

SEGMENTASI PELANGGAN AGEN *BEAUTY LOCAL BRAND* MENGGUNAKAN ALGORITMA K-MODES BERBASIS PYTHON

Mochammad Ainul Yaqin¹⁾, Nisa Ayunda^{2)*}, Sujarwo³⁾, dan M. Ali Murtadho⁴⁾

^{1, 2, 3)}Program Studi Matematika, Universitas Pesantren Tinggi Darul 'Ulum

⁴⁾Program Studi Sistem Informasi, Universitas Pesantren Tinggi Darul 'Ulum
alamat institusi

e-mail: ayaqin115@gmail.com¹⁾, nisaayunda@mipa.unipdu.ac.id²⁾, jarwo@mipa.unipdu.ac.id³⁾,
alimurtadho@ft.unipdu.ac.id⁴⁾

ABSTRAK

Pertumbuhan *beauty local brand* di Indonesia kini mengalami pertumbuhan yang pesat. Salah satu *brand local* kecantikan yang telah berkembang, yaitu *brand MS Glow*, menyusun strategi pemasaran dengan sistem stokist. Sistem stokist dijalankan dengan memperbanyak distributor, agen, reseller, serta memaksimalkan promosi melalui iklan. Ketidaktepatan stok produk dan penurunan interaksi konten di media sosial merupakan permasalahan inti yang dihadapi oleh distributor, agen, dan juga reseller *brand* kecantikan lokal. Permasalahan tersebut dapat diatasi apabila manajemen mampu mengidentifikasi karakteristik pelanggannya. Oleh karena itu, pada penelitian ini dilakukan segmentasi pelanggan salah satu agen *brand* kecantikan lokal, yaitu Agen *Ami Ms Glow Peterongan*, dengan Algoritma K-Modes. Algoritma ini dapat digunakan untuk menangani data beratribut kategorik. Hal ini sesuai dengan data pelanggan yang memiliki tipe atribut kategorik. Hasil dari penelitian ini yaitu pelanggan agen tersebut terbagi menjadi 3 segmen pelanggan yang didominasi oleh segmen pertama, dengan karakteristik pelanggan: berusia 26 – 35 tahun; tidak memakai KB dan tidak sedang hamil; berjenis kulit berminyak; tidak berjerawat di luar menstruasi; tidak sedang berjerawat; memiliki bekas jerawat dan komedo; tidak berflek hitam; serta tidak mengalami kulit yang terkadang merah dan gatal, berjumlah 193 pelanggan, segmen kedua berjumlah 76 pelanggan, dan segmen ketiga berjumlah 98 pelanggan. Penentuan jumlah stok dan target konten untuk pemasaran dapat menyesuaikan hasil segmentasi pelanggan tersebut.

Kata Kunci: Segmentasi Pelanggan, K-Modes, Python, *Beauty Local Brand*.

ABSTRACT

The growth of local beauty brands in Indonesia is now experiencing rapid growth. One of the local beauty brands that has developed, namely the *MS Glow* brand, has developed a marketing strategy with a stockist system. The stockist system is run by increasing distributors, agents, resellers, and maximizing promotions through advertising. The inaccuracy of product stock and the decrease in content interaction on social media are the core problems faced by distributors, agents, and also resellers of local beauty brands. These problems can be overcome if management is able to identify the characteristics of its customers. Therefore, in this study, customer segmentation was carried out by one of the local beauty brand agents, namely *Agent Ami Ms Glow Peterongan*, with the K-Modes Algorithm. This algorithm can be used to handle categorical attribute data. This is in accordance with customer data that has a categorical attribute type. The results of this study are that the agent's customers are divided into 3 customer segments which are dominated by the first segment, with customer characteristics: aged 26-35 years; not using family planning and not pregnant; oily skin type; no acne outside of menstruation; not having acne; have acne scars and blackheads; no black spots; and do not experience skin that is sometimes red and itchy, totaling 193 customers, the second segment totaling 76 customers, and the third segment totaling 98 customers. Determining the amount of stock and target content for marketing can adjust the results of the customer segmentation.

Keywords: Customer Segmentation, K-Modes, Python, *Beauty Local Brand*.

I. PENDAHULUAN

Dewasa ini, industri kecantikan di Indonesia mengalami perkembangan cukup pesat. Berdasarkan data yang dirilis oleh Nielsen dan Euromonitor, pertumbuhan penjualan produk kecantikan di Indonesia pada tahun 2017 adalah sebesar 11,99%, yang memberikan kontribusi sebesar USD 1,4 miliar atau setara Rp 19 triliun, nilai ini diperkirakan

akan terus meningkat setiap tahunnya [20]. Hal ini tak lepas juga dari banyaknya *brand* kecantikan yang berkembang di Indonesia. Salah satunya yaitu *MS Glow*, yang merupakan *brand* kecantikan dan *skincare* lokal yang berdiri sejak tahun 2013. Pada bulan Agustus 2021, *MS Glow* meraih rekor MURI sebagai perusahaan kosmetik dengan jaringan penjualan terbanyak di Indonesia, yakni sebanyak 78.147 jaringan penjualan [22]. Dalam jaringan penjualannya, *MS Glow*

dibagi ke dalam 4 tingkatan level, yakni Reseller, Member, Agen, dan Distributor.

