

Klasifikasi Pendapatan Driver Shopee Food Menggunakan Metode K-Means

Ariq Azhar Zhafari¹, Supatman²

^{1,2}Informatika, Fakultas Teknologi Informasi, Universitas Mercubuana Yogyakarta
ariqazhar8563@gmail.com¹

ABSTRACT

PT Shopee Internasional Indonesia is a company that has e-commerce features and also has a food and food delivery service known as Shopee Food which started operating in April 2020. At the beginning of 2021, Shopee Food started fulfilling food and drink orders, thereby attracting many driver partners to deliver it to consumers. This is because the payment system is easy to use. The income received by drivers through ShopeePay is accumulated and properly recorded in the Shopee Food driver account. Therefore, it is necessary to classify a driver's income every week to find out the highest and lowest income. The aim of the research carried out is grouping to minimize the objective function set in the clustering process, which basically tries to minimize variation within one cluster and maximize variation between clusters, and to classify several incomes based on the K-Means method. The research method used is the K-Means clustering method which is used to group data into several groups (clusters) based on similar features. Data collection was carried out for 5 months (May-October) 2023. Then the data obtained was identified in which week the income data was the highest and lowest. So the clustering is carried out which is displayed in the diagram by marking 4 features to make it easier to read the results, namely red (low), blue (medium), black (highest), and blue box (average). The results obtained based on diagrams (visual data) produce the highest data in the 1st month with income of IDR 39,000.00, while the lowest data is in the 4th month with IDR 6,400.00 and with a small data ratio. The conclusion obtained from the research results is that you get the highest income, namely in month 1 with an income of IDR 39,000.00.

Keywords : clustering, k-means, shopeefood.

ABSTRAK

PT Shopee Internasional Indonesia merupakan perusahaan yang memiliki fitur *e-commerce* sekaligus memiliki layanan pesan antar makanan dan minuman yang dikenal dengan Shopee Food mulai beroperasi pada April 2020. Pada awal tahun 2021, Shopee Food mulai memenuhi pesanan makanan dan minuman sehingga menarik banyak mitra *driver* untuk mengantarkannya ke konsumen. Hal ini dikarenakan sistem pembayaran yang mudah dilakukan. Pendapatan yang diterima driver melalui ShopeePay terakumulasi dan tercatat dengan baik pada akun driver Shopee Food. Maka dari itu perlunya dilakukan klasifikasi pendapatan pada seorang *driver* setiap minggunya untuk mengetahui pendapatan tertinggi dan terendah. Tujuan dari penelitian yang dilakukan adalah pengelompokan untuk meminimalkan *objective function* yang di set dalam proses *clustering*, yang pada dasarnya berusaha untuk meminimalkan variasi dalam satu *cluster* dan memaksimalkan variasi antar *cluster*, dan untuk mengklasifikasikan beberapa pendapatan berdasarkan metode K-Means. Metode penelitian yang digunakan yaitu metode K-Means *clustering* yang digunakan untuk mengelompokkan data menjadi beberapa kelompok (*cluster*) berdasarkan kesamaan fitur. Pengambilan data dilakukan selama 5 bulan (mei-oktober) 2023. Kemudian data yang diperoleh diidentifikasi pada minggu ke-berapa yang data pendapatannya tertinggi dan terendah. Sehingga dilakukan pengclasteran yang ditampilkan dalam diagram dengan ditandai 4 fitur untuk

memudahkan dalam pembacaan hasilnya yaitu merah (rendah), biru (sedang), hitam (tertinggi), dan kotak biru (rata-rata). Hasil yang diperoleh berdasarkan diagram (visual data) menghasilkan data tertinggi pada bulan ke 1 dengan pendapatan sebesar Rp 39,000.00 sedangkan untuk data terendah berada di bulan ke 4 dengan Rp 6,400.00 serta dengan rasio data yang sedikit. Kesimpulan yang didapatkan dari hasil penelitian adalah mendapatkan pendapatan tertinggi yaitu pada bulan 1 dengan pendapatan Rp 39,000.00.

Kata kunci : *clustering, k-means, shopee food.*

PENDAHULUAN

PT Shopee International Indonesia merupakan perusahaan yang memiliki fitur *e-commerce* sekaligus memiliki layanan pesan antar makanan dan makanan yang dikenal dengan Shopee Food mulai beroperasi pada April 2020. Saat itu, Shopee Food hanya melayani pembelian makanan beku, minuman ringan, aneka kue, dan makanan olahan saja. Namun pada awal tahun 2021, Shopee Food mulai memenuhi pesanan makanan dan minuman bekerja sama dengan berbagai industri makanan dan minuman dan menarik banyak mitra *driver* untuk mengantarkannya ke konsumen (Vania dan Simbolon, 2021). Shopee Food memiliki dua sistem pembayaran yaitu, *Cash on Delivery* dan *Shopeepay*. *Cash On Delivery* (COD) adalah sistem pembayaran dimana konsumen membayar *driver* Shopee Food secara tunai saat mengambil pesannya. Pembayaran *ShopeePay* merupakan uang elektronik yang diterima ketika konsumen melakukan pembayaran di aplikasi Shopee menggunakan akun saldo *ShopeePay* milik konsumen.

Dana yang diterima oleh *driver* Shopee Food melalui metode atau sistem pembayaran *ShopeePay* akan otomatis masuk ke akun *driver* Shopee Food. Menurut Fikri, (2021) menunjukkan bahwa penggunaan *ShopeePay* sebagai dompet digital berpengaruh terhadap perilaku konsumtif mahasiswa, memiliki nilai hubungan sebesar 0,678 yang artinya signifikan. Selanjutnya, rata-rata mahasiswa menggunakan *ShopeePay* untuk mendapatkan diskon, gratis ongkos kirim dan *cashback* saat bertransaksi menggunakan pembayaran *ShopeePay*.

Penerimaan metode pembayaran *ShopeePay* pada Shopee Food memberikan dampak yang baik bagi *driver* Shopee Food dalam pengelolaan pendapatan yang diterima, karena jumlah pendapatan yang diterima melalui *ShopeePay* terakumulasi dan tercatat dengan baik pada akun *driver* Shopee Food. Pencairan dana *ShopeePay* cukup mudah karena dapat dilakukan melalui rekening bank dan juga dapat ditransfer ke akun saldo *ShopeePay* pada aplikasi Shopee. Keberadaan sistem pembayaran digital yang inovatif merupakan instrumen penting dalam sebuah bisnis, meskipun terkadang terdapat beberapa resiko tertentu yang dapat muncul.

Cluster analysis merupakan salah satu metode *object mining* yang bersifat tanpa latihan (*unsupervised analysis*), sedangkan *K-Means Cluster analysis* merupakan salah satu metode *cluster analysis* non hirarki yang berusaha untuk mempartisi objek yang ada

kedalam satu atau lebih *cluster* atau kelompok objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu *cluster* yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam *cluster* yang lain. Tujuan dari penelitian yang dilakukan adalah pengelompokan untuk meminimalkan *objective function* yang di set dalam proses *clustering*, yang pada dasarnya berusaha untuk meminimalkan variasi dalam satu *cluster* dan memaksimalkan variasi antar *cluster*, dan untuk mengklasifikasikan beberapa pendapatan berdasarkan metode K-Means.

TINJAUAN LITERATUR

K-Means merupakan salah satu algoritma *clustering* dengan metode partisi yang berbasis titik pusat (*centroid*). Algoritma ini pertama kali diusulkan oleh McQueen (1996) dan dikembangkan oleh Hartigan dan Wong tahun 1975. K-Means mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang cepat dan efisien (Irwansyah dan faisal, 2015). Metode K-Means sangat sederhana untuk dijalankan dan diterapkan, cukup cepat, mudah dipergunakan, dan mudah beradaptasi umum digunakan dalam berbagai aplikasi kecil hingga menengah. Secara ilmiah, K-Means menjadi salah satu metode yang paling penting dalam penggunaan *data mining*.

Metode K-Means *clustering* merupakan salah satu dari pengelompokan metode non hirarki yang tujuannya mengelompokkan objek yang diawali dengan mengidentifikasi data yang akan di *cluster*. Pada pertama kali iterasi, titik utama setiap *cluster* yang ditetapkan secara bebas. Lalu dihitung antar jarak data dengan tiap titik utama pada *cluster*. Untuk melakukan perhitungan jarak data ke-*i* pada pusat cluster ke-*j* yang diberi nama ($D(1_2)$) (Tanjung, *et al.*, 2021).

METODE PENELITIAN

K-Means Algorithm

Penggunaan metode penelitian yang digunakan menggunakan metode K-Means merupakan salah satu algoritma *clustering* yang digunakan untuk mengelompokkan data menjadi beberapa kelompok (*cluster*) berdasarkan kesamaan fitur. Tujuan utama dari algoritma ini adalah untuk meminimalkan *varians* dalam setiap *cluster*. Pengambilan data dilakukan selama 5 bulan (Mei-Oktober) 2023. Kemudian data yang diperoleh diidentifikasi pada minggu ke-berapa yang data pendapatannya tertinggi dan terendah. Kemudian dilakukan pengklasteran yang ditampilkan dalam diagram dengan ditandai 4 cluster yaitu tinggi dengan warna hitam, sedang dengan warna biru, rendah dengan warna merah, dan ada rata rata dengan kotak warna biru.

