

TESIS

**ANALISIS KOMBINASI ALGORITMA K-MEANS CLUSTERING DAN
TOPSIS UNTUK MENENTUKAN PENDEKATAN STRATEGI
MARKETING BERDASARKAN BACKGROUND TARGET AUDIENS**



Disusun oleh:

Nama : Nurus Sarlifatul Ngaeni
NIM : 22.55.2353
Konsentrasi : Business Intelligence

**PROGRAM STUDI S2 INFORMATIKA
PROGRAM PASCASARJANA UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

TESIS

**ANALISIS KOMBINASI ALGORITMA K-MEANS CLUSTERING DAN
TOPSIS UNTUK MENENTUKAN PENDEKATAN STRATEGI
MARKETING BERDASARKAN BACKGROUND TARGET AUDIENS**

**ANALYSIS OF COMBINATION OF K-MEANS CLUSTERING
ALGORITHM AND TOPSIS TO DETERMINE A MARKETING
STRATEGY APPROACH BASED ON TARGET AUDIENCE
BACKGROUND**

Diajukan untuk memenuhi salah satu syarat memperoleh derajat Magister



Disusun oleh:

Nama : Nurus Sarifatul Ngaeni
NIM : 22.55.2353
Konsentrasi : Business Intelligence

**PROGRAM STUDI S2 INFORMATIKA
PROGRAM PASCASARJANA UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

HALAMAN PENGESAHAN

**ANALISIS KOMBINASI ALGORITMA K-MEANS CLUSTERING DAN TOPSIS
UNTUK MENENTUKAN PENDEKATAN STRATEGI MARKETING
BERDASARKAN BACKGROUND TARGET AUDIENS**

**ANALYSIS OF COMBINATION OF K-MEANS CLUSTERING ALGORITHM AND
TOPSIS TO DETERMINE A MARKETING STRATEGY APPROACH BASED ON
TARGET AUDIENCE BACKGROUND**

Dipersiapkan dan Disusun oleh

Nurus Sarifatul Ngaeni

22.55.2353

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tesur
Program Studi S2 Informatika
Program Pascasarjana Universitas AMIKOM Yogyakarta
pada hari Senin, 3 Juni 2024

Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer

Yogyakarta, 3 Juni 2024

Rektor

Prof. Dr. M. Suyanto, M.M.
NIK. 190302001

HALAMAN PERSETUJUAN

**ANALISIS KOMBINASI ALGORITMA K-MEANS CLUSTERING DAN TOPSIS
UNTUK MENENTUKAN PENDEKATAN STRATEGI MARKETING
BERDASARKAN BACKGROUND TARGET AUDIENS**

**ANALYSIS OF COMBINATION OF K-MEANS CLUSTERING ALGORITHM AND
TOPSIS TO DETERMINE A MARKETING STRATEGY APPROACH BASED ON
TARGET AUDIENCE BACKGROUND**

Dipersiapkan dan Disusun oleh

Nurus Sarifatul Ngaeni

22.55.2353

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tesis
Program Studi S2 Informatika
Program Pascasarjana Universitas AMIKOM Yogyakarta
pada hari Senin, 3 Juni 2024

Pembimbing Utama

Prof. Dr. Kusriani, M.Kom

NIK. 190302106

Anggota Tim Penguji

Dr. Kumara Ari Yuana, ST., MT

NIK. 190302575

Pembimbing Pendamping

Kusnawi, S.Kom., M.Eng

NIK. 190302112

M. Hanafi, S.Kom., M.Eng., Ph D

NIK. 190302024

Prof. Dr. Kusriani, M.Kom

NIK. 190302106

Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer

Yogyakarta, 3 Juni 2024

Direktur Program Pascasarjana

Prof. Dr. Kusriani, M.Kom

NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN TESIS

Yang bertandatangan di bawah ini,

Nama mahasiswa : Nurus Sarifatul Ngaeni
NIM : 22.55.2353
Konsentrasi : Business Intelligence

Menyatakan bahwa Tesis dengan judul berikut:
**ANALISIS KOMBINASI ALGORITMA K-MEANS CLUSTERING DAN TOPSIS
UNTUK MENENTUKAN PENDEKATAN STRATEGI MARKETING
BERDASARKAN BACKGROUND TARGET AUDIENS**

Dosen Pembimbing Utama : Prof. Dr. Kusriani, M.Kom
Dosen Pembimbing Pendamping : Kusnawi, S.Kom., M.Eng

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi

Yogyakarta, 3 Juni 2024
Yang Menyatakan,



Nurus Sarifatul Ngaeni

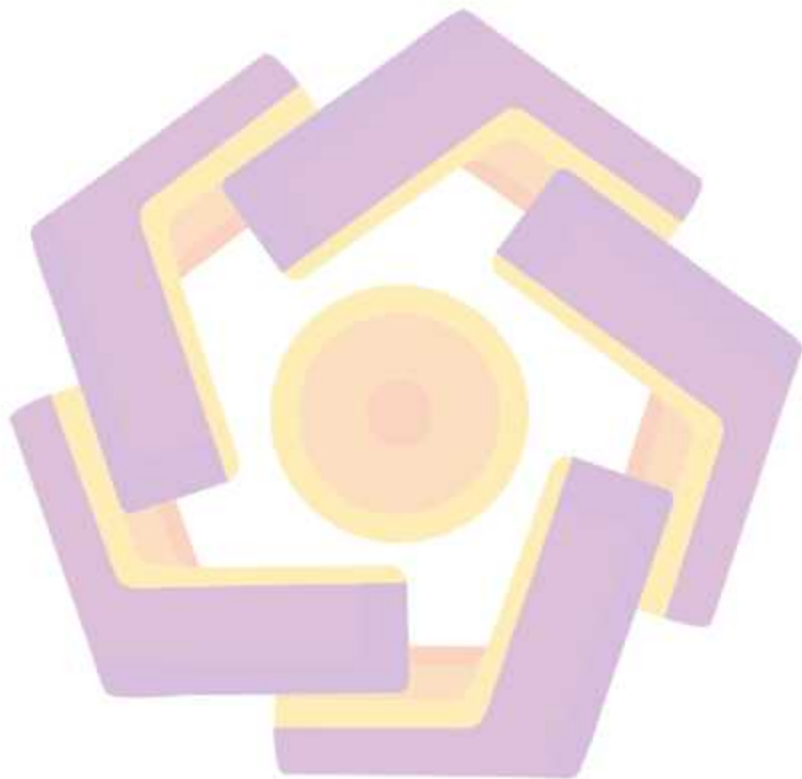
HALAMAN PERSEMBAHAN

Alhamdulillah, rasa syukur saya panjatkan kepada Allah SWT untuk semua hal baik yang dititipkan sampai hari ini. Karya penelitian ini saya persembahkan untuk semua pihak yang mendukung selesainya tesis ini. Terimakasih kepada suami, orangtua serta Ananda yang menjadi pengingat sekaligus menyemangatkan untuk menuntaskan amanah ini sebaik mungkin. Semoga Allah berikan balasan kebaikan dan pahala selalu teriring untuk mereka. Aamiin.



HALAMAN MOTTO

Selalu yakin dan percaya Alloh selalu memberikan yang terbaik untuk setiap hamba-Nya. Bismillah, ambil setiap kesempatan dan tunaikan sebaik mungkin dan semampu kita.



KATA PENGANTAR

Alhamdulillah, rasa syukur senantiasa dipanjatkan kepada Allah SWT yang telah memberikan segala nikmat. Tidak hanya bermodal kerja keras, dengan pertolongan Allah, doa dan dukungan orang – orang terkasih alhamdulillah penelitian dengan judul “Analisis Kombinasi Algoritma K-Means Clustering dan TOPSIS Untuk Menentukan Pendekatan Strategi Marketing Berdasarkan Background Target Audiens” dapat terselesaikan tepat waktu. Atas terselesainya penelitian ini, dengan sepenuh hati kami ingin menyampaikan rasa terimakasih kepada:

1. Suami, Orangtua, Anak beserta keluarga besar Bani Chomsin dan Bani Karso
2. Ibu Prof. Dr. Kusriani, M.Kom selaku Direktur Program Magister Informatika Universitas Amikom Yogyakarta sekaligus sebagai pembimbing pertama.
3. Bapak Kusnawi, S.Kom, M.Eng selaku pembimbing kedua.
4. Seluruh Dosen Pasca Sarjana beserta Admisi Magister Informatika Universitas Amikom Yogyakarta serta Rekan seperjuangan mahasiswa yang memberikan motivasi dalam penyelesaian penelitian.

Dalam proses penulisan tesis ini tentu tidak luput dari hambatan dan tantangan, namun berkat arahan dan bimbingan dari berbagai pihak pada akhirnya tesis ini bisa diselesaikan dengan baik. Sebagai penulis tentu saja kami menyadari penelitian ini belum bisa dikatakan sempurna.

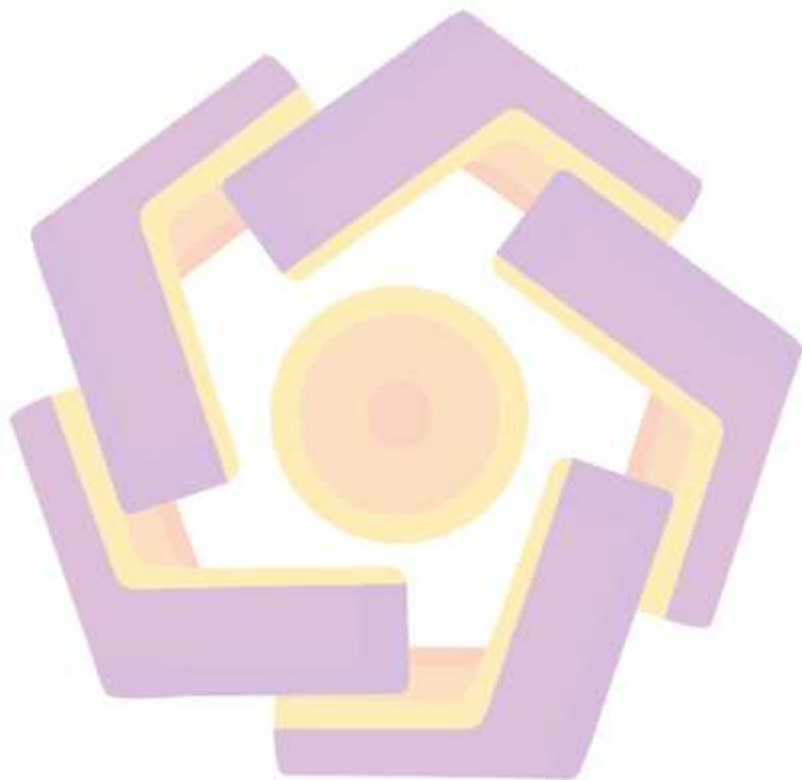
Yogyakarta, 3 Juli 2024
Penulis

DAFTAR ISI

HALAMAN JUDUL.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PERNYATAAN KEASLIAN TESIS.....	v
HALAMAN PERSEMBAHAN.....	vi
HALAMAN MOTTO.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
INTISARI.....	xiv
<i>ABSTRACT</i>	xv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah.....	5
1.3. Tujuan Penelitian.....	6
1.4. Manfaat Penelitian.....	6
BAB II TINJAUAN PUSTAKA.....	7
2.1. Tinjauan Pustaka.....	7
2.2. Keaslian Penelitian.....	12
2.3. Landasan Teori.....	20
BAB III METODE PENELITIAN.....	28
3.1. Jenis, Sifat, dan Pendekatan Penelitian.....	28
3.2. Metode Pengumpulan Data.....	28
3.3. Metode Analisis Data.....	28

3.4. Alur Penelitian.....	30
BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....	32
4.1 Pengumpulan Data	32
4.2 Inisiasi Data.....	33
4.3 Pemanggilan Dataset	37
4.4 Pembersihan.....	38
4.5 Perhitungan K-Means Clustering.....	39
4.5.1 Menghitung Centroid Setiap Cluster.....	39
4.5.2 Menghitung Nilai Rata – rata Cluster.....	40
4.5.3 Menghitung Nilai Silhouette Score.....	42
4.6 Penghitungan TOPSIS (Technique for Order Preference by Similarity to Ideal Solution)	43
4.6.1 Menentukan Keputusan Matriks Ternormalisasi.....	43
4.6.2 Menentukan Keputusan Matriks Ternormalisasi Terbobot.....	44
4.6.3 Menentukan Matriks Solusi Ideal Positif dan Negatif.....	45
4.6.4 Menghitung Jarak Alternatif.....	46
4.6.5 Menghitung Preferensi Dari Solusi Ideal.....	47
4.6.6 Melakukan Perangkingan.....	48
4.7 Decision Tree.....	49
4.7.1 Menggabungkan Hasil Cluster dan TOPSIS.....	49
4.7.2 Import Data Decision Tree.....	50
4.7.3 Menggabungkan Variabel.....	52
4.7.4 Rekomendasi Promosi Untuk Masing – Masing Cluster.....	53
BAB V PENUTUP.....	56

5.1. Kesimpulan	56
5.2. Saran	57
LAMPIRAN.....	63

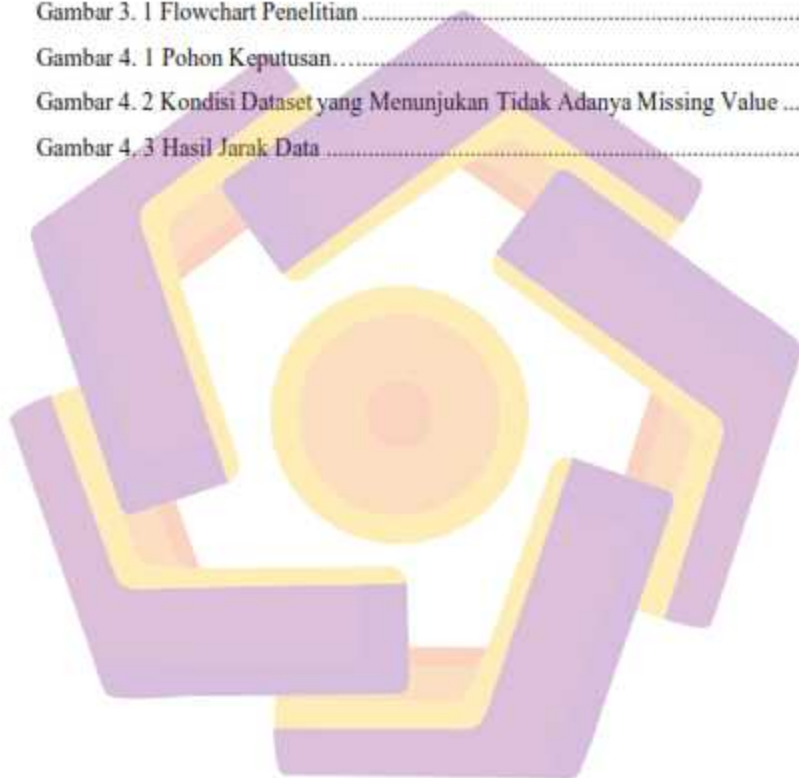


DAFTAR TABEL

Tabel 2.1. Matriks Literatur Review dan Posisi Penelitian	12
Tabel 4. 1 Data Pendaftar Mahasiswa Baru	33
Tabel 4. 2 Kriteria Data	34
Tabel 4. 3 Strategi Promosi.....	36
Tabel 4. 4 Hasil Pemanggilan Dataset pada Google Collaboratory.....	38
Tabel 4. 5 Hasil Perhitungan Iterasi 1	40
Tabel 4. 6 Nilai Rata – rata Iterasi 1	40
Tabel 4. 7 Perhitungan Iterasi ke 7	41
Tabel 4. 8 Nilai Rata – rata Iterasi 7	41
Tabel 4. 9 Matriks yang Telah Dinormalisasi	44
Tabel 4. 10 Matriks Ternormalisasi Terbobot.....	45
Tabel 4. 11 Hasil Perhitungan Jarak Alternatif.....	47
Tabel 4. 12 Hasil Perhitungan Preferensi Solusi Ideal	48
Tabel 4. 13 Hasil Perangkingan	49
Tabel 4. 14 Hasil Penggabungan Cluster dan TOPSIS	50
Tabel 4. 15 Hasil Import Data Decision Tree.....	51
Tabel 4. 16 Penggabungan Hasil Variabel.....	52
Tabel 4. 17 Hasil Rekomendasi Untuk Setiap Klaster	53

DAFTAR GAMBAR

Gambar 2. 1 Proses Data Mining Untuk Knowledge Discovery.....	22
Gambar 2. 2 Algoritma K-Means Clustering	23
Gambar 3. 1 Flowchart Penelitian	30
Gambar 4. 1 Pohon Keputusan.....	37
Gambar 4. 2 Kondisi Dataset yang Menunjukkan Tidak Adanya Missing Value	38
Gambar 4. 3 Hasil Jarak Data	39



INTISARI

Promosi merupakan agenda tahunan STIMIK Tunas Bangsa Banjarnegara. Tujuan dari kegiatan promosi ini yakni untuk menarik mahasiswa baru yang lebih banyak lagi setiap tahunnya. Disisi lain promosi kampus menemui kendala untuk memetakan data pendaftar tahun – tahun sebelumnya sehingga pertimbangan untuk kebijakan promosi baru berdasarkan data sekolah asal alumni atau mahasiswa. Dengan menggunakan algoritma K-Means Clustering, data pendaftar bisa dikelompokkan sesuai dengan latar belakang yang direpresentasikan melalui atribut asal sekolah, pekerjaan orangtua dan daerah asal. Kemudian data diproses menggunakan DSS dengan metode TOPSIS untuk mendapatkan referensi prioritas jenis pemasaran untuk masing – masing cluster. Hasil penghitungan nilai silhouette coefficient untuk lima kluster diperoleh skor 0,426. Sedangkan pada proses perankingan menggunakan metode TOPSIS diperoleh rank pertama terdapat pada cluster 0 dengan skor 0,994110. Tahapan lebih lanjut menggunakan metode Decision Tree untuk mendapatkan output berupa rekomendasi jenis promosi untuk masing – masing cluster. Sebagai contoh, cluster 0 direkomendasikan menggunakan jenis promosi dengan kode P1, P2, P3, P8 dan P9.

