

## PENERAPAN MACHINE LEARNING UNTUK PENENTUAN SEGMENTASI MAHASISWA BARU DENGAN METODE K MODES

Devi Sugianti<sup>1)</sup>, Hermanus Wim Hapsoro<sup>2)</sup>, Wahyu Setianto<sup>3)</sup>

STMIK Widya Pratama<sup>123)</sup>

devi.sugianti9807@gmail.com<sup>1)</sup>, wimhapsoro06@gmail.com<sup>2)</sup>, wahyu.s8106@gmail.com<sup>3)</sup>

### Abstrak

Strategi marketing sangat diperlukan, untuk mendapatkan mahasiswa sebanyak-banyaknya. Kelangsungan perguruan tinggi bergantung terhadap pembiayaan dari mahasiswa. Dibutuhkan strategi marketing untuk mempromosikan keunggulan dari setiap program studi, agar penerimaan mahasiswa baru semakin meningkat. Data yang perlu dianalisa adalah data pendaftaran mahasiswa baru di Sistem Pendaftaran Mahasiswa Baru SIPENMARU. Langkah langkah penelitian yang dilakukan : identifikasi masalah, pengumpulan data, pemrosesan data dan pemodelan. Data yang dikumpulkan dari tahun 2019 sampai dengan tahun 2021 sebanyak 1.219 data pendaftar dengan variabel sebanyak 109. Saat pemrosesan data dilakukan yang digunakan ada 5 variabel yaitu : jenis kelamin, asal kota, pekerjaan ayah, asal sekolah dan program studi. Pada fase pemodelan menggunakan clustering untuk mengelompokkan data dengan karakteristik yang sama, teknik cluster dengan mengelompokkan dengan tipe data bersifat kategorikal menggunakan K modes. Cluster dibagi menjadi 2 cluster, dari hasil eksperimen didapatkan hasil untuk program studi Teknik informatika di dominasi berjenis kelamin Laki-laki, asal kota Pekalongan, pekerjaan ayah PNS dan asal sekolah SMA. Sedangkan untuk program studi Sistem Informasi di dominasi jenis kelamin perempuan, asal kota Batang, pekerjaan ayah wiraswasta, asal sekolah SMK. Dari hasil penelitian didapatkan dua cluster untuk sebagai segemetasi mahasiswa baru yang akan di sasar oleh tim marketing

**Kata kunci:** Segmentasi, Mahasiswa Baru, K modes

### 1. Pendahuluan

Setiap awal tahun akademik untuk mendapatkan mahasiswa baru, maka perlu melakukan analisa data mahasiswa untuk menentukan strategi marketing (Chasanah dan Widiyono 2017). Di awal tahun 2020 adanya pandemi memberi dampak yang luas yang harus dihadapi oleh Perguruan Tinggi untuk kelangsungannya, karena sebagian besar Perguruan Tinggi bergantung terhadap pendapatan pembiayaan dari mahasiswa (Yuyun, Juhad dan Abdul 2022). Strategi marketing diperlukan untuk mempromosikan keunggulan dari setiap program studi, agar penerimaan mahasiswa baru semakin meningkat. (Abriyanto dan Damastuti 2019). Data mining dipilih sebagai salah satu cara untuk dapat menganalisa dan

menentukan strategi marketing agar tepat sasaran (Gupitha 2018). *Knowledge discovery database* (KDD) merupakan teknik yang dilakukan oleh data mining dengan adanya data yang melimpah untuk dijadikan sebuah informasi yang penting. (Asroni, Fitri dan Prasetyo 2018).