Salah satu agen MS Glow yang berada di Kabupaten Jombang yaitu Ami MS Glow Peterongan, yang merupakan salah satu agen resmi dari MS Glow Store yang terletak di Jl. Brawijaya No. 86, Peterongan, Kecamatan Peterongan, Kabupaten Jombang. Agen Ami MS Glow Peterongan sudah berkembang cukup pesat. Hal ini dapat dilihat dari omset Agen Ami MS Glow Peterongan yang mencapai ratusan juta per bulan dengan jumlah pelanggan yang sudah mencapai ribuan pelanggan. Meskipun begitu, Ami MS Glow Peterongan memiliki beberapa permasalahan mendasar pada kegiatan usahanya. Berdasarkan wawancara peneliti dengan pemilik Agen Ami MS Glow Peterongan, terdapat dua permasalahan yang sedang dihadapi oleh manajemen agen. Yang pertama yaitu ketidaktepatan dalam melakukan penyetokan produk. Indikasinya yaitu terdapat beberapa jenis barang yang sepi peminatnya, seperti paket MS Kids, MS Slim, *body soap*, dan lain-lain, sehingga membuat barang menjadi lama tidak terjual dan bahkan ada yang tidak terjual. Masalah yang kedua yaitu tingkat interaksi beberapa konten di media sosial menurun. Manajemen Agen Ami MS Glow Peterongan memanfaatkan media sosial untuk menjalankan strategi pemasaran, yaitu dengan cara membuat konten informatif seputar dunia kecantikan. Tetapi, pada beberapa waktu terakhir, terdapat beberapa konten yang memiliki tingkat interaksi yang kecil. Kedua permasalahan Ami MS Glow Peterongan tersebut dapat diatasi apabila manajemen agen mampu mengidentifikasi serta memahami tipe dan karakteristik pelanggannya. Apabila karakteristik pelanggan dipahami dengan baik, maka manajemen Agen Ami MS Glow Peterongan dapat menyetok barang sesuai dengan kebutuhan pelanggan serta konten informatif di media sosial akan lebih tepat sasaran.

Masalah pemahaman karakteristik pelanggan dapat diselesaikan menggunakan metode segmentasi pelanggan. Segmentasi pelanggan adalah upaya pengelompokan pelanggan berdasarkan karakteristik yang sama antar pelanggan [14]. Metode yang umumnya digunakan untuk menyelesaikan masalah segmentasi yaitu metode *K-Means*, tetapi metode ini hanya bekerja pada *dataset* dengan tipe atribut numerik [13]. Pada penelitian ini, data yang diolah adalah data isian formulir konsultasi keluhan yang dialami oleh pelanggan, yang mana *dataset* yang diolah tersebut atributnya bertipe kategorik, sehingga dibutuhkan metode yang dapat mengatasi data bertipe kategorik. Algoritma *K-Modes* adalah metode yang dikembangkan dari algoritma *K-Means* yang dapat diterapkan pada data bertipe kategorik. Kelebihan dari algoritma *K-Modes* yaitu dapat menghasilkan klaster yang lebih stabil dengan waktu komputasi yang lebih singkat dibandingkan dengan algoritma *K-Means* [15]. Penelitian ini bertujuan untuk melakukan segmentasi

pelanggan Agen Ami MS Glow Peterongan menggunakan Algoritma *K-Modes*. Hasil dari penelitian ini diharapkan dapat membantu manajemen Agen Ami Ms Glow Peterongan dalam mengenali karakteristik pelanggannya, sehingga nantinya akan berguna dalam penerapan strategi pemasaran dan perencanaan stok barang Agen Ami MS Glow Peterongan.

II. PANDUAN UNTUK MEMPERSIAPKAN ARTIKEL

A. Segmentasi Pelanggan

Segmentasi pelanggan adalah usaha untuk mengelompokkan pelanggan berdasarkan kebutuhan [19]. Tujuannya yaitu memberikan produk tunggal tertentu kepada sekumpulan pelanggan yang telah diidentifikasi memiliki kebutuhan tertentu, sehingga strategi pemasaran perusahaan dapat diterapkan dengan lebih efektif dan ekonomis. Asumsi yang mendasari segmentasi pelanggan yaitu sebagai berikut [4]:

- Tidak semua karakteristik konsumen mirip satu sama lain.
- Subkelompok konsumen dengan latar belakang, perilaku, nilai, dan kebutuhan sama dapat dikenali.
- Subkelompok lebih kecil dan lebih homogen dibandingkan pasar secara keseluruhan.
- Lebih mudah memuaskan sekumpulan kecil konsumen yang sama dibanding berupaya memuaskan kumpulan besar konsumen yang beragam.

Pentingnya melakukan segmentasi pelanggan adalah sebagai berikut [9]:

- Memungkinkan perusahaan kecil untuk menciptakan ceruk pasar dan mempertahankannya karena mampu menyelaraskan kompetensinya dengan sasaran pasar.
- Mampu mengidentifikasi celah atau gap di pasar, yaitu segmen yang selama ini belum atau tidak terlayani secara baik.
- Mampu mengidentifikasi segmen yang sudah menurun di pasar.
- Memungkinkan perusahaan untuk menyesuaikan produk atau jasanya dengan kebutuhan pasar.
- Perusahaan akan kehilangan keunggulan kompetitif apabila tidak melakukan segmentasi.

Pada intinya, segmentasi pasar adalah proses pengelompokan pasar sebuah produk atau jasa yang heterogen ke dalam beberapa segmen, yang mana tiap segmen cenderung mempunyai kesamaan dalam hal kebutuhan, keinginan, perilaku, serta respon terhadap program pemasaran [21].

B. Clustering

Clustering atau klasterisasi adalah salah satu metode pada data mining yang tujuannya untuk mengelompokkan data atau objek ke dalam kelompok-kelompok. *Cluster* adalah kumpulan objek data yang

memiliki kesamaan karakteristik antar data dalam satu *cluster* dan memiliki ketidaksamaan karakteristik terhadap objek data yang berbeda *cluster* [17]. *Clustering* merupakan pembelajaran yang tidak terawasi dimana tidak memerlukan target luaran atau yang sering disebut *unsupervised learning* [16].