Kata Kunci: Sistem Pendukung Keputusan; Pemasaran; Perguruan Tinggi; K-Means Clustering; TOPSIS

ABSTRACT

The promotion is an annual agenda for STIMIK Tunas Bangsa Banjarnegara. The aim of this promotional activity is to attract more new students every year. On the other hand, campus promotion encounters obstacles in mapping applicant data from previous years so that considerations for new promotion policies are based on data from the school of origin of alumni or students. By using the K-Means Clustering algorithm, applicant data can be grouped according to the background represented through the school origin attribute, parents' occupation and place of origin. Then the data is processed using DSS with the TOPSIS method to obtain priority references for marketing types for each cluster. The results of calculating the silhouette coefficient value for the five clusters obtained a score of 0.426. Meanwhile, in the ranking process using the TOPSIS method, the first rank was found in cluster 0 with a score of 0.994110. Further stages use the Decision Tree method to obtain output in the form of recommendations for promotion types for each cluster. For example, cluster 0 is recommended to use promotion types with codes P1, P2, P3, P8 and P9.

Keyword: Decision Support System; Marketing; College; K-Means Clustering; TOPSIS



BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Bagian Strategi pemasaran dikatakan cocok mempunyai indikator berupa tercapainya target (Assylla & Nugraha, 2022). Promosi sendiri merupakan bagian dari strategi yang dirancang secara matang oleh perusahaan atau lembaga untuk mencapai tujuan yang disepakati Bersama (Winasis et al., 2022). Melakukan inovasi dalam pemasaran salah satunya bisa dipersiapkan serta dilakukan secara tersistem dengan bantuan teknologi. Tidak terkecuali dalam bidang manajemen pemasaran itu sendiri yang bisa membantu proses analisa berdasarkan data yang dimiliki (Bofinger et al., 2022). Untuk mengetahui kelompok – kelompok data, analisa bisa menggunakan K-Means Clustering (Sakur et al., 2022). Penggunaan teknologi untuk mencari solusi atau strategi sesuai dengan data yang dimiliki berdasarkan nilai yang paling dekat dengan tujuan serta nilai yang paling jauh bisa menggunakan metode TOPSIS (Khomsatun et al., 2020).

Penerimaan Mahasiswa Baru (PMB) di kampus STIMIK Tunas Bangsa Banjarnegara merupakan agenda tahunan yang dilaksanakan secara rutin. Tim marketing telah melakukan berbagai jenis pemasaran, termasuk melalui sosial media, pemasangan spanduk, sosialisasi, kerjasama lembaga, beasiswa, dan rekomendasi alumni. Promosi dilakukan hanya berdasarkan program rutin yang sudah dijalankan sebelumnya. Mengingat terbatas sumber daya yang ada, untuk meningkatkan efektivitas dan efisiensi pemasaran, tim memanfaatkan data mining dengan menggunakan metode K-Mean Clustering dan metode TOPSIS. Pendekatan

ini memungkinkan identifikasi kelompok mahasiswa berdasarkan asal sekolah, pekerjaan orangtua, dan daerah asal. Hasil analisis selanjutnya diproses menggunakan metode TOPSIS untuk mendapatkan urutan kluster terbaik. Untuk mendapatkan keluaran berupa rekomendasi jenis promosi masing – masing cluster, peneliti menggunakan decision tree berdasarkan rumusan yang disetujui oleh Lembaga. Kombinasi metode ini memberikan STIMIK Tunas Bangsa rekomendasi strategi yang sesuai untuk masing – masing kluster, dengan harapan dapat menjadi dasar untuk penentuan kebijakan, termasuk penawaran program dan layanan yang sesuai dengan preferensi target audien. Pendekatan ini juga memungkinkan pengelolaan sumber daya dan anggaran pemasaran dapat difokuskan pada kelompok-kelompok yang memiliki potensi konversi tinggi.

Data mining sudah banyak digunakan sebagai teknik pengolahan data, dimana data berasal dari pendaftar mahasiswa baru tahun – tahun sebelumnya. Pola atau informasi yang didapatkan dari hasil analisa inilah yang nantinya bisa dijadikan sebagai acuan dalam proses menentukan kebijakan pemasaran. Pengolahan data mahasiswa baru untuk keperluan pemasaran sudah banyak diamati. Khoironi melakukan penelitian terkait pemanfaatan data mining untuk menyeleksi kandidat penerima beasiswa menggunakan algoritma K-Mean Clustering yang dilanjutkan dengan proses perankingan menggunakan metode AHP(Khoironi, 2020).

Pengamatan lain dilakukan oleh Oki Oktaviarna Tensao dengan menerapkan data mining dengan algoritma K-Means Clustering untuk membuat kelompok – kelompok mahasiswa baru di STMIK Primakara. Tujuan dari

penelitian ini yakni untuk menentukan kebijakan seputar strategi marketing yang paling tepat untuk setiap kelompok mahasiswa. Namun penelitian ini baru berupa pengelompokkan, artinya belum bisa memberikan usulan jenis marketing apa yang paling sesuai untuk masing – masing kelompok siswa(Oki Oktaviarna Tensao et al., 2022).

Pengamatan lainnya dilakukan oleh Hidayat yang menerapkan algoritma K-Means Clustering untuk mengklasifikasi data jamaah umrah di AET Travel Indonesia. Variabel penelitiannya mencakup usia, jenis kelamin dan daerah asal. Hasil pengelompokkan yang didapatkan dalam penelitian ini dijadikan sebagai pengetahuan baru yang mempermudah manager menyusun strategi pemasaran di AET Travel Indonesia. Penelitian mengenai strategi pemasaran di Pendidikan Tinggi dilakukan oleh SUTRISNO & Andriyani yang menggunakan metode pengambilan keputusan berbasis atribut ganda (MADM) dengan metode Simple Additive Weighting (SAW). Tujuan dari penelitian ini yakni berupa rekomendasi asal jurusan apa saja yang paling prospek dijadikan sebagai target promosi(Hidayat, 2022).

Algoritma K-Means Clustering kemudian dikombinasikan dengan metode SAW untuk mendapatkan rekomendasi lokasi serta jenis promosi yang paling tepat. Penelitian yang dilakukan oleh Kasri dan Jati ini bisa dikembangkan lagi dengan melibatkan beberapa faktor yang mempengaruhi proses pendaftaran mahasiswa baru untuk mendapatkan wawasan yang lebih mendalam terkait optimasi strategi pemasaran. Penelitian selanjutnya dilakukan oleh Andry yang mengembangkan sistem pendukung keputusan untuk menentukan lokasi promosi menggunakan

metode Analytical Hierarchy Process (AHP). Tujuan penelitian ini yakni untuk mendapatkan rekomendasi lokasi promosi yang tepat sasaran sehingga manajemen bisa melakukan pemasaran secara lebih efektif dan efisien (Kasri & Jati, 2020).

Penelitian mengenai strategi promosi juga dilakukan oleh Nanda Ayu Rahmalinda dimana data mahasiswa baru dianalisa menggunakan metode K-Means Clustering untuk menyelaraskan pemilihan strategi promosi (Rahmalinda & Jananto, 2022). Penelitian lain yang mengkombinasikan dua metode dilakukan oleh Suarnatha yang menggunakan asosiasi algoritma apriori dan perangkingan item set terlaris menggunakan TOPSIS. Kedua metode yang digunakan dirasa membuat hasil analisa menjadi lebih bermanfaat (Suarnatha et al., 2022).

Beberapa kajian yang diuraikan di atas berkaitan dengan tantangan yang dihadapi membuat peneliti menggunakan beberapa metode untuk menentukan jenis pemasaran yang direkomendasikan. Dalam penelitian yang dilakukan oleh Amaliya menyatakan bahwa kombinasi kedua metode antara TOPSIS dan K-Means bisa dibuktikan memberikan hasil yang lebih optimal dalam hal pengelompokan metode pembelajaran berdasarkan bidang keilmuan (Amaliya Hari Nafisah, 2023). Dengan memanfaatkan metode K-Mean Clustering, maka peneliti bisa melakukan identifikasi kelompok – kelompok mahasiswa berdasarkan latar belakang target audiens. Yakni dilihat dari asal sekolah, pekerjaan orangtua, dan daerah asalnya. Selanjutnya hasil analisa akan diolah lagi menggunakan metode TOPSIS untuk membantu proses pemeringkatan strategi marketing sesuai dengan latar belakang target audien. Untuk mendapatkan *output* berupa rekomendasi jenis pemasaran untuk masing – masing cluster, digunakan juga penggabungan hasil clustering dan

perangkingan berdasarkan pohon keputusan yang dibuat. Kombinasi kedua metode ini memungkinkan STIMIK Tunas Bangsa Banjarnegara mendapatkan rekomendasi strategi yang tepat untuk masing – masing cluster. Latar belakang audiens ini bisa dijadikan sebagai dasar untuk penentuan kebijakan berupa penawaran program serta layanan yang paling sesuai dengan preferensi target audiens. Integrasi dari kedua algoritma dapat dijadikan sebagai bagian dari referensi pengembangan integrasi dua algoritma di ranah pemasaran atau promosi, khususnya di lingkup perguruan tinggi. Hasil dari penelitian yang dilakukan nantinya juga dapat dijadikan sebagai model referensi yang bisa diadaptasi oleh instansi lain sesuai dengan karakteristik variabel yang dimiliki.

2.1 Rumusan Masalah

Berdasarkan uraian yang dipaparkan pada bagian latar belakang, maka peneliti membuat rumusan masalah ke dalam beberapa poin berikut:

- a. Bagaimana mengkombinasikan algoritma K-Means Clustering dengan metode TOPSIS untuk menghasilkan alternatif jenis pemasaran yang disarankan?
- b. Sejauh mana tingkat akurasi penelitian dengan kombinasi algoritma K-Means dan TOPSIS dalam menentukan pendekatan strategi marketing di STIMIK Tunas Bangsa Banjarnegara?
- c. Faktor apa saja yang mempengaruhi tingkat akurasi dalam penerapan kombinasi Algoritma K-Means Clustering dengan Metode TOPSIS?

3.1 Tujuan Penelitian

Berikut ini tujuan penelitian yang diharapkan tercapai:

- a. Menerapkan kombinasi algoritma machine learning dalam konteks pemasaran kampus yang dapat menunjang keberhasilan promosi.
- b. Memberikan wawasan dan informasi mengenai kelompok – kelompok pendaftar serta jenis promosi yang direkomendasikan untuk setiap cluster.
- c. Membantu manajemen menentukan kebijakan marketing yang lebih efektif dan efisien.

4.1 Manfaat Penelitian

Berikut ini beberapa manfaat penelitian yang diharapkan oleh peneliti:

- a. Integrasi algoritma K-Means Clustering dan metode TOPSIS bisa menjadi bagian dari pengembangan integrasi dua metode berbeda yang efektif dalam menganalisa data pemasaran.
- b. Kombinasi kedua metode ini bisa menjadi referensi wawasan riset terkait penelitian strategi pemasaran di tingkat perguruan tinggi menggunakan kombinasi algoritma K-Means Clustering dan metode TOPSIS.
- c. Hasil dari penelitian ini bisa digunakan sebagai model dan dapat diadaptasi oleh perguruan tinggi lain dalam hal pengambilan keputusan terkait pemasaran sesuai dengan kebutuhan serta karakteristik institusi.

BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Pustaka

Beberapa penelitian sebelumnya yang sudah dilakukan terkait pemanfaatan algoritma machine learning dalam ranah perguruan tinggi dan pemasaran. Sebagai pertimbangan yakni penelitian yang mengkombinasikan algoritma clustering dan AHP sebagai tindak lanjut penentu keputusan yang diteliti oleh (Khoironi, 2020) dengan judul “Kombinasi Metode K-Means dan Analytic Hierarchy Process Dalam Menentukan Penerima Beasiswa”. Kriteria dalam menentukan penerima beasiswa yakni nilai akademik dan kondisi ekonomi berupa penghasilan orangtua. Namun yang menjadi kendala yakni Nilai antara satu siswa dan siswa lainnya mempunyai tingkat kemiripan yang cukup dekat. Dengan menggunakan K-Mean Clustering, dibuatlah kelompok - kelompok siswa yang mempunyai kemiripan nilai. Tahapan selanjutnya yakni menggunakan metode AHP untuk proses perankingan. Evaluasi metode KA menggunakan metode Borda dengan bobot 0,511628 sedangkan untuk metode AHP mendapatkan bobot 0,511628. Kesimpulan dari penelitian ini yakni metode KA mempunyai tingkat relevansi yang lebih baik.

Oki Oktaviarna Tensao et al., (2022) dalam penelitiannya yang berjudul “Implementasi Data Mining dengan Algoritma K-Means Clustering untuk Analisis Calon Mahasiswa Baru di STIMIK Primakara : Strategi Promosi yang Efektif dan Efisien”. Dalam penelitian ini menggunakan atribut berupa wilayah asal sekolah, program studi dan gelombang pendaftaran. Ketiga atribut

tersebut dinilai kurang lengkap untuk diproses menjadi cluster karena masih ada kemungkinan faktor yang berpengaruh pada keputusan pendaftaran, misalnya saja faktor gambaran kondisi ekonomi pendaftar.

Hidayat, (2022) dalam penelitiannya yang berjudul “Implementasi Metode K-Means Clustering dalam Klasifikasi Data Jamaah Umrah: Studi Kasus AET Travel Indonesia” menggunakan variabel berupa usia, jenis kelamin, dan daerah asal pengguna jasa. Hasil penelitian terbantu beberapa cluster rentang usia pendaftar yang paling potensial, yakni mulai usia 41 – 70 tahun, 21 – 40 tahun dan 1 – 20 tahun. Pihak manajemen menentukan sendiri strategi yang sekiranya paling relevan untuk masing – masing cluster. Peneliti mengharapkan adanya penambahan dimensi analisis berupa kombinasi antara algoritma K-Means Clustering dengan metode decision support system untuk mengoptimalkan strategi marketing berdasarkan latar belakang pengguna jasa.

