitian ini adalah kecamatan, kelurahan atau desa, jarak tempuh pergi ke sekolah.

## 1.4 Tujuan Penelitian

Tujuan dari penelitian Tugas Akhir ini adalah Mengelompokkan wilayah/zona yang memiliki jumlah pendaftar siswa baru menggunakan metode k-means, sehingga dapat memudahkan proses administrasi dan pelayanan Pendidikan.

## 1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Membantu proses pendaftaran siswa baru menjadi lebih efisien dan efektif.
2. Memudahkan orang tua/wali siswa dalam melakukan pendaftaran siswa baru.
3. Membantu sekolah dalam mengatur jumlah pendaftar siswa baru yang seimbang di setiap wilayah/zona.
4. Menjadi referensi bagi peneliti selanjutnya dalam melakukan penelitian tentang pendaftaran siswa baru menggunakan metode k-means.

## BAB 2

### KAJIAN PUSTAKA

#### Kajian Metode

##### 2.1.1 Data Mining

Data mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Informasi tersebut dihasilkan dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat pada basis data (Hafizah, Tegiono, and Maya 2019). Data mining terutama digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut *Knowledge Discovery Database* (KDD). Data mining adalah proses pencarian secara otomatis informasi yang berguna dalam tempat penyimpanan data berukuran besar. Teknik data mining digunakan untuk memeriksa basis data berukuran besar sebagai cara untuk menemukan pola yang baru dan berguna.

Sehingga dapat disimpulkan bahwa data mining adalah sebuah proses untuk mencari informasi yang tidak diketahui secara manual terdapat dalam sebuah basis data dan merubahnya menjadi sebuah informasi atau pengetahuan baru yang berguna bagi penggunanya.

##### 2.1.2 *Knowledge Discovery Database* (KDD)

Istilah data mining dan *Knowledge Discovery Database* (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu database besar. *Knowledge Discovery Database* (KDD) merupakan sebuah proses mengubah data mentah menjadi data yang bermanfaat berbentuk informasi. *Knowledge Discovery Database* (KDD) memiliki Langkah-langkah yang dilakukan secara berurutan.

*Knowledge Discovery Database* (KDD) yang terdiri (Manurung and Hasugian 2019) sebagai berikut:

a. *Data Cleaning*

Langkah ini digunakan untuk menghapus data yang tidak digunakan dan data yang tidak konsisten.

#### Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

b. *Data Integration*

Langkah ini memungkinkan adanya penggabungan dari beberapa sumber data.

c. *Data Selection*

Langkah ini adalah seleksi data dimana merupakan proses menganalisis data yang relevan dari dalam database.

d. *Data Transformation*

Langkah ini mengubah data ke dalam bentuk yang sesuai untuk diolah dengan menganalisis ringkasan atau jumlah agregasi.

e. *Data Mining*

Langkah ini merupakan hal penting di mana metode cerdas diterapkan guna mengolah pola-pola data.

f. *Patterns Evaluation*

Langkah ini mengidentifikasi pola-pola menarik yang menjelaskan mengenai ukuran dasar pengetahuan yang ada.

g. *Knowledge Presentation*

Proses untuk menyajikan atau memaparkan pengetahuan didapatkan dari hasil data yang telah diolah.

### 2.7.3 Tugas Utama *Data Mining*

Data Mining memiliki enam tugas utama yang dapat dilakukan (Tampubolon et al. 2013), sebagai berikut:

a. *Description* (Deskripsi)

Tujuan deskripsi untuk mengidentifikasi pola yang muncul secara berulang pada data sekaligus mengubah polanya menjadi aturan dan kriteria yang mudah dimengerti. Aturan dan kriteria yang mudah dimengerti ini berguna untuk meningkatkan pengetahuan (*knowledge*) pada sebuah sistem. Deskripsi sering dibutuhkan dalam teknik postprocessing guna melakukan validasi dan menjelaskan hasil dari proses data mining. Postprocessing adalah proses yang digunakan untuk



memastikan hasil valid dan berguna untuk digunakan pihak berkepentingan dalam data tersebut.

b. *Prediction* (Prediksi)

Prediksi hampir sama dengan klasifikasi, akan tetap data diklasifikasikan berdasarkan perilaku atau nilai yang diperkirakan pada masa yang akan datang. Contoh : Prediksi harga cabai untuk tiga bulan ke depan, prediksi harga saham, prediksi jumlah pengunjung lokasi wisata. Dalam prediksi metode klasifikasi dan estimasi dapat digunakan guna menghasilkan sebuah data yang valid.

c. *Estimation* (Estimasi)

Estimasi hampir sama dengan prediksi, namun *variable* target dari estimasi mengarah ke numerik dari pada kategori. Model ini dibangun menggunakan *record* lengkap yang menyediakan nilai dari *variable* target sebagai nilai prediksi. Pada peninjauan selanjutnya estimasi nilai dari *variable* target dibuat berdasarkan nilai *variable* prediksi. Contoh : Dilakukan estimasi terhadap tekanan darah sistolik pasien rumah sakit berdasarkan umur, jenis kelamin, berat badan, dan sodium darah. Hubungan tekanan darah sistolik dan *variable* prediksi dalam proses pembelajaran menghasilkan model estimasi.