Data mining merupakan bagian dari *Machine learning* dengan menggunakan teknik atau metode untuk mencari pola atau informasi menarik dari data terpilih (Ramdhani dan Perdana 2021). Berbagai persoalan dapat diselesaikan dengan *Machine learning* dengan menggunakan sekumpulan data training atau pengalaman di masa lalu (*past experience*) (Primartha 2018). Terdapat tiga kelompok dari algoritma *machine learning* : *Supervised Learning*, *Unseupervised Learning*, dan *Reinforecement Learning*

(Primartha, Algoritma Machine Learning 2021). *Unsupervised learning* digunakan untuk menganalisa *cluster*, untuk menarik kesimpulan dari *dataset* (Nurhayati, Busman dan Pradono 2019). *Cluster* memisahkan data yang memiliki kemiripan karakteristik yang sama untuk berkumpul dalam *cluster* yang sama, dan memisahkan data-data dengan karakteristik yang berbeda untuk berkumpul dalam satu *cluster* (Prasetyo 2014). Teknik cluster untuk atribut yang bersifat kategorikal dengan menggunakan K Modes (Yulianto, Andreas dan Mulyani 2021). Dalam penelitian yang dilakukan oleh (Malikhatin, Rusgiyono dan Asih 2021) bahwa dapat mempermudah dalam komputasi dengan jumlah cluster sesuai yang diinginkan, jumlah cluster terbaik setelah dilakukan validasi Dunn Index didapatkan 2 cluster. Metode K Modes juga di gunakan untuk proses penentuan penerimaan bantuan langsung tunai (BLT) pengelompokkan data BLT dengan menggunakan 17 kriteria, dengan hasil pengujian dengan 160 data yang diuji didapatkan 96,875%. (Aprilia, Agus dan Rudhistiar 2021). Metode K-Modes dapat digunakan untuk klasterisasi penduduk lanjut usia Sumatera Selatan, hasilnya adalah 4 cluster dengan data yang digunakan adalah sensus penduduk Sumatera selatan yang berusia > 60 tahun dengan sampel 47.358 jiwa (Selva dan Pratama 2017). Untuk pengelompokkan tingkat minat belanja online di provinsi daerah istimewa Yogyakarta dengan metode K mode menghasilkan 5 cluster yang terbaik, dengan data yang diambil sebanyak 2415 record menggunakan 11 variabel. (Astri, et al. 2021). Metode K mode juga digunakan untuk pengelompokkan resep masakan daerah dengan mendaftarkan resep sebanyak 355 resep dan dikelompokkan menjadi 4 kelompok yang lebih umum dan 8 kelompok yang lebih spesifik (Indriani dan Budiman 2017)

Dari uraian diatas, dapat dirumuskan bahwa bagaimana penerapan *machine learning* untuk penentuan segmentasi mahasiswa baru dengan metode K modes agar dapat menentukan strategi marketing untuk menjangkau mahasiswa baru

## 2. Metode Penelitian

Kerangka pemikiran masalah dapat dilihat pada gambar 1



Gambar 1. Kerangka Pemikiran Penelitian

### 2.1 Identifikasi Permasalahan

Pada fase identifikasi permasalahan bahwa untuk mendapatkan mahasiswa sebanyak-banyaknya di perguruan tinggi dengan tantangan keadaan pandemi seperti ini membuat perguruan tinggi untuk membuat strategi marketing sebaik-baiknya dan seefisien mungkin dapat mengenai masyarakat. Saat ini strategi marketing yang tidak dapat dilakukan adalah mendatangi sekolah-sekolah untuk melakukan paparan, karena adanya pandemi. Strategi marketing yang masih bisa dilakukan adalah dengan cara menyebarkan brosur, spanduk, radio, dan media sosial.

### 2.2 Fase Pengumpulan Data

Pada fase pengumpulan data, mengumpulkan data pendaftaran dari 4 Program Studi pada tahun 2019 sampai dengan 2021 dengan jumlah pendaftar 1.219 pendaftar. Ada 109 variabel yang dilakukan untuk penginputan data pendaftaran..

### 2.3 Pemrosesan Data

Data mahasiswa yang mendaftar akan diinput dalam sistem pendaftaran mahasiswa baru SIPENMARU. Semua data pendaftar di input dengan jumlah variabel sebanyak 109 variabel, dari data pribadi, data orang tua, data asal sekolah. Adapun yang menjadi bahan pertimbangan dengan menggunakan 5 variabel untuk mendukung K Modes. Variabel yang digunakan adalah jenis kelamin, asal kota, pekerjaan ayah, asal sekolah, program studi

Dalam menentukan strategi marketing mahasiswa baru dapat menggunakan variabel

sebagai berikut: jenis kelamin, asal kota, pekerjaan ayah, asal sekolah, program studi.