SUTRISNO & Andriyani, (2021) dalam penelitiannya yang berjudul “Penerapan Metode Pengambilan Keputusan Berbasis Atribut Ganda (MADM) dengan Metode Simple Additive Weighting (SAW) untuk Menentukan Target Promosi Berdasarkan Asal Jurusan Calon Mahasiswa di Sekolah Menengah”. Tujuan dari penelitian ini yakni untuk memberikan rekomendasi target promosi yang paling tepat untuk menaikkan jumlah pendaftar berdasarkan asal jurusan. Berdasarkan hasil penelitian, disimpulkan bahwa alternatif ke 4 merupakan target yang paling potensial.

Sedangkan predikat target yang paling tidak potensial tertuju pada alternatif ke 6.

Kasri & Jati, (2020) dalam penelitiannya yang berjudul “Kombinasi K Means Clustering dan Metode Simple Additive Weighting (SAW) untuk Menentukan Lokasi dan Strategi Promosi Penerimaan Mahasiswa”. Tujuan dari penelitian ini yakni menerapkan algoritma K-Means Clustering yang dikombinasikan dengan metode Simple Additive Weighting (SAW) untuk mengoptimalkan strategi promosi berdasarkan lokasi prioritas serta jenis marketing yang disarankan untuk masing – masing cluster. Kriteria data yang digunakan yakni meliputi data kecamatan, pekerjaan orangtua, dan program studi yang dipilih. Untuk penelitian berikutnya bisa diperluas lagi dengan menggunakan variabel yang mempengaruhi pendaftaran mahasiswa baru. Integrasi data dari sumber yang berbeda akan menghasilkan wawasan baru dalam optimalisasi strategi promosi perguruan tinggi.

Andry et al., (2021) dalam penelitiannya yang berjudul “Sistem Pendukung Keputusan Penentuan Lokasi Promosi Penerimaan Siswa Baru Di MTS S.Hubbul Wathon Modal Bangsa” yang mengembangkan Decision Support System menggunakan metode Analytical Hierarchy Process (AHP) untuk menghasilkan informasi yang berupa rekomendasi lokasi promosi. Peneliti menggunakan lima kriteria yang terdiri dari jarak ke sekolah, waktu tempuh, jumlah sekolah lain, Kawasan jalan serta jumlah siswa dalam kurun waktu tiga tahun sebelumnya. Diharapkan pengembangan pada penelitian selanjutnya untuk melakukan evaluasi secara menyeluruh terkait

efektivitas dan efisiensi kinerja sistem. Pengembangan juga bisa dilakukan berdasarkan umpan balik dari pengguna serta perkembangan kebutuhan sekolah.

Wijayanto et al., (2023) dalam penelitiannya yang berjudul “Penetapan Lokasi Pemasangan Spanduk Penerimaan Siswa Baru di SMK Bina Informatika Menggunakan Metode Analytical Hierarchy Process (AHP)” yang digunakan sebagai pembantu keputusan alternatif lokasi pemasangan spanduk. Dalam penelitian ini dihasilkan lima alternatif lokasi teratas untuk pemasangan spanduk pendaftaran siswa baru.

Rusliyawati et al., (2020) dalam penelitiannya yang berjudul “Penerapan Metode Simple Additive Weighting (SAW) dalam Penentuan Model Social Customer Relationship Management (SCRM) sebagai Strategi Bisnis di Universitas Rusliyawati” yang menganalisis dan menghitung karakteristik dari masing – masing social media. Tujuan dari penelitian ini yakni untuk menggunakan metode SAW dalam alternatif pengambilan keputusan terkait model SCRM untuk optimasi strategi pemasaran.

Rismayadi et al., (2021) dalam penelitiannya yang berjudul “Penggunaan Algoritma K-Means Clustering untuk Menentukan Strategi Pemasaran di CV. Integreet Konstruksi” dengan tujuan untuk mengelompokkan data penjualan yang diwakili menggunakan variabel data penjualan. Output dari berupa cluster – cluster dengan kategori penjualan tertinggi, rata – rata dan paling rendah. Hasil klusteriasi ini dijadikan sebagai pertimbangan dalam menentukan strategi pemasaran sehingga bisa meningkatkan omset penjualan.

2.2 Keaslian Penelitian

Tabel 2.1. Matriks Literatur Review dan Posisi Penelitian

Analisis Kombinasi Algoritma K-Means Clustering dan TOPSIS Untuk Menentukan Pendekatan Strategi Marketing Berdasarkan Background Target Audiens

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
1	Kombinasi Metode K-Means dan Analytic Hierarchy Process Dalam Menentukan Penerima Beasiswa	Khouroni, MTI Amikom Yogyakarta, 2020	Menentukan rekomendasi penerima beasiswa yang tepat. Menentukan penerima beasiswa menggunakan Metode K-Means dan Analytical Hierarchy Process	Metode clustering dengan algoritma K-Means dapat menyeleksi kandidat yang dibutuhkan dan dapat membantu mengoptimalkan proses perangkaian di tahap AHP. Hasil evaluasi menggunakan metode borda menyatakan pada metode KA mendapatkan bobot 0,511628 dan pada metode AHP mendapatkan bobot 0,488372, maka dapat disimpulkan metode KA dinilai lebih relevan untuk diterapkan pada data tersebut	Untuk memaksimalkan hasil dari penelitian ini perlu diadakan penelitian lebih mendalam lagi dengan membandingkannya dengan algoritma clustering dan algoritma sistem pendukung keputusan lainnya.	Pada penelitian yang dirujuk menggunakan metode K-Means yang dikombinasikan dengan AHP sedangkan penelitian yang penulis lakukan menggunakan kombinasi K-Means dan metode TOPSIS.
2	Implementasi Data Mining dengan Algoritma K-Means Clustering untuk Analisis Calon Mahasiswa Baru di STMIK Primakara:	Oki Oktaviarna Tensao, I Nyoman Yudi Anggara Wijaya, Ketut Queena Fredlina. INFORMASI (Jurnal Informatika dan	Tujuan penelitian adalah untuk mencerpakan data mining dengan algoritma K-Means Clustering untuk menganalisis dan	Dari penelitian ini, dapat disimpulkan bahwa penerapan data mining dengan algoritma K-Means Clustering telah berhasil dalam menganalisis dan mengelompokkan data calon mahasiswa baru di STMIK	Pertama, penggunaan algoritma K-Means Clustering dapat Menghasilkan hasil yang kurang optimal jika jumlah cluster yang diinginkan tidak diketahui	Penelitian yang akan dilakukan lebih fokus pada penggunaan algoritma K-Means Clustering lalu ditindaklanjuti dengan kombinasi dengan

Tabel 2.1. Matriks Literatur Review dan Posisi Penelitian

Analisis Kombinasi Algoritma K-Means Clustering dan TOPSIS Untuk Menentukan Pendekatan Strategi Marketing Berdasarkan Background Target Audiens

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	Strategi Promosi yang Efektif dan Efisien.	Sistem Informasi) Volume 14 No.1/Mei/2022	mengelompokkan data calon mahasiswa baru pada STMIK Primakara, guna menentukan strategi promosi yang lebih efektif dan efisien.	Primakara. Hasil clustering menunjukkan pola pendaftaran mahasiswa baru yang dapat digunakan untuk menentukan strategi promosi yang lebih efektif dan efisien. Wilayah Denpasar Timur dan Gianyar menjadi fokus utama promosi, dengan strategi pemasaran yang disesuaikan dengan hasil cluster. Metode CRISP-DM dan perangkat lunak RapidMiner digunakan dalam penelitian ini, dan hasilnya dapat digunakan sebagai dasar untuk pengambilan keputusan dalam strategi promosi di masa depan.	sebelumnya. Kedua, penelitian ini hanya menggunakan tiga atribut (wilayah sekolah asal, program studi, dan gelombang pendaftaran untuk melakukan clustering, sehingga terdapat kemungkinan bahwa faktor-faktor lain yang juga berpengaruh pada keputusan pendaftaran mahasiswa tidak dipertimbangkan. Ketiga, penelitian ini hanya fokus pada strategi promosi berdasarkan hasil clustering, namun tidak mempertimbangkan faktor-faktor eksternal lainnya yang juga dapat memengaruhi keputusan pendaftaran mahasiswa, seperti tren ekonomi, persaingan dengan perguruan tinggi lain, dan faktor-faktor sosial. Oleh karena itu, penelitian ini perlu diperluas dengan mempertimbangkan faktor-faktor tambahan	metode TOPSIS sedangkan pada penelitian yang dirujuk berfokus pada proses clusterisasi pendafiar.

Tabel 2.1. Matriks Literatur Review dan Posisi Penelitian

Analisis Kombinasi Algoritma K-Means Clustering dan TOPSIS Untuk Menentukan Pendekatan Strategi Marketing Berdasarkan Background Target Audiens

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
					yang dapat memengaruhi keputusan pendaftaran mahasiswa.	
3	Implementasi Metode K-Means Clustering dalam Klasifikasi Data Jamaah Umrah: Studi Kasus AET Travel Indonesia	Taufik Hidayat Sistem Informasi dan Teknologi, Volume 4, Nomor 1, tahun 2022	Tujuan penelitian ini adalah untuk mengelompokkan data jemaah umrah dari AET Travel Indonesia ke dalam subset berdasarkan variabel seperti usia, jenis kelamin, dan daerah asal menggunakan metode K-Means Clustering.	Setelah dilakukan perhitungan data yang telah dilakukan melalui software RapidMiner diperoleh anggota kelompok sangat disukai dari rentang usia mulai 41 sampai 70 tahun, kelompok disukai dengan rentang usia mulai 21 sampai 40 tahun dan kelompok kurang disukai mulai usia 1 sampai 20 tahun. Dengan di dapatkan hasil pengelompokan ini maka di dapatkan pengetahuan baru yang dapat memudahkan manajer dalam strategi pemasaran pada AET Travel Indonesia.	Penelitian ini dapat diperkaya dengan menyertakan analisis terhadap kendala atau tantangan yang dihadapi selama implementasi metode K-Means Clustering. Selain itu, penelitian ini dapat lebih bermanfaat dengan memberikan wawasan mengenai arah penelitian masa depan atau bidang-bidang yang dapat dieksplorasi lebih lanjut terkait klasifikasi data jamaah umrah menggunakan metode clustering.	Penelitian yang akan dilakukan yakni menambahkan dimensi analisis dengan menggabungkan algoritma K-Means Clustering dan TOPSIS untuk menentukan pendekatan strategi pemasaran berdasarkan latar belakang audiens sedangkan pada penelitian yang dirujuk menggunakan menggunakan K-Means untuk proses clusterisasi calon jamaah.
4	Penerapan Metode Pengambilan Keputusan Berbasis Atribut Ganda (MADM) dengan Metode Simple Additive Weighting (SAW) untuk	B.T. Sutrisno SP dan Widyastuti Andriyani [4] Media Publikasi: Jurnal SIMETRIS [4] Tahun: 2020	Tujuan penelitian ini adalah untuk memberikan alternatif keputusan kepada para pengambil kebijakan dalam hal penerimaan	Dari penelitian ini, dapat disimpulkan bahwa penerapan Metode Simple Additive Weighting (SAW) dalam kerangka Multiple Attribute Decision Making (MADM) dapat memberikan rekomendasi target promosi yang tepat berdasarkan	Saran untuk kedepannya, penelitian ini masih dalam bentuk perhitungan, sehingga dapat dikembangkan lebih lanjut dalam bentuk aplikasi yang dapat dengan mudah diakses	Penelitian yang dirujuk berfokus untuk mendapatkan alternatif keputusan berupa target promosi berdasarkan jurusan calon mahasiswa sedangkan penelitian yang akan

Tabel 2.1. Matriks Literatur Review dan Posisi Penelitian

Analisis Kombinasi Algoritma K-Means Clustering dan TOPSIS Untuk Menentukan Pendekatan Strategi Marketing Berdasarkan Background Target Audiens

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	Menentukan Target Promosi Berdasarkan Asal Jurusan Calon Mahasiswa di Sekolah Menengah		mahasiswa untuk menentukan target promosi yang tepat guna meningkatkan jumlah mahasiswa baru. Metode yang digunakan adalah Multiple Attribute Decision Making (MADM) dengan metode Simple Additive Weighting (SAW) untuk menentukan asal jurusan calon mahasiswa di Sekolah Menengah yang menjadi target promosi.	asal jurusan calon mahasiswa di Sekolah Menengah. Hasil penelitian menunjukkan bahwa alternatif ke-4 merupakan asal jurusan di Sekolah yang paling prospek menjadi target promosi, sementara alternatif ke-6 merupakan yang paling tidak prospek. Dengan demikian, model ini dapat membantu pengambil kebijakan dalam menentukan persyaratan calon mahasiswa yang akan mendaftar. Saran untuk penelitian selanjutnya adalah mengembangkan model ini dalam bentuk aplikasi yang dapat diakses dengan mudah oleh para pengambil kebijakan.	para pengambil kebijakan secara cepat, sehingga dapat memutuskan secara tepat terkait persyaratan PMB Fakultas Kesehatan Universitas Jenderal Achmad Yani Yogyakarta.	dilakukan yakni menggunakan variabel yang merepresentasikan latar belakang pendaftar.
5	Kombinasi K-Means Clustering dan Metode Simple Additive Weighting (SAW) untuk Menentukan Lokasi dan Strategi Promosi Penerimaan Mahasiswa Baru.	Muhammad Ali Kasri, Handaru Jati, KHAZANAH INFORMATIKA, 2020	Tujuan penelitian ini adalah untuk menggunakan kombinasi metode K-Means Clustering dan Simple Additive Weighting (SAW) untuk menentukan lokasi dan strategi	Kombinasi K-Means Clustering dan metode Simple Additive Weighting (SAW) efektif digunakan untuk menentukan lokasi dan strategi promosi dalam penerimaan mahasiswa baru. Metode K-Means digunakan untuk pengelompokan data dan menentukan lokasi promosi, sementara metode SAW	Penelitian ini juga dapat diperluas dengan mempertimbangkan faktor-faktor lain yang dapat memengaruhi penerimaan mahasiswa baru, seperti faktor ekonomi, sosial, atau geografis. Integrasi data dari sumber yang berbeda	Peneliti yang akan dilakukan menambahkan variabel asal sekolah untuk memperkuat hasil rekomendasi jenis marketing yang sesuai sedangkan pada penelitian yang dirujuk hanya menggunakan

Tabel 2.1. Matriks Literatur Review dan Posisi Penelitian

Analisis Kombinasi Algoritma K-Means Clustering dan TOPSIS Untuk Menentukan Pendekatan Strategi Marketing Berdasarkan Background Target Audiens

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			promosi yang optimal untuk penerimaan mahasiswa baru. Metode ini diharapkan dapat membantu dalam mengoptimalkan waktu, tenaga, dan biaya yang diperlukan dalam proses pemasaran universitas.	membantu dalam pemeringkatan alternatif berdasarkan kriteria tertentu. Hasil penelitian menunjukkan bahwa kombinasi kedua metode ini dapat menghasilkan rekomendasi lokasi promosi yang sesuai dengan kriteria tertentu, namun juga memiliki kelemahan tertentu. Oleh karena itu, kombinasi ini dapat menjadi landasan untuk pengembangan strategi pemasaran yang lebih efektif dalam mengoptimalkan waktu, tenaga, dan biaya.	juga dapat memberikan wawasan yang lebih mendalam dalam menentukan lokasi dan strategi promosi yang optimal.	variabel pekerjaan orang tua untuk merepresentasikan kondisi ekonomi.
6	Sistem Pendukung Keputusan Penentuan Lokasi Promosi Penerimaan Siswa Baru Di MTS S.Hubbul Wathon Modal Bangsa	Andry, Yani Maulita, Suci Ramadan, Seminar Nasional informatika (SENATIKA), 2021	Tujuan dari penelitian ini adalah untuk mengembangkan Sistem Pendukung Keputusan (SPK) menggunakan metode Analytical Hierarchy Process (AHP) guna menentukan lokasi promosi penerimaan siswa	Sistem Pendukung Keputusan (SPK) dengan metode Analytical Hierarchy Process (AHP) dapat membantu sekolah dalam memilih lokasi promosi yang efektif dan efisien untuk penerimaan siswa baru. Sistem ini dirancang dengan antarmuka yang memudahkan pengguna dalam menginput data kriteria, subkriteria, dan alternatif. Menu Analisa pada sistem pendukung keputusan menggunakan metode	Disarankan untuk terus mengembangkan dan menerapkan sistem pendukung keputusan (SPK) menggunakan metode Analytical Hierarchy Process (AHP) untuk menentukan lokasi promosi yang strategis dalam penerimaan siswa baru di sekolah menengah. Selain itu, disarankan untuk	Penelitian yang sudah dilakukan fokus pada menggunakan metode untuk mendapatkan rekomendasi lokasi (sekolah) yang strategis sedangkan penelitian yang akan dilakukan lebih menggali informasi terkait latar belakang pendaftar.

