

# SENASTIKA Universitas Malikussaleh

---

## KLASTERISASI DAN SEGMENTASIPELANGGAN DENGAN MENGGUNAKAN ALGORITMA ORDERING POINTS TO IDENTIFY THE CLUSTERING STRUCTURE (OPTICS) DAN DBSCAN

Aryo Wibisono Putra<sup>1</sup>, Zulham Ariyandi<sup>2</sup>, Revi Diwana<sup>3</sup>, Supriyatna<sup>4</sup>, Munirul Ula<sup>5</sup>

<sup>1,2,3,4,5</sup>Program Studi Teknik Informatika, Universitas Malikussaleh, Lhokseumawe, Aceh

Email: <sup>1</sup>[aryo.210170132@mhs.unimal.ac.id](mailto:aryo.210170132@mhs.unimal.ac.id) , <sup>2</sup>[zulham.210170152@mhs.unimal.ac.id](mailto:zulham.210170152@mhs.unimal.ac.id) ,

<sup>3</sup>[revi.210170153@mhs.unimal.ac.id](mailto:revi.210170153@mhs.unimal.ac.id) , <sup>4</sup>[supriyatna.210170133@mhs.unimal.ac.id](mailto:supriyatna.210170133@mhs.unimal.ac.id) , <sup>5</sup>[munirulula@unimal.ac.id](mailto:munirulula@unimal.ac.id)

### Abstrak

Pengelompokan pada customer merupakan salah satu kiat yang bertujuan untuk menganalisa pengaruh beberapa faktor yang menentukan para pengunjung pada sebuah Mall. Faktor atau fitur yang digunakan dalam menganalisa pengaruh tersebut adalah pendapatan pertahun dan juga nilai pembelanjaan. Penelitian ini menggunakan metode *clustering* dengan algoritma *Ordering points to identify the clustering structure* (OPTICS). Hasil Analisa dari penelitian ini menunjukkan bahwa terdapat cluster yang dihasilkan diantaranya pelanggan dengan pendapatan rendah dan pengeluaran rendah, pelanggan dengan pendapatan tinggi dan pengeluaran tinggi, pelanggan dengan pendapatan tinggi dan pengeluaran rendah. Perbandingan antara algoritma OPTICS dengan algoritma DBSCAN menunjukkan bahwa hasil klusterisasi yang dihasilkan oleh algoritma OPTICS lebih unggul. Dan hasil akhir yang didapat dari penerapan algoritma *Ordering points to identify the clustering structure* (OPTICS) adalah para pengunjung yang memiliki penghasilan tinggi dan nilai pembelanjaan rendah dapat menjadi perhatian para pihak mall untuk meningkatkan faktor agar para pengunjung dapat menjadi lebih royal dan pengunjung ini dapat menjadi *revenue* bagi pihak mall di masa yang akan mendatang.

**Keywords:** *OPTICS, DBSCAN, clustering, pelanggan, pendapatan, pengeluaran*

---

## 1. PENDAHULUAN

Pusat perbelanjaan atau yang biasa disebut dengan mall merupakan salah satu pusat perbelanjaan yang dirancang dengan konsep modern dengan mengikuti perkembangan zaman saat ini. Shopping mall merupakan suatu arena yang memiliki arti tempat yang luas dalam sebuah bangunan yang terdiri dari berbagai macam toko, baik toko pakaian, toko buku, toko handphone, tempat makan/nongkrong dan bahkan juga timezone yang dilengkapi dengan area parkir yang luas [1].

Fungsi dari mall sendiri sebenarnya sama dengan pasar-pasar tradisional lainnya yaitu bertemunya para pedagang dengan para pembeli yang sedang mencari kebutuhan yang mereka butuhkan [2]. Yang membedakan hanyalah konsep yang ditawarkan oleh pihak mall yang mana hal tersebut memiliki daya Tarik yang cukup tinggi untuk para customer untuk mengunjungi mall dengan konsep yang modern. Yang mana hal tersebut jauh berbeda dengan pasar tradisional yang hanya seadanya dengan yang terpenting roda perekonomian mereka berjalan[3]

Yang cukup menarik, tak hanya menjadi tempat pusat perbelanjaan saja, mall juga sekarang telah menjadi salah satu tempat yang ditujukan pada customer untuk berliburan atau rekreasi [4]. Hal tersebut membuat penggabungan dari fungsi utama menjadi beberapa fungsi lainnya sangat menguntungkan dari kedua belah pihak, pelanggan mendapatkan kepuasan hati terhadap hal yang dicari dan dalam sisi lain pihak mall juga mendapatkan keuntungan dari betahnya pelanggan dalam mendapatkan keuntungan yang lebih. Maka dari itu, pembangunan konsep mall tidak hanya mementingkan penyediaan toko saja, tetapi harus memilih konsep yang menyenangkan dan membuat para customer betah berlama disana [5].