Clustering

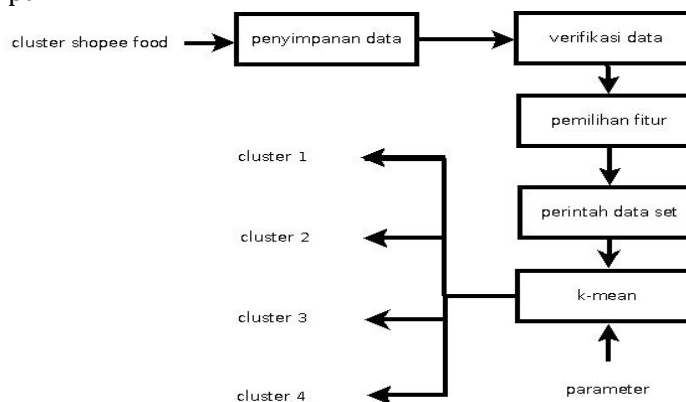
Clustering adalah salah satu alat untuk menganalisis data, yang memecahkan permasalahan pengelompokan data. Objek dari *clustering* adalah kasus pendistribusian (orang, objek, peristiwa, angka, atau data lainnya) ke dalam kelompok, sedemikian rupa hingga mendapatkan derajat tingkat keterhubungan antar anggota cluster yang berbeda. Dengan cara ini, masing-masing *cluster* menguraikan, dalam kaitan dengan kumpulan atau koleksi data, *class* dimana milik anggotanya. *Cluster* disebut juga data atau item yang dikelompokkan berdasarkan pilihan konsumen ataupun hubungan logis.

HASIL DAN PEMBAHASAN

Metode *cluster* ini meliputi *sequential threshold*, *parallel threshold* dan *optimizing threshold*. *Sequential threshold* melakukan pengelompokan dengan terlebih dahulu memilih satu objek dasar yang akan dijadikan nilai awal cluster, kemudian semua *cluster* yang ada dalam jarak terdekat dengan cluster ini akan bergabung, lalu dipilih *cluster* kedua dan semua objek yang mempunyai kemiripan dengan cluster ini akan digabungkan, demikian seterusnya sehingga terbentuk beberapa *cluster* dengan keseluruhan objek yang terdapat didalamnya.

Jika diberikan sekumpulan objek (X_1, X_2, \dots, X_n) maka algoritma K-Means *Cluster analysis* akan mempartisi X dalam k buah *cluster*, setiap *cluster* memiliki *centroid* dari objek-objek dalam *cluster* tersebut. Pada tahap awal algoritma K-Means *Cluster analysis* dipilih secara acak k buah objek sebagai *centroid*, kemudian jarak antara objek dengan *centroid* dihitung dengan menggunakan jarak *euclidean*, objek ditempatkan dalam *cluster* yang terdekat dihitung dari titik tengah *cluster*.

Centroid baru ditetapkan jika semua objek sudah ditempatkan dalam *cluster* terdekat. Proses penentuan *centroid* dan penempatan objek dalam *cluster* diulangi sampai nilai *centroid* konvergen (*centroid* dari semua *cluster* tidak berubah lagi) (Ediyanto dan Satyahadewi, 2013). Proses *machine learning* yang digunakan dengan mengambil data 5 bulan dilakukan uji dari pendapatan *driver* Shopee Food pada data minggu ke 1 sampai 4.



Gambar 1. Flowchart langkah pengolahan data.

Berdasarkan (Gambar 1) proses dimulai dari melakukan penyimpanan data, dimana proses awal ini sangat penting karena perlunya untuk mengumpulkan data yang akan diproses lalu menyimpannya. Proses selanjutnya yaitu verifikasi data, di proses ini bertujuan untuk memastikan data valid sesuai yang ada di lapangan. Selanjutnya ada proses pemilihan fitur dimana dari data yang sudah di verifikasi sesuai dengan data lapangan, maka akan dilakukan pemilihan fitur dengan memilah data dimana nantinya untuk diuji. Setelah melakukan pemilihan fitur maka akan dilakukan perintah dataset, dengan melakukan perintah dataset ini maka akan memusatkan titik pusat dari data yang akan diuji.

```
#mengambil data shopee
datashopee = pd.read_csv('data_skripsi_final_1_kmean_1_fix.csv')
df=pd.DataFrame(datashopee,columns=['nomor','total_waktu','biaya_antar','seminggu'])
X=np.asarray(datashopee)
x_train=X[:,1:]
x_train

array([[4.8611110e-03, 6.4000000e+03],
       [4.8611110e-03, 6.4000000e+03],
       [4.8611110e-03, 8.8000000e+03],
       ...,
       [5.1388889e-02, 8.8000000e+03],
       [5.1388889e-02, 9.6000000e+03],
       [5.1388889e-02, 1.9200000e+04]])
```

Gambar 2. Perintah dataset.