Tabel 2.1. Matriks Literatur Review dan Posisi Penelitian

Analisis Kombinasi Algoritma K-Means Clustering dan TOPSIS Untuk Menentukan Pendekatan Strategi Marketing Berdasarkan Background Target Audiens

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			baru di sebuah sekolah menengah. Sistem ini diharapkan dapat membantu sekolah dalam memilih lokasi promosi yang efektif dan efisien serta membantu dalam manajemen dan penerimaan siswa.	AHP menampilkan data analisa nilai prioritas kriteria dan subkriteria, serta hasil perankingan.	melakukan evaluasi menyeluruh terhadap efektivitas dan efisiensi sistem dalam membantu manajemen penerimaan siswa di sekolah. Selanjutnya, disarankan untuk mempertimbangkan peningkatan atau modifikasi sistem berdasarkan umpan balik pengguna dan perkembangan kebutuhan sekolah.	
7	Penetapan Lokasi Pemasangan Spanduk Penerimaan Siswa Baru di SMK Bina Informatika Menggunakan Metode Analytical Hierarchy Process (AHP)	Santosa Wijayanto, idealis: Indonesia Journal Information System, 2023	Penelitian ini bertujuan untuk membantu pengambil keputusan sekolah dalam menetapkan pemilihan lokasi pemasangan spanduk penerimaan siswa baru secara efektif dan efisien.	Hasil perhitungan bobot prioritas kriteria dan alternatif menunjukkan bahwa kriteria C5, C6, dan C7 memiliki nilai prioritas yang sama dan tertinggi. Selain itu, penelitian ini memberikan rekomendasi 5 lokasi teratas untuk pemasangan spanduk, dengan Jl. Cendrawasih Raya menjadi pilihan teratas. Dengan demikian, penelitian ini dapat membantu sekolah dalam menetapkan pemilihan lokasi pemasangan spanduk secara efektif dan efisien.	Penelitian ini juga dapat menjadi dasar untuk pengembangan metode penentuan lokasi pemasangan spanduk yang lebih efektif dan efisien di masa depan. Diharapkan penelitian ini dapat memberikan kontribusi positif bagi pengambil keputusan dalam menentukan lokasi pemasangan spanduk promosi penerimaan siswa baru di sekolah.	Penelitian yang akan dilakukan menggunakan metode TOPSIS untuk mendapatkan output berupa rekomendasi jenis pemasaran sedangkan penelitian yang dirujuk fokus pada rekomendasi lokasi pemasangan spanduk menggunakan metode AHP.

Tabel 2.1. Matriks Literatur Review dan Posisi Penelitian

Analisis Kombinasi Algoritma K-Means Clustering dan TOPSIS Untuk Menentukan Pendekatan Strategi Marketing Berdasarkan Background Target Audiens

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
8	Penerapan Metode Simple Additive Weighting (SAW) dalam Penentuan Model Social Customer Relationship Management (SCRM) sebagai Strategi Bisnis di Universitas Rusliyawati	Rusliyawati, Proses 6thICITB Indonesia, 2020	Tujuan penelitian ini adalah untuk mengembangkan aplikasi berbasis website yang dapat menganalisis dan menghitung karakteristik masing-masing model Social Customer Relationship Management (SCRM) sebagai strategi bisnis di Universitas Rusliyawati melalui pendekatan Simple Additive Weighting (SAW). Selain itu, penelitian ini juga bertujuan untuk menggunakan metode SAW dalam pengambilan keputusan untuk memilih model SCRM yang dapat mendukung strategi bisnis.	Hasil implementasi penilaian media sosial alternatif dengan ketentuan kriteria yang telah ditentukan dan populasi penilai mahasiswa program studi sistem informasi angkatan 2017, 2018 dan 2019 menunjukkan media sosial youtube dengan skor 0,888 dan media sosial yang direkomendasikan untuk digunakan sebagai strategi bisnis bagi institusi pendidikan tinggi.	Melakukan studi lanjutan untuk memperluas cakupan penelitian dengan mempertimbangkan lebih banyak model SCRM dan media sosial yang relevan. Melakukan analisis perbandingan antara metode SAW dengan metode pengambilan keputusan lainnya untuk memilih model SCRM.	Penelitian yang akan dilakukan menggunakan kombinasi metode K-Means Clustering dan TOPSIS dengan dataset mahasiswa baru untuk mendapatkan rekomendasi jenis pemasarandengan pada penelitian yang dirujuk bertujuan mendapatkan rekomendasi media sosial untuk optimasi pemasaran.

Tabel 2.1. Matriks Literatur Review dan Posisi Penelitian

Analisis Kombinasi Algoritma K-Means Clustering dan TOPSIS Untuk Menentukan Pendekatan Strategi Marketing Berdasarkan Background Target Audiens

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
9	Penggunaan Algoritma K-Means Clustering untuk Menentukan Strategi Pemasaran di CV. Integreet Konstruksi	Ali Akbar, dkk. JURNAL RESPONSIF, 2021	Tujuan penelitian ini adalah untuk menggunakan algoritma K-Means Clustering dalam menentukan strategi pemasaran yang tepat bagi CV. Integreet Konstruksi. Dengan menggunakan data penjualan dari bulan Agustus - November 2019, penelitian ini bertujuan untuk mengelompokkan data penjualan berdasarkan lokasi, jenis pekerjaan, dan sumbernya.	Berdasarkan hasil pengolahan data dapat diketahui bahwa cluster 1 adalah kelompok dengan rata-rata penjualan paling tinggi dan paling tepat untuk dijadikan sebagai dasar dalam penentuan strategi pemasaran, cluster 2 adalah kelompok dengan rata-rata penjualan paling rendah, sehingga kurang tepat untuk dijadikan dasar dalam penentuan strategi penjualan, dan cluster 3 adalah kelompok data yang berada di tengah-tengah dan masih bisa dipertimbangkan untuk dijadikan dasar penentuan strategi pemasaran.	Untuk penelitian selanjutnya, bisa menambahkan atribut nilai proyek agar perusahaan dapat mengetahui data penjualan yang mana yang berpotensi meningkatkan omset pendapatan yang paling banyak.	Penelitian yang akan dilakukan menggunakan kombinasi dua metode untuk mendapatkan output berupa alternatif strategi pemasaran sedangkan penelitian yang dirujuk berupa proses clusterisasi yang nantinya dijadikan sebagai panduan bagi manajemen untuk menyusun strategi pemasaran secara manual.

2.3 Landasan Teori

2.3.1 Data Mining

Perkembangan teknologi membuat pengumpulan dan penyimpanan data berkembang pesat. Hampir semua organisasi mempunyai sekumpulan data dalam jumlah besar bahkan gunung data yang belum diolah. Ekstraksi data hingga analisa perlu dilakukan menggunakan alat dan teknik yang bisa membuat sekumpulan data ini mempunyai nilai guna lebih banyak. Data Mining adalah campuran dari beragam metode analisis data menggunakan beragam algoritma bisa menjadi solusi untuk masalah ini.

Data Mining menjadi salah satu alternatif penyelesaian untuk menemukan informasi atau wawasan baru dalam suatu basis data. Sebuah organisasi data umumnya akan mempunyai banyak sekali data, disinilah peluang analisa data bisa dilakukan untuk menggali informasi yang bermanfaat. Teknik data mining yang umum digunakan yakni mulai dari Classification, Association dan Clustering (Kusnawi, 2007).

Data mining merupakan serangkaian langkah yang dilakukan untuk proses eksplorasi nilai tambah kumpulan data berupa pengetahuan atau wawasan baru yang sebelumnya tidak bisa digali secara manual. Penting untuk diperhatikan bahwa istilah “mining” ini sendiri merujuk pada proses mendapatkan nilai lebih yang didapatkan melalui sebagian besar materi dasar. Inilah sebabnya Data Mining menjadi akar dalam beragam disiplin ilmu yang mencakup Artificial Intelligence (kecerdasan buatan), Machine Learning (Pembelajaran Mesin), statistika hingga rumpun basis data. Data mining inilah yang mengambil peran sebagai media untuk

menggalai wawasan atau pengetahuan baru menggunakan metode tertentu (Bahar et al., 2016).

Data mining adalah sebuah proses ekstraksi pola yang didalamnya menggunakan sejumlah data dalam skala besar yang bersifat sangat kompleks, tersirat serta belum bisa diambil sebagai wawasan atau pengetahuan baru. Dengan adanya Data Mining, maka kumpulan data dalam jumlah yang besar ini bisa mempunyai nilai guna lebih bagus. Pengetahuan yang didapatkan nantinya bahkan bisa digunakan sebagai alat prediksi dengan tingkat kepastian yang cukup spesifik (Marsono et al., 2021).

Pemanfaatan rumpun ilmu Data Mining ini juga bisa dijadikan sebagai alat untuk memberikan alternatif atau rekomendasi keputusan. Misalnya saja dalam konteks business solution untuk mengatasi masalah atau tantangan menggunakan infrastruktur di bidang teknologi saat ini (Anisyia, 2020).

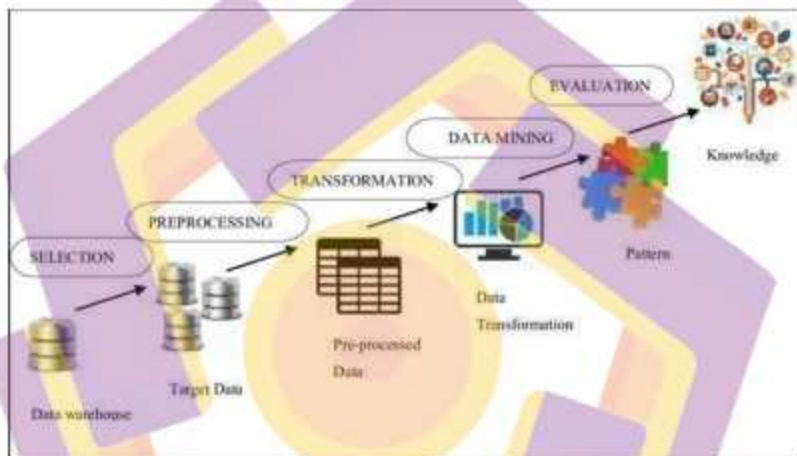
Berdasarkan pengertian di atas, dapat disimpulkan bahwa Data Mining ini adalah proses analisa dataset yang dilakukan secara sistematis dan komprehensif. Pemanfaatan Data Mining bisa digunakan untuk mengidentifikasi serta menganalisa pola dan tren tertentu. Selain itu, pengetahuan baru juga bisa didapatkan dan dijadikan sebagai alternatif pendukung keputusan atau optimasi di banyak sektor.

2.3.2 Proses Data Mining Untuk Knowledge Discovery

Pemanfaatan Data Mining untuk menganalisa sumber data bisa dilakukan dengan beragam metode. Beberapa jenis metode data mining yang bisa digunakan yaitu KDD, CRISP-DM, Clustering, Klasifikasi dan lainnya. Knowledge Discovery

in Database Process (KDD) merupakan salah satu jenis metode data mining. Fayyed et al. (1996) mendefinisikan KDD adalah proses analisa data untuk mengetahui informasi berharga maupun mengenali pola – pola tertentu. (Sakur et al., 2022)

Proses KDD ini meliputi seleksi data, preprocessing data, interpretasi serta evaluasi. Alur proses KDD ini ditunjukkan melalui gambar 1 berikut:



Gambar 2.1. Proses Data Mining Untuk Knowledge Discovery

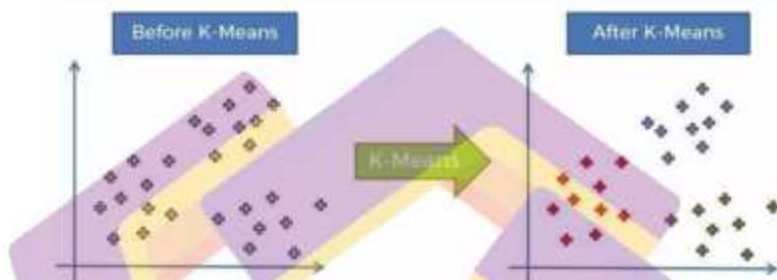
2.3.3 Algoritma K-Means Clustering

Algoritma K-Means Clustering ini digunakan untuk membuat cluster - cluster yang mempunyai kemiripan satu dengan lainnya. Pengelompokan data menggunakan metode K-Means populer digunakan untuk penambangan data di bidang analisis klaster. (Ahlawat et al., n.d.)

K-Means Clustering banyak digunakan dalam analisis data karena beberapa pertimbangan. Selain mudah dan populer, teknik pengelompokan relatif lebih

cepat bahkan untuk kumpulan data yang terbilang besar sekalipun. (Raval & Jani, 2016)

Cara kerja metode ini ditunjukkan melalui gambar 2 berikut :



Gambar 2.2. Algoritma K-Means Clustering

Sumber : <https://medium.com/@kurniasp/k-means-clustering-using-scikit-learn-in-python-51073b6f51e5>

K-Means Clustering adalah salah satu jenis metode dalam Data Mining yang digunakan untuk membuat clustering, dimana objek yang mempunyai kemiripan berdasarkan karakteristik tertentu yang sudah ditetapkan. Hasil dari pengelompokan data ini yakni dijadikan sebagai acuan dalam menentukan strategi tertentu, misalnya saja dalam hal strategi pemasaran (Bahar et al., 2016).

K-Means merupakan metode clustering non hirarki, dimana cara kerjanya yakni memisahkan data dengan kemiripan tertentu ke dalam cluster yang sama. Kemudian untuk data yang mempunyai karakteristik berbeda akan dimasukkan ke dalam cluster lain (Agusta Yudi, 2007). Untuk tahapan melakukan clustering menggunakan algoritma K-Mean Clustering yang dipaparkan oleh Agusta Yuda :

1. Peneliti menentukan jumlah cluster yang nantinya akan dibentuk
2. Pilih centroid awal secara random

3. Menghitung alokasi data berdasarkan jumlah cluster yang dibentuk. Kedekatan antara satu objek dengan objek lainnya akan ditentukan oleh jarak. Artinya mana objek yang paling dekat akan menentukan suatu data masuk ke dalam cluster tertentu.
4. Menghitung nilai pusat cluster (centroid) yang merupakan nilai rata – rata atas semua objek data dalam sebuah cluster.
5. Lakukan kembali alokasi masing – masing objek menggunakan centroid baru sampai nilai pusat cluster tidak berubah.
6. Apabila nilai pusat cluster masih berubah atau ada objek yang berpindah grup maka proses akan dikembalikan pada menghitung jarak antara objek ke centroid.