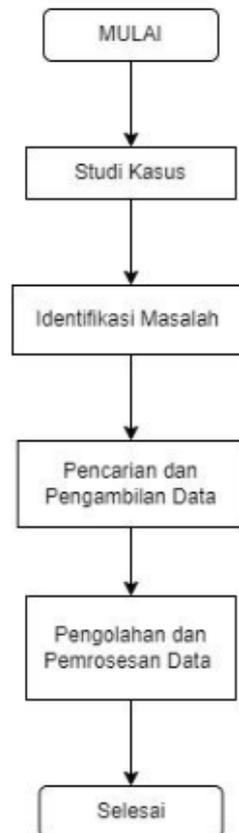
Namun, perbelanjaan di mall telah membuat tingkat status sosial yang cukup signifikan dalam kehidupan bersosial [6]. Perubahan gaya hidup masyarakat juga dapat ditentukan dari berapa jumlah penghasilan yang didapatkan dalam setiap bulan maupun setiap tahun. Semakin tinggi pendapatan yang diperoleh disetiap tahunnya, maka gaya hidup yang akan dijalani juga akan berubah sesuai dengan keadaan status sosial yang didapatkan atas besaran gaji yang diperoleh [7].

Dalam hal tersebut, pada penelitian kali ini akan membahas hal terkait mulai dari Analisa pendapatan dan juga seberapa besar jumlah pengeluaran dalam perbelanjaan setiap customer. Hal tersebut juga memberikan sebuah informasi pada pihak mall dalam mengembangkan strategi marketing dalam pengolahan mall tersebut. Yang mana dari data tersebut dapat memberikan nilai akurasi dalam melihat revenue pada pihak mall untuk keberlangsungan kemajuan dari pihak mereka.

## 2. METODE PENELITIAN

Metode penelitian yang digunakan adalah bersifat kuantitatif. Dimana penulis mencari data yang diperoleh dan juga bersumber dari internet (*internet searching*). Yang mana informasi dan data yang didapat dihasilkan dari situs yang menyediakan dataset untuk dilakukan Analisa lebih lanjut dalam metode dan juga algoritma yang telah dipilih untuk menganalisa data yang telah ada. Penganalisaan ini khusus dilakukan pada dataset yang telah diperoleh dari web dengan dataset yang berisi informasi mengenai para customer mall.

Dalam menganalisa, penulis menggunakan situs yaitu Google Colaboratory. Dimana situs ini dapat digunakan dalam menganalisa dataset yang ada. Dengan Bahasa pemograman yang digunakan yaitu phyton dan metode yang digunakan adalah clustering serta algoritma yang mendukung dalam Analisa ini adalah Ordering points to identify the clustering structure (OPTICS).



Gambar 1. Alur Penelitian

### 2.1 Pengumpulan Data

Pada penelitian ini data yang diperoleh bersumber dari suatu situs yaitu situs [kaggle.com](https://www.kaggle.com/), dimana situs ini dapat memberikan berbagai dataset yang dibutuhkan sesuai dengan keinginan para peneliti. Dalam dataset ini, terdapat kurang lebih 200 data yang menunjukkan data para customer yang memiliki data spending score dan juga anual income yang menjadi data rujukan dalam penelitian ini. Kemudian, data ini akan dilakukan pemrosesan

sesuai dengan algoritma yang telah dipilih, sehingga dapat memberikan data yang dibutuhkan kepada pihak mall dalam menentukan revenue pada penjualan mereka.

## 2.2 OPTICS

Algoritma OPTICS (Ordering Points To Identify the Clustering Structure) adalah metode pengelompokan berbasis kepadatan yang dirancang untuk mengatasi beberapa keterbatasan dari algoritma DBSCAN [8]. OPTICS adalah algoritma yang digunakan untuk mengidentifikasi struktur clustering dalam data dengan kepadatan yang bervariasi. Algoritma ini tidak hanya mengelompokkan data, tetapi juga menghasilkan urutan titik berdasarkan jarak jangkauan (reachability distance), yang memungkinkan analisis lebih lanjut terhadap struktur cluster yang ada [2].