d. *Classification* (Klasifikasi)

Klasifikasi adalah proses untuk menemukan suatu model atau fungsi guna mendeskripsikan dan membedakan data ke dalam beberapa kelas. Contoh : Penggolongan pendapatan dapat dibedakan menjadi tiga bagian, pendapat tinggi, pendapatan sedang, dan pendapatan rendah.

e. *Clustering* (Pengklastran)

Pengklastran adalah pengelompokan data tanpa berdasarkan kelas data tertentu ke dalam kelas objek yang memiliki kesamaan. *Cluster* adalah kumpulan *record* yang memiliki kemiripan satu sama lain dan memiliki ketidakmiripan satu sama lain dalam cluster lain. Tujuan dari pengklastran ini adalah menghasilkan pengelompokan objek yang mirip satu sama lain ke dalam sebuah kelompok. Semakin besar kemiripan dari

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

sebuah objek cluster dan semakin besar perbedaan tiap cluster maka kualitas analisis semakin baik.

f. *Association* (Asosiasi)

Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam satu waktu. Asosiasi berguna mengungkap aturan untuk mengukur hubungan dua atau lebih atribut.

## 2.1.4 Clustering

### 1. Pengertian Clustering

*Clustering* adalah salah satu sub bab dari data mining dan merupakan proses di mana sampel yang sama dibagi menjadi kelompok-kelompok yang disebut *cluster*. Setiap cluster termasuk sampel di mana anggota yang mirip satu sama lain dan berbeda dengan sampel yang tersedia dari kelompok lain (Pradana, Nurcahyo, and Saputro 2021). Analisa *cluster* merupakan Teknik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimiliki. Analisis cluster mengklasifikasikan objek sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam *cluster* yang sama.

### 2. Jenis-Jenis Clustering

Dalam data mining terdapat jenis-jenis metode yang dapat digunakan, berikut jenis-jenis metode clustering:

a. *Centroid-Based Clustering*

*Centroid-based Clustering* adalah metode *clustering* dengan cara mengelompokkan data-data ke dalam sebuah *clusters* yaitu *nonhierarchical clusters*. *Centroid-based Clustering* merupakan jenis metode yang sangat peka terhadap *outlier*. *Centroid-based Clustering* merupakan salah satu yang menerapkan algoritma iteratif dalam *clustering*, dimana sebuah cluster dibentuk dari jarak terdekat antara titik data ke pusat *cluster*. Pusat *centroid* dibentuk dengan mempertimbangkan beberapa hal sehingga jarak titik data menjadi minimum dengan pusat cluster. Algoritma yang populer dari *Centroid-based Clustering* adalah *K-Means Clustering*.

#### Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

b. *Density-Based Clustering*

*Density-based Clustering* merupakan metode dengan menghubungkan area dengan kepadatan yang sama ke dalam satu kelompok. Dengan metode ini *cluster* dibentuk berdasarkan kepadatan dari masing-masing data point. Wilayah yang padat atau memiliki data yang banyak akan dianggap sebagai satu *cluster*. Sedangkan wilayah yang memiliki data sedikit dianggap sebagai *outlier*. Algoritma yang termasuk *Density-based Clustering* adalah DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*), OPTICS (*Ordering Points to Identify Clustering Structure*), dan HDBSCAN (*Hierarchical Density-Based Spatial Clustering of Applications with Noise*).

c. *Distribution-Based Clustering*

*Distribution-based Clustering* merupakan metode dengan mengasumsikan data terdiri dari sebuah distribusi. Bertambahnya jarak dari pusat distribusi, maka probabilitas suatu titik termasuk ke dalam kelompok distribusi akan berkurang. Metode ini bekerja dengan baik pada data sintesis dan cluster dengan ukuran yang beragam. Algoritma yang termasuk *Distribution-based Clustering* adalah *Expectation-maximization*.

d. *Hierachical Clustering*

*Hierachical Clustering* merupakan tipe yang mirip dengan *centroid-based Clustering*, dikarenakan *Hierachical Clustering* mendefinisikan cluster berdasarkan jarak terdekat antara data *point*. Pada dasarnya data point yang lebih dekat akan memiliki perilaku yang sama dibandingkan dengan data *point* yang lebih jauh. Pengelompokan akan direpresentasikan dengan menggunakan dendogram.

2.1.5 *K-Means Clustering*

1. *Pengertian K-Means Clustering*

*K-means clustering* merupakan salah satu metode *cluster analysis* non hirarki yang berusaha untuk mempartisi objek yang ada ke dalam satu atau lebih cluster berdasarkan karakteristiknya (Nur Khormarudin 2016). Sehingga objek yang mempunyai karakteristik yang sama dapat

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dikelompokkan dalam satu *cluster* yang sama. Sedangkan objek yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam cluster lain (Talakua, Leleury, and Taluta 2017). K-means merupakan algoritma klusterisasi yang memiliki dasar sederhana dengan meminimalkan *Sum of Square Error* (SSE) antara objek-objek data dengan sejumlah *k centroid*. K-means merupakan algoritma yang menetapkan nilai-nilai *cluster* (*k*) secara acak, sementara nilai tersebut menjadi pusat dari *cluster* yang disebut *centroid*. Kemudian menghitung jarak setiap data yang ada terhadap masing-masing *centroid* dan ditemukan jarak paling dekat dari setiap data dengan *centroid* hingga nilai *centroid* tidak berubah (stabil).