Terdapat dua metode dalam *Clustering*, yaitu metode hirarki dan metode nonhirarki. Metode hirarki adalah metode yang dimulai dengan mengelompokkan beberapa objek yang memiliki kesamaan paling dekat. Kemudian diteruskan ke objek lain yang memiliki kedekatan lainnya. Sehingga, *cluster* akan membentuk hirarki dari yang paling mirip sampai yang tidak mirip. Sedangkan metode nonhirarki dimulai dengan menentukan terlebih dahulu jumlah *cluster* baru dilakukan proses *clustering*. Metode hirarki digunakan apabila jumlah klaster belum diketahui, sedangkan pada metode nonhirarki bertujuan mengelompokkan objek ke dalam k buah *cluster* [7].

C. Analisis Cluster

Analisis klaster adalah kumpulan dari beberapa teknik pengolahan data multivariabel yang bertujuan untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya [6]. Hasil dari klaster yang terbentuk harus menunjukkan kesamaan internal yang tinggi dalam satu klaster dan keragaman yang tinggi antar klaster. Prinsip dari analisis klaster adalah proses untuk mereduksi sekumpulan objek besar menjadi lebih kecil yang disebut klaster [5].

D. Algoritma K-Modes

Algoritma *K-Modes* diperkenalkan oleh Huang pada tahun 1997. Algoritma ini merupakan hasil dari modifikasi algoritma *clustering K-Means* yang mana hanya bisa digunakan untuk mengelompokkan data numerik, sedangkan algoritma *K-Modes* dapat digunakan untuk mengelompokkan data kategorik [10].

Data dengan karakteristik yang sama akan tergabung ke dalam satu kelompok dan yang berbeda akan tergabung ke dalam kelompok yang lain. Tujuan algoritma ini untuk meminimalkan variasi dalam satu kelompok dan memaksimalkan variasi antar kelompok. Kelebihan dari algoritma *K-Modes* yaitu dapat menghasilkan klaster yang lebih stabil dengan waktu komputasi yang lebih singkat dibandingkan dengan algoritma *K-Means* [15]. Karena merupakan pengembangan dari algoritma *K-Means*, algoritma ini termasuk ke dalam metode *clustering* nonhirarki.

Beberapa modifikasi yang dilakukan terhadap algoritma *K-Means* adalah sebagai berikut [16]:

- Menggunakan ukuran pencocokan ketidaksamaan sederhana dari fitur data berjenis kategorikal.
- Menggunakan modus (nilai yang paling sering muncul) sebagai ganti mean klaster.

- Untuk mencari modus dari sekumpulan nilai, digunakan metode berbasis frekuensi.

Misal x dan y adalah dua data bertipe kategorik, x sebagai pusat *cluster* dan y sebagai data yang akan dibandingkan dengan pusat *cluster*. Ukuran ketidaksamaan antara x dan y bisa diukur dengan jumlah ketidakcocokan nilai dari fitur yang berkorespondensi dari dua data. Data tersebut semakin mirip jika semakin kecil nilai ketidakcocokannya. Metrik seperti ini biasa disebut dengan pencocokan sederhana (*simple matching*). Rumus untuk menghitung jarak dalam *K-Modes* adalah sebagai berikut [23],

$$d(X, Y) = \sum_{j=1}^r \in (x_j, y_j) \quad \dots\dots\dots (1)$$

Di mana :

$d(x, y)$ = Jarak data x ke y

x_j = Nilai fitur ke- j dari x

y_j = Nilai fitur ke- j dari y

r = Jumlah fitur

Dengan :

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

Langkah-langkah dalam melakukan klasterisasi dengan algoritma *K-Modes* sebagai berikut [16]:

- Memilih k modus awal sebagai titik pusat, satu untuk setiap klaster.
- Menghitung jarak masing-masing data terhadap semua titik pusat klaster. Alokasikan setiap objek ke klaster terdekat menggunakan persamaan (1).
- Memperbarui modus (sebagai pusat klaster baru) dari setiap klaster dengan nilai kategori yang sering muncul pada setiap klaster.
- Mengulangi langkah b dan c apabila tidak terdapat data yang berpindah klaster atau posisi pusat klaster tidak berubah.

Gambar diagram alir untuk algoritma *K-Modes* dapat dilihat pada Gambar 1 [15].

E. Python

Python merupakan salah satu bahasa pemrograman terpopuler di dunia [12]. Python banyak digunakan dalam bidang analisis data, data mining, *Machine Learning* dan kecerdasan buatan. Popularitas Python ini salah satunya adalah tersedianya banyak *library* (modul siap pakai) yang mendukung kegiatan analisis data dan *Machine Learning*, seperti Matplotlib, Numpy, Pandas, Scikit-learn, dan *library* lainnya.



Gambar 1. Diagram alir algoritma K-Modes

III. METODE PENELITIAN

A. Tahapan Penelitian

Adapun tahapan-tahapan yang dilaksanakan dalam penelitian ini yaitu sebagai berikut.

1) Identifikasi Masalah

Pada tahap ini dilakukan analisis terhadap permasalahan yang akan dijadikan topik penelitian ini. Penelitian ini mengambil studi kasus pada Agen Ami MS Glow Peterongan dengan permasalahan yang diangkat yaitu segmentasi pelanggan Agen Ami MS Glow Peterongan berdasarkan keluhan yang dialami pelanggan.