Setelah melakukan pemilihan fitur maka akan dilakukan perintah dataset, dengan melakukan perintah data set ini maka akan mengelompokan data yang sesuai dengan perintah yang diberikan. Pada (Gambar 2) memunculkan koordinat dari pengelompokan 4 data yang akan digunakan untuk di *clusterkan*.

```
#Proses klasterisasi dengan KMeans
kmeans = KMeans(n_clusters=4)
kmeans.fit(x_train)

C:\Users\ACER\AppData\Local\Packages\Python
super()._check_params_vs_input(x, defau

▼ KMeans
KMeans(n_clusters=4)
```

Gambar 3. Clustering data menjadi 4 fitur.

menunjukkan hasil yang sudah tertampil pada grafik berikut penjelasan dari *pengclusteran* sebagai berikut:

Cluster 1 : Tinggi (hitam)

Cluster 2 : Sedang (biru)

Cluster 3 : Kecil (merah)

Cluster 4 : Rata-rata (kotak biru)

Berdasarkan dari pengclusteran yang telah dibuat mendapatkan hasil di bulan 1 yang tertinggi di Rp 39,000.00 di cluster 1. Untuk pendapatan sedang dari Rp 8,000.00 - Rp 12,000.00 hampir di semua bulan ada akan tetapi hasil pada bulan ke 4 paling kecil dari bulan yang lain. Pendapatan kecil dimulai dari Rp 6,400.00 - Rp 8,000.00 dari seluruh bulan memiliki jenis pendapatan kecil ini akan tetapi ada 2 bulan yang memiliki pendapatan kecil diantara kelima bulan yaitu bulan ke 4 dan ke 5.

KESIMPULAN DAN SARAN

Hasil penelitian clusterisasi yang berjudul “KLAUSTERISASI PENDAPATAN DRIVER SHOPEE FOOD MENGGUNAKAN METODE K-MEANS” pada data Hasil pendapatan driver dapat ditarik kesimpulan bahwa hasil pendapatan tertinggi ada pada bulan ke 1 sedangkan untuk hasil terendah berada di data pada bulan ke 4.

Dari penelitian ini tentu masih banyak terdapat kekurangan sehingga analisis *cluster* diatas perlu dilakukan lagi penelitian lanjutan agar *clusterisasi* data pendapatan driver shopee food akurat.

DAFTAR PUSTAKA

- Ediyanto, M. N. M., & Satyahadewi, N. (2013). Pengklasifikasian Karakteristik Dengan Metode K-Means Cluster Analysis. Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya, 2(02).
- Fikri, A. (2021). Pengaruh Penggunaan Shopeepay sebagai Dompot Digital terhadap Perilaku Konsumtif Mahasiswa FEB USU (Doctoral dissertation, Universitas Sumatera Utara).
- Handoko, K. (2016). Penerapan Data Mining dalam Meningkatkan Mutu Pembelajaran Menggunakan Metode K-MEANS Clustering. Jurnal Nasional Teknologi dan Sistem Informasi, 2(3), 31-40.
- Irwansyah, E., & Faisal, M. (2015). Advanced Clustering: Teori dan Aplikasi. Deepublish.
- Rubin, V. L., Chen, Y., & Conroy, N. K. (2015). Deception detection for news: three types of fakes. Proceedings of the Association for Information Science and Technology, 52(1), 1-4.
- Tanjung, F. A., Windarto, A. P., & Fauzan, M. (2021). Penerapan Metode K-Means Pada Pengelompokan Pengangguran Di Indonesia. Jurasik (Jurnal Riset Sistem Informasi Dan Teknik Informatika), 6(1), 61-74.

Vania, I. ., & Simbolon, R. . (2021). PENGARUH PROMO SHOPEEFOOD TERHADAP MINAT BELI PENGGUNA SHOPEE (Di Daerah Tangerang Selatan). Jurnal Ekonomis, 14(2b).