

Mengidentifikasi Strategi Promosi pada Jasa Penjualan Saldo Digital menggunakan Pendekatan *Clustering*

Hamka Surya Nugraha ^{1,*}, Hilman Mutaqin ¹, Adittia Fathah ¹, Christina Juliane ¹

¹ Program Studi Sistem Informasi, STMIK LIKMI Bandung, Indonesia

* Correspondence: hamka@sttcipasung.ac.id

Copyright: © 2023 by the authors

Received: 17 Desember 2022 | Revised: 22 Desember 2022 | Accepted: 7 Januari 2023 | Published: 20 Juni 2023

Abstrak

Pesatnya perkembangan teknologi informasi telah memberikan dampak terhadap berbagai bidang, salah satunya adalah telepon seluler. Semakin banyaknya masyarakat yang memiliki telepon seluler, berdampak pada kebutuhan produk digital khususnya saldo digital yang semakin meningkat. Tujuan penelitian ini untuk mengidentifikasi penjualan saldo digital agar menjadi masukan khususnya strategi promosi konter ZAR. Penelitian ini menggunakan metode *clustering* dan Algoritma *K-Means* agar didapatkan beberapa cluster jenis transaksi berdasarkan besaran nominal harga penjualan, *silhouette* digunakan untuk mencari jumlah *cluster*. Dataset yang digunakan pada penelitian ini adalah penjualan saldo digital konter ZAR dari bulan November 2021 sampai bulan Oktober 2022. Penelitian ini menghasilkan jumlah *cluster* sebanyak 2 buah. *Cluster* pertama dengan penjualan nominal harga rendah sebanyak 49.213 data, sedangkan *cluster* kedua dengan penjualan nominal harga tinggi sebanyak 3.076 data. Hasil penelitian menunjukkan bahwa algoritma *k-means* dapat digunakan untuk *clustering* penjualan saldo digital konter ZAR dan penjualan saldo digital konter ZAR cukup baik namun ada data di cluster nominal harga tinggi. Pemilik konter ZAR dapat membuat strategi promosi dengan memberikan potongan harga dibawah 142.000 sehingga pelanggan yang membeli dengan nominal harga tinggi pindah ke nominal harga rendah.

Kata kunci: *clustering; k-means; saldo digital; silhouette*

Abstract

The rapid development of information technology has impacted various fields, one of which is cellular telephony. The increasing number of people who have cell phones impacts the need for digital products, especially digital balances, which are increasing. This study aims to identify digital balance sales so that they become input, especially the ZAR counter-promotion strategy. This study uses the clustering method and the K-Means Algorithm to obtain several clusters of transaction types based on the nominal sales price. The silhouette is used to find the number of clusters. The dataset used in this study is the sale of digital ZAR counterbalances from November 2021 to October 2022. This research resulted in a total of 2 clusters. First cluster with low price nominal sales was 49,213 data, while second with high nominal price sales was 3,076. The study results show that the k-means algorithm can be used to cluster sales of ZAR counter digital balances, and ZAR counter digital balances are quite good. Still, there is data in clusters with high nominal prices. ZAR counter owners can create promotional strategies by providing discounts below 142,000 so that customers who buy at a high nominal price move to a lower nominal price.

Keyword: *clustering; k-means; digital balance; silhouette*

PENDAHULUAN

Pesatnya perkembangan dunia teknologi informasi telah memberikan dampak yang baik terhadap berbagai bidang (Hermawan et al., 2022), salah satunya adalah alat komunikasi



elektronik. Perkembangan alat komunikasi elektronik yang semakin murah dan berkembang dengan pesat seperti telepon seluler menjadikan kebutuhan saldo digital seperti saldo dompet digital, saldo angkutan *online*, token listrik maupun saldo pulsa semakin berkembang (Hadikristanto & Setyaningsih, 2020). Kebutuhan terhadap saldo digital yang semakin berkembang membuat transaksi menjadi lebih banyak sehingga data yang dibuat juga semakin banyak. Salah satu contohnya adalah konter ZAR. Jumlah transaksi setiap bulannya lebih dari 4500 transaksi. Penjual hanya mencatat transaksi, namun tidak tahu data tersebut harus diapakan. Kondisi tersebut menyebabkan penjual kebingungan untuk menentukan strategi promosi yang tepat. Adanya pengumpulan data bisa memberikan solusi dan manfaat bagi pemilik usaha (Wulandari et al., 2022) seperti penentuan strategi promosi. Padahal promosi memiliki peran yang penting bagi pengusaha (Nurdiyansyah et al., 2018).

Agar strategi promosi yang dibuat bisa tepat maka perlu melakukan analisa data untuk mengetahui karakteristik konsumen berdasarkan nominal pembelian. *Data mining* merupakan proses penemuan pola yang tidak terlihat yang menghasilkan pengetahuan yang tidak diketahui sebelumnya dari kumpulan data (Pramadhana, 2021; Takdirillah, 2020). *Data mining* merupakan sebuah aktifitas yang bisa digunakan untuk mempelajari kebiasaan pelanggan sehingga pemilik usaha dapat mengembangkan strategi bisnis yang baik (Purwani et al., 2022; Sihombing et al., 2021). Objek dalam *data mining* adalah data yang sangat banyak serta kompleks (Haryatmi & Pramita Hervianti, 2021; Rahayu et al., 2022; Uska et al., 2020). *Data preprocessing* adalah proses untuk mempersiapkan data mentah sebelum diproses lebih lanjut (Aqila & Bijaksana, 2020). Teknik *Clustering* merupakan suatu pendekatan penting yang digunakan untuk mencari keseragaman di dalam data serta menempatkan data yang seragam ke dalam kelompok-kelompok (Syahputra et al., 2018). Pengelompokan atau *clustering* merupakan suatu cara yang digunakan untuk membuat pengelompokan data kedalam sebuah himpunan data atau klaster data (Nishom & Fathoni, 2018). Membagikan jumlah data yang sangat besar kedalam grup data yang lebih kecil akan sangat membantu dalam meringkas data serta untuk berbagai penggunaan data mining. Sebuah pengertian secara intuitif adalah diberikan jumlah dataset lalu kelompokan data tersebut ke dalam grup yang mengandung kemiripan data poin. Ada banyak variasi model yang dikembangkan untuk analisis clustering dimana model yang berbeda akan bekerja lebih baik tergantung dari perbedaan skenario dan tipe data (Richie & Halim Agung, 2020). Algoritma yang dipakai pada penelitian adalah *K-Means*. Pendekatan yang dipakai adalah *Silhouette Method*.