2) Studi Literatur

Pada tahap ini dilakukan penggalian informasi yang dapat dijadikan referensi dalam pelaksanaan penelitian. Sumber informasi didapatkan dari narasumber atau pemilik Agen Ami Ms Glow Peterongan, buku, buku elektronik, penelitian terdahulu seperti pada *paper* dan jurnal, tugas akhir, artikel atau dokumen lainnya yang terkait dengan segmentasi pelanggan. Referensi yang diperlukan yaitu yang berhubungan dengan segmentasi pelanggan, *clustering*, algoritma K-Modes dan penerapannya dengan bahasa Python.

3) Pengumpulan Data

Tahap selanjutnya yaitu mengumpulkan data-data yang dibutuhkan sebagai pendukung utama dalam pelaksanaan penelitian ini. Jenis data yang digunakan dalam penelitian ini yaitu data primer,

berupa data isian formulir konsultasi keluhan pelanggan Agen Ami MS Glow Peterongan yang diambil dalam rentang waktu 1 Januari hingga 28 Februari 2022. Sumber data didapatkan dari manajemen agen Ami Ms Glow Peterongan. Instrumen penelitian yang digunakan adalah formulir konsultasi keluhan yang diisi oleh pelanggan sebelum melakukan pembelian produk untuk yang pertama kalinya di Agen Ami MS Glow Peterongan.

4) Data Cleaning

Pada tahap ini, dilakukan pembersihan terhadap data, yakni menghilangkan data yang tidak lengkap serta beberapa kolom yang tidak diperlukan untuk proses olah data. Sehingga *dataset* hanya akan berisi atribut yang dibutuhkan untuk proses olah data.

5) Transformasi Data

Sebelum melakukan proses *clustering*, terlebih dahulu harus melakukan transformasi data. Dari *dataset* keluhan pelanggan, data atribut usia pelanggan adalah data yang perlu dilakukan transformasi, yakni mengubah data usia ke dalam kategori-kategori usia.

6) Clustering

Pada tahap ini data sudah siap digunakan untuk proses pengolahan data, yaitu proses *clustering* atau segmentasi. Pada penelitian ini, algoritma *clustering* yang digunakan adalah algoritma K-Modes. Proses implementasi algoritma K-Modes menggunakan bahasa pemrograman Python yang dijalankan di *platform* Google Colaboratory Notebook.

7) Interpretasi

Tahap selanjutnya yaitu melakukan analisis terhadap pola yang dihasilkan dari proses *clustering* menggunakan algoritma K-Modes.

8) Hasil & Kesimpulan

Tahap terakhir pada penelitian ini yaitu mendapatkan hasil segmentasi pelanggan dan penarikan kesimpulan dari kesemua proses yang telah dilaksanakan dalam penelitian dari awal hingga akhir.

IV. HASIL DAN PEMBAHASAN

Data yang diperoleh pada proses pengumpulan data adalah sebanyak 402 *record* data. Berupa data primer yang didapat dari isian formulir konsultasi pelanggan Ami Ms Glow Peterongan, sehingga perlu dilakukan penyiapan terlebih dahulu sebelum dapat digunakan pada proses olah data.

A. Data Cleaning

Pada proses ini dilakukan pembersihan data dari data yang tidak lengkap serta dari atribut yang tidak memenuhi syarat untuk pengolahan data. Secara keseluruhan, atribut data pada formulir konsultasi pelanggan Agen Ami MS Glow Peterongan bertipe campuran antara kategorik dan numerik. Jadi, sebelum

diimplementasikan pada algoritma K-Modes perlu dilakukan penyesuaian pada dataset agar dataset hanya terdiri dari atribut data kategorik saja karena algoritma K-Modes hanya bekerja pada data kategorik [18]. Setelah dilakukan penyesuaian, atribut data yang digunakan untuk proses pengolahan data yaitu sebagai berikut: usia, berjerawat di luar siklus menstruasi, ada tidaknya jerawat, bekas jerawat, flek hitam, komedo, jenis kulit, sedang hamil atau tidak, kulit terkadang merah dan gatal atau tidak, dan menjalankan program KB atau tidak. Setelah melalui proses pembersihan data, dari 402 record data hanya digunakan sebanyak 367 record data dalam proses olah data.

B. Transformasi Data

Dari hasil pembersihan data/*Data Cleaning*, semua atribut dataset bernilai kategorik, tetapi ada satu atribut data yang tidak bernilai kategorik, yaitu atribut Usia yang menghasilkan data bernilai numerik. Maka perlu dilakukan transformasi data usia menjadi data kategorik agar dapat diimplementasikan pada algoritma *K-Modes*. Caranya yaitu dengan mengelompokkan data usia ke dalam beberapa kelompok atau kategori usia. Kategori usia menurut Departemen Kesehatan RI tahun 2009 adalah sebagai berikut [1]:

- Masa balita = 0 – 5 tahun.
- Masa kanak-kanak = 6 – 11 tahun.
- Masa remaja awal = 12 – 16 tahun.
- Masa remaja akhir = 17 – 25 tahun.
- Masa dewasa awal = 26 – 35 tahun.
- Masa dewasa akhir = 36 – 45 tahun.
- Masa lansia awal = 46 – 55 tahun.
- Masa lansia akhir = 56 – 65 tahun.
- Masa manula = 65 tahun ke atas

Gambar 2 menampilkan dataset yang telah melewati proses pembersihan data dan transformasi data pada atribut usia.