## 2.3 Pre-processing Data

Preprocessing data adalah proses persiapan dan transformasi data mentah (raw data) menjadi format yang lebih terstruktur dan siap analisis [9]. Proses ini sangat penting dalam analisis data dan machine learning karena membantu menghasilkan data yang berkualitas dan reliabel untuk keperluan analisis lebih lanjut. [10]. Dari data yang akan diteliti terdapat beberapa kelompok yang akan dihasilkan mulai dari cluster dengan pendapatan tinggi dan spending score tinggi, cluster dengan pendapatan rendah dengan spending score rendah, cluster dengan pendapatan tinggi dan spending score rendah dan lain sebagainya.

## 2.4 Permodelan

Permodelan dapat dilakukan setelah proses pre-processing data telah selesai dilakukan. Pada penelitian ini, dilakukan penerapan algoritma OPTICS (Ordering Points To Identify the Clustering Structure) untuk melakukan segmentasi pelanggan pada dataset Mall Customers. Algoritma OPTICS dipilih karena kemampuannya dalam mengidentifikasi cluster berbasis kepadatan (density-based clustering), yang cocok untuk data dengan distribusi yang tidak seragam atau variatif. OPTICS mampu menangani cluster dengan kepadatan yang berbeda-beda, memberikan fleksibilitas dalam menemukan pola pada pelanggan dengan karakteristik belanja yang beragam [11]

Permodelan ini menghasilkan beberapa cluster utama yang dapat diinterpretasikan sebagai kelompok pelanggan dengan karakteristik yang berbeda-beda. Misalnya, terdapat cluster pelanggan dengan pendapatan tinggi dan spending score tinggi, yang merupakan kelompok target utama bagi strategi pemasaran, serta cluster pelanggan dengan pendapatan tinggi namun spending score rendah, yang memberikan peluang untuk meningkatkan engagement mereka melalui program pemasaran yang lebih efektif. Hasil ini menunjukkan bahwa algoritma OPTICS berhasil mengidentifikasi pola perilaku pelanggan berdasarkan data yang ada, yang dapat digunakan untuk pengambilan keputusan bisnis yang lebih tepat sasaran.

## 2.4 Evaluasi

Pada bagian ini, algoritma OPTICS digunakan untuk mengelompokkan data pelanggan dari dataset Mall Customers. Evaluasi dilakukan dengan menganalisis hasil visualisasi dan metrik performa dari algoritma yang telah diterapkan. Hasil clustering diinterpretasikan untuk memberikan insight mengenai perilaku pelanggan berdasarkan Annual Income (k\$) dan Spending Score (1-100). Algoritma OPTICS berhasil mengelompokkan pelanggan ke dalam beberapa cluster berdasarkan pola kepadatan (density-based clustering). Pada hasil visualisasi, cluster yang terbentuk menunjukkan adanya pengelompokan pelanggan berdasarkan pola pembelian dan tingkat pendapatan mereka. Sebelum menerapkan algoritma OPTICS pada data, standarisasi data perlu dilakukan agar fitur-fitur yang memiliki skala berbeda menjadi setara.

$$z = \frac{x - \mu}{\sigma}$$

Di mana:

- $x$  adalah nilai asli dari fitur
- $\mu$  adalah nilai rata-rata dari fitur tersebut

- $\sigma$  adalah standar deviasi dari fitur tersebut.

Untuk mendapatkan hasil dari standar deviasi, dapat menggunakan rumus sebagai berikut:

$$\sqrt{\frac{\sum (X_i - \mu)^2}{n}}$$

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Deskriptif Analisis

Dataset Mall Customers terdiri dari data pelanggan dengan beberapa atribut seperti CustomerID, Gender, Age, Annual Income (k\$), dan Spending Score (1-100). Atribut Annual Income (k\$) menunjukkan pendapatan tahunan pelanggan, sementara Spending Score (1-100) mengukur tingkat pengeluaran pelanggan di pusat perbelanjaan dengan skala dari 1 hingga 100, di mana nilai yang lebih tinggi menunjukkan pelanggan yang lebih sering berbelanja.