2. Keuntungan *K-Means*

Adapun keuntungan dari metode *K-means* (Aldillah 2018) antara lain:

- a. K-means di beberapa aplikasi disebut dengan segmentasi data, hal ini dikarenakan pengelompokan mempartisi set data besar ke dalam grup sesuai dengan kesamaannya.
- b. K-means digunakan juga untuk mendeteksi *outlier*.
- c. K-means mempartisi sekumpulan objek data ke dalam himpunan bagian, sehingga banyak digunakan dalam aplikasi seperti intelijen bisnis, pengenalan pola gambar, pencarian *web*, biologi, dan keamanan.

3. Kekurangan *K-Means*

Selain memiliki keuntungan dalam penggunaannya, K-means juga memiliki beberapa kekurangan antara lain:

1. *K-means* tidak dapat menjamin *konvergen* pada optimum global.
2. *K-means* sering terjebak pada optimum lokal, dimana *centroid* akhir yang dihasilkan tidak benar-benar menjadi pusat *cluster* yang sesungguhnya.
3. *Output K-means* bergantung pada *centroid* awal yang ditentukan secara acak.



Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## Penelitian Terkait

Beberapa penelitian terkait yang akan menjadi acuan penulis mengenai implementasi *K-Means clustering* sebagai berikut:

**Tabel 1 Penelitian Terkait**

No	Peneliti	Tahun	Judul	Hasil
1	S. Ashari, S. Khansa, C. H. M. Surudin, and I. N. Isnainiyah	2018	Klustering Jumlah Penduduk Kota Bandung Berdasarkan Jenis Kelamin Per Kecamatan	Menerapkan proses klustering dengan metode k-means sesuai jumlah penduduk di kota bandung tersebut dengan maksud untuk mendapatkan kluster atas padat atau tidaknya jumlah penduduk di kota bandung perkecamatan(Ashari et al. 2018).
2	E. Desi, S. Aliyah, S. Lestari, and W. Dari	2022	Implementasi Algoritma K-Means Untuk Penerimaan Siswa Baru Di SMANPAS Berdasarkan Nilai Rapot dan Hasil Tes	Hasil dari penelitian ini akan menghasilkan 4 orang pelajar baru akan diterima berdasarkan perhitungan K-Means dengan titik acuan atau Cluster dengan nilai 91 untuk nilai rapot dan 88 untuk hasil tes ujian, 2 Orang pelajar baru sebagai cadangan dengan Cluster nilai 70 untuk rapot dan 70 untuk hasil tes ujian, dan 2 Orang pelajar yang tidak diterima di SMANPAS dengan Cluster nilai 60 untuk rapot dan 70 untuk hasil tes ujian(Desi et al. 2022).

<p>3. © Hak cipta milik UIN Suska Riau</p>	<p>S. Badruddin et al.</p>	<p>2022</p>	<p>Model Klasterisasi Wilayah Kota Berdasarkan Indeks Infrastruktur Untuk Perencanaan Struktur Ruang Berkelanjutan</p>	<p>Hasil dari penelitian adalah Analytical Hierarchy Process (AHP), diketahui prioritas pembangunan adalah pembangunan jalan, ruang terbuka hijau, listrik, air minum, telekomunikasi, air limbah, drainase, dan terakhir persampahan. Hasil analisis SWOT menunjukkan bahwa grafik terletak pada bagian kekuatan dan peluang. Sehingga kebijakan yang diambil ke depan adalah fokus utama pada pembangunan infrastruktur persampahan, air limbah, ruang terbuka hijau, dan telekomunikasi (Badruddin, Achmad, and Fadhy 2022).</p>
<p>4</p>	<p>D. Firdaus</p>	<p>2017</p>	<p>Penggunaan Data mining dalam kegiatan pembelajaran</p>	<p>Hasil yang diharapkan dari proses data mining akan dipisahkan menjadi pengumpulan data, transformasi data, dan analisis data. Aturan asosiasi, klasifikasi, dan clustering algoritma data mining yang dieksplorasi dalam analisis data untuk computer dibantu system pembelajaran (Firdaus 2017).</p>
<p>State Islamic University of Sultan Syarif Kasim Riau</p>	<p>E. Dianawati, P. P. Yanti, and Y. Suryandari</p>	<p>2019</p>	<p>Klustering Jumlah Penumpang pada Halte Bus Rapid Transit Kota Tangerang</p>	<p>Hasil klustering memiliki sifat yang hamper sama antara data dalam satu kluster yaitu jumlah penumpang yang masuk per bulan, sehingga dengan melakukan karakteristik klustering, pengetahuan tentang padatnya jadwal penumpang dalam satu hari dapat membantu pihak terkait untuk mengantisipasi kepadatan tersebut (Dianawati, Yanti, and Suryandari 2019).</p>