Sampel data yang akan diolah dapat dilihat pada tabel 1

Tabel 1. Data mahasiswa

| NO | Jenis Kelamin | Asal Kota  | Pekerjaan Ayah | Asal Sekolah | progd |
|----|---------------|------------|----------------|--------------|-------|
| 1  | L             | PEMALANG   | PNS            | SMK          | TI    |
| 2  | L             | PEMALANG   | PNS            | SMA          | TI    |
| 3  | L             | PEKALONGAN | PNS            | SMA          | TI    |
| 4  | P             | PEMALANG   | WIRASWASTA     | SMA          | TI    |
| 5  | L             | BATANG     | SWASTA         | SMK          | SI    |
| 6  | P             | PEKALONGAN | SWASTA         | SMA          | SI    |
| 7  | P             | PEKALONGAN | SWASTA         | SMK          | SI    |
| 8  | L             | PEKALONGAN | SWASTA         | SMK          | TI    |
| 9  | L             | PEKALONGAN | SWASTA         | SMA          | TI    |
| 10 | P             | PEMALANG   | WIRASWASTA     | SMK          | SI    |
| 11 | P             | PEKALONGAN | PNS            | SMK          | SI    |
| 12 | P             | BATANG     | GURU           | SMK          | KA    |
| 13 | P             | BATANG     | SWASTA         | SMA          | KA    |
| 14 | P             | PEMALANG   | WIRASWASTA     | SMA          | MI    |
| 15 | L             | PEKALONGAN | GURU           | SMA          | SI    |
| 16 | P             | PEKALONGAN | PNS            | SMK          | TI    |
| 17 | P             | BATANG     | WIRASWASTA     | SMK          | SI    |
| 18 | L             | BATANG     | GURU           | SMA          | MI    |

#### 2.4 Pemodelan

Pemodelan yang digunakan *K-mode*, *K mode* pertama kali diperkenalkan oleh Huang tahun 1998. *K-mode* merupakan pengembangan dari *K-means* dengan data bersifat kategorikal. *K-Modes* melakukan modifikasi pada *K-Means* sebagai berikut:

1. Menggunakan ukuran pencocokan ketidakmiripan sederhana pada fitur data bertipe kategorikal
2. Mengamati mean cluster dengan modus (nilai yang paling sering muncul)

3. Menggunakan metode berbasis frekuensi untuk mencari modus dari sekumpulan nilai (Prasetyo 2014)

Langkah-langkah clustering dengan *K modes*:

1. Pilih *k* data sebagai inisialisasi centroid (modus), satu untuk setiap cluster.
2. Alokasikan data ke cluster dengan modulusnya terdekat menggunakan persamaan

$$d(X, Y) = \sum_{j=1}^r \epsilon(X_j, Y_j) \dots \dots \dots (1)$$

*r* adalah jumlah fitur, sedangkan  $\epsilon()$  adalah nilai pencocokan seperti pada persamaan berikut :

$$\in (x_j, y_j) = \begin{cases} 0 & (x_j = y_j) \\ 1 & (x_j \neq y_j) \end{cases} \dots\dots\dots(2)$$

- Perbarui modus (sebagai centroid) setiap cluster dengan nilai kategori yang sering muncul pada setiap cluster.

Ulangi langkah 2 dan 3 selama masih memenuhi syarat (1) masih ada data yang berpindah cluster atau (2) perubahan nilai fungsi objektif masih dibawah ambang batas yang ditentukan

**.3. Hasil dan Pembahasan**

Dari tabel 1. Data mahasiswa untuk dapat dikelompokkan dengan menggunakan metode K-Modes dengan Langkah-langkah yang dilakukan sebagai berikut

- Inisialisasi

Dipilih 2 data sebagai Modes awal secara acak. Data yang diambil adalah record ke 3 dan ke 10.