index	CustomerID ▲	Gender	Age	Annual Income (k\$)	Spending Score (1-100)
0	1	Male	19	15	39
1	2	Male	21	15	81
2	3	Female	20	16	6
3	4	Female	23	16	77
4	5	Female	31	17	40
5	6	Female	22	17	76
6	7	Female	35	18	6
7	8	Female	23	18	94
8	9	Male	64	19	3
9	10	Female	30	19	72
10	11	Male	67	19	14
11	12	Female	35	19	99
12	13	Female	58	20	15
13	14	Female	24	20	77
14	15	Male	37	20	13
15	16	Male	22	20	79

Gambar 2. Tampilan Data

#### 3.2 Pre-processing Data

##### 1. Memeriksa missing values

Memeriksa missing values adalah proses untuk mengidentifikasi dan menganalisis data yang hilang dalam sebuah dataset [12]. Proses ini penting untuk memastikan kualitas data sebelum dilakukan analisis lebih lanjut.

```
[14] print('Null Values :', X.isnull().values.sum())
Null Values : 0
```

Gambar 3. Memeriksa Missing Values

##### 2. Data Training dan Data Testing

Data training digunakan untuk melatih model clustering [13]. Pada konteks dataset mall customers, data ini akan digunakan untuk mengidentifikasi pola dan hubungan antar fitur seperti umur, pendapatan tahunan, skor pengeluaran, dan jenis kelamin. Sedangkan data testing digunakan untuk mengevaluasi kinerja model yang telah dilatih [14]. Pada dataset mall customers, data ini akan digunakan untuk memprediksi apakah model dapat

membagi pelanggan dengan akurasi yang tinggi berdasarkan fitur-fitur yang ada. Dalam hal ini fitur yang digunakan adalah annual income dan spending score.

```

Data Training (X_train):
  Age  Annual Income (k$)  Spending Score (1-100)
79    49                  54                      42
197   32                  126                     74
38    36                  37                       26
24    54                  28                       14
122   40                  69                       58

Data Testing (X_test):
  Age  Annual Income (k$)  Spending Score (1-100)
95    24                  60                      52
15    22                  20                      79
30    60                  30                       4
158   34                  78                       1
128   59                  71                      11

```

Gambar 4. Data Training dan Testing Annual Income

```

Label Training (y_train):
79    42
197   74
38    26
24    14
122   58
Name: Spending Score (1-100), dtype: int64

Label Testing (y_test):
95    52
15    79
30    4
158   1
128   11
Name: Spending Score (1-100), dtype: int64

```

Gambar 5. Data Training dan Testing Spending Score

### 3. Standarisasi Data

Ini bertujuan untuk mengubah skala fitur sehingga semua fitur memiliki pengaruh yang seimbang dalam analisis [15]. Standarisasi biasanya dilakukan untuk memastikan bahwa semua fitur berada dalam rentang yang sama, sedangkan standarisasi mengubah data menjadi distribusi dengan rata-rata 0 dan deviasi standar 1.

```

Annual Income (k$)  Spending Score (1-100)
0    -1.227776      -0.312376
1    -1.070369      0.949142
2    -0.912961     -1.303568
3    -0.598147      0.828997
4    -0.440740     -0.282340
5     0.031481      0.798961
6     0.188889     -1.303568
7     0.975924      1.339611
8     1.133331     -1.393676
9     1.920367      0.678816

```

Gambar 6. Standarisasi Data

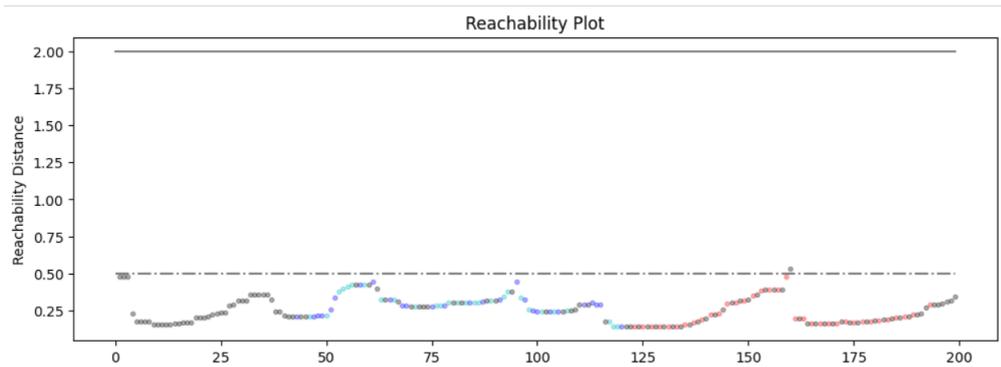
### 3.3 Implementasi Algoritma OPTICS pada Dataset Mall Customers

Dataset pelanggan mall yang berisi informasi penting seperti identitas pelanggan, gender, usia, pendapatan tahunan, dan skor pengeluaran digunakan untuk menerapkan algoritma OPTICS dalam penelitian ini. Dataset ini terdiri dari 200 baris dan 5 kolom, dan dapat diakses melalui platform Kaggle. Sebelum penggunaan algoritma clustering, analisis eksploratif dilakukan untuk memahami distribusi data dan menemukan outlier.

