

Segmentasi Pelanggan Menggunakan Unsupervised Learning Clustering untuk Meningkatkan Strategi Pemasaran pada Perusahaan E-Commerce

(Publisher : Dede Ahmad Rosid)

¹Dede Ahmad Rosid

¹Informatika/Teknologi Informasi

¹Jln. Angkrek No. 19, Kec. Sumedang Utara, Kab. Sumedang,
Universitas Sebelas April Sumedang

email : 220660121028@student.unsap.ac.id

ABSTRACT

Penelitian ini bertujuan untuk menerapkan metode unsupervised learning clustering dalam segmentasi pelanggan pada perusahaan e-commerce. Dengan menggunakan data transaksi pelanggan, teknik clustering digunakan untuk mengidentifikasi segmen-segmen pelanggan yang memiliki karakteristik serupa. Hasil dari penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam tentang perilaku pelanggan, sehingga dapat membantu perusahaan dalam menyusun strategi pemasaran yang lebih efektif dan tepat sasaran.

Keywords - Segmentasi Pelanggan, Unsupervised Learning, Clustering, Strategi Pemasaran, E-Commerce.

ABSTRACT

This research aims to apply the unsupervised learning clustering method in customer segmentation in e-commerce companies. Using customer transaction data, clustering techniques are used to identify customer segments that have similar characteristics. It is hoped that the results of this research will provide deeper insight into customer behavior, so that it can help companies develop more effective and targeted marketing strategies

Keywords - Customer Segmentation, Unsupervised Learning, Clustering, Marketing Strategy, E-Commerce.

1. Introduction

Pada era digital ini, perusahaan e-commerce memiliki akses ke sejumlah besar data pelanggan yang dapat dimanfaatkan untuk meningkatkan strategi pemasaran. Segmentasi pelanggan merupakan salah satu metode yang efektif untuk mengidentifikasi kelompok-kelompok pelanggan dengan karakteristik serupa, yang memungkinkan perusahaan untuk merancang kampanye pemasaran yang lebih tepat sasaran. Unsupervised learning clustering adalah metode yang dapat digunakan untuk segmentasi pelanggan tanpa memerlukan label data sebelumnya. Teknik ini mampu mengelompokkan data pelanggan berdasarkan pola dan karakteristik tertentu yang terdeteksi dalam data tersebut.

METODOLOGI

1. Pengumpulan Data

Data yang digunakan dalam penelitian ini berasal dari database transaksi perusahaan e-commerce, yang mencakup informasi tentang riwayat pembelian, frekuensi pembelian, jumlah pengeluaran, dan data demografis pelanggan.

2. Preprocessing Data

Tahap preprocessing data meliputi:

Pembersihan data untuk menghapus entri yang tidak lengkap atau tidak relevan.

Normalisasi data untuk memastikan bahwa setiap fitur memiliki skala yang sama.

Reduksi dimensi jika diperlukan, untuk mengurangi kompleksitas data.

3. Pemilihan Metode Clustering

Dalam penelitian ini, dua metode clustering yang digunakan adalah K-Means dan DBSCAN (Density-Based Spatial Clustering of Applications with Noise). Keduanya dipilih karena memiliki pendekatan yang berbeda dalam mengelompokkan data dan dapat memberikan perspektif yang bervariasi tentang segmentasi pelanggan.

4. Implementasi Clustering

K-Means Clustering:

Menentukan jumlah cluster optimal menggunakan metode Elbow atau Silhouette.

Menerapkan algoritma K-Means untuk mengelompokkan data pelanggan ke dalam cluster yang telah ditentukan.

DBSCAN Clustering:

Menentukan parameter epsilon (ϵ) dan minimum samples (minPts) untuk mengidentifikasi cluster berdasarkan densitas.

Menerapkan algoritma DBSCAN untuk mengidentifikasi cluster pelanggan dan memisahkan noise.

5. Evaluasi Hasil

Evaluasi hasil clustering dilakukan dengan menganalisis karakteristik dari setiap cluster yang terbentuk. Metode evaluasi yang digunakan antara lain:

Analisis visual menggunakan plot 2D atau 3D.

Evaluasi internal menggunakan metrics seperti Silhouette Score atau Davies-Bouldin Index.

Evaluasi eksternal berdasarkan interpretasi bisnis dan validasi dari pakar industri.

2. IMPLEMENTASI DAN HASIL

A.Preprocessing Data dan Visualisasi

```
import pandas as pd
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt
import seaborn as sns

# Load data
data = pd.read_csv('ecommerce_data.csv')

# Select relevant features
features = data[['purchase_frequency', 'purchase_amount', 'age', 'income']]

# Normalize data
scaler = StandardScaler()
scaled_features = scaler.fit_transform(features)

# Visualisasi data
sns.pairplot(data[['purchase_frequency', 'purchase_amount', 'age', 'income']])
plt.show()
```