Tabel 2. Inisialisasi awal Centroid

| CENTROID | Jenis Kelamin | Asal Kota  | Pekerjaan Ayah | Pendidikan | Progdi |
|----------|---------------|------------|----------------|------------|--------|
| C1       | L             | PEKALONGAN | PNS            | SMA        | TI     |
| C2       | P             | PEMALANG   | WIRASWASTA     | SMK        | SI     |

- Iterasi ke 1

Hitung ketidakmiripan setiap data ke centroid (modus) dengan persamaan (1). Contoh perhitungan jarak ke setiap centroid pada data pertama:

$$d(x_i, c_i) = \in (x_{11}, c_{11}) + \in (x_{12}, c_{12}) + \in (x_{13}, c_{13}) + \in (x_{14}, c_{14}) + \in (x_{15}, c_{15})$$

$$= \in (l, l) + \in (Pemalang, Pekalongan) + \in (PNS, PNS) + \in (SMK, SMA) + \in (TI, TI)$$

$$= 0 + 1 + 0 + 1 + 0$$

$$= 2$$

$$d(x_i, c_2) = \in (x_{11}, c_{21}) + \in (x_{12}, c_{22}) + \in (x_{13}, c_{23}) + \in (x_{14}, c_{24}) + \in (x_{15}, c_{25})$$

$$d(x_i, c_2) = \in (x_{11}, c_{21}) + \in (x_{12}, c_{22}) + \in (x_{13}, c_{23}) + \in (x_{14}, c_{24}) + \in (x_{15}, c_{25})$$

$$= \in (l, p) + \in (Pemalang, Pemalang) + \in (PNS, Wiraswasta) + \in (SMK, SMK) + \in (TI, SI)$$

$$= 1 + 0 + 1 + 0 + 1$$

$$= 3$$

Tabel 3. Hasil perhitungan ketidakmiripan setaip data ke centroid dengan iterasi 1

| DATA KE-I | JARAK KE CENTROID |   | TERDEKAT | CLUSTER YANG DIKUTI |
|-----------|-------------------|---|----------|---------------------|
|           | 1                 | 2 |          |                     |
| 1         | 2                 | 3 | 2        | 1                   |
| 2         | 1                 | 3 | 1        | 1                   |

| DATA<br>KE-I | JARAK KE<br>CENTROID |   | TERDEKAT | CLUSTER<br>YANG<br>DIKUTI |
|--------------|----------------------|---|----------|---------------------------|
|              | 1                    | 2 |          |                           |
| 3            | 0                    | 4 | 0        | 1                         |
| 4            | 3                    | 1 | 1        | 2                         |
| 5            | 4                    | 3 | 3        | 2                         |
| 6            | 3                    | 2 | 2        | 2                         |
| 7            | 4                    | 2 | 2        | 2                         |
| 8            | 2                    | 4 | 2        | 1                         |
| 9            | 1                    | 4 | 1        | 1                         |
| 10           | 5                    | 0 | 0        | 2                         |
| 11           | 3                    | 2 | 2        | 2                         |
| 12           | 5                    | 3 | 3        | 2                         |
| 13           | 4                    | 3 | 3        | 2                         |
| 14           | 4                    | 1 | 1        | 2                         |
| 15           | 2                    | 3 | 2        | 1                         |
| 16           | 2                    | 3 | 2        | 1                         |
| 17           | 5                    | 1 | 1        | 2                         |
| 18           | 3                    | 4 | 3        | 1                         |

Selanjutnya dihitung modus yang baru untuk setiap cluster berdasarkan data yang bergabung pada setiap clusternya. Untuk cluster 1, ada 8 data dan untuk Tabel 4. Anggota C1

cluster 2 ada 10 data yang bergabung ke dalamnya dapat dilihat pada tabel 4 dan tabel 5:

| NO    | Jenis Kelamin | Asal Kota  | Pekerjaan Ayah | Pendidikan | progdi |
|-------|---------------|------------|----------------|------------|--------|
| 1     | L             | PEMALANG   | PNS            | SMK        | TI     |
| 2     | L             | PEMALANG   | PNS            | SMA        | TI     |
| 3     | L             | PEKALONGAN | PNS            | SMA        | TI     |
| 8     | L             | PEKALONGAN | SWASTA         | SMK        | TI     |
| 9     | L             | PEKALONGAN | SWASTA         | SMA        | TI     |
| 15    | L             | PEKALONGAN | GURU           | SMA        | SI     |
| 16    | P             | PEKALONGAN | PNS            | SMK        | TI     |
| 18    | L             | BATANG     | GURU           | SMA        | MI     |
| Modus | L             | PEKALONGAN | PNS            | SMA        | TI     |