Analisis ini mencakup penghitungan statistik deskriptif untuk variabel numerik dan visualisasi menggunakan boxplot untuk menemukan outlier, terutama dalam kolom pendapatan tahunan.

Selanjutnya, algoritma OPTICS diterapkan dengan menggunakan parameter yang sesuai, yaitu epsilon ( $\epsilon$ ) dan minimum points ( $\text{minPts}$ ). Epsilon digunakan untuk menentukan jarak jangkauan (reachability distance), sementara  $\text{minPts}$  menunjukkan jumlah titik yang harus ada di sekitar core point agar dapat dianggap sebagai bagian dari cluster [8]. Dengan pendekatan ini, algoritma OPTICS mampu mengidentifikasi struktur clustering dalam data dengan kepadatan yang bervariasi, tidak hanya mengelompokkan data tetapi juga menghasilkan urutan titik berdasarkan jarak jangkauan.

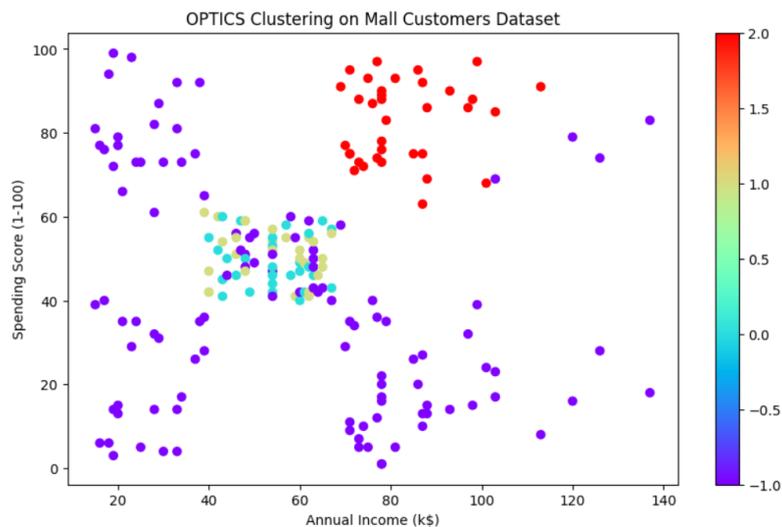
## 1. Reachability Distance



Gambar 7. Reachability Distance

Reachability distance pada dataset Mall Customers membantu kita memahami bagaimana pelanggan dikelompokkan berdasarkan fitur Annual Income (k\$) dan Spending Score (1-100). Titik-titik dengan reachability distance yang rendah merupakan bagian dari cluster dengan kepadatan tinggi, sementara titik-titik dengan reachability distance yang tinggi mungkin berada di tepi cluster atau dianggap sebagai outlier. Algoritma OPTICS dengan metrik ini memungkinkan identifikasi cluster yang lebih fleksibel dibandingkan algoritma lain seperti DBSCAN, terutama ketika densitas antar cluster bervariasi.