	jerawat.diluar.mens	sedang.berjerawat	bekas.jerawat	flek.hitam	komedo	sedang.hamil	jenis.kulit	klt.kdang.merah.n.gatal	kb	usia_bin
0	tidak	tidak	ada	tidak ada	ada	tidak	berminyak	iya	tidak	36-45
1	tidak	tidak	tidak ada	ada	tidak ada	tidak	kering	iya	tidak	46-55
2	tidak	tidak	ada	ada	ada	tidak	normal	iya	tidak	36-45
3	iya	tidak	ada	ada	ada	tidak	berminyak	tidak	iya	36-45
4	tidak	tidak	ada	ada	ada	tidak	kombinasi	tidak	iya	36-45
...	...	...	...	...	...	...	...	...	...	...
362	tidak	tidak	tidak ada	ada	tidak ada	tidak	kombinasi	iya	iya	26-35
363	tidak	iya	tidak ada	tidak ada	ada	iya	normal	tidak	tidak	26-35
364	tidak	iya	tidak ada	tidak ada	ada	iya	normal	tidak	tidak	26-35
365	iya	tidak	tidak ada	ada	ada	iya	normal	tidak	tidak	26-35
366	tidak	iya	ada	ada	tidak ada	tidak	berminyak	tidak	iya	26-35

367 rows x 10 columns

Gambar 2. Dataset dengan pengelompokan atribut usia

C. Clustering

Pada tahap ini mulai dilakukan proses implementasi algoritma K-Modes pada dataset. Proses implementasi algoritma dilakukan menggunakan bahasa Python yang dijalankan di platform Google Colaboratory Notebook. Dalam algoritma K-Modes, penentuan titik awal dapat dilakukan secara acak,

namun hal itu dapat mengakibatkan iterasi yang tidak terprediksi jumlah dan akurasi. Pada penelitian ini, penentuan titik awal menggunakan metode Cao yang mana metode ini pusat kluster yang dihasilkan akan berdampak secara langsung terhadap informasi kluster di akhir iterasi [24]. Metode perbandingan nilai cost setiap jumlah kluster digunakan untuk menentukan jumlah kluster yang optimal. Nilai k yang optimal terletak pada titik ketika nilai cost k tidak mengalami penurunan signifikan dengan penambahan nilai k [11]. Untuk perbandingannya digunakan k mulai dari 1 sampai 6. Hasil perhitungan cost k seperti pada gambar 3 dan didapatkan hasil kluster optimal pada k = 3.

```

Init: initializing centroids
Init: initializing clusters
Starting iterations...
Run 1, iteration: 1/100, moves: 0, cost: 1455.0
Init: initializing centroids
Init: initializing clusters
Starting iterations...
Run 1, iteration: 1/100, moves: 43, cost: 1213.0
Run 1, iteration: 2/100, moves: 54, cost: 1185.0
Run 1, iteration: 3/100, moves: 6, cost: 1185.0
Init: initializing centroids
Init: initializing clusters
Starting iterations...
Run 1, iteration: 1/100, moves: 90, cost: 1067.0
Run 1, iteration: 2/100, moves: 30, cost: 1043.0
Run 1, iteration: 3/100, moves: 7, cost: 1043.0
Init: initializing centroids
Init: initializing clusters
Starting iterations...
Run 1, iteration: 1/100, moves: 82, cost: 1007.0
Run 1, iteration: 2/100, moves: 36, cost: 985.0
Run 1, iteration: 3/100, moves: 6, cost: 985.0
Init: initializing centroids
Init: initializing clusters
Starting iterations...
Run 1, iteration: 1/100, moves: 93, cost: 934.0
Run 1, iteration: 2/100, moves: 14, cost: 934.0
Init: initializing centroids
Init: initializing clusters
Starting iterations...
Run 1, iteration: 1/100, moves: 81, cost: 899.0
Run 1, iteration: 2/100, moves: 45, cost: 867.0
Run 1, iteration: 3/100, moves: 12, cost: 867.0

```

Gambar 3 Perhitungan cost k

Setelah menentukan metode inisialisasi dan jumlah kluster yang optimal, tahap selanjutnya yaitu melakukan prediksi pengelompokan pelanggan. Proses pengelompokan dapat dilihat pada gambar 4. Pada gambar 4 menampilkan *syntax* untuk melakukan proses *clustering*/pengelompokan dengan menggunakan k = 3. Hasil keluarannya yaitu berupa array kluster setiap record data seperti pada gambar 5.

```

[22] #Choosing K=3
km_cao = KModes(n_clusters=3, init = "Cao", n_init = 1, verbose=1)
fitClusters_cao = km_cao.fit_predict(ami)

Init: initializing centroids
Init: initializing clusters
Starting iterations...
Run 1, iteration: 1/100, moves: 90, cost: 1067.0
Run 1, iteration: 2/100, moves: 30, cost: 1043.0
Run 1, iteration: 3/100, moves: 7, cost: 1043.0

```

Gambar 4 Proses *Clustering* dengan k = 3

Informasi dari masing-masing kluster/segmen dapat dimunculkan dengan perintah '`cluster_n.info()`'. Gambar 6 adalah keluaran informasi dari kluster/segmen pertama. Dari 367 record data, kluster/segmen pertama memiliki anggota sebanyak 193 anggota. Gambar 7 adalah keluaran informasi dari