2.3.4 Sistem Pendukung Keputusan

Sistem pendukung keputusan merupakan sistem pemodelan yang dimanfaatkan untuk memberikan rekomendasi hasil atau keputusan baik dalam situasi yang bersifat semi terstruktur ataupun tidak terstruktur. Dalam proses pengambilan keputusan tentunya dibutuhkan kualitas informasi yang relevan dalam artian informasi atau data yang digunakan memang berkaitan serta mempunyai pengaruh dalam pengambilan keputusan di bidang tersebut (Andrianto et al., 2017). Salah satu contoh penggunaan sistem pendukung keputusan yakni untuk menilai kinerja karyawan, dimana proses penilaian kinerja karyawan sangat rentan mengalami subjektivitas. Dengan menggunakan sistem pendukung keputusan yang dirancang secara kompeten, diharapkan subjektivitas dalam penilaian kinerja ini bisa diminimalisir. (Kusrini, 2007)

Pendekatan Pengambilan Keputusan Multi-Kriteria (MCDM) melibatkan proses analisis dan evaluasi keputusan yang mempertimbangkan berbagai kriteria atau faktor dalam rangka mencapai solusi optimal. Pendekatan ini memungkinkan penanganan kompleksitas keputusan dengan mempertimbangkan berbagai aspek yang relevan, sehingga dapat memberikan pandangan yang komprehensif dalam mengambil keputusan. (Vafaei et al., 2021)

2.3.5 TOPSIS (Technique For Others To Reference by Similarity to Ideal Solution)

Sebagai metode yang digunakan dalam pengambilan keputusan, TOPSIS ini pertama kali diperkenalkan oleh Yoon dan Hwang pada tahun 1981 untuk menangani pengambilan keputusan bersifat multikriteria. Metode TOPSIS sendiri sudah banyak diaplikasikan untuk keperluan pengambilan keputusan di beragam sektor. Beberapa contohnya yakni diterapkan pada aplikasi yang menunjang keputusan terkait investasi, penilaian performa karyawan, evaluasi pemasaran bahkan terkait proses perancangan robot.

Pendekatan TOPSIS ini banyak digunakan logikanya mudah dimengerti serta proses komputasinya terbilang mudah. Untuk mencari pilihan alternatif terbaik bagi masing – masing kriteria bahkan dalam digambarkan dalam bentuk matematika sederhana. Prosedur perbandingan yang digunakan pada metode TOPSIS juga melibatkan adanya bobot kepentingan. (García-Cascales & Lamata, 2012)

TOPSIS merupakan salah satu metode yang digunakan untuk pengambilan keputusan bersifat multikriteria atau alternatif pilihan yang merupakan alternatif

dari jarak paling kecil dari solusi positif dan jarak paling besar dari solusi ideal negatif dilihat dari sudut pandang geometris menggunakan jarak Euclidean. Dan yang menjadi catatan bahwa alternatif dengan jarak paling kecil yang berasal dari solusi ideal positif tidak selalu harus memiliki jarak terbesar dari solusi ideal yang negative. Metode TOPSIS mempertimbangkan kondisi dimana dengan jarak terhadap solusi ideal positif seta jarak terhadap solusi ideal negative secara bersamaan. Keluaran berupa solusi dalam metode TOPSIS didapatkan dengan cara menentukan kedekatan relative suatu alternatif pada solusi ideal positif. Proses perankingan alternatif akan dilakukan berdasarkan prioritas nilai kedekatan suatu alternatif pada solusi ideal positif. Nantinya daftar alternatif yang sudah diranking inilah yang akan dijadikan sebagai referensi keputusan dalam kebijakan tertentu. (Nata et al., 2017)

Metode TOPSIS populer digunakan untuk pengambilan keputusan yang dilakukan secara praktis serta minim subjektivitas. Penggunaan metode TOPSIS menggunakan konsep yang terbilang sederhana serta mudah dipahami. Selain itu proses komputasinya efisien serta metode ini mempunyai kemampuan dalam mengukur kinerja relative dari setiap alternatif keputusan. (Servanda, 2018). Berikut ini tahapan metode TOPSIS melalui Langkah – Langkah berikut.

1. Membuat matrik keputusan yang sudah melewati proses normalisasi menggunakan persamaan.
2. Membuat matrik keputusan ternormalisasi dengan nilai bobot menggunakan persamaan

3. Membuat matrik dengan solusi ideal positif dan matrik solusi ideal negatif dengan menggunakan persamaan
4. Menentukan nilai preferensi untuk setiap pilihan alternatif menggunakan persamaan.
5. Melakukan perankingan pada setiap alternatif pilihan yang mana nantinya nilai V_i yang lebih besar berarti alternatif A_i lebih diprioritaskan.

2.3.6 Decision Tree

Decision tree merupakan salah satu metode pengambilan keputusan yang menggunakan susunan opsi menjadi bentuk bercabang. Seperti namanya, decision tree tampak seperti bentuk pohon yang mempunyai banyak cabang opsi keputusan. Bisa dikatakan, dengan menggunakan metode decision tree maka peneliti bisa mendapatkan beberapa opsi dari pilihan keputusan. Untuk menggunakan decision tree, dimulai dari membuat simpul lalu selanjutnya cabang opsi yang terdiri dari beberapa opsi keputusan (Jannah et al., 2023).

Metode Decision Tree merupakan suatu struktur kerja yang dapat digunakan untuk proses pengambilan keputusan atau kebijakan. Untuk mendapatkan sebuah keputusan dibutuhkan juga variabel penelitian yang berfungsi sebagai bahan pertimbangan untuk pengambilan keputusan dan kebijakan tertentu. Dalam lingkup data science, decision tree ini bisa dimanfaatkan untuk mendapatkan keputusan yang efektif (Herlambang et al., 2023)

BAB III

METODE PENELITIAN

3.1 Jenis, Sifat, dan Pendekatan Penelitian

Pendekatan penelitian yang dilakukan merupakan penelitian kuantitatif dengan cara studi experimental dimana peneliti akan melakukan experiment dan analisa terhadap dataset yang digunakan. Sedangkan sifat penelitian yakni experimental dimana peneliti akan mengalisis penerapan kombinasi kedua metode untuk mendapatkan rekomendasi lokasi serta jenis pemasaran yang sesuai.

3.2 Metode Pengumpulan Data

Dataset penelitian merupakan data pendaftar atau mahasiswa baru yang didapatkan secara studi dokumen dan wawancara melalui bagian pusat administrasi kampus STIMIK Tunas Bangsa Banjarnegara.

3.3 Metode Analisis Data

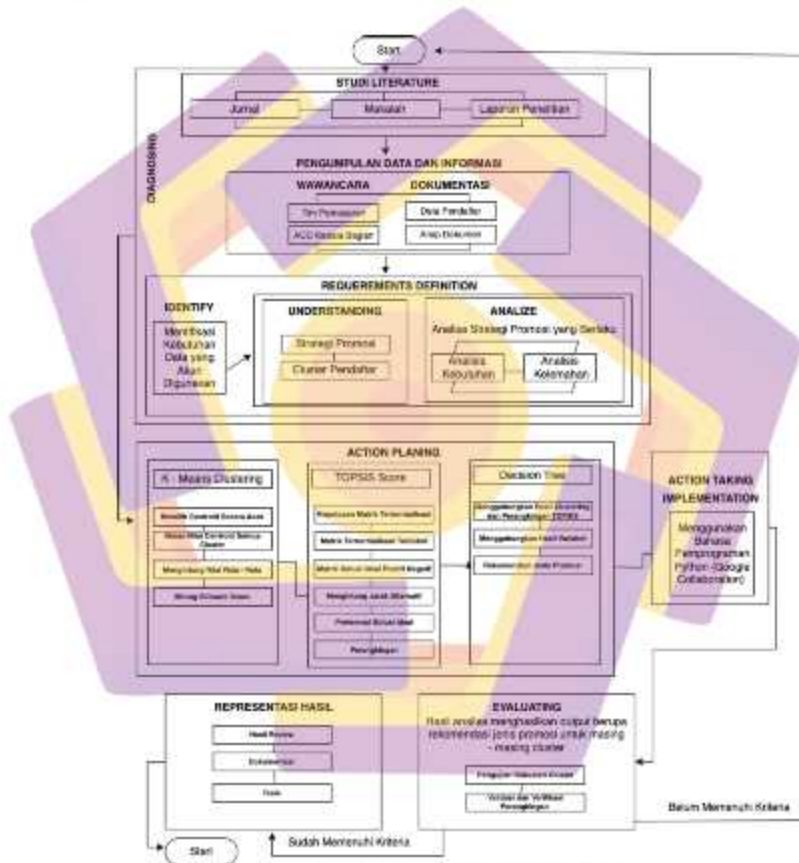
Peneliti menggunakan Bahasa pemrograman python yang diakses melalui google collab. Dataset yang digunakan yakni data mahasiswa baru mulai dari tahun 2012 hingga 2023. Penggunaan data pada penelitian ini yakni meliputi variabel asal sekolah sebagai representasi dari latar belakang audiens, pekerjaan orangtua sebagai gambaran kondisi ekonomi keluarga. Sedangkan daerah asal menjadi varibel yang dijadikan sebagai tolak ukur jarak antara sekolah dan tempat tinggal lalu program studi sebagai gambaran pilihan pendaftar. Skenario yang dilakukan

yakni peneliti melakukan inisiasi pada setiap variabel data. Nilai inisiasi data untuk setiap subkriteria ditentukan berdasarkan ketentuan dari bagian pemasaran. Misalnya untuk variabel lokasi, peneliti akan menggunakan kecamatan sebagai jarak kilometer dari kampus. Semakin dekat jaraknya maka nilai inisiasinya lebih tinggi.

Setelah cluster terbentuk, peneliti akan melakukan proses perangkingan menggunakan metode TOPSIS. Setiap kriteria akan diberikan bobot sesuai dengan tingkat kepentingannya. Selanjutnya membuat matrik keputusan pada semua kluster yang sudah dinormalisasi. Menentukan matriks keputusan berdasarkan solusi ideal positif dan keputusan solusi ideal negative. Mulai menghitung matriks setiap kluster dengan cara mengalikan matriks keputusan dan bobot kriterianya. Jarak Euclidean masing – masing kluster perlu dihitung dengan solusi ideal positif berbobot serta solusi ideal negative berbobot. Selanjutnya skor proksimitas relative dihitung menggunakan rumus TOPSIS. Pemeringkatan bisa dilakukan dengan cara melihat hasil terbaik ke yang terburuk berdasarkan perolehan skor proksimitas relative.

3.4 Alur Penelitian

Alur penelitian yang akan dilakukan tergambar melalui flow chart pada gambar 3.1 berikut:



Gambar 3.1. Flowchart Penelitian

Tahap awal penelitian, dikumpulkan dataset lalu masuk pada tahap preprocessing dimana data akan dibersihkan dari nilai yang tidak valid atau ganda. Selain itu, peneliti juga akan melakukan konversi format data agar dapat diolah

menggunakan bahasa pemrograman python serta melakukan normalisasi jika diperlukan. Peneliti selanjutnya menerapkan algoritma K-Mean Clustering dengan menentukan jumlah cluster, setelah itu memilih centroid awal secara acak. Selanjutnya yakni menghitung jarak data objek ke centroid dan menghitung nilai rata – rata dari masing – masing cluster. Jika objek berpindah group atau cluster atau terjadi adanya perubahan nilai centroid maka proses harus dikembalikan pada tahap menghitung jarak data ke centroid. Dan tahapan selanjutnya yakni peneliti memastikan apakah objek yang terdapat pada cluster sudah bisa dikatakan baik dengan menggunakan metode silhouette koefisien. Penggunaan siluet koefisien ini bertujuan untuk memastikan kualitas dan kekuatas cluster yang terbentuk.

Setelah peneliti mendapatkan centroid akhir pada perhitungan K-Means maka tahapan selanjutnya yakni melakukan perangkingan menggunakan metode TOPSIS. Tujuan dari proses ini yakni untuk memberikan bobot pada setiap atribut yang terdapat pada hasil pengelompokkan. Langkah pertama yakni melakukan normalisasi bobot agar total bobot untuk masing – masing atribut. Pemberian bobot pada masing – masing atribut dengan mencerminkan tingkat kepentingannya dalam menentukan strategi pemasaran. Setelah itu, tahapan selanjutnya yakni menghitung skor akhir dengan masing – masing cluster dengan mengalikan bobot dan nilai atribut serta menjumlahkan hasilnya. Untuk mendapatkan hasil rekomendasi jenis promosi, peneliti menggunakan decision tree yang di dalamnya menggabungkan hasil data K-Means dan rangking TOPSIS.

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

Pada bab ini akan dijelaskan mengenai analisa dan implementasi dari rancangan tahapan alur penelitian. Terdapat beberapa tahapan proses mulai dari persiapan dataset, penghitungan menggunakan algoritma K-Means, Perangkingan menggunakan TOPSIS dan evaluasi.

4.1 Pengumpulan Data

Dataset yang digunakan dalam penelitian ini adalah data pendaftar mahasiswa baru yang didapat dalam format excel. Agar dapat diolah ke dalam bahasa pemrograman python dilakukan rekapitulasi kembali terhadap data dan disimpan dalam format csv. Dataset yang diperoleh menyajikan beberapa data seperti nama mahasiswa, alamat, sekolah asal, program studi, pekerjaan orangtua, waktu daftar dan lainnya.

Dalam penelitian ini hanya mengambil data pendaftar meliputi daerah asal, sekolah asal, program studi dan pekerjaan orangtua. Penelitian ini tidak mengambil keseluruhan data pendaftar dari dataset sehingga data alamat dipisahkan secara manual untuk memperoleh data berupa kabupaten atau daerah asal pendaftar. Data asal sekolah juga dipisahkan secara manual untuk mendapatkan informasi sekolah asal secara umum. Data pendaftar yang akan digunakan terlihat dalam tabel 4.1 berikut.

Tabel 4.1. Data Pendaftar Mahasiswa Baru

No	Nama	Asal Sekolah	Pekerjaan Orangtua	Program Studi	Daerah Asal
1	MAHASISWA 1	SMK	Pedagang	S1 Sistem Informasi	Banjarnegara
2	MAHASISWA 2	SMK	Pedagang	S1 Sistem Informasi	Banjarnegara
3	MAHASISWA 3	SMA	Petani	S1 Sistem Informasi	Banjarnegara
4	MAHASISWA 4	SMA	Petani	S1 Sistem Inomasi	Banjarnegara
5	MAHASISWA 5	SMA	Petani	S1 Sistem Informasi	Banjarnegara
6	MAHASISWA 6	SMA	Petani	S1 Sistem Informasi	Banjarnegara
7	MAHASISWA 7	SMA	Petani	S1 Sistem Informasi	Banjarnegara
8	MAHASISWA 8	SMK	Petani	S1 Sistem Informasi	Banjarnegara
9	MAHASISWA 9	SMA	Petani	S1 Sistem Informasi	Banjarnegara
10	MAHASISWA 10	MA/MAN	PNS	S1 Sistem Informasi	Banjarnegara
11	MAHASISWA 11	SMA	Guru Swasta	S1 Sistem Informasi	Banjarnegara
12	MAHASISWA 12	SMA	Guru Swasta	S1 Teknik Informatika	Banjarnegara
13	MAHASISWA 13	SMK	Guru Swasta	S1 Teknik Informatika	Banjarnegara
14	MAHASISWA 14	SMK	Guru Swasta	S1 Teknik Informatika	Banjarnegara
1110	MAHASISWA 1110	SMK	Petani	S1 Teknik Informatika	Banjarnegara

Jumlah data yang digunakan dalam penelitian ini ada 1110 pendaftar yang melakukan pendaftaran selama periode 2012 – 2023 di kampus STIMIK Tunas Bangsa Banjarnegara.

4.2 Inisiasi Data

Untuk memudahkan penambangan data, maka kriteria data diinisiasi. Inisiasi data dan bobot untuk masing – masing kriteria tergambar pada tabel 4.2 berikut.