## 2. Visualisasi Hasil Clustering

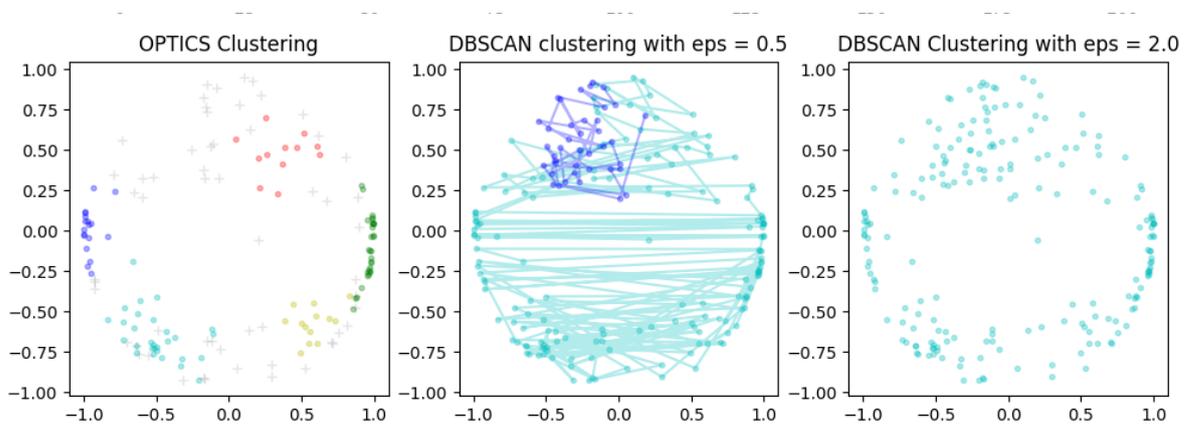


Gambar 8. Visualisasi Hasil Clustering

Berdasarkan visualisasi clustering dari dataset Mall Customers, yang menampilkan pembagian pelanggan ke dalam beberapa cluster menggunakan algoritma OPTICS, kita dapat menginterpretasikan beberapa aspek penting tentang pola pengelompokan pelanggan berdasarkan fitur Annual Income (k\$) dan Spending Score (1-100). Hasil clustering ini divisualisasikan dalam dua dimensi, di mana setiap titik mewakili satu pelanggan, dan titik-titik yang dikelompokkan bersama menunjukkan bahwa mereka memiliki karakteristik serupa.

Cluster padat adalah kelompok pelanggan yang berada dekat satu sama lain, baik dari sisi Annual Income (k\$) maupun Spending Score. Pelanggan dalam cluster yang sama cenderung memiliki pola pendapatan dan perilaku belanja yang serupa. Dari data di atas, mungkin ada cluster pelanggan dengan pendapatan tahunan menengah dan pengeluaran yang konsisten, atau cluster pelanggan dengan penghasilan tinggi dan skor pengeluaran yang juga tinggi. Ada juga kemungkinan terbentuknya cluster yang lebih longgar, di mana titik-titik pelanggan masih cukup dekat dalam hal jarak densitas, tetapi tidak sepadat cluster lain. Ini mungkin mewakili kelompok pelanggan yang memiliki variasi lebih besar dalam Annual Income (k\$) atau Spending Score (1-100), namun masih cukup dekat untuk dianggap berada dalam satu cluster.

### 3. Perbandingan Algoritma OPTICS dengan Algoritma DB-SCAN



Gambar 9. Perbandingan Algoritma OPTICS dengan DBSCAN

Berdasarkan perbandingan antara algoritma OPTICS dengan DBSCAN (dengan nilai epsilon 0.5 dan 2.0) pada dataset Mall Customers, terdapat beberapa kesimpulan penting yang dapat diambil. Algoritma OPTICS menunjukkan keunggulan dalam hal fleksibilitas dan akurasi dalam mengidentifikasi cluster dengan densitas yang bervariasi. OPTICS tidak bergantung pada satu nilai epsilon tetap, sehingga mampu menemukan cluster dengan ukuran dan kepadatan yang berbeda serta memisahkan outlier dengan lebih akurat.

Di sisi lain, DBSCAN dengan epsilon 0.5 menghasilkan clustering yang sangat ketat, di mana hanya beberapa cluster kecil yang terbentuk, dan sebagian besar titik data dianggap sebagai outlier. Hal ini disebabkan oleh jangkauan radius pencarian yang terlalu kecil, sehingga banyak titik yang tidak memenuhi syarat untuk dimasukkan dalam cluster, meskipun beberapa di antaranya sebenarnya berdekatan. Sementara itu, DBSCAN dengan epsilon 2.0 berhasil menemukan lebih banyak cluster karena jangkauannya lebih luas. Namun, algoritma ini juga cenderung menggabungkan titik-titik yang seharusnya berada di cluster terpisah menjadi satu, sehingga cluster yang dihasilkan kurang presisi. Meskipun jumlah outlier berkurang, kualitas cluster menjadi lebih longgar dibandingkan dengan OPTICS.

Secara keseluruhan, algoritma OPTICS lebih unggul dalam mengelompokkan data dengan variasi kepadatan yang tinggi, seperti yang terlihat pada dataset Mall Customers. DBSCAN dengan epsilon kecil cenderung terlalu ketat, sedangkan dengan epsilon besar, cluster menjadi kurang akurat. Oleh karena itu, OPTICS memberikan hasil clustering yang lebih seimbang dan informatif.