klaster/segmen kedua yang memiliki anggota sebanyak 76 anggota. Gambar 8 adalah keluaran informasi dari klaster/segmen ketiga yang memiliki anggota sebanyak 98 anggota.

```
[23] fitClusters_cao
```

```
array([[0, 1, 1, 0, 0, 0, 0, 2, 1, 0, 0, 2, 0, 2, 0, 2, 2, 0, 0, 0, 2, 0,
        2, 1, 0, 0, 1, 0, 2, 0, 0, 0, 1, 0, 2, 2, 0, 1, 2, 0, 2, 0, 2, 1,
        2, 0, 0, 1, 2, 2, 0, 2, 0, 0, 0, 0, 2, 2, 0, 0, 0, 0, 0, 1, 0, 0,
        1, 0, 0, 0, 2, 2, 0, 0, 2, 0, 0, 0, 2, 0, 2, 0, 2, 0, 0, 2, 2,
        0, 2, 2, 0, 0, 2, 0, 0, 0, 1, 0, 2, 0, 2, 1, 2, 1, 0, 2, 2, 1, 0,
        0, 0, 2, 2, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 2, 0, 0, 1, 1, 0,
        1, 0, 1, 0, 1, 0, 0, 2, 0, 0, 0, 2, 0, 2, 0, 1, 1, 0, 0, 0, 0, 0,
        2, 0, 0, 0, 0, 0, 1, 2, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1,
        1, 0, 2, 1, 1, 1, 2, 0, 2, 2, 1, 0, 0, 1, 0, 2, 0, 0, 0, 2, 2, 0,
        2, 0, 2, 0, 0, 1, 0, 1, 2, 0, 0, 1, 2, 0, 0, 0, 2, 2, 2, 0, 0,
        0, 2, 0, 2, 0, 2, 0, 2, 0, 2, 2, 0, 1, 0, 2, 0, 0, 0, 0, 1, 1,
        1, 0, 1, 1, 0, 0, 0, 0, 2, 2, 0, 0, 0, 1, 2, 2, 0, 0, 0, 0, 0,
        2, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0,
        0, 1, 1, 0, 1, 0, 1, 0, 2, 2, 2, 1, 1, 2, 0, 2, 2, 2, 1, 0, 0,
        2, 2, 1, 1, 2, 2, 0, 2, 0, 2, 0, 2, 1, 1, 2, 2, 2, 2, 0, 1, 0, 0,
        2, 2, 0, 2, 0, 2, 2, 1, 0, 1, 1, 1, 1, 2, 1, 1, 0, 0, 0, 1,
        0, 0, 0, 0, 0, 2, 0, 1, 1, 1, 1, 0, 0, 0, 0], dtype=uint16)
```

Gambar 5. Hasil keluaran proses *clustering*

```
[28] cluster_0.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 193 entries, 0 to 366
Data columns (total 11 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                -
0   jerawat.diluar.mens                  193 non-null    object
1   sedang.berjerawat                  193 non-null    object
2   bekas.jerawat                      193 non-null    object
3   flek.hitam                         193 non-null    object
4   komedo                             193 non-null    object
5   sedang.hamil                       193 non-null    object
6   jenis.kulit                       193 non-null    object
7   klt.kdang.merah.n.gatal            193 non-null    object
8   kb                                  193 non-null    object
9   usia_bin                           193 non-null    category
10  cluster_predicted                  193 non-null    uint16
dtypes: category(1), object(9), uint16(1)
memory usage: 15.9+ KB
```

Gambar 6 Informasi segmen pertama

```
[29] cluster_1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 76 entries, 1 to 362
Data columns (total 11 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                -
0   jerawat.diluar.mens                  76 non-null    object
1   sedang.berjerawat                  76 non-null    object
2   bekas.jerawat                      76 non-null    object
3   flek.hitam                         76 non-null    object
4   komedo                             76 non-null    object
5   sedang.hamil                       76 non-null    object
6   jenis.kulit                       76 non-null    object
7   klt.kdang.merah.n.gatal            76 non-null    object
8   kb                                  76 non-null    object
9   usia_bin                           76 non-null    category
10  cluster_predicted                  76 non-null    uint16
dtypes: category(1), object(9), uint16(1)
memory usage: 6.4+ KB
```

Gambar 7 Informasi segmen kedua

```
[30] cluster_2.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 98 entries, 7 to 357
Data columns (total 11 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                -
0   jerawat.diluar.mens                  98 non-null    object
1   sedang.berjerawat                  98 non-null    object
2   bekas.jerawat                      98 non-null    object
3   flek.hitam                         98 non-null    object
4   komedo                             98 non-null    object
5   sedang.hamil                       98 non-null    object
6   jenis.kulit                       98 non-null    object
7   klt.kdang.merah.n.gatal            98 non-null    object
8   kb                                  98 non-null    object
9   usia_bin                           98 non-null    category
10  cluster_predicted                  98 non-null    uint16
dtypes: category(1), object(9), uint16(1)
memory usage: 8.2+ KB
```

Gambar 8 Informasi segmen ketiga

Tabel 1a. Hasil pusat klaster

Atribut					
Clust er	Jerawat di luar menstruasi	Sedang berjerawat	Bekas jerawat	Flek hitam	Komedo
0	Tidak	Tidak	Ada	Tidak	Ada
1	Tidak	Tidak	Tidak	Ada	Tidak
2	Iya	Iya	Ada	Tidak	Ada

Tabel 1b. Hasil pusat klaster

Atribut					
Cluster	Sedang hamil	Jenis kulit	Kulit terkadang merah dan gatal	KB	Usia
0	Tidak	Berminyak	Tidak	Tidak	26-35
1	Tidak	Kering	Iya	Tidak	36-45
2	Tidak	Berminyak	Iya	Tidak	17-25

Berdasarkan hasil pengolahan data pada dataset keluhan pelanggan Agen Ami MS Glow Peterongan menggunakan algoritma *K-Modes*, didapatkan 3 klaster atau segmen pelanggan. Tabel 4.1 memperlihatkan tiga pusat klaster yang dihasilkan dalam proses *clustering* menggunakan algoritma *K-Modes*. Seperti halnya yang dipaparkan oleh Indriani dan Budiman [11], Yulianita dan Istiawan [23], Badruttamam dkk [3] dalam penelitian mereka mengenai implementasi algoritma *K-Modes*, ketiga pusat klaster yang terbentuk ini merupakan representasi dari masing-masing klaster atau segmen pelanggan. Artinya karakteristik semua anggota yang berada di setiap klaster atau segmen pelanggan tersebut dapat diwakili oleh pusat klasternya.