Tabel 4.2. Kriteria Data

KRITERIA	SUBKRITERIA	NILAI
<i>Asal Sekolah (V1)</i>	SMK	8
	SMA	7
	MA/MAN	6
	PKBM	5
<i>Pekerjaan Orangtua (V2)</i>	Pedagang	8
	PNS	7
	Guru Swasta	6
	Karyawan	5
	Petani	4
	Buruh	3
	Yang Lain	2
<i>Program Studi (V3)</i>	SI Sistem Informasi	1
	SI Informatika	2
<i>Daerah Asal (V4)</i>	Banjarnegara (1 - 10 km)	8
	Purbalingga (33 km)	7
	Wonosobo (42 km)	6
	Batang (58 km)	5
	Purwokerto (68 km)	4
	Pekalongan (95 km)	3
	Blora (300 km)	2

Dasar pemberian bobot atau nilai pada subkriteria penelitian yakni berdasarkan tingkat prioritas yang ditetapkan oleh bagian pemasaran Kampus STIMIK Tunas Bangsa Banjarnegara. Kriteria dalam menentukan nilai setiap subkriteria merupakan keputusan kepala bagian beserta tim marketing perguruan tinggi STIMIK Tunas Bangsa Banjarnegara.

Inisiasi subkriteria asal sekolah diambil dari seluruh data pendaftar terbanyak dan evaluasi tingkat pemahaman mahasiswa mengikuti proses pembelajaran. Untuk lulusan SMK mempunyai nilai yang paling tinggi karena berdasarkan evaluasi yang dilakukan, tingkat pemahaman mahasiswa baru yang berasal dari SMK lebih baik dibandingkan mahasiswa dari asal sekolah lain. Nilai paling rendah yakni 5 digunakan untuk bobot variabel asal sekolah PKBM. Hasil evaluasi mahasiswa baru

yang berasal dari PKBM membutuhkan waktu dan adaptasi lebih banyak karena sebelumnya belum terbiasa menggunakan perangkat komputer.

Inisiasi pekerjaan orangtua menggunakan nilai 2 hingga 8, dimana tingkat subkriteria paling tinggi diurut dari pekerjaan yang mempunyai kondisi ekonomi lebih baik. Standar ini dilihat berdasarkan tingkat kelancaran administrasi mahasiswa selama kuliah di Kampus STIMIK Tunas Bangsa Banjarnegara. Inisiasi variabel program studi berasal dari dua jurusan yang saat ini tersedia di Kampus STIMIK Tunas Bangsa Banjarnegara. Untuk nilai 2 diberikan pada program studi dengan jumlah mahasiswa lebih banyak. Saat ini program studi yang paling banyak jumlah mahasiswanya yakni S1 – Informatika sedangkan untuk program studi S1 – Sistem Informasi mempunyai nilai atau bobot 1.

Inisiasi variabel asal daerah menggunakan satuan jarak kilometer (km) dimana nilai tertinggi diberikan pada jarak yang paling dekat dengan jangkauan kampus. Pemilihan wilayah ini merupakan daerah atau kota asal para pendaftar mahasiswa baru. Dengan pertimbangan efisiensi anggaran, tim pemasaran kampus memberikan bobot tertinggi pada subkriteria dengan satuan jarak 1 -10 km yang mana merupakan jangkauan dalam kota Banjarnegara. Untuk bobot 7 diberikan pada subkriteria dengan satuan jarak 33 km yang merupakan kota Purbalingga. Untuk bobot 6 diberikan pada subkriteria dengan satuan jarak 42 km yang merupakan kota Wonosobo. Untuk bobot 5 diberikan pada subkriteria dengan satuan jarak 58 km yang merupakan kota Batang. Untuk bobot 4 diberikan pada subkriteria dengan satuan jarak 68 km yang merupakan kota Purwokerto. Untuk bobot 3 diberikan pada subkriteria dengan satuan jarak 95 km yang

merupakan kota Pekalongan. Untuk bobot 2 diberikan pada pada subkriteria dengan satuan jarak 300 km yang merupakan kota Blora dimana satuan jarak kota ini terbilang paling jauh dari lokasi Kampus STIMIK Tunas Bangsa Banjarnegara.

Untuk data strategi promosi diambil dari media promosi yang telah dilakukan oleh tim pemasaran kampus. Untuk rumusan mengenai media promosi ini dilakukan dengan menganalisa target pasar atau audiens. Dari sembilan media promosi yang digunakan, tujuan utamanya yakni untuk meningkatkan brand awareness atau tingkat kesadaran serta membangun citra atau reputasi di mata audiens. Muara dari proses promosi ini yakni untuk meningkatkan jumlah pendaftar.

Strategi promosi disajikan pada gambar tabel 4.3, hal ini sesuai dengan teknik promosi yang dilakukan oleh perguruan tinggi.

Tabel 4.3. Strategi Promosi

<i>KODE</i>	<i>JENIS PROMOSI</i>
<i>P1</i>	Instagram
<i>P2</i>	Facebook
<i>P3</i>	Brosur
<i>P4</i>	Spanduk/Flyer
<i>P5</i>	Sosialisasi ke Sekolah
<i>P6</i>	Beasiswa
<i>P7</i>	Kerjasama Lembaga
<i>P8</i>	Seminar
<i>P9</i>	Rekomendasi Teman/Alumni

Peneliti melakukan klasifikasi data pemasaran yang diperoleh dengan cara wawancara dengan ketua departemen, tim marketing, serta persetujuan ketua yayasan. Data yang disusun berupa kriteria dan alternatif yang akan digunakan dalam penyusunan hirarki permasalahan. Gambar hirarki ini bertujuan untuk

mencari solusi dalam menentukan jenis media pemasaran yang akan digunakan nantinya. Berikut gambar hirarki permasalahan yang digunakan untuk mencari solusi penentuan jenis promosi seperti ditunjukkan pada gambar 4.1 berikut.



Gambar 4.1. Pohon Keputusan

4.3 Pemanggilan Dataset

Proses selanjutnya yakni memanfaatkan Google Drive untuk menyimpan dataset dan Google Collaboratory untuk menjalankan program dengan menggunakan bahasa pemrograman python. Untuk mengolah dataset dimulai dengan pemanggilan dengan perintah sebagai berikut:

```

from google.colab import drive
drive.mount('/content/drive')

path_to_file =
'/content/drive/MyDrive/dataset_riset_nurus_01.csv'
data = pd.read_csv(path_to_file)
  
```

Perintah di atas akan menampilkan dataset seperti pada tabel 4.4 Dari gambar tersebut akan muncul dataset yang sudah diinisiasi melalui Google Collaboratory.

Hal ini karena dataset dalam format csv yang diunggah sudah dilakukan proses inisiasi terlebih dahulu.

Tabel 4.4. Hasil Pemanggilan Dataset pada Google Collaboratory

Mahasiswa	Asal Sekolah	Pekerjaan Orangtua	Program Studi	Daerah Asal
0	1	8	1	8
1	2	8	1	8
2	3	7	1	8
3	4	7	1	8
4	5	7	1	8

4.4 Pembersihan

Proses pembersihan data dilakukan untuk mencegah kesalahan pada proses asosiasi seperti adanya redundansi. Kondisi lain yakni memastikan tidak ada data hilang serta tipe data yang kurang sesuai. Proses ini dilakukan dengan perintah sebagai berikut:

```
data.isnull().sum()
#tidak terdapat missing value pada seluruh variabel
```

Berdasarkan code di atas, maka akan menampilkan kondisi dataset yang tersedia sebagai berikut.

```
mahasiswa      0
asal_sekolah   0
pekerjaan_orangtua  0
program_studi  0
daerah_asal    0
dtype: int64
```

Gambar 4.2. Kondisi Dataset yang Menunjukkan Tidak Adanya Missing Value
 Gambar 4.2 di atas menunjukkan bahwa seluruh variabel penelitian yang terdiri dari data mahasiswa, asal sekolah, pekerjaan orangtua, program studi dan daerah asal tidak ada data kosong atau missing value.

4.5 Perhitungan K-Means Clustering

Perhitungan K-Means dimulai dengan menentukan centroid awal secara acak. Berdasarkan wawancara dengan bagian pemasaran, pemilihan centroid awal yaitu nilai $k=5$. Nilai ini muncul berdasarkan jumlah lokasi promosi yang akan didatangi oleh tim penerimaan mahasiswa baru nantinya.

4.5.1 Menghitung Centroid Setiap Cluster

Selanjutnya yakni menghitung jarak data ke setiap centroid. Perhitungan dimulai dari data ke-1 hingga data ke-1110 untuk mengetahui matriks jarak C1, C2, C3, C4, C5.

- **Cluster 1**

$$C1 = \sqrt{(x1 - c1x1)^2 + (x2 - c1x2)^2 + (x3 - c1x3)^2 + (x4 - c1x4)^2}$$

$$= \sqrt{(8 - 8)^2 + (8 - 4)^2 + (1 - 2)^2 + (8 - 8)^2}$$

$$= 4,123$$
- **Cluster 2**

$$C2 = \sqrt{(x1 - c2x1)^2 + (x2 - c2x2)^2 + (x3 - c2x3)^2 + (x4 - c2x4)^2}$$

$$= \sqrt{(8 - 5)^2 + (8 - 4)^2 + (1 - 2)^2 + (8 - 8)^2}$$

$$= 5,447$$
- **Cluster 3**

$$C3 = \sqrt{(x1 - c3x1)^2 + (x2 - c3x2)^2 + (x3 - c3x3)^2 + (x4 - c3x4)^2}$$

$$= \sqrt{(8 - 8)^2 + (8 - 4)^2 + (1 - 2)^2 + (8 - 3)^2}$$

$$= 6,401$$
- **Cluster 4**

$$C4 = \sqrt{(x1 - c4x1)^2 + (x2 - c4x2)^2 + (x3 - c4x3)^2 + (x4 - c4x4)^2}$$

$$= \sqrt{(8 - 8)^2 + (8 - 3)^2 + (1 - 1)^2 + (8 - 8)^2}$$

$$= 3,000$$
- **Cluster 5**

$$C5 = \sqrt{(x1 - c5x1)^2 + (x2 - c5x2)^2 + (x3 - c5x3)^2 + (x4 - c5x4)^2}$$

$$= \sqrt{(8 - 8)^2 + (8 - 4)^2 + (1 - 1)^2 + (8 - 8)^2}$$

$$= 4,000$$

Gambar 4.3. Hasil Jarak Data

Perbandingan dan pemilihan jarak terdekat dari data dengan cluster yang paling dekat. Jarak digunakan sebagai parameter yang bisa menunjukkan suatu kelompok berdasarkan cluster terdekat.

4.5.2 Menghitung Nilai Rata – rata Cluster

Kemudian nilai rata – rata dari masing – masing cluster akan dijadikan sebagai centroid baru pada perhitungan iterasi. Setelah nilai centroid baru didapatkan, maka perhitungan iterasi dilakukan hingga tidak ada nilai centroid yang berubah pada masing – masing cluster.

Tabel 4.5. Hasil Perhitungan Iterasi 1

Mhs	C1	C2	C3	C4	C5	Jarak Terdekat	Cluster
1	4,123	5,477	6,481	3,000	4,000	3,000	4
2	4,123	5,477	6,481	3,000	4,000	3,000	4
3	1,414	3,000	5,196	1,414	1,000	1,000	5
4	1,414	3,000	5,196	1,414	1,000	1,000	5
5	1,414	3,000	5,196	1,414	1,000	1,000	5
6	1,414	3,000	5,196	1,414	1,000	1,000	5
7	1,414	3,000	5,196	1,414	1,000	1,000	5
8	1,000	3,742	5,099	1,000	0,000	0,000	5
9	1,414	3,000	5,196	1,414	1,000	1,000	5
1110	0,000	3,606	5,000	1,414	1,000	0,000	1

Tabel 4.6. Nilai Rata – rata Iterasi 1

Data ke	Cluster	Centroid Awal			
		V1	V2	V3	V4
478	1	8	4	2	8
339	2	5	4	2	6
220	3	8	4	2	3
567	4	8	5	1	8
412	5	8	4	1	8

Pada penelitian ini, hasil dari perhitungan jarak untuk masing – masing data dengan centroid pada perhitungan iterasi 1 terlihat pada tabel 4.5. Proses perhitungan iterasi berhenti pada iterasi 7, dimana nilai centroid pada masing – masing cluster pada iterasi ke 7 tidak berubah. Tabel 4.6 merupakan tahap perhitungan terakhir dengan nilai centroid untuk masing – masing cluster.

Cluster 1 mempunyai jumlah mahasiswa sebanyak 158, cluster 2 berjumlah 454 mahasiswa, cluster 3 berjumlah 85 mahasiswa, cluster 4 berjumlah 232 dan cluster 5 sebanyak 181 mahasiswa.

Tabel 4.7. Perhitungan Iterasi ke 7

Mhs	C1	C2	C3	C4	C5	Jarak Terdekat	Cluster
1	3,965	4,888	5,628	1,604	3,799	1,604	4
2	3,965	4,888	5,628	1,604	3,799	1,604	4
3	1,095	2,906	4,836	2,803	0,527	0,527	5
4	1,095	2,906	4,836	2,803	0,527	0,527	5
5	1,095	2,906	4,836	2,803	0,527	0,527	5
6	1,095	2,906	4,836	2,803	0,527	0,527	5
7	1,095	2,906	4,836	2,803	0,527	0,527	5
8	1,194	3,381	4,956	2,888	0,590	0,590	5
9	1,095	2,906	4,836	2,803	0,527	0,527	5
1110	0,652	3,384	4,952	2,861	1,161	0,652	1

Tabel 4.8. Nilai Rata – rata Iterasi 7

Cluster	Centroid Baru				Hasil Cluster
	V1	V2	V3	V4	
1	7,388	4,213	2,000	7,931	158
2	6,008	4,442	1,488	5,349	454
3	6,913	5,110	1,520	3,323	85
4	7,259	6,721	1,578	7,769	232
5	7,465	4,239	1,000	7,937	181

Dari hasil iterasi 7 terbentuk anggota cluster dengan nilai rata – rata seperti pada tabel 4.7 Untuk nilai rata – rata cluster 1 nilai asal sekolah yakni 7,388, nilai untuk variabel pekerjaan orangtua 4,213, nilai untuk program studi 2,000 dan daerah asal mendapai nilai 7,931. Hasil perhitungan hingga cluster 5 selengkapnya terlihat pada tabel 4.8 di atas.

4.5.3 Menghitung Nilai Silhouette Score

Untuk mengukur seberapa baik objek ditempatkan dalam suatu kluster yang sudah dilakukan, maka peneliti mencari nilai Silhouette Score. Proses penghitungannya dilakukan berdasarkan jarak antara objek dengan kluster yang berdekatan. Cara kerja metode Silhouette Coefficient ini mengukur jarak dari masing – masing sampel data dengan kluster terdekat serta kluster terdekat kedua. Setelah proses tersebut selesai lalu rasio dari jarak data ini akan dihitung. Panduan penghitungan nilai Silhouette Score berkisar antara -1 hingga maksimal 1 dengan ketentuan semakin tinggi nilai Silhouette Score maka hasil pengelompokkan kluster dianggap semakin baik. Berikut interpretasi hasil perhitungan menggunakan Silhouette Score :

- Hasil nilai mendekati 1 artinya objek data sudah berada pada kluster yang tepat serta mempunyai jarak yang baik dengan kluster lainnya.
- Hasil nilai mendekati 0 artinya objek data berada di batas antara dua kluster.
- Hasil nilai mendekati -1 artinya objek data kemungkinan besar berada pada kluster yang kurang tepat bahkan salah.

Untuk menghitung nilai Silhouette Score melalui Google Collaboration menggunakan code perintah sebagai berikut.

```
silhouette_scores = silhouette_score(data2, cluster,
sample_size=len(data2), random_state=42)
silhouette_scores
```

Hasil score yang didapatkan yakni 0,4279098918854162. Berdasarkan ketentuan, nilai Silhouette Score 0,427 termasuk dalam kategori kualitas cluster yang baik. Interpretasi nilai ini masih bisa dikatakan cukup baik, namun tidak termasuk luar biasa karena masih mempunyai ruang untuk perbaikan.