**Hak Cipta Diindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

<p>© Hak cipta milik UIN Suska Riau</p>	<p>M. G. Pradana, A. C. Nurcahyo, and P. H. Saputro</p>	<p>2020</p>	<p>Penerapan Metode K-Means Klustering untuk Menentukan Kepuasan Pelanggan K-Means Clustering Method to Determine Customer Satisfaction</p>	<p>Teknik klustering yaitu memisahkan dan mengelompokkan data berdasarkan kluster. Hasil dari penelitian ini terbagi menjadi 2 kluster yaitu kluster 0 yaitu puas dan kluster 1 yaitu tidak puas ataupun netral(Pradana, Nurcahyo, and Saputro 2021).</p>
<p>State Islamic University of Sultan Syarif Kasim Riau</p>	<p>H. K. Candra, M. Bahit, and B. Sabella</p>	<p>2021</p>	<p>Penerapan Metode Klustering Fuzzy C- Means Untuk Penentuan Peminatan Pemilihan Jurusan Pada Sekolah Menengah Tingkat Atas</p>	<p>Hasil pengklasteran algoritma K-Means disebabkan karena (a) Algoritma K- means melakukan pengklasteran data siswa berdasarkan kemiripan pola data (nilai) dalam kelompok- kelompok yang telah ditetapkan, dan tidak terikat pada suatu aturan atau nilai-nilai variable tertentu. (b) metode pengklasteran siswa yang di gunakan di SMA 13 banjarmasin dalam penentuan jurusan yaitu mengelompokkan siswa berdasarkan kesamaan nilai dalam kelompok- kelompok yang telah di tetapkan, tetapi terikat pada suatu aturan atau nilai variable tertentu, yaitu nilai standar minimum (nilai kriteria ketuntasan minimum/ KKM) untuk masuk pada kelompok tertentu.(Santosa, Fu'adah, and Rizal 2023)</p>

**Hak Cipta Diindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

<p>∞ © Hak cipta milik UIN Suska Riau</p> <p>10 State Islamic University of Sultan Syarif Kasim Riau</p>	Rahmawati and T. Arifin	2020	Penerapan Algoritma K-Means Untuk Pengelompokan Siswa Lolos Snmptn Di Sman 8 Bandung	Algoritma k-means dengan tujuan untuk memperoleh informasi mengenai perguruan tinggi mana yang paling banyak diminati oleh siswa siswinya. Hasil penelitian berdasarkan validasi menggunakan metrik davis bouldin index maka diperoleh 2 cluster yang di bentuk sebanyak 2 kluster dengan hasil 0,558(Arifin, Sanjaya, and Sanjaya 2020).
	H. Harliana	2020	Penerapan Algoritma Clustering Untuk Mengelompokkan Kenakalan Siswa	Hasil pengujian pada 23 dataset maka di dapatkan kluster ke-1 adalah kluster dimana siswa memiliki kenakalan umum paling banyak(Harliana 2020).
	R. K. Aldi Nurzahputra, Much Aziz Muslim	2016	Online Fuzzy C-Means Clustering for Lecturer Performance Assessment Based on National and International Journal Publication	Hasil penelitian mengelompokkan kemampuan siswa pintar, siswa sedang dan siswa kurang pintar. Sehingga siswa mana saja yang akan diberi pelajaran tambahan agar dapat mencapai nilai standar kelulusan ujian nasional(Aldi, Aziz, and Roni 2016).
	H. Sariangseh, W. Wanayumini, and R. Rosnelly	2021	Penentuan Kelas Menggunakan Algoritma K Medoids Untuk Clustering Siswa Tunagrahita	Hasilnya pengelompokkan didapat 3 cluster, yang memiliki jumlah siswa paling tinggi ialah kelas tunagrahita sedang dan cluster yang rendah ialah tunagrahita ringan, pihak sekolah luar biasa Muzdalifah dapat mempersiapkan kelas berdasarkan pengelompokkan untuk aktivitas belajar mengajar(Sariangseh, Wanayumini, and Rosnelly 2021).