### 3.4 Diskusi

Hasil clustering dari algoritma OPTICS menunjukkan bahwa pelanggan dapat dibagi ke dalam beberapa kelompok berdasarkan pendapatan tahunan dan kebiasaan belanja mereka (spending score). Pola umum yang dihasilkan oleh cluster adalah Cluster dengan pendapatan tinggi dan spending score tinggi. Dimana kelompok ini diidentifikasi sebagai pelanggan premium yang kemungkinan besar adalah pelanggan loyal dengan potensi belanja yang signifikan. Mereka mungkin lebih terbuka terhadap produk atau layanan premium, sehingga dapat menjadi target utama untuk program loyalitas atau promosi eksklusif.

Cluster dengan pendapatan rendah dan spending score rendah juga menjadi hasil cluster yang dihasilkan oleh data yang telah diteliti. Dimana kelompok ini menunjukkan pelanggan yang berbelanja lebih sedikit dan mungkin memiliki keterbatasan finansial. Strategi pemasaran untuk kelompok ini harus fokus pada harga kompetitif, promosi diskon, atau layanan yang menarik bagi konsumen dengan anggaran lebih terbatas.

Selanjutnya adalah cluster dengan pendapatan tinggi dan spending score rendah. Dalam cluster ini pelanggan menunjukkan peluang besar untuk ditingkatkan. Meskipun mereka memiliki kemampuan finansial, spending score yang rendah menunjukkan bahwa mereka mungkin belum terlibat secara penuh dengan mall. Upaya promosi dan strategi pemasaran yang lebih personal bisa diterapkan untuk mengubah mereka menjadi pelanggan yang lebih aktif dalam berbelanja.

Hasil dari clustering OPTICS pada dataset Mall Customers dapat diaplikasikan dalam strategi bisnis mall untuk meningkatkan performa penjualan dan kepuasan pelanggan. Melakukan segmentasi Pelanggan Hasil clustering dapat digunakan untuk menciptakan segmentasi pelanggan yang lebih spesifik. Setiap segmen dapat ditargetkan dengan strategi pemasaran yang lebih relevan, misalnya dengan memberikan promosi khusus bagi pelanggan dalam cluster pendapatan tinggi dengan spending score rendah untuk meningkatkan engagement mereka.

Selain melakukan segmentasi pelanggan, peningkatan layanan juga sangat penting dilakukan. Peningkatan Layanan dengan memahami kebutuhan dan karakteristik pelanggan di setiap cluster, mall dapat merancang pengalaman belanja yang lebih personal, seperti penawaran produk yang lebih disesuaikan, program loyalitas, dan layanan khusus untuk pelanggan premium.

Selanjutnya dari data outlier yang telah diketahui dalam visualisasi data yang telah dilakukan, identifikasi outlier memberikan kesempatan untuk mempelajari lebih dalam pelanggan yang tidak sesuai dengan pola umum. Dengan memperhatikan kebutuhan unik mereka, mall dapat merancang program pemasaran yang lebih efektif dan personal untuk meningkatkan loyalitas mereka agar dapat memberikan dampak yang lebih baik dalam penjualan mall.

## 4. KESIMPULAN

Penelitian ini membahas penerapan algoritma OPTICS dan DBSCAN pada dataset Mall Customers untuk mengidentifikasi pola clustering berdasarkan fitur Annual Income (k\$) dan Spending Score (1-100). Hasil analisis menunjukkan bahwa algoritma OPTICS memiliki keunggulan dalam mengelompokkan data dengan kepadatan yang bervariasi, menghasilkan cluster yang lebih akurat dan outlier yang teridentifikasi dengan baik. OPTICS memungkinkan identifikasi segmen pelanggan yang lebih kompleks karena algoritma ini tidak bergantung pada satu nilai epsilon, sehingga mampu menangkap cluster dengan ukuran dan bentuk yang berbeda dalam distribusi data.

Secara keseluruhan, OPTICS terbukti lebih unggul dalam menangani data dengan variasi kepadatan yang tinggi seperti pada dataset Mall Customers. Algoritma ini tidak hanya menghasilkan clustering yang lebih baik, tetapi juga memberikan visualisasi yang lebih informatif tentang pola distribusi pelanggan. Oleh karena itu, OPTICS lebih direkomendasikan untuk digunakan dalam kasus analisis pelanggan yang melibatkan variasi signifikan dalam perilaku pembelian dan pendapatan, serta dalam identifikasi segmen pelanggan untuk keperluan pemasaran yang lebih efektif dan strategis.