Berdasarkan gambar di atas (gambar 4.24), didapatkan informasi dari ketiga pusat kluster sebagai berikut,

- a) Kluster 0 atau segmen pelanggan pertama didominasi oleh pelanggan yang berkarakteristik berusia 26 – 35 tahun, tidak memakai KB dan tidak sedang hamil, berjenis kulit berminyak, tidak berjerawat di luar menstruasi, tidak sedang berjerawat, memiliki bekas jerawat dan komedo, tidak memiliki flek hitam, serta tidak mengalami kulit yang terkadang merah dan gatal.
- b) Kluster 1 atau segmen pelanggan kedua didominasi oleh pelanggan yang berkarakteristik berusia 36 – 45 tahun, tidak memakai KB dan tidak sedang hamil, berjenis kulit kering, tidak berjerawat di luar menstruasi, tidak sedang berjerawat, tidak memiliki bekas jerawat dan tidak berkomedo, memiliki flek hitam, serta kulit yang terkadang merah dan gatal.
- c) Kluster 2 atau segmen pelanggan ketiga didominasi oleh pelanggan yang berkarakteristik pelanggan berusia 17 – 25 tahun, yang tidak memakai KB dan tidak sedang hamil, berjenis kulit berminyak, berjerawat di luar menstruasi dan sedang berjerawat, memiliki bekas jerawat dan berkomedo, tidak memiliki flek hitam, serta kulit yang terkadang merah dan gatal.

Sekarang kita dapat mengetahui karakteristik dari masing-masing segmen pelanggan Agen Ami MS Glow Peterongan. Pengetahuan dan pemahaman mengenai karakteristik pelanggan sangat diperlukan oleh pihak manajemen Ami MS Glow Peterongan untuk dapat menyelesaikan permasalahan yang tengah dihadapi, yaitu ketidaktepatan stok produk dan penurunan interaksi konten di media sosial. Dari ketiga segmen pelanggan tersebut dapat diketahui bahwa segmen pertama yang paling banyak jumlah anggotanya dibandingkan dengan segmen kedua dan ketiga. Untuk itu, hendaknya segmen pertama dengan segala karakteristiknya dapat lebih diprioritaskan oleh pihak manajemen. Sebagai gambaran, dalam pembelian stok produk dapat diperbanyak jenis produk yang sesuai dengan tipe dan karakteristik segmen pertama. Iklan promosi di media sosial bisa lebih dikhususkan untuk usia 26 – 35 tahun (usia pelanggan dari segmen pertama). Dan untuk konten yang dibuat di media sosial dapat diperbanyak untuk konten yang berhubungan dengan karakteristik segmen pertama.

V. KESIMPULAN

Berdasarkan uraian hasil analisis dan pembahasan di atas, diperoleh kesimpulan bahwa algoritma *K-Modes* dapat digunakan untuk melakukan segmentasi pelanggan Agen Ami Ms Glow Peterongan. Hasil segmentasi pelanggan Agen Ami Ms Glow Peterongan menggunakan algoritma *K-Modes* menghasilkan 3 segmen pelanggan dengan segmen

pertama yang mendominasi secara signifikan. Dari ketiga segmen pelanggan yang dihasilkan, segmen pertama memiliki jumlah anggota terbesar, disusul segmen ketiga, kemudian segmen kedua. Untuk menyelesaikan permasalahan ketidaktepatan stok produk dan penurunan interaksi konten di media sosial yang dihadapi manajemen hendaknya lebih memperhatikan segmen pertama dengan segala karakteristiknya. Sebagai gambaran, dalam pembelian stok produk dapat diperbanyak jenis produk yang sesuai dengan segmen pertama, iklan promosi di media sosial bisa lebih dikhususkan untuk usia 26 – 35 tahun (usia pelanggan segmen pertama), serta memperbanyak jenis konten di media sosial yang berhubungan dengan karakteristik segmen pertama.