**Hak Cipta Diindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



<p>20 © Hak cipta milik UIN Suska Riau</p>	<p>E. Fammaldo and L. Hakim</p>	<p>2019</p>	<p>Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Tingkat Kesejahteraan Keluarga Untuk Program Kartu Indonesia Pintar</p>	<p>Hasil dari penelitiannya akan berupa laporan 3 tabel yang dimana tiap tabel berisi data keluarga yang tergolong miskin, sederhana dan kaya, dimana golongan sederhana ini nantinya bisa ditinjau kembali apakah berhak atau tidak mendapatkan bantuan program Kartu Indonesia Pintar(Fammaldo and Hakim 2019).</p>
<p>13 © Suska Riau</p>	<p>R. Hidayat</p>	<p>2022</p>	<p>Clustering Menggunakan Algoritma K-Means untuk Mengelompokan Wilayah Rawan Kejahatan di Wilayah Kabupaten Solok</p>	<p>Hasil penelitian ini dengan mengelompokkan wilayah yang rawan tindak kejahatan, dapat memberikan pengetahuan baru atau informasi baru bagi warga yang tinggal disatu wilayah di Kabupaten Solok agar lebih meningkat kewaspadaan agar terhindar dari tindak kejahatan sehingga dapat menurunnya tingkat kriminalitas di wilayah Kabupaten Solok(Hidayat 2022).</p>
<p>14 State Islamic University of Sultan Syarif Kasim Riau</p>	<p>Sinaga, Sulaiman Sembiring, Rahmat W, Sumarno, S</p>	<p>2022</p>	<p>Penerapan Algoritma Naive Bayes untuk Klasifikasi Prediksi Penerimaan Siswa Baru</p>	<p>Hasil dari penelitian ini berdasarkan pengujian sebanyak 30 data testing yang diolah menggunakan Rapid Miner, diperoleh tingkat akurasi sebesar 76,67% yaitu 26 siswa diterima dan sebanyak 4 tidak diterima. Kesimpulannya bahwa proses prediksi berupa klasifikasi dengan menggunakan Naive Bayes dapat lebih cepat dan akurat serta menghasilkan tingkat akurasi yang tinggi(Sinaga, Sembiring, and Sumarno 2022).</p>

**Hak Cipta Diindungi Undang-Undang**

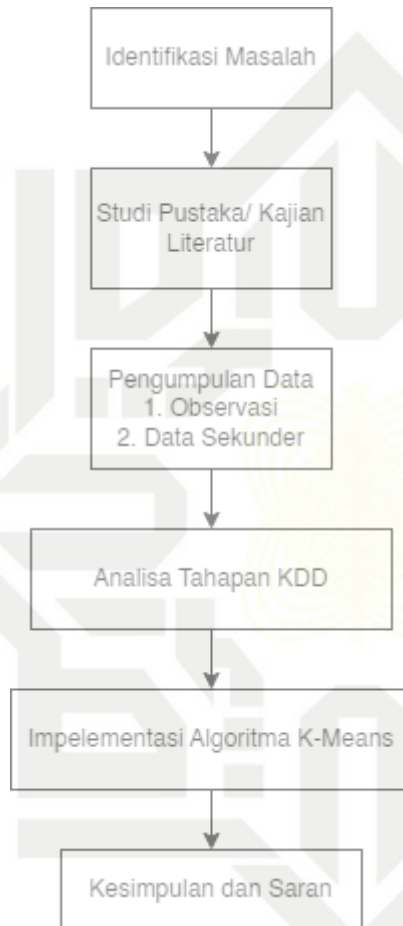
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**Hak Cipta Diindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## BAB 3 METODOLOGI PENELITIAN

Metode penelitian merupakan alur yang di susun secara ilmiah yang dijadikan pedoman dalam melakukan penelitian, hal ini bertujuan untuk tercapainya tujuan yang diharapkan. Alur dalam penelitian ini digambarkan



sebagai berikut :

**Gambar 1 Desain Penelitian**

### 3.1 Mengidentifikasi Masalah

Pada tahap ini dilakukan identifikasi terhadap permasalahan yang dijadikan topik penelitian. Perumusan masalah dilakukan untuk mendapatkan solusi yang diharapkan dari suatu penelitian. Perumusan masalah yang akan diteliti adalah mengelompokkan jalur zonasi pendaftaran siswa baru pada SMAN (Sekolah Menengah Atas Negeri) 2 Bangkinang menggunakan metode K-Means Clustering.

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

### 3.2 Studi Pustaka / Kajian Literatur

Studi Pustaka merupakan cara pengumpulan data dengan cara mengambil data atau informasi melalui berbagai sumber seperti buku, jurnal, artikel, dan melalui internet. Pengumpulan data melalui internet harus dipastikan bahwa sumber data atau informasi yang digunakan berasal dari pihak yang berkompeten dan terpercaya. Studi Pustaka dilakukan sebagai acuan data yang berhubungan dengan penelitian yang akan dilakukan.

### 3.3 Pengumpulan Data

Pengumpulan data merupakan tahap yang dilakukan guna mendapatkan berbagai informasi secara lengkap dan rinci guna mencapai tujuan penelitian. Setelah data yang dibutuhkan telah terkumpul, maka dapat dilakukan transformasi data dan normalisasi data. Adapun cara yang dapat dilakukan untuk pengumpulan data, sebagai berikut :

#### 1. Observasi

Observasi merupakan pengumpulan data dengan cara melakukan tanya jawab secara langsung dengan pihak yang bersangkutan terhadap penelitian yang akan dilakukan. Pada penelitian ini observasi dilakukan oleh peneliti dengan bapak Suardiman, S. E selaku Kepala Tata Usaha SMAN 2 Bangkinang.

Observasi ini dilakukan untuk mengetahui hal-hal mengenai berbagai informasi mengenai jalur pendaftaran siswa baru yang diterapkan pada sekolah tersebut, berupa jalur zonasi dan lingkungan zonasi yang menjadi rujukan penerimaan siswabarbaru.