K-Means merupakan teknik komputasi yang dilakukan dengan efisien, serta menjadi salah satu dari algoritma *clustering* yang cukup banyak digunakan untuk metode klasterisasi (Anggarwal, 2015; Ashari et al., 2022). *K-Means* merupakan algoritma tanpa pengawasan yang digunakan untuk penyelesaian masalah *clustering* (Widiawati, 2021). K' dalam *K-Means clustering* merupakan banyaknya jumlah grup atau kluster. Tujuan dari metodologi data mining ini adalah untuk melakukan observasi dari tiap nilai atribut individu dan membandingkan atribut tersebut dengan nilai *means*, atau dengan kata lain nilai rata-rata, pengamatan dilakukan dengan grup lain yang memiliki potensi agar dapat menemukan grup yang mirip satu sama lain (North, 2012). *K-Means* merupakan algoritma yang mudah serta efektif untuk menemukan kluster didalam data (Larose & Larose, 2005).

Metode *Silhouette* digunakan untuk mencari jumlah *cluster* agar implementasi *K-Means clustering* menjadi optimal. Metode ini menghitung koefisien *Silhouette* dari setiap titik data untuk mengukur seberapa mirip sebuah titik dengan *cluster*-nya sendiri. Nilai siluet berkisar antara 1 sampai -1, dimana semakin tinggi nilainya menandakan bahwa objek tersebut lebih cocok dengan *cluster*-nya sendiri dan tidak cocok dengan cluster lainnya (Nuiiaa et al., 2022). Dengan diketahui kesamaan dan pengelompokan dalam data yang berarti karakteristik konsumen maka strategi promosi dapat lebih mudah dirumuskan sehingga strategi akan lebih tepat sasaran (Dana et al., 2018).

Orange adalah sebuah aplikasi *open source* penambang data. Aplikasi Orange dapat digunakan untuk analisis serta visualisasi data. Hal ini membuat aplikasi tersebut bisa digunakan untuk pemilihan eksperimen, pemodelan peramalan, dan sistem rekomendasi. Aplikasi Orange selalu disukai dalam faktor inovatif, kualitas dan kendalan. Aplikasi Orange mempermudah penggunaannya melakukan eksperimen dengan data serta proses data analisis secara intuitif (Wiguna & Rifai, 2021).

Penelitian yang pernah dilakukan oleh (Astuti et al., 2022) menggunakan algoritma *K-Means*. Menghasilkan tiga *cluster* yang dibuat yaitu 114 produk yang laku, 5 kurang laku, dan 14 tidak laku. Penelitian lain juga pernah dilakukan oleh (Palembang & Palembang, 2021) menggunakan algoritma *K-Means* menghasilkan 3 *cluster* dengan kategori tingkat harga cabai rawit tinggi ada 10 provinsi, tingkat harga cabai rawit rendah ada 12 provinsi dan tingkat harga cabai rawit sedang dengan 12 provinsi. Sedangkan penelitian yang dibuat oleh (Siburian et al., 2019) menggunakan algoritma *K-Means* mengelompokkan harga eceran beras berdasarkan 33 provinsi. Dari penelitian tersebut menghasilkan 3 *clustering* dengan kategori *cluster* tinggi ada 11 provinsi, *cluster* sedang ada 11 provinsi dan *cluster* rendah ada 11 provinsi. Penelitian yang dibuat oleh (Indriyani & Irfiani, 2019) menggunakan algoritma *K-Means* menghasilkan 3 *cluster* yaitu paling laris, cukup laris dan kurang laris. 2 jenis produk masuk kedalam kategori sangat laris, 8 jenis produk masuk kedalam kategori cukup laris, dan 18 jenis produk masuk kedalam kategori kurang laris. Penelitian yang dibuat oleh (Mardalius, 2018) menggunakan algoritma *k-means* mengelompokkan aksesoris yang laku dengan 4 data, kurang laku dengan 10 data dan tidak laku dengan 12 data. Penelitian sebelumnya menggunakan jumlah data yang digunakan sebanyak 26 sampai 133 data sedangkan pada penelitian ini jumlah yang digunakan mencapai 52.388 data.

Penelitian sebelumnya analisis *clustering* dilakukan dengan algoritma *K-Means* dan pendekatan *Euclidean Distance*. Pada penelitian ini menggunakan pendekatan berbeda, yaitu *silhouette* serta menggunakan jumlah data yang lebih banyak yaitu sebanyak 52.388 sehingga pola yang terbentuk bisa digunakan sebagai acuan untuk konter lain karena polanya lebih generik. Perangkat lunak yang digunakan pada penelitