

Clustering Data Pelanggan Menggunakan K-Means untuk Segmentasi Pasar pada Sentra UMKM di Kota Banda Aceh

Clustering Customer Data Using K-Means for Market Segmentation in Micro, Small, and Medium Enterprise (MSME) Hubs in Banda Aceh

M. Bayu Wibawa¹, Mahendar Dwi Payana², Rizal Fahmi³

¹²³Program Studi Informatika Universitas Ubudiyah Indonesia, Banda Aceh

*Koresponding Penulis : ¹mbayuw@uui.ac.id, ²mahendar@uui.ac.id, ³Rizalf23@gmail.com

Abstract

Segmentasi pasar merupakan elemen penting dalam merumuskan strategi pemasaran yang efektif, terutama bagi pelaku Usaha Mikro, Kecil, dan Menengah (UMKM) yang ingin memperluas jangkauan pasar dan meningkatkan keterlibatan pelanggan. Penelitian ini bertujuan untuk menerapkan algoritma K-Means clustering dalam mengelompokkan data pelanggan dari sentra UMKM di Kota Banda Aceh ke dalam beberapa segmen pasar yang teridentifikasi. Data yang digunakan mencakup informasi demografis pelanggan, frekuensi pembelian, dan nilai transaksi, yang diperoleh melalui survei dan catatan penjualan dari beberapa sentra UMKM terpilih. Metode penelitian melibatkan tahapan preprocessing data, penentuan jumlah kluster optimal menggunakan Metode Elbow, serta proses clustering menggunakan algoritma K-Means. Hasil penelitian menunjukkan bahwa pelanggan dapat dikelompokkan secara efektif ke dalam beberapa segmen yang memiliki karakteristik berbeda, seperti pelanggan loyal dengan nilai transaksi tinggi, pelanggan sesekali yang sensitif terhadap harga, serta pelanggan baru dengan tingkat keterlibatan rendah. Temuan ini memberikan wawasan strategis bagi pelaku UMKM dalam merancang kampanye pemasaran yang lebih tepat sasaran, menyesuaikan penawaran produk, serta mengalokasikan sumber daya secara lebih efisien. Dengan memanfaatkan segmentasi berbasis data, UMKM di Banda Aceh dapat meningkatkan daya saing dan responsivitas terhadap kebutuhan pasar di era ekonomi digital.

Kata Kunci: K-Means, Clustering, Segmentasi Pasar, Data Pelanggan, UMKM, Banda Aceh

Abstract

Market segmentation plays a vital role in shaping marketing strategies, particularly for Micro, Small, and Medium Enterprises (MSMEs) seeking to expand their market reach and optimize customer engagement. This study aims to apply the K-Means clustering algorithm to classify customer data from MSME centers in Banda Aceh into distinct market segments. The data used includes customer demographics, purchase frequency, and transaction values, which were collected through surveys and sales records from selected MSME hubs. The methodology involves preprocessing the dataset, determining the optimal number of clusters using the Elbow Method, and performing clustering with K-Means. The results indicate that the customers can be effectively grouped into several meaningful segments, each with unique characteristics such as high-value loyal customers, price-sensitive occasional buyers, and new or low-engagement customers. **Keywords: K-Means, Clustering, Market Segmentation, Customer Data, MSMEs, Banda Aceh**

PENDAHULUAN

Usaha Mikro, Kecil, dan Menengah (UMKM) memiliki peran penting dalam mendorong pertumbuhan ekonomi nasional, termasuk di Kota Banda Aceh. Dalam era digital dan kompetitif saat ini, UMKM dituntut untuk mampu memahami kebutuhan dan perilaku konsumen secara lebih mendalam agar dapat bersaing secara efektif. Salah satu pendekatan strategis yang dapat diterapkan untuk mencapai hal tersebut adalah segmentasi pasar, yaitu proses pengelompokan konsumen ke dalam beberapa segmen berdasarkan karakteristik tertentu, seperti perilaku pembelian, preferensi produk, dan nilai transaksi. Segmentasi pasar tidak hanya membantu pelaku usaha dalam menyusun strategi pemasaran yang lebih tepat sasaran, tetapi juga memungkinkan peningkatan efisiensi operasional serta loyalitas pelanggan. Namun, sebagian besar pelaku UMKM masih melakukan segmentasi secara manual atau berdasarkan intuisi, yang rentan terhadap kesalahan dan kurang optimal. Di sinilah pendekatan berbasis data dan teknologi menjadi sangat relevan.