#### Data Sekunder

Data sekunder adalah data yang dikumpulkan oleh pihak lain untuk tujuan lain, tetapi dapat digunakan dalam penelitian atau analisis. Data ini tidak dikumpulkan secara langsung oleh peneliti atau pengguna data. Penelitian ini menggunakan data sekunder dalam konteks zonasi pendaftaran siswa baru pada SMA yang mencakup data siswa baru, data sekolah terdahulu, data tempat tinggal/domisili atau data dari sumber lain yang relevan dengan proses zonasi tersebut. Data sekunder dapat digunakan dalam

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

tahap pemilihan data dan diproses sebelum diterapkan pada tahapan KDD menggunakan algoritma K-Means.

### Analisa Tahapan KDD (*Knowledge Discovery Database*)

Berikut merupakan penerapan KDD dalam penelitian ini :

1. Data Cleaning

Pada penelitian ini menggunakan data Penerimaan Siswa yang telah lulus dan diterima pada Sekolah Menengah Atas Negeri (SMAN) 2 Bangkinang dari Tahun Pelajaran 2018-2023 jumlah  $\pm 1000$  data. Tahapan ini proses pembersihan data yang tidak lengkap atau *noise* seperti data tempat tinggal, data tempat tinggal yang diluar zona, data siswa baru, data asal sekolah dan data jarak tempuh siswa pergi ke Sekolah. Jumlah data yang tidak lengkap sekitar  $\div 300$  data.

2. Data Integration

Pada penelitian ini peneliti melakukan penggabungan data, yaitu menggabungkan data asal sekolah, data tempat tinggal dan data jarak tempuh pergi siswa ke Sekolah. Tahapan ini proses penggabungan data yang telah dibersihkan menjadi data induk yang akan diproses selanjutnya, data yang telah digabungkan sekitar  $\pm 800$  data terhitung dari Tahun Pelajaran 2018-2023.

3. Data Selection

Pada tahap ini peneliti melakukan seleksi terhadap atribut pada data, setelah melakukan seleksi, peneliti memerlukan atribut dari data asal sekolah, tempat tinggal atau domisili tempat tinggal dan data jarak tempuh siswa pergi ke sekolah.

4. Data Transformation

Tahapan ini data ditransformasi menjadi angka satu hingga sekian menurut banyaknya data yang digunakan, hal ini bertujuan untuk memudahkan peneliti dalam melakukan pengelompokan.

5. Data Mining

Pada tahap ini dilakukan pengolahan data berdasarkan tahapan pada algoritma K-Means. K-means Clustering memiliki empat Langkah utama



Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

sebagai berikut:

- a. Dipilih beberapa k objek secara acak sebagai *centroid* awal.
- b. Data yang tidak termasuk *centroid* digabungkan ke *cluster* terdekat berdasarkan ukuran jarak tertentu.
- c. Setiap *centroid* diperbarui berdasarkan rata-rata dari objek yang ada di dalam setiap *cluster*.
- d. Langkah ke dua dan ke tiga dilakukan berulang (iterasi) hingga *centroid* stabil yaitu, semua *centroid* yang dihasilkan dalam iterasi baru sama dengan semua *centroid* yang dihasilkan pada iterasi sebelumnya.

Berikut merupakan langkah-langkah algoritma *K-Means Clustering* :

- a. Penentuan *cluster* awal

Dalam penentuan n buah pusat cluster awal dilakukan acak guna merepresentasikan urutan dan input. Pusat awal cluster didapatkan dari data sendiri bukan dengan menentukan titik baru, yaitu dengan acak pusat awal dari data.

- b. Perhitungan jarak dengan pusat *cluster*

Untuk mengukur jarak antar data dengan pusat cluster digunakan Euclidian distance, sebagai berikut :

$$d(x,y) = \sum_{i=0}^n (x_i - y_i)^2$$

Penjelasan :

$x_i$  = Data pada cluster ke i

$y_i$  = Data pada cluster awal

- c. Pengelompokan data

Jarak hasil perhitungan akan dilakukan perbandingan dan dipilih jarak terdekat antara data dengan pusat cluster, hasil perhitungan ini menunjukkan bahwa data tersebut beradadalam satu kelompok dengan pusat cluster terdekat. Adapun cara dalam pengelompokan pilih nilai jarak tiap pusat cluster dengan data, kemudian cari nilai jarak terkecil, setelah itu kelompokkan data denganpusat cluster yang memiliki jarak terkecil.

- d. Penentuan pusat *cluster* baru

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Pusat cluster baru digunakan untuk melakukan iterasi selanjutnya, iterasi ini terus berjalan jika hasil yang didapatkan belum konvergen. Adapun langkah dalam penentuan pusat *cluster* cari jumlah anggota tiap *cluster*, setelah itu hitung pusat baru dengan rumus sebagai berikut :

$$X_{centroid\_baru} = \frac{x_1 + x_2 + x_3 \dots + x_n}{\sum n}$$

Penjelasan :

$$x_1 + x_2 + x_3 \dots + x_n = \text{Data yang termasuk dalam satu cluster}$$

$$\sum n = \text{Jumlah anggota dalam satu cluster}$$

- e. Menentukan hasil *cluster* telah konvergen

Dalam K-Means clustering, data akan mengalami beberapa iterasi, cluster konvergen yaitu data yang dihasilkan iterasi baru telah sama dengan iterasi sebelumnya. Selain menggunakan kesamaan iterasi terbaru dengan iterasi sebelumnya, penentuan cluster dapat menggunakan batas maksimal iterasi. Proses iterasi akan berhenti jika telah memenuhi batas maksimal iterasi. Batas maksimal iterasi dapat ditentukan sendiri oleh peneliti, misalnya peneliti hanya ingin perulangan iterasi dilakukan sebanyak 10 kali. Jika selama 10 kali iterasi tidak mendapatkan data yang sama, maka iterasi terakhir yang digunakan untuk melakukan pengelompokan.

## BAB 5

### PENUTUP

#### 5.1 Kesimpulan

Berdasarkan hasil penelitian K-Means *Clustering* untuk pendaftaran siswa baru melalui wilayah atau zona disimpulkan bahwa penerapan algoritma k-means clustering pada penentuan wilayah atau zona, menghasilkan sebuah informasi mengenai data siswa baru yang telah diterima dan lolos masuk ke SMAN 2 Bangkinang. Dari 912 data siswa yang diteliti terdapat 706 siswa yang merupakan anggota Cluster 0 atau dapat diartikan masuk zona 1 dan 206 siswa yang masuk ke dalam Cluster 1 yang dapat diartikan sebagai zona 2, dari wilayah atau zona yang masuk kategori penerimaan siswa baru pada sekolah tersebut.

#### 5.2 Saran

Adapun saran dari penulis untuk penelitian selanjutnya adalah sebagai berikut:

1. Untuk penelitian selanjutnya disarankan untuk menggunakan data yang mencakup daerah yang lebih luas, seperti data siswa di tingkat kabupaten, dan lain sebagainya untuk mendapatkan manfaat yang lebih luas lagi.
2. Diharapkan untuk penelitian selanjutnya, dapat menerapkan algoritma ini kedalam sebuah sistem informasi sehingga dapat digunakan oleh kalangan mahasiswa dan tenaga pengajar.



## DAFTAR PUSTAKA

- Aldi, Nurzahputra, Muslim Much Aziz, and Kurniawan Roni. 2016. "Online Fuzzy C-Means Clustering for Lecturer Performance Assesment Based on National and International Journal Publication." *International Conference on Mathematics, Science, And Education* 2016(Icmse).
- Aldifah, Rizma. 2018. "STRATEGI PENGEMBANGAN AGRIBISNIS JAGUNG DI INDONESIA (National Maize Agribusiness Development Strategy)." *Analisis Kebhijaksanaan Pertanian* Vol. 15(No. 1): 43–66.  
<http://dx.doi.org/10.21082/akp.v15n1.2017.43-66>.
- Arifin, Toni, Adhirajasa Reswara Sanjaya, and Adhirajasa Reswara Sanjaya. 2020. "PENERAPAN ALGORITMA K-MEANS UNTUK PENGELOMPOKAN SISWA LOLOS SNMPTN." 2(2): 184–90.
- Ashari, Sarah, Salsabilah Khansa, Calvin Habib Maulana Surudin, and Ika Nurlaili Isnainiyah. 2018. "KLUSTERING JUMLAH PENDUDUK KOTA BANDUNG BERDASARKAN JENIS KELAMIN PER KECAMATAN PADA TAHUN 2012 DENGAN METODE K-MEANS Sarah Ashari , Salsabilah Khansa , Calvin Habib Maulana Surudin , Ika Nurlaili Isnainiyah." *Seminar Nasional Informatika, Sistem Informatika Dan Keamanan Siber (SEINASI-KESI)*: 22–28.
- Badriuddin, Sabri, Ashfa Achmad, and Noer Fadhly. 2022. "Model Klasterisasi Wilayah Kota Berdasarkan Indeks Infrastruktur Untuk Perencanaan Struktur Ruang Berkelanjutan." *Jurnal Arsip Rekayasa Sipil dan Perencanaan* 5(3): 222–35.
- Desa Efani, Siti Aliyah, Silvia Lestari, and Wulan Dari. 2022. "Implementasi Algoritma K-Means Untuk Penerimaan Siswa Baru Di SMANPAS Berdasarkan Nilai Rapot Dan Hasil Tes." *It (Informatic Technique) Journal* 10(1): 01.
- Dianawati, Erni, Putri Previa Yanti, and Yulia Suryandari. 2019. "Klustering Jumlah Penumpang Pada Halte Bus Rapid Transit Kota Tangerang." *Jurnal Sistem Cerdas* 2(3): 163–72.
- Fammaldo, Eric, and Lukman Hakim. 2019. "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Tingkat Kesejahteraan Keluarga Untuk Program Kartu Indonesia Pintar." *Jurnal Ilmiah Teknologi Infomasi Terapan* 5(1): 23–31.





Firdaus, Diky. 2017. "Penggunaan Data Mining Dalam Kegiatan Sistem Pembelajaran Berbantuan Komputer." 6: 91–97.