Tabel 5. Anggota C2

| NO | Jenis Kelamin | Asal Kota  | Pekerjaan Ayah | Pendidikan | Progdi |
|----|---------------|------------|----------------|------------|--------|
| 4  | P             | PEMALANG   | WIRASWASTA     | SMA        | TI     |
| 5  | L             | BATANG     | SWASTA         | SMK        | SI     |
| 6  | P             | PEKALONGAN | SWASTA         | SMA        | SI     |

| NO    | Jenis Kelamin | Asal Kota  | Pekerjaan Ayah | Pendidikan | Progdi |
|-------|---------------|------------|----------------|------------|--------|
| 7     | P             | PEKALONGAN | SWASTA         | SMK        | SI     |
| 10    | P             | PEMALANG   | WIRASWASTA     | SMK        | SI     |
| 11    | P             | PEKALONGAN | PNS            | SMK        | SI     |
| 12    | P             | BATANG     | GURU           | SMK        | KA     |
| 13    | P             | BATANG     | SWASTA         | SMA        | KA     |
| 14    | P             | PEMALANG   | WIRASWASTA     | SMA        | MI     |
| 17    | P             | BATANG     | WIRASWASTA     | SMK        | SI     |
| Modus | P             | BATANG     | WIRASWASTA     | SMK        | SI     |

Modus yang didapatkan dari 2 cluster tersebut adalah centroid baru yang didapat

Tabel 6 Centorid baru dari modus yang didapatkan dari inisialisasi awal

Tabel 6. Centroid baru

| CENTROID | Jenis Kelamin | Asal Kota  | Pekerjaan Ayah | Pendidikan | Progdi |
|----------|---------------|------------|----------------|------------|--------|
| C1       | L             | PEKALONGAN | PNS            | SMA        | TI     |
| C2       | P             | BATANG     | WIRASWASTA     | SMK        | SI     |

### 3. Iterasi 2

Hitung ketidakmiripan setiap data ke centroid baru dengan persamaan (1).

Maka didapatkan seperti tabel 7 dibawah ini:

Tabel 7 Hasil perhitungan ketidakmiripan setaip data ke centroid dengan iterasi 2

| DATA<br>KE-I | JARAK KE CENTROID |   | TERDEKAT | CLUSTER<br>YANG<br>DIKUTI |
|--------------|-------------------|---|----------|---------------------------|
|              | 1                 | 2 |          |                           |
| 1            | 2                 | 4 | 2        | 1                         |
| 2            | 1                 | 5 | 1        | 1                         |
| 3            | 0                 | 5 | 0        | 1                         |
| 4            | 3                 | 3 | 3        | 2                         |
| 5            | 4                 | 2 | 2        | 2                         |
| 6            | 3                 | 3 | 3        | 2                         |
| 7            | 4                 | 2 | 2        | 2                         |
| 8            | 2                 | 4 | 2        | 1                         |
| 9            | 1                 | 5 | 1        | 1                         |
| 10           | 5                 | 1 | 1        | 2                         |
| 11           | 3                 | 2 | 2        | 2                         |
| 12           | 5                 | 2 | 2        | 2                         |
| 13           | 4                 | 3 | 3        | 2                         |
| 14           | 4                 | 3 | 3        | 2                         |

| DATA<br>KE-I | JARAK KE CENTROID |   | TERDEKAT | CLUSTER<br>YANG<br>DIIKUTI |  |
|--------------|-------------------|---|----------|----------------------------|--|
|              | 1                 | 2 |          |                            |  |
| 15           | 2                 | 4 | 2        | 1                          |  |
| 16           | 2                 | 3 | 2        | 1                          |  |
| 17           | 5                 | 0 | 0        | 2                          |  |
| 18           | 3                 | 4 | 3        | 1                          |  |

Dari hasil perhitungan ketidakmiripan setiap data dengan centroid pada iterasi 2

maka didapatkan untuk anggota C1 dan C2 adalah sebagai berikut:

Tabel 8 Anggota untuk C1

| NO    | Jenis Kelamin | Asal Kota  | Pekerjaan Ayah | Pendidikan | progdi |
|-------|---------------|------------|----------------|------------|--------|
| 1     | L             | PEMALANG   | PNS            | SMK        | TI     |
| 2     | L             | PEMALANG   | PNS            | SMA        | TI     |
| 3     | L             | PEKALONGAN | PNS            | SMA        | TI     |
| 8     | L             | PEKALONGAN | SWASTA         | SMK        | TI     |
| 9     | L             | PEKALONGAN | SWASTA         | SMA        | TI     |
| 15    | L             | PEKALONGAN | GURU           | SMA        | SI     |
| 16    | P             | PEKALONGAN | PNS            | SMK        | TI     |
| 18    | L             | BATANG     | GURU           | SMA        | MI     |
| Modus | L             | PEKALONGAN | PNS            | SMA        | TI     |

Tabel 9. Anggota untuk C2

| NO    | Jenis Kelamin | Asal Kota  | Pekerjaan Ayah | Pendidikan | Progdi |
|-------|---------------|------------|----------------|------------|--------|
| 4     | P             | PEMALANG   | WIRASWASTA     | SMA        | TI     |
| 5     | L             | BATANG     | SWASTA         | SMK        | SI     |
| 6     | P             | PEKALONGAN | SWASTA         | SMA        | SI     |
| 7     | P             | PEKALONGAN | SWASTA         | SMK        | SI     |
| 10    | P             | PEMALANG   | WIRASWASTA     | SMK        | SI     |
| 11    | P             | PEKALONGAN | PNS            | SMK        | SI     |
| 12    | P             | BATANG     | GURU           | SMK        | KA     |
| 13    | P             | BATANG     | SWASTA         | SMA        | KA     |
| 14    | P             | PEMALANG   | WIRASWASTA     | SMA        | MI     |
| 17    | P             | BATANG     | WIRASWASTA     | SMK        | SI     |
| Modus | P             | BATANG     | WIRASWASTA     | SMK        | SI     |

Karena sudah tidak ada perubahan untuk anggota yang ada di C1 dan C2 maka, dapat disimpulkan bahwa proses perhitungan dihentikan Dari hasil perhitungan dengan

menggunakan metode K modes untuk kelompok ke 1 dengan data yang sering muncul adalah Jenis kelamin =laki-laki, asal kota =Pekalongan, Pekerjaan ayah =

PNS, Pendidikan = SMA program studi TI, sedangkan untuk kelompok ke 2 dengan data yang sering muncul jenis kelamin = Perempuan, asal kota = Batang, pekerjaan ayah = wiraswasta, pendidikan = SMK, program studi = Sistem Informasi.

#### 4. Kesimpulan dan Saran

##### 4.1 Kesimpulan

*K Modes* dapat digunakan sebagai pemodelan *machine learning* untuk penentuan segmentasi mahasiswa baru agar dapat menjangkau mahasiswa baru. Metode *K mode* dapat digunakan sebagai pemodelan untuk memecahkan kasus segmentasi mahasiswa baru dengan variabel yang berupa kategorikal. Untuk program studi Teknik informatika di dominasi berjenis kelamin Laki-laki, asal kota Pekalongan, pekerjaan ayah PNS dan asal sekolah SMA. Sedangkan untuk program studi Sistem Informasi di dominasi jenis kelamin perempuan, asal kota Batang, pekerjaan ayah wiraswasta, asal sekolah SMK.

##### 4.2 Saran

Sistem clustering untuk penentuan strategi marketing dengan menggunakan metode *K modes* bisa dilakukan dengan menggunakan metode clustering yang lain untuk sebagai pembanding. Untuk variabel yang di olah hanya 5 variabel, yaitu jenis kelamin, asal kota, pekerjaan ayah, asal sekolah dan program studi. dapat Ditambahkan variabel yang lain, karena variabel yang ada di pendaftaran sejumlah 109 variabel. Dengan keterbatasan metode *K Modes* dengan variabel yang bersifat kategorikal, maka dapat dibuat dengan metode *K Means* yang dapat menampung data dengan fitur nominal atau ordinal