4.6 Penghitungan TOPSIS (Technique for Order Preference by Similarity to Ideal Solution)

Setelah proses clustering dilakukan, tahapan selanjutnya yakni melakukan penghitungan TOPSIS untuk menganalisis keputusan bersifat multi-kriteria dan mendapatkan alternatif terbaik untuk masing – masing cluster.

4.6.1 Menentukan Keputusan Matriks Ternormalisasi

Tahapan pertama yang dilakukan yakni melakukan penghitungan nilai sum squared untuk setiap kolom. Setelah itu baru melakukan perhitungan normalisasi. Berikut ini perintah yang digunakan untuk menghitung sum squared dan normalisasi melalui Google Collaboration.

```
# Menghitung nilai sum squared untuk setiap kolom
sum_squared = np.sqrt(np.sum(average_values**2, axis=0))

# Menghitung normalisasi
normalized_matrix = average_values / sum_squared
normalized_matrix.head()
```

Kode perintah di atas digunakan untuk menghitung nilai sum squared setiap kolom pada matriks. Setelah itu matriks akan dinormalisasi dengan cara membaginya dengan menggunakan nilai sum squared. Pada tabel 4.9 berikut merupakan matriks yang telah dinormalisasi dan siap digunakan dalam proses analisa selanjutnya.

Tabel 4.9. Matriks yang Telah Dinormalisasi

Cluster	Asal Sekolah	Pekerjaan Orangtua	Program Studi	Daerah Asal
0	0.432909	0.340209	0.452638	0.255383
1	0.481326	0.351764	0.447421	0.587403
2	0.475689	0.558110	0.464696	0.578390
3	0.442764	0.538595	0.452739	0.257930
4	0.398176	0.393486	0.417156	0.434249

4.6.2 Menentukan Keputusan Matriks Ternormalisasi Terbobot

Matriks ternormalisasi terbobot dilakukan pada kriteria yang digunakan sebagai variabel penelitian. Setelah melakukan normalisasi matriks keputusan, selanjutnya yakni memberikan bobot pada setiap kriteria penelitian. Berikut code perintah yang digunakan dalam proses menentukan keputusan matriks ternormalisasi terbobot.

```
asal_sekolah = 4/normalized_matrix['asal_sekolah'] #4: penting
pekerjaan_orangtua = 3/normalized_matrix['pekerjaan_orangtua']
#3: cukup penting
program_studi = 2/normalized_matrix['program_studi'] #2: tidak
penting
daerah_asal = 5/normalized_matrix['daerah_asal'] #5: sangat
penting

normalized_weighted = pd.DataFrame({'asal_sekolah':
asal_sekolah, 'pekerjaan_orangtua': pekerjaan_orangtua,
'program_studi': program_studi, 'daerah_asal': daerah_asal})
normalized_weighted
```


Code perintah di atas nilai bobot preferensi dari setiap kriteria berdasarkan tingkat kepentingan atau prioritas yang ditetapkan. Nilai bobot tertinggi yang digunakan pada variabel daerah asal karena patokan jarak menjadi pertimbangan penting ketika melakukan promosi. Hasil perhitungan matriks ternormalisasi terbobot terlihat melalui tabel 4.10 berikut.

Tabel 4.10. Matriks Ternormalisasi Terbobot

Cluster	Asal Sekolah	Pekerjaan Orangtua	Program Studi	Daerah Asal
0	9.238.545	8.665.295	4.418.244	19.563.100
1	8.310.380	8.528.457	4.470.068	8.512.042
2	8.408.860	5.375.287	4.303.891	8.644.687
3	9.034.153	5.570.049	4.417.554	19.385.139
4	10.045.821	7.624.162	4.794.373	11.514.130

4.6.3 Menentukan Matriks Solusi Ideal Positif dan Negatif

Dalam proses ini tujuannya yakni mencari nilai maksimum dan minimum dari kriteria asal sekolah, pekerjaan orangtua, program studi, daerah asal untuk membentuk matriks solusi ideal positif dan solusi ideal negatif. Berikut ini code perintah untuk menentukan matriks solusi ideal positif negatif.

```
#1. Menentukan matriks dari solusi ideal positif
max_sekolah = max(asal_sekolah)
max_pekerjaan = max(pekerjaan_orangtua)
max_studi = max(program_studi)
max_daerah = max(daerah_asal)

#2. Menentukan matriks dari solusi ideal negatif
min_sekolah = min(asal_sekolah)
min_pekerjaan = min(pekerjaan_orangtua)
min_studi = min(program_studi)
min_daerah = min(daerah_asal)
```

4.6.4 Menghitung Jarak Alternatif

Untuk menghitung jarak setiap alternatif dengan solusi ideal positif dan solusi ideal negatif dapat dilakukan dengan perintah sebagai berikut.

```
#1. Menghitung jarak alternatif dengan ideal positif
dist_positif = ((asal_sekolah -
max_sekolah)**2)+((pekerjaan_orangtua -
max_pekerjaan)**2)+((program_studi -
max_studi)**2)+((daerah_asal - max_daerah)**2)
#jika melihat dari rumusnya seperti ini:
#dist_positif = (((asal_sekolah -
max_sekolah)**2)+((pekerjaan_orangtua -
max_pekerjaan)**2)+((program_studi -
max_studi)**2)+((daerah_asal - max_daerah)**2))**0.05

#2. Menghitung jarak alternatif dengan ideal negatif
dist_negatif = ((asal_sekolah -
min_sekolah)**2)+((pekerjaan_orangtua -
min_pekerjaan)**2)+((program_studi -
min_studi)**2)+((daerah_asal - min_daerah)**2)
#jika melihat dari rumusnya seperti ini:
#dist_negatif = (((asal_sekolah -
min_sekolah)**2)+((pekerjaan_orangtua -
min_pekerjaan)**2)+((program_studi -
min_studi)**2)+((daerah_asal - min_daerah)**2))**0.05

distance = pd.DataFrame({'Positif': dist_positif,
'Negatif': dist_negatif})
distance
```

Performa setiap cluster dapat dievaluasi dengan cara melihat nilai hasil jarak alternatif. Misalnya untuk nilai jarak lebih kecil lebih dekat ke solusi ideal positif dan nilai jarak yang lebih besar lebih dekat dengan solusi ideal negatif, maka kondisi ini menunjukkan hasil mendekati kondisi ideal. Berlaku juga sebaliknya jika jarak dengan nilai yang lebih besar lebih dekat dengan solusi ideal positif dan nilai jarak yang lebih kecil lebih dekat dengan solusi ideal negatif. Pada tabel 4.11

terlihat hasil perhitungan jarak alternatif dari lima cluster, dimana untuk cluster 0 mempunyai jarak ke solusi ideal positif 0.792940 dan jarak solusi ideal negatif 133.824663.

Tabel 4.11. Hasil Perhitungan Jarak Alternatif

Cluster	Positif	Negatif
0	0.792940	133.824.663
1	125.261.524	9.970.096
2	132.956.115	0.027293
3	10.777.684	118.798.034
4	65.869.884	17.322.290

4.6.5 Menghitung Preferensi Dari Solusi Ideal

Tahapan selanjutnya yakni menghitung skor TOPSIS untuk masing – masing cluster. Untuk menghitung preferensi dari solusi ideal masing – masing cluster menggunakan perintah berikut.

```
topsis_score = dist_negatif/(dist_negatif+dist_positif)
topsis = pd.DataFrame({'topsis_score': topsis_score})
topsis
```

Proses ini nantinya bisa menghasilkan output berupa klasifikasi cluster – cluster berdasarkan seberapa dekat mereka dengan solusi ideal negatif. Jika cluster mendapatkan skor mendekati 1 artinya cluster tersebut mendekati solusi ideal negatif, sedangkan cluster yang mendekati 0 menunjukkan cluster tersebut lebih dekat dengan solusi ideal positif.

Tabel 4.12. Hasil Perhitungan Preferensi Solusi Ideal

Cluster	Topsis Score
0	0.994110
1	0.073726
2	0.000205
3	0.916824
4	0.208220

Hasil representasi numerik pada tabel 4.12 menunjukkan kondisi cluster 0 memperoleh skor TOPSIS mencapai 0.994110. Interpretasi skor TOPSIS untuk cluster ini tergolong sangat tinggi karena mendekati 1. Dengan kata lain cluster ini paling dekat dengan solusi ideal negatif dalam konteks pengambilan keputusan yang diimplementasikan. Sedangkan kluster yang mendapatkan hasil TOPSIS paling rendah yakni cluster 2 dengan skor 0.000205 atau paling mendekati nilai 0 dan lebih dekat dengan solusi ideal positif.

4.6.6 Melakukan Perangkingan

Dari skor TOPSIS yang sudah diperoleh, proses selanjutnya yakni melakukan perangkingan untuk lima cluster. Proses perangkingan ini dapat dilakukan menggunakan kode perintah berikut.

```
topsis['rank']=topsis['topsis_score'].rank(ascending =
False)
topsis
```

Perangkingan dilakukan dengan cara memberikan peringkat dari semua skor TOPSIS dengan ketentuan melihat dari nilai tertinggi hingga nilai terendah. Pada tabel 4.13 terlihat rangking 1 jatuh pada cluster 0, disusul cluster 3, cluster 4, cluster 1 dan yang paling rendah yakni cluster 2 dengan score 0.000205.

Tabel 4.13. Hasil Perangkingan

Cluster	Topsis Score	Rank
0	0.994110	1
1	0.073726	4
2	0.000205	5
3	0.916824	2
4	0.208220	3

4.7 Decision Tree

Untuk mendapatkan hasil analisa berupa rekomendasi jenis marketing untuk masing – masing cluster, maka perlu menggabungkan hasil kluster dan skor TOPSIS.

4.7.1 Menggabungkan Hasil Cluster dan TOPSIS

Untuk menggabungkan hasil cluster dan perangkingan TOPSIS, sebelumnya perlu melakukan penggabungan DataFrame lalu konversi tipe data menjadi interget. Berikut ini bahasa perintah yang digunakan untuk menggabungkan hasil cluster dan TOPSIS.

```
cluster5_topsis = pd.merge(round(average_values), topsis,
on='cluster', how='inner')

#convert float to integer
cluster5_topsis['asal_sekolah'] = cluster5_topsis
['asal_sekolah'].astype(int)
cluster5_topsis['pekerjaan_orangtua'] = cluster5_topsis
['pekerjaan_orangtua'].astype(int)
cluster5_topsis['program_studi'] = cluster5_topsis
['program_studi'].astype(int)
cluster5_topsis['daerah_asal'] = cluster5_topsis
['daerah_asal'].astype(int)
cluster5_topsis['rank'] = cluster5_topsis
['rank'].astype(int)

cluster5_topsis
```


Pada gambar 4.14 terlihat hasil penggabungan cluster dan TOPSIS. Dengan catatan data hasil clustering mendapat tambahan data hasil score dan ranking TOPSIS.

Tabel 4.14. Hasil Penggabungan Cluster dan TOPSIS

cluster	asal sekolah	pekerjaan orangtua	program studi	daerah asal	topsis score	rank
0	7	4	2	3	0.994110	1
1	7	4	2	8	0.073726	4
2	7	7	2	8	0.000205	5
3	7	6	2	4	0.936824	2
4	6	5	1	6	0.208220	3

Interpretasi dari penggabungan hasil cluster dan TOPSIS memberikan wawasan bagaimana masing – masing cluster diurutkan berdasarkan seberapa baik data ini mendekati solusi ideal negatif yang dilengkapi dengan data karakteristik atribut. Misalnya untuk cluster 0 mempunyai nilai atribut yang terdiri dari asal sekolah bernilai 7 (tinggi), pekerjaan orangtua 4 (sedang) program studi 2 (sedang), dan untuk daerah asal bernilai 3 (rendah). Sedangkan untuk skor TOPSIS terbilang sangat tinggi yakni 0.994110 atau mendekati nilai 1.

4.7.2 Import Data Decision Tree

Pohon keputusan sebelumnya sudah dibuat melalui spreadsheet, agar bisa dijalankan melalui Google Collab maka perlu diunggah terlebih dahulu. Ketentuan jenis promosi dijalankan berdasarkan beberapa kriteria berikut.

1. Jika jarak daerah asalnya dekat (7-8) maka dapat dilakukan promosi melalui brosur, spanduk, sosialisasi ke sekolah
2. Jika jarak daerah asalnya sedang (5-6) maka dapat dilakukan promosi melalui instagram, facebook, brosur, spanduk

3. Jika jarak daerah asalnya jauh (2-3) maka dapat dilakukan promosi melalui instagram, facebook, seminar, rekomendasi teman/alumni
4. Jika pekerjaan orang tua adalah pedagang, petani, buruh, atau guru swasta maka dapat memberikan beasiswa
5. Jika orang tua PNS maka bisa promosi melalui kerjasama Lembaga

Berikut ini kode perintah yang dijalankan untuk melakukan import data pohon keputusan melalui Google Collaboration.

```
#Dari file spreadsheet ambil sheet yang detail tree
file2 = ('/content/drive/MyDrive/dataset_penelitian.csv')

#Membaca data
data_tree = pd.read_csv(file2, encoding = "ISO-8859-1")
data_tree.head() #menampilkan data dari baris pertama
hingga baris kelima
```

Dataset seperti pada tabel 4.15 Dari gambar tersebut akan muncul dataset yang sudah diinisiasi melalui Google Collaboratory. Hal ini karena dataset dalam format csv yang diunggah sudah dilakukan proses inisiasi terlebih dahulu. Perintah di atas akan menampilkan data pohon keputusan yang sudah diimport ke Google Collaboration seperti pada gambar 4.15 berikut.

Tabel 4.15. Hasil Import Data Decision Tree

	daerah_asal	asal_sekolah	pekerjaan_orangtua	program_studi	kode_promosi
0	8	8	8	1	P3,P4,P5,P6
1	8	8	8	2	P3,P4,P5,P6
2	8	8	7	1	P3,P4,P5,P7
3	8	8	7	2	P3,P4,P5,P7
4	8	8	6	1	P3,P4,P5

Varibel data yang muncul yakni daerah asal, asal sekolah, pekerjaan orangtua, program studi, dan kode promosi. Misalnya untuk data baris pertama dimana daerah asal mendapatkan nilai 8, asal sekolah 8, pekerjaan orangtua 8, program studi 1 dan kode promosi yang muncul yakni P3, P4, P5, P6.

4.7.3 Menggabungkan Variabel

Empat variabel penelitian yang terdiri dari daerah asal, asal sekolah, pekerjaan orangtua, program studi akan digabungkan menjadi satu yang akan menjadi nilai pada kolom 'combined'. Hal ini bertujuan untuk mempermudah pemrosesan data dalam hal menentukan kode promosi. Berikut kode perintah yang digunakan untuk menggabungkan empat variabel.

```
data_tree['combined']=data_tree['asal_sekolah'].astype(str)
+ data_tree['pekerjaan_orangtua'].astype(str) +
data_tree['program_studi'].astype(str)+
data_tree['daerah_asal'].astype(str)
data_tree.head()
```

Pada gambar 4.16 terlihat kolom baru yang merupakan data gabungan dari nilai – nilai empat variabel penelitian yang sudah melalui proses konversi menjadi tipe data interger.

Tabel 4. 16 Penggabungan Hasil Variabel

	Daerah asal	Asal sekolah	Pekerjaan orangtua	Program studi	Kode promosi	combined
0	8	8	8	1	P3,P4,P5,P6	8818
1	8	8	8	2	P3,P4,P5,P6	8828
2	8	8	7	1	P3,P4,P5,P7	8718
3	8	8	7	2	P3,P4,P5,P7	8728
4	8	8	6	1	P3,P4,P5	8618

4.7.4 Rekomendasi Promosi Untuk Masing – Masing Cluster

Tahapan selanjutnya yakni memproses data untuk mendapatkan output berupa rekomendasi jenis promosi untuk masing – masing cluster. Pertama yakni dengan menggabungkan data nilai dari empat variabel hasil K-Means dan nilai perangkingan TOPSIS menjadi satu kolom baru. Kemudian tipe data pada kolom baru tersebut dikonversi menjadi tipe data interger. Hasil datanya kemudian digabungkan dengan dataframe lain yang relevan. Tahapan proses ini bisa dijalankan menggunakan bahasa perintah berikut.

```
cluster5_topsis['combined']=cluster5_topsis['asal_sekolah'].
astype(str) +
cluster5_topsis['pekerjaan_orangtua'].astype(str) +
cluster5_topsis['program_studi'].astype(str)+
cluster5_topsis['daerah_asal'].astype(str)

cluster5_topsis['combined'] = cluster5_topsis
['combined'].astype(int)
cluster5_topsis

promosi = pd.merge(cluster5_topsis, data_trees,
on='combined', how='inner')
promosi
```

Output berupa rekomendasi jenis promosi untuk masing – masing cluster bisa terlihat Pada tabel 4.17 berikut.