Hafizah, Tugiono, and Widiarti Rista Maya. 2019. "Penerapan Data Mining Dalam Memprediksi Jumlah Penumpang Pada CV . Surya Mandiri Sukses Dengan Menggunakan Metode Regresi Linier." *Jurnal Teknologi Informasi dan Sistem Komputer TGD* 2(1): 54–61.

Haris Ihsanil Huda, and Maya Utami Dewi. 2018. "Pendaftaran Siswa Baru Online Di Smk Negeri 4 Kendal." *Elkom : Jurnal Elektronika dan Komputer* 11(2): 56–66.

Harliana, Harliana. 2020. "Penerapan Algoritma K-Means Clustering Untuk Mengelompokkan Kenakalan Siswa." *Journal of Information System Management (JOISM)* 2(1): 19–22.

Hidayat, Rahmat. 2022. "Clustering Menggunakan Algoritma K-Means Untuk Mengelompokkan Wilayah Rawan Kejahatan Di Wilayah Kabupaten Solok." *Jurnal Ilmiah Mahasiswa Manajemen, Bisnis dan Akuntansi (JIMMBA)* 4(5): 646–54.

Manurung, Oktaviani, and Penda Sudarto Hasugian. 2019. "Analisa Algoritma Apriori Untuk Peminjaman Buku Pada Perpustakaan SMA 1 Silima Pungga-Pungga Parongil." *Remik* 4(1): 154–60.

Nur Khormarudin, Agus. 2016. "Teknik Data Mining: Algoritma K-Means Clustering." *Jurnal Ilmu Komputer*: 1–12.  
<https://ilmukomputer.org/category/datamining/>.

Pradana, Musthofa Galih, Azriel Christian Nurcahyo, and Pujo Hari Saputro. 2021. "Penerapan Metode K-Means Klustering Untuk Menentukan Kepuasan Pelanggan." *Creative Information Technology Journal* 7(1): 42.

Santosa, Atharizky Ade, R Yunendah Nur Fu'adah, and Syamsul Rizal. 2023. "Deteksi Penyakit Pada Tanaman Padi Menggunakan Pengolahan Citra Digital Dengan Metode Convolutional Neural Network." *Journal of Electrical and System Control Engineering* 6(2): 98–108.

Sariingsah, Husin, Wanayumini Wanayumini, and Rika Rosnelly. 2021. "Penentuan Kelas Menggunakan Algoritma K Medoids Untuk Clustering Siswa Tunagrahita." *Jurnal Media Informatika Budidarma* 5(1): 83.

- Hak Cipta Diinundungi Undang-Undang
1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
    - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
    - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
  2. Diarangi mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Sinaga, Sulaiman, Rahmat W Sembiring, and S Sumarno. 2022. "Penerapan Algoritma Naive Bayes Untuk Klasifikasi Prediksi Penerimaan Siswa Baru."

*Journal of Machine ...* 1(1): 55–64.

<https://journal.fkpt.org/index.php/malda/article/view/162%0Ahttps://journal.fkpt.org/index.php/malda/article/download/162/115>.

Talataua, Mozart W., Zeth A. Leleury, and A. W. Taluta. 2017. "Analisis Cluster Dengan Menggunakan Metode K-Means Untuk Pengelompokan Kabupaten/Kota Di Provinsi Maluku Berdasarkan Indikator Indeks Pembangunan Manusia Tahun 2014." *BAREKENG: Jurnal Ilmu Matematika dan Terapan* 11(2): 119–28.

Tambubolon, Kennedi et al. 2013. "Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan." : 93–106.

Yandri, Rudy, Fernando Siboro, and Dina Yantika Wusni Simanjuntak. 2019. "Model Sistem Pendaftaran Sekolah Untuk Monitoring Pemetaan Wilayah Sebagai Penunjang Keputusan Penerimaan Siswa Baru." *SENSI Journal* 5(1): 73–85.

Zulham, Zulham, and Buyung Solihin Hasugian. 2022. "Pengelompokan Siswa Dalam Menentukan Penerima Beasiswa Berdasarkan Prestasi Akademik Dengan Algoritma K-Means." *Warta Dharmawangsa* 16(3): 231–41.

## DAFTAR RIWAYAT HIDUP

### Biodata Diri

Nama : Ikhsanul Wardana  
 Jenis Kelamin : Laki-laki  
 Tempat Tanggal Lahir : Simpang Kubu, 05 Juli 1995  
 Agama : Islam  
 Kewarganegaraan : Indonesia  
 Alamat : Dusun II Metro Lestari Desa Simpang Kubu Kec. Kampar Kab. Kampar, Provinsi Riau  
 Asal : Dusun II Metro Lestari Desa Simpang Kubu Kec. Kampar Kab. Kampar, Provinsi Riau  
 Email : 11750115190

### Riwayat Pendidikan

2000 – 2006 SD MUHAMMADIYAH 048 BANGKINANG  
 2006 – 2009 MTs Muallimin Muhammadiyah Bangkinang  
 2010 – 2013 MA Muallimin Muhammadiyah Bangkinang  
 2017 – 2024 S1 Teknik Informartika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang  
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:  
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.  
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.  
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.