#### DAFTAR PUSTAKA

- Abriyanto, Arif, and Natalia Damastuti. 2019. "Segmentasi Mahasiswa dengan Unsupervised Algoritma Guna Membangun Strategi Marketing Penerimaan Mahasiswa." *Insand Comtech* 10-18.
- Aprilia, Yolanda Putri Kartikasari, Yosep Pranoto Agus, and Deddy Rudhistiar. 2021. "Penerapan Metode K-Modes untuk Proses Penentuan Penerimaan Bantuan Langsung Tunai (BLT)." *JATI* 389-397.
- Asroni, Hidayatul Fitri, and Eko Prasetyo. 2018. "Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokan data Calon Mahasiswa Baru di Universitas Muhammadiyah Yogyakarta (Studi Kasus: Fakultas Kedokteran dan Ilmu Kesehatan, dan Fakultas Ilmu Sosial dan Ilmu Politik)." *Semesta Teknik* 60-64.
- Astri, Alyeska Az-zahra, Almira Marsaoly Fajriyati, Intan Lestyani Putri, Roghibah Salsabila, and Wa Ode Madjida Zuhayeni. 2021. "Penerapan Algoritma K Modes Clustering dengan Validasi Davies Bouldin Index pada Pengelompokan tingkat Minat Balanja Onlien di Provisi Daerah Istimewa Yogyakarta ." *Jurnal Matematika dan Statistika* 24-36.
- Chasanah, Tria Titiani, and Widiyono. 2017. "penentuan strategi promosi penerimaan mahasiswa baru dengan algoritma clustering k -means." *Ic-Tech* 39-44.

- Gupitha, Rino. 2018. "Penentuan Strategi Marketing Sekolah Menengah Kejuruan Terpadu Lampang Subang Menggunakan Metode K means Clustering." *Global* 17-24.
- Indriani, Fatma, and Irawan Budiman. 2017. "K Modes Clustering untuk Mengetahui Jenis Masakan yang Populer Pada Website Resep Online (Studi Kasus Banjar di Cookpad.Com)." *JTIK* 290-296.
- Malikhatin, Hanik, Agus Rusgiyono, and Di Maruddani Asih. 2021. "Penerapan K Modes Clustering dengan Validasi Dunn Index Pada Pengelompokan Karakteristik Calon TKI Menggunakan R-GUI." *Gaussian* 359-366.
- Nurhayati, Busman, and Rayi iswara Pradono. 2019. "Pengembangan Algoritma Unsupervised Learning Technique Pada Big Data Analysis di Media Sosial Sebagai Media Promosi Online Bagi Masyarakat." *Teknik Informatika* 79-96.
- Prasetyo, Eko. 2014. *Data Mining Mengolah Data Menjadi Informasi menggunakan Matlab*. Yogyakarta: Andi.
- Primartha, Rifkie. 2021. *Algoritma Machine Learning*. Bandung: Informatika Bandung.
- . 2018. *Belajar Machine Learning Teori dan Praktek*. Bandung: Informatika Bandung.
- Ramdhani, Mokhammad Raharjo, and Agus Windarto Perdana. 2021. "Penerapan Machine Learning dengan Konsep Data Mining Rough Set (Prediksi Tingkat Pemahaman Mahasiswa terhadap Matakuliah)." *Media Informatika Budidarma* 317-326.
- Selva, Fithri Jumeilah, and Dicky Pratama. 2017. "Klasterisasi Penduduk Lanjut Usia Sumatera Selatan Menggunakan Algoritma K Modes." *Technology Acceptance Model* 85-889.
- Yulianto, Heribertus, Felix Sutanto Andreas, and sri Mulyani. 2021. "Pengelompokan Mahasiswa Berbasis Categorical Variabels Menggunakan Metode K-Modes Clustering." *Sendiu* 424-429.
- Yuyun, Baiq Hairani, Moh Juhad, and Sam'an Majid Abdul. 2022. "Strategi Perguruan Tinggi Dalam Menjaring Mahasiswa Baru di Era Pandemi Tahun akademik 2020/2021 (Studi di STIA Muhammadiyah Selong)." *Mentari Publika* 195-204.