Salah satu metode yang banyak digunakan untuk analisis segmentasi berbasis data adalah algoritma K-Means clustering. K-Means merupakan teknik unsupervised learning yang efektif dalam mengelompokkan data ke dalam kluster-kluster berdasarkan kemiripan fitur. Dengan menerapkan metode ini pada data pelanggan UMKM, pelaku usaha dapat mengidentifikasi pola yang tersembunyi dan membentuk kelompok pelanggan yang homogen, sehingga strategi bisnis dapat disesuaikan dengan lebih tepat. Penelitian ini dilakukan untuk menerapkan algoritma K-Means dalam proses segmentasi data pelanggan dari sentra-sentra UMKM di Kota Banda Aceh. Dengan memanfaatkan data seperti frekuensi pembelian, nilai transaksi, dan karakteristik demografis, diharapkan penelitian ini dapat memberikan kontribusi nyata dalam pengambilan keputusan berbasis data oleh pelaku UMKM lokal.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksploratif berbasis unsupervised machine learning, yaitu algoritma K-Means Clustering, untuk melakukan segmentasi terhadap data pelanggan pada sentra UMKM di Kota Banda Aceh.

1. Jenis dan Sumber Data

Data yang digunakan merupakan data primer dan sekunder yang meliputi:

- Data primer: Diperoleh melalui survei langsung kepada pelanggan UMKM, dengan variabel seperti usia, jenis kelamin, frekuensi kunjungan, rata-rata pembelian per transaksi, dan jenis produk yang dibeli.
- Data sekunder: Diambil dari catatan transaksi dan basis data penjualan yang dimiliki oleh UMKM setempat.

2. Tahapan Penelitian

Penelitian dilakukan melalui beberapa tahapan utama sebagai berikut:

a. Pengumpulan Data

Data dikumpulkan dari beberapa UMKM yang tergabung dalam sentra industri lokal di Banda Aceh, khususnya pada sektor kuliner, kerajinan, dan fesyen.

b. Pra-pemrosesan Data (Data Preprocessing)

- Pembersihan data dari nilai kosong (missing values)

- Normalisasi data numerik menggunakan metode Min-Max Scaling
 - Konversi data kategorik (jika ada) ke bentuk numerik (one-hot encoding atau label encoding)
 - c. Penentuan Jumlah Kluster
Untuk menentukan jumlah kluster optimal, digunakan Metode Elbow, yang memvisualisasikan nilai *Within-Cluster Sum of Squares (WCSS)* terhadap jumlah kluster. Titik “tekukan” pada grafik digunakan sebagai indikasi jumlah kluster ideal.
 - d. Proses Clustering dengan K-Means
Algoritma K-Means dijalankan menggunakan jumlah kluster yang telah ditentukan. Setiap data pelanggan akan dikelompokkan ke dalam salah satu kluster berdasarkan kemiripan atribut.
 - e. Evaluasi dan Interpretasi Kluster
Setelah proses clustering, masing-masing kluster dianalisis untuk mengidentifikasi karakteristik dominan, seperti kluster pelanggan dengan frekuensi pembelian tinggi, pelanggan baru, atau pelanggan dengan preferensi produk tertentu. Visualisasi dilakukan menggunakan diagram sebar (scatter plot) dan heatmap.
3. Alat dan Teknologi
- Penelitian ini dilakukan menggunakan:
- Bahasa pemrograman: Python
 - Pustaka: Pandas, Scikit-learn, Matplotlib, Seaborn
 - Alat bantu analisis: Jupyter Notebook

HASIL DAN PEMBAHASAN

Penelitian ini berhasil mengelompokkan data pelanggan dari beberapa sentra UMKM di Kota Banda Aceh menggunakan algoritma K-Means Clustering. Data yang digunakan terdiri atas 1.000 entri pelanggan, dengan atribut seperti frekuensi pembelian, rata-rata nilai transaksi, usia pelanggan, dan jenis produk yang dibeli.

1. Penentuan Jumlah Kluster Optimal

Proses penentuan jumlah kluster menggunakan metode Elbow menunjukkan bahwa titik optimal terjadi pada $k = 3$. Hal ini mengindikasikan bahwa data pelanggan dapat dikelompokkan secara efektif ke dalam tiga segmen utama.

2. Hasil Clustering

Berdasarkan hasil clustering, diperoleh tiga kelompok pelanggan dengan karakteristik sebagai berikut:

- Kluster 1: Pelanggan Loyal Bernilai Tinggi
 - Frekuensi pembelian tinggi (lebih dari 4 kali/bulan)
 - Rata-rata transaksi besar ($> \text{Rp}500.000$)
 - Sering membeli produk premium dan kerajinan khas
 - Rata-rata usia: 35–50 tahun
 - Rekomendasi: Berikan layanan khusus, program loyalitas, dan diskon eksklusif
- Kluster 2: Pelanggan Potensial
 - Frekuensi pembelian sedang (2–3 kali/bulan)
 - Transaksi menengah ($\text{Rp}200.000 - \text{Rp}500.000$)

- Umumnya membeli produk fesyen dan makanan
- Rata-rata usia: 25–34 tahun
- Rekomendasi: Tawarkan promosi berkala, edukasi produk melalui media social
- Klaster 3: Pelanggan Baru dan Tidak Aktif
 - Frekuensi pembelian rendah (≤ 1 kali/bulan)
 - Transaksi kecil ($< \text{Rp}200.000$)
 - Umumnya membeli produk sekali tanpa pengulangan
 - Rata-rata usia: 18–24 tahun
 - Rekomendasi: Tingkatkan pendekatan digital, retargeting iklan, dan pemberian insentif awal

3. Visualisasi Klaster

Visualisasi menggunakan scatter plot dua dimensi berdasarkan frekuensi pembelian dan nilai transaksi menunjukkan pemisahan klaster yang cukup jelas. Klaster loyal cenderung terkonsentrasi di area atas kanan (nilai transaksi dan frekuensi tinggi), sedangkan pelanggan baru berada di area kiri bawah.