## DAFTAR PUSTAKA

- [1] J. Ekonomi Bisnis, V. Febrianti, P. Komarudin, dan U. Islam Kalimantan MAB, "Analisis strategi pemasaran dalam meningkatkan traffic pengunjung pada shopping centre Citimall Kuala Kapuas," *J. Ekonomi Bisnis dan Akuntansi*, vol. 2, no. 1, 2023.
- [2] T. M. Dista dan F. F. Abdulloh, "Clustering pengunjung mall menggunakan metode K-Means dan Particle Swarm Optimization," *Jurnal Media Informatika Budidarma*, vol. 6, no. 3, p. 1339, Jul. 2022, doi:

10.30865/mib.v6i3.4172.

- [3] M. M. Pausy dan S. Nugroho, "Fungsi rekreatif shopping mall di Jakarta," 2020. [Online]. Available: <https://jurnal.ft.uns.ac.id/index.php/senthong/index>.
- [4] A. F. Hapsari, "The Unique Shopping Family Mall sebagai strategi pencitraan Galeria Mall Yogyakarta."
- [5] F. D. S. Rompis dan S. 2 Abstrak, "Hibridisasi fungsi pasar tradisional dan mall. Suatu studi imajinatif," *The Function Hybridization of Traditional Market and Mall. An Imaginative Study*.
- [6] Subagio, "Pengaruh lingkungan sosial terhadap perilaku konsumtif mahasiswa Fakultas Pendidikan Olahraga dan Kesehatan (FPOK) IKIP Mataram," *Jurnal Ilmu Sosial dan Pendidikan*, vol. 3, no. 3, 2019.
- [7] V. G. Ranika, A. Mutrofin, dan E. C. A. Nathania, "Analisis preferensi mall culture dalam gaya hidup konsumtif," *Wissen: Jurnal Ilmu Sosial dan Humaniora*, vol. 2, no. 3, pp. 89–96, Jun. 2024, doi: 10.62383/wissen.v2i3.175.
- [8] S. H. Hastuti, A. Septiani, H. Hendrayani, dan W. P. Nurmayanti, "Penerapan metode OPTICS dan ST-DBSCAN untuk klasterisasi data kesehatan," *Edumatic: Jurnal Pendidikan Informatika*, vol. 8, no. 1, pp. 252–261, Jun. 2024, doi: 10.29408/edumatic.v8i1.25765.
- [9] D. Rifaldi, A. Fadlil, dan Herman, "Teknik preprocessing pada text mining menggunakan data tweet 'Mental Health'," *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 3, no. 2, pp. 161–171, Apr. 2023, doi: 10.51454/decode.v3i2.131.
- [10] F. Alghifari dan D. Juardi, "Penerapan data mining pada penjualan makanan dan minuman menggunakan metode algoritma Naïve Bayes."
- [11] M. Namvar, M. R. Gholamian, dan S. Khakabi, "A two phase clustering method for intelligent customer segmentation," *ISMS 2010 - UKSim/AMSS 1st International Conference on Intelligent Systems, Modelling and Simulation*, 2019, pp. 215–219, doi: 10.1109/ISMS.2010.48.
- [12] E. P. Anggita, "Penerapan FP-Growth dalam menganalisis data penjualan di Toko X," 2023. [Online]. Available: [www.researchgate.net](http://www.researchgate.net).
- [13] P. A. Jusia, F. M. Irfan, dan S. Dinamika Bangsa Jambi, "Clustering data untuk rekomendasi penentuan jurusan perguruan tinggi menggunakan metode K-Means," 2019.
- [14] N. L. P. Dewi, I. N. Purnama, dan N. W. Utami, "Penerapan data mining untuk clustering penilaian kinerja dosen menggunakan algoritma K-Means (studi kasus: STMIK Primakara)," *Jurnal Ilmiah Teknologi Informasi Asia*, vol. 16, no. 2, 2022.
- [15] F. Sukmana, B. Firmansyah, dan W. Sa'adah, "Implementasi ISO 9126 dan Fishbone Analisis pada sistem perpustakaan sekolah di UPT SD Negeri 27 Gresik," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 8, no. 1, pp. 345–534, Feb. 2023, doi: 10.29100/jipi.v8i1.3305.