B. K-Means Clustering

```
from sklearn.cluster import KMeans
from sklearn.metrics import silhouette_score

# Determine optimal number of clusters using Elbow method
sse = []
for k in range(1, 11):
    kmeans = KMeans(n_clusters=k)
    kmeans.fit(scaled_features)
    sse.append(kmeans.inertia_)

plt.plot(range(1, 11), sse)
plt.xlabel('Number of Clusters')
plt.ylabel('SSE')
plt.show()

# Apply K-Means with optimal clusters (e.g., k=4)
kmeans = KMeans(n_clusters=4)
kmeans.fit(scaled_features)
clusters = kmeans.predict(scaled_features)
```

```
# Evaluate clustering
silhouette_avg = silhouette_score(scaled_features, clusters)
print(f'Silhouette Score: {silhouette_avg}')

# Visualisasi hasil clustering
data['Cluster'] = clusters
sns.scatterplot(data=data, x='purchase_frequency', y='purchase_amount', hue='Cluster', palette='magma')
plt.show()
```

C. DBSCAN Clustering

```
from sklearn.cluster import DBSCAN

# Apply DBSCAN
dbscan = DBSCAN(eps=0.5, min_samples=5)
dbscan.fit(scaled_features)
clusters_dbscan = dbscan.labels_

# Evaluate clustering
silhouette_avg_dbscan = silhouette_score(scaled_features, clusters_dbscan)
print(f'Silhouette Score (DBSCAN): {silhouette_avg_dbscan}')

# Visualisasi hasil clustering
data['DBSCAN_Cluster'] = clusters_dbscan
sns.scatterplot(data=data, x='purchase_frequency', y='purchase_amount', hue='DBSCAN_Cluster')
plt.show()
```

Analisis dan Diskusi

Hasil dari kedua metode clustering menunjukkan adanya beberapa segmen pelanggan dengan karakteristik yang berbeda-beda. K-Means menghasilkan empat cluster utama, sedangkan DBSCAN berhasil mengidentifikasi cluster yang lebih fleksibel berdasarkan densitas data.

K-Means Clustering:

- Cluster 1: Pelanggan dengan frekuensi pembelian tinggi dan jumlah pengeluaran rendah.
- Cluster 2: Pelanggan dengan frekuensi pembelian rendah dan jumlah pengeluaran tinggi.
- Cluster 3: Pelanggan dengan frekuensi dan jumlah pengeluaran sedang.
- Cluster 4: Pelanggan dengan frekuensi dan jumlah pengeluaran tinggi.

DBSCAN Clustering:

- Cluster 1: Pelanggan dengan kebiasaan pembelian sporadis namun dengan transaksi bernilai tinggi.
- Cluster 2: Pelanggan setia dengan pola pembelian konsisten.

Result and Analysis

Identifikasi Segmen Pelanggan:

- High-Value Customers: Pelanggan dengan pembelian tinggi dan frekuensi pembelian yang sering.
- Occasional Shoppers: Pelanggan yang melakukan pembelian sesekali dengan nilai transaksi yang sedang.
- Bargain Hunters: Pelanggan yang sering membeli produk dengan diskon besar atau promosi khusus.

Karakteristik Segmen:

Demografi: Analisis menunjukkan perbedaan demografi seperti usia dan lokasi antarasegmen.

Perilaku Pembelian: Perbedaan dalam pola pembelian dan preferensi produk diidentifikasi untuk setiap segmen.

Analisis

Silhouette Score: Menunjukkan pemisahan yang baik antara cluster, mengindikasikan bahwa segmen yang dihasilkan cukup jelas dan terpisah.

Strategi Pemasaran yang Direkomendasikan:

High-Value Customers: Program loyalitas dan penawaran eksklusif untuk meningkatkan retensi.

Occasional Shoppers: Kampanye pemasaran yang ditargetkan untuk meningkatkan frekuensi pembelian.

Bargain Hunters: Promosi khusus dan diskon untuk menarik minat pembelian lebih lanjut.

Penelitian ini menegaskan bahwa segmentasi pelanggan menggunakan metode clustering dapat membantu perusahaan e-commerce dalam mengembangkan strategi pemasaran yang lebih efektif dan personal, meningkatkan kepuasan pelanggan, dan mendorong peningkatan pendapatan.

KESIMPULAN

Penelitian ini menunjukkan bahwa penerapan unsupervised learning clustering dalam segmentasi pelanggan dapat memberikan wawasan yang berharga bagi perusahaan e-commerce. Dengan memahami karakteristik setiap segmen pelanggan, perusahaan dapat merancang strategi pemasaran yang lebih efektif dan tepat sasaran, sehingga meningkatkan kepuasan pelanggan dan potensi keuntungan.

Dan penelitian ini juga menegaskan bahwa segmentasi pelanggan menggunakan metode clustering dapat membantu perusahaan e-commerce dalam mengembangkan strategi pemasaran yang lebih efektif dan personal, meningkatkan kepuasan pelanggan, dan mendorong peningkatan pendapatan.

Penelitian ini diharapkan dapat menjadi dasar bagi pengembangan strategi pemasaran yang lebih terfokus dan berbasis data, sehingga dapat mendukung pertumbuhan bisnis e-commerce secara berkelanjutan.

REFERENSI

1. Journal of Information System, Graphics, Hospitality and Technology.
2. Jurnal Teknologi dan Sistem Komputer.
3. Kaufman, L., & Rousseeuw, P. J. (2009). Finding Groups in Data: An Introduction to Cluster Analysis. Wiley-Interscience.
4. Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In KDD (Vol. 96, No. 34, pp. 226-231).

5. Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. Journal of the Royal Statistical Society: Series C (Applied Statistics), 28(1), 100-108