Hasil clustering menunjukkan bahwa UMKM dapat mengidentifikasi secara lebih akurat jenis pelanggan yang dimilikinya. Hal ini sangat penting untuk strategi pemasaran yang lebih efektif. Misalnya:

- Pelanggan di Klaster 1 dapat menjadi target untuk strategi retensi dan upselling.
- Pelanggan Klaster 2 merupakan target potensial untuk dikembangkan menjadi pelanggan loyal.
- Pelanggan Klaster 3 memerlukan pendekatan yang lebih edukatif dan promosi agresif.

Selain itu, pendekatan berbasis data ini lebih objektif dibandingkan pendekatan konvensional yang mengandalkan intuisi pelaku usaha.

KESIMPULAN

Penelitian ini membuktikan bahwa algoritma K-Means Clustering efektif digunakan untuk segmentasi pasar berbasis data pelanggan pada sentra UMKM di Kota Banda Aceh. Berdasarkan data yang telah dianalisis, pelanggan berhasil dikelompokkan ke dalam tiga segmen utama dengan karakteristik yang berbeda-beda, yaitu pelanggan loyal bernilai tinggi, pelanggan potensial, dan pelanggan baru/tidak aktif. Segmentasi ini memberikan informasi yang sangat berguna bagi pelaku UMKM untuk menyusun strategi pemasaran yang lebih tepat sasaran, meningkatkan efisiensi promosi, dan mengoptimalkan pelayanan terhadap pelanggan. Dengan memahami perilaku dan kebiasaan konsumen melalui analisis klaster, UMKM dapat merancang pendekatan yang sesuai untuk mempertahankan pelanggan lama sekaligus menarik pelanggan baru.

SARAN

1. Peningkatan Akurasi dengan Data Lokal: Disarankan untuk mengembangkan model yang dilatih dengan dataset yang lebih besar dan memperhatikan unsur bahasa daerah atau slang lokal agar sistem lebih sensitif terhadap konteks budaya setempat.

2. Pemanfaatan Sistem Berbasis Data : Pelaku UMKM disarankan mulai membangun sistem pencatatan data pelanggan yang lebih terstruktur agar dapat diterapkan analisis serupa secara berkala.
3. Integrasi dengan Strategi Pemasaran Digital : Hasil segmentasi sebaiknya dijadikan dasar dalam merancang konten promosi, penawaran khusus, dan kampanye digital yang disesuaikan dengan karakteristik tiap klaster pelanggan.
4. Peningkatan Kualitas Layanan : Klaster pelanggan loyal perlu dijaga dengan pemberian insentif khusus seperti program loyalitas, sementara pelanggan potensial perlu ditingkatkan keterikatannya melalui layanan yang lebih personal.
5. Pengembangan Sistem Otomatisasi : Disarankan untuk mengembangkan sistem berbasis aplikasi sederhana yang dapat digunakan UMKM dalam melakukan analisis data pelanggan secara otomatis, bahkan tanpa latar belakang teknis yang tinggi.

DAFTAR PUSTAKA

- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Elsevier.
- Jati, G. N., & Lestari, S. (2020). Segmentasi Pelanggan Menggunakan Algoritma K-Means Clustering pada Data Transaksi Penjualan. *Jurnal Teknik Informatika dan Sistem Informasi*, 7(1), 45–52.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer.
- Putri, A. Y., & Saputra, R. (2021). Analisis Segmentasi Pasar Menggunakan K-Means Clustering pada UMKM Bidang Kuliner. *Jurnal Sistem Informasi*, 15(3), 233–240. <https://doi.org/10.21009/jsi.153.06>
- Wibowo, A., & Nugroho, R. A. (2019). Penerapan K-Means Clustering dalam Segmentasi Konsumen. *Jurnal Ilmu Komputer dan Aplikasi*, 11(2), 88–95.
- Rokach, L., & Maimon, O. (2005). Clustering Methods. In *Data Mining and Knowledge Discovery Handbook* (pp. 321–352). Springer.
- Suhartono, D., & Setiawan, A. (2018). Penerapan Data Mining untuk Segmentasi Pelanggan pada Toko Ritel Menggunakan K-Means. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, 5(4), 511–518.
- Sugiyono. (2019). *Metode Penelitian Kuantitatif, Kualitatif dan R&D*. Alfabeta.