DAFTAR PUSTAKA

- [1] Amin, M.A., & Juniati, D. (2017) Klasifikasi Kelompok Umur Manusia Berdasarkan Analisis Dimensi Fraktal *Box Counting* dari Citra Wajah dengan Deteksi Tepi Canny. *Jurnal Ilmiah Matematika.*, 2(6), (Online), (<https://media.neliti.com/media/publications/249455-none-23b6a822.pdf>), diakses pada 12 Januari 2022.
- [2] Ashish. (2019). “Bank Customer Clustering (K-Modes Clustering)”. Dalam Kaggle. <https://www.kaggle.com/code/ashydv/bank-customer-clustering-k-modes-clustering/notebook>. Diakses pada 11 Januari 2022.
- [3] Badruttamam, A., Sudarmo, & Maruddani, D.A.I. (2020). Penerapan Analisis Kluster *K-Modes* dengan Validasi Davies Bouldin Index dalam Menentukan Karakteristik Kanal Youtube di Indonesia. *Jurnal Gaussian.*, (Online), 9(3): 263-272, (<https://ejournal3.undip.ac.id/index.php/gaussian/article/view/28907>), diakses pada 11 Januari 2022.
- [4] Blythe, J. (2005). *Essential of Marketing*, Edisi Ketiga. Harlow, Essex: Pearson Education Limited.
- [5] Fahmi, M.F. (2016). *Sistem Informasi Geografis (SIG) Sebaran Daerah Aliran Sungai (DAS) Menggunakan Segmentasi Berbasis Algoritma K-Modes Clustering dan Davies-Bouldin Index*. Tesis tidak diterbitkan. Surabaya: FTI, ITS.
- [6] Hair, J. F., Black, W. C., Babin, B. J. dan Anderson, R. E. (2014). *Multivariate Data Analysis 7th*. USA: Pearson.
- [7] Han, J., & Kamber, M. (2006). *Data Mining: Concept and Techniques 2nd*. United States of America: Elsevier.
- [8] Harani, N.H., Cahyo, P., & Fikri, A.N. (2020). Segmentasi Pelanggan Produk *Digital Service* Indihome Menggunakan Algoritma K-Means Berbasis Python. *Jurnal Manajemen Informatika (JAMIKA)*, (Online), 10(2): 133-146, (<https://ojs.unikom.ac.id/index.php>), diakses 11 Januari 2022.
- [9] Hooley, G., Piercy, N.F., dan Nicolaud, B. (2008). *Marketing Strategy and Competitive Positioning*, Edisi keempat. Harlow, Essex: Pearson Education Limited.
- [10] Huang, Z., (1997). *A Fast Clustering Algorithm to Cluster Very Large Categorical Data Sets in Data Mining*. Canberra: Cooperative Research Centre for Advanced Computational Systems.
- [11] Indriani, F., & Budiman, I. (2017). K-Modes Clustering untuk Mengetahui Jenis Masakan Daerah yang Populer pada Website Resep Online (Studi Kasus: Masakan Banjar di Cookpad.com). *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, (Online), 4(4): 290-296, (<https://jtiik.ub.ac.id/index.php/jtiik>), diakses pada 11 Januari 2022
- [12] Kumar. (2018). “2018 Kaggle Machine Learning & Data Science Survey”. Dalam Kaggle. <https://www.kaggle.com/code/sudhirn17/data-science-survey-2018/notebook>. Diakses pada 11 Januari 2022.

- [13] Kurniawan, D. (2020). *Pengenalan Machine Learning dengan Python*. Jakarta: PT Elex Media Komputindo.
- [14] Morissan. (2010). *Periklanan: Komunikasi Pemasaran Terpadu*. Jakarta: Kharisma Putra Utama.
- [15] Nduru, E. K., Buulolo, E., & Pristiwanto. (2018). *Implementasi Algoritma K-Modes untuk Menentukan Strategi Marketing STMIK Budi Darma*. Makalah disajikan dalam Konferensi Nasional Teknologi Informasi dan Komputer (KOMIK), Medan, Oktober 2018. Dalam *Ejournal STMIK Budidarma*, (Online), (<https://ejournal.stmik-budidarma.ac.id/index.php/komik/article/download/903/778>), diakses pada 11 Januari 2022.
- [16] Prasetyo, E. (2014). *Data Mining Mengolah Data Menjadi Informasi dengan Matlab*. Yogyakarta: Andi.
- [17] Rohmawati, N., Sofi, D., & Mohamad, J. (2015). Implementasi Algoritma K-Means dalam Pengklasteran Mahasiswa Pelamar Beasiswa. *Jurnal Ilmiah Teknologi Informasi Terapan, 1(UTAMA)*, 63.
- [18] Sharma, N. & Ashok, S. (2015). K-Modes Clustering Algorithm for Categorical Data. *International Journal of Computer Applications.*, 127(17), (Online), (<https://www.ijcaonline.org/research/volume127/number17/sharma-2015-ijca-906708.pdf>), diakses pada 11 Januari 2022.
- [19] Smith, W.R. (1956). Product Differentiation and Market Segmentation as Alternative Marketing Strategies. *Journal of Marketing*, 20(3), 3-8.
- [20] Supriyatna, I. & Djailani, M. F. (2022). “Pertumbuhan Industri Kosmetik Peluang Cuan Baru”. Dalam *SUARA*. (<https://www.suara.com/bisnis/2021/10/01/165627/pertumbuhan-industri-kosmetik-peluang-cuan-baru>). Diakses pada 11 Januari 2022.
- [21] Tjiptono, F. (2015). *Strategi Pemasaran*, Edisi Keempat. Yogyakarta: CV. Andi Offset (Penerbit Andi).
- [22] Wartaekonomi (2021, 24 Agustus). Pecahkan Rekor! MS Glow Sabet Rekor Muri, Bosnya Sampai Gemetaran Menerima Penghargaan Ini. Diakses pada 20 November 2021, dari (<https://wartaekonomi.co.id/read356730/pecahkan-rekor-ms-glow-sabet-rekor-muri-bosnya-sampai-gemetaran-menerima-penghargaan-ini>
- [23] Yulianita, T., & Istiawan, D. (2017). “Implementasi Algoritma K-Modes untuk Penentuan Prioritas Rehabilitasi Daerah Aliran Sungai Berdasarkan Parameter Lahan Kritis”. *The Sixth University Research Colloquium 2017 Universitas Muhammadiyah Magelang* : 429-439.
- [24] Yulianton, H., Sutanto, F. A., & Mulyani, S. (2021). *Pengelompokan Mahasiswa Berbasis Categorical Variables Menggunakan Metode K-Modes Clustering*. Makalah disajikan dalam Seminar Nasional Multi Disiplin Ilmu (SENDIU), Semarang, 28 Juli 2021. Sendi Unisbank, (Online), (<https://www.unisbank.ac.id/ojs>), diakses pada 11 Januari 2022.