Tabel 4.17. Hasil Rekomendasi Untuk Setiap Kluster

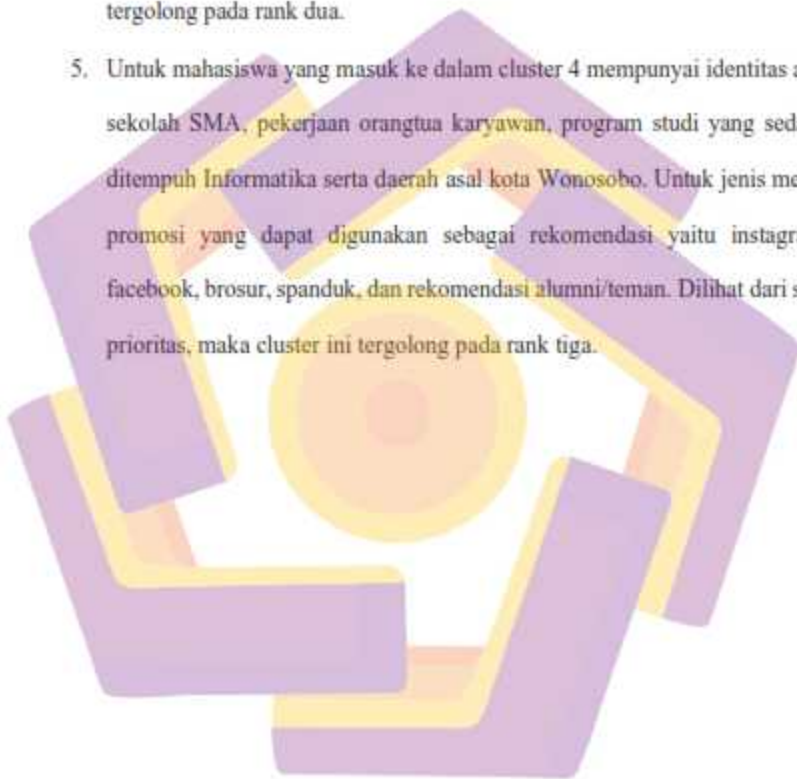
	asal_sekolah	pekerjaan_orangtua	program_studi	daerah_asal	topsis_score	rank	combined	kode_promosi
0	7	4	2	8	0.994110	1	7428	P1,P2,P6,P8,P9
1	7	4	2	8	0.073726	4	7428	P3,P4,P5,P6
2	7	7	2	8	0.000205	5	7728	P3,P4,P5,P7
3	7	6	2	4	0.916824	2	7624	P1,P2,P6,P8,P9
4	6	5	1	6	0.208220	3	6516	P1,P2,P3,P4,P9

Berikut ini interpretasi jenis promosi yang dapat dilakukan pada setiap cluster.

1. Untuk mahasiswa yang masuk ke dalam cluster 0 mempunyai identitas asal sekolah SMK, pekerjaan orangtua petani, program studi yang sedang ditempuh Sistem Informasi serta daerah asal kota Pekalongan. Untuk jenis media promosi yang dapat digunakan sebagai rekomendasi yaitu instagram, facebook, beasiswa, seminar, dan rekomendasi alumni/teman. Dilihat dari segi prioritas, maka cluster ini tergolong pada rank pertama yang paling diprioritaskan.
2. Untuk mahasiswa yang masuk ke dalam cluster 1 mempunyai identitas asal sekolah SMK, pekerjaan orangtua petani, program studi yang sedang ditempuh Sistem Informasi serta daerah asal kota Banjarnegara. Untuk jenis media promosi yang dapat digunakan sebagai rekomendasi yaitu brosur, spanduk, sosialisasi ke sekolah-sekolah, dan beasiswa. Dilihat dari segi prioritas, maka cluster ini tergolong pada rank empat.
3. Untuk mahasiswa yang masuk ke dalam cluster 2 mempunyai identitas asal sekolah SMK, pekerjaan orangtua PNS, program studi yang sedang ditempuh Sistem Informasi serta daerah asal kota Banjarnegara. Untuk jenis media promosi yang dapat digunakan sebagai rekomendasi yaitu brosur, spanduk, sosialisasi ke sekolah-sekolah, dan kerjasama Lembaga. Dilihat dari segi prioritas, maka cluster ini tergolong pada rank lima yang mana menjadi urutan paling akhir dari lima klaster.
4. Untuk mahasiswa yang masuk ke dalam cluster 3 mempunyai identitas asal sekolah SMK, pekerjaan orangtua PNS (Pegawai Negeri Sipil), program

studi yang sedang ditempuh Sistem Informasi serta daerah asal kota Banjarnegara. Untuk jenis media promosi yang dapat digunakan sebagai rekomendasi yaitu instagram, facebook, beasiswa, seminar, dan rekomendasi alumni/teman. Dilihat dari segi prioritas, maka cluster ini tergolong pada rank dua.

5. Untuk mahasiswa yang masuk ke dalam cluster 4 mempunyai identitas asal sekolah SMA, pekerjaan orangtua karyawan, program studi yang sedang ditempuh Informatika serta daerah asal kota Wonosobo. Untuk jenis media promosi yang dapat digunakan sebagai rekomendasi yaitu instagram, facebook, brosur, spanduk, dan rekomendasi alumni/teman. Dilihat dari segi prioritas, maka cluster ini tergolong pada rank tiga.



BAB V

PENUTUP

5.1. Kesimpulan

Pembahasan pada bagian hasil menunjukkan beberapa poin berikut :

- a. Hasil klasterisasi menunjukan dari lima klaster yang terbentuk mempunyai kemiripan atau karakteristik yang serupa. Sedangkan pada proses TOPSIS menghasilkan urutan prioritas klaster tertinggi hingga terendah dari kelima klaster yang terbentuk.
- b. Perhitungan Silhouette Score menunjukkan klaster yang dihasilkan oleh K-Means masuk dalam kategori cukup baik dalam artian terdapat pemisahan yang jelas dari masing – masing data yang berbeda atau tidak ada tumpang tindih data dalam klaster yang sama.
- c. Beberapa faktor yang berpengaruh pada hasil akurasi penelitian ini yakni jumlah cluster, penentuan kriteria yang relevan dan dapat mewakili karakteristik audiens serta pemberian bobot yang tepat.

5.2. Saran

Saran memuat berbagai usulan atau pendapat yang sebaiknya dikaitkan oleh penelitian sejenis. Saran dibuat berdasarkan kelemahan, pengalaman, kesulitan, kesalahan, temuan baru yang belum diteliti dan berbagai kemungkinan arah penelitian selanjutnya. Saran yang disusun harus didasarkan atas hasil penelitian yang dilakukan.



Daftar Pustaka

- Agusta Yudi. (2007). K-Means – Penerapan, Permasalahan dan Metode Terkait. *Jurnal Sistem Dan Informatika*, 3(Februari), 47–60.
- Ahlawat, R., Sahay, S., Sabitha, S., & Bansal, A. (n.d.). *Analysis of factors affecting enrollment pattern in Indian universities using k-means clustering*.
- Amaliya Hani Nafisah. (2023). *Clustering Bidang Keilmuan Menggunakan Kombinasi*. 04, 405–413.
- Andrianto, C. B., Kusriani, K., & Fatta, H. Al. (2017). Analisis Sistem Pendukung Keputusan Penerima Beasiswa Di Smp Muhammadiyah 2 Kalasan. *Respati*, 12(34), 46–60. <https://doi.org/10.35842/jtir.v12i34.101>
- Andry, Yani Maulita, & Suci Ramadani. (2021). *Sistem Pendukung Keputusan Penentuan Lokasi Promosi Penerimaan Siswa Baru Di MTS. S. Hubbul Wathan Modal Bangsa*.
- Anisya. (2020). *DATA MINING DALAM PREDIKSI PASOKAN KELAPA SAWIT*. 3(1). <https://doi.org/10.31869/rtj.v3i1.1687>
- Assylla, S., & Nugraha. (2022). Perancangan Strategi Pemasaran dengan Pendekatan Analisis SWOT dan Metode TOPSIS. *Jurnal Riset Teknik Industri*, 129–140. <https://doi.org/10.29313/jrti.v2i2.1283>
- Bahar, A., Pramono, B., & Sagala, L. H. S. (2016). Penentuan strategi penjualan alat-alat tattoo di studio sonyxattoo menggunakan metode. *SemanTIK*, 2(2),

75–86.

- Bolinger, M. T., Josefy, M. A., Stevenson, R., & Hitt, M. A. (2022). Experiments in Strategy Research: A Critical Review and Future Research Opportunities. In *Journal of Management* (Vol. 48, Issue 1).
<https://doi.org/10.1177/01492063211044416>
- Garcia-Cascales, M. S., & Lamata, M. T. (2012). On rank reversal and TOPSIS method. *Mathematical and Computer Modelling*, 56(5–6), 123–132.
<https://doi.org/10.1016/j.mcm.2011.12.022>
- Herlambang, H. P., Saputra, F., Prasetyo, M. H., Puspitasari, D., & Nurlaela, D. (2023). Perbandingan Klasifikasi Tingkat Penjualan Buah di Supermarket dengan Pendekatan Algoritma Decision Tree, Naive Bayes dan K-Nearest Neighbor. *Jurnal INSAN - Journal of Information System Management Innovation*, 3(1), 21–28. <https://doi.org/10.31294/jinsan.v3i1.2097>
- Hidayat, T. (2022). Klasifikasi Data Jamaah Umroh Menggunakan Metode K-Means Clustering. *Jurnal Sistim Informasi Dan Teknologi*, 19–24.
<https://doi.org/10.37034/jsisfotek.v4i1.115>
- Jannah, S. S. N., Firdya Riska, N., Alamsyah, M., & Indrayatna, F. (2023). Pengklasifikasian Penyakit Jantung dengan Metode Decision Tree. *Seminar Nasional Statistika Aktuaria*, 2–7. <http://prosidingnsna.statistics.unpad.ac.id>
- Kasri, M. A., & Jati, H. (2020). *Kombinasi K-Means dan Simple Additive Weighting dalam Menentukan Lokasi dan Strategi Pemasaran Universitas*. 2, 132–141.
- Khoironi. (2020). *Kombinasi Metode K-Means Dan Analytic Hierarchy Process*

Dalam Menentukan Penerima Beasiswa. Magister Teknik Informatika, Universitas AMIKOM Yogyakarta.

Khomsatun, K., Ikhsan, D., Ali, M., & Kursini, K. (2020). Sistem Pengambilan Keputusan Pemilihan Lahan Tanam Di Kabupaten Wonosobo Dengan K-Means Clustering Dan Topsis. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 9(1), 55.

<https://doi.org/10.23887/janapati.v9i1.23073>

Kusnawi. (2007). *Pengantar Solusi Data Mining. Seminar Nasional Teknologi. 2007*(November), 1-9.

Kursini, K. (2007). *Konsep dan Aplikasi Sistem Pendukung Keputusan* (Fl. Sigit Suyantoro (ed.)). Penerbit Andi.

https://books.google.co.id/books?id=RhEMeAAAQBAJ&pg=PA76&hl=id&source=gbs_selected_pages&cad=1#v=onepage&q&f=false

Marsono, M., Saripurna, D., & Zunaidi, M. (2021). Analisis Data Mining Pada Strategi Penjualan Produk PT Aquasolve Sanaria Dengan Menggunakan Metode K-Means Clustering. *J-SISKO TECH (Jurnal Teknologi Sistem Informasi Dan Sistem Komputer. TGD)*, 4(1), 127.

<https://doi.org/10.53513/jsk.v4i1.60>

Nata, A., Soedijono, B., & Al Fatta, H. (2017). *PENENTUAN WILAYAH PROMOSI PENERIMAAN SISWA BARU DENGAN METODE TOPSIS.*

Oki Oktaviarna Tensao, I Nyoman Yudi Anggara Wijaya, & Ketut Queen Fredlina. (2022). Analisa Data Mining dengan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru Pada

- STMIK Primakara. *INFORMASI (Jurnal Informatika Dan Sistem Informasi)*, 14(1), 1–17. <https://doi.org/10.37424/informasi.v14i1.135>
- Rahmalinda, N. A., & Jananto, A. (2022). Penerapan Metode K-Means Clustering Dalam Menentukan Strategi Promosi Berdasarkan Data Penerimaan Mahasiswa Baru. *Jurnal Tekno Kompak*, 16(2), 163–175.
- Raval, U. R., & Jani, C. (2016). International Journal of Computer Science and Mobile Computing Implementing & Improvisation of K-means Clustering Algorithm. In *International Journal of Computer Science and Mobile Computing* (Vol. 5, Issue 5). www.ijcsmc.com
- Rismayadi, A. A., Fatonah, N. N., & Junianto, E. (2021). ALGORITMA K-MEANS CLUSTERING UNTUK MENENTUKAN STRATEGI PEMASARAN DI CV. INTEGREET KONSTRUKSI. *JURNAL RESPONSIF*, 3(1), 30–36. <http://ejurnal.ars.ac.id/index.php/jti>
- Rusliyawati, Damayanti, Agus Wantoro, F. N. P. (2020). Penerapan Metode Simple Additive Weighting (SAW) Dalam Penentuan Social Customer Relationship Management (SCRM). 249–257.
- Sakur, S. B. H., Silangen, M., & Tuwohingide, D. (2022). Penerapan Algoritma K-Means Cluster dan Metode TOPSIS pada Pemilihan Mahasiswa kunjungan Industri. *Jurnal Ilmiah Teknik Informatika Dan Sistem Informasi*, 11(3), 851–860. <http://ojs.stmik-banjarbaru.ac.id/index.php/jutisi/article/view/1045>
- Servanda, Y. (2018). ANALISIS PENENTUAN PRIORITAS MEDIA PROMOSI PERGURUAN TINGGI MENGGUNAKAN KOMBINASI AHP DAN TOPSIS.

2(2).

Suarnatha, I. P. D., Agus, I. M., & Gunawan, O. (2022). Jurnal Computer Science and Information Technology (CoSciTech) manusia. *CoSciTech*, 3(2), 73–80.

SUTRISNO, B. T., & Andriyani, W. (2021). PENERAPAN MADM DENGAN METODE SAW UNTUK MENENTUKAN TARGET PROMOSI BERDASARKAN ASAL JURUSAN DI SEKOLAH. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 11(2), 480–492.
<https://doi.org/10.24176/simet.v11i2.4784>

Vafaei, N., Ribeiro, R. A., & Camarinha-Matos, L. M. (2021). Assessing Normalization Techniques for Simple Additive Weighting Method. *Procedia Computer Science*, 199, 1229–1236.
<https://doi.org/10.1016/j.procs.2022.01.156>

Wijayanto, S., Informasi, S., Pamulang, U., & Selatan, T. (2023). *BINA INFORMATIKA MENGGUNAKAN METODE ANALYTICAL*. 6, 156–164.

Winasis, C. L. R., Widiyanti, H. S., & Hadibrata, B. (2022). Determinasi Keputusan Pembelian: Harga, Promosi Dan Kualitas Produk (Literature Review Manajemen Pemasaran). *Jurnal Ilmu Manajemen Terapan*, 3(4), 452–462. <https://doi.org/10.31933/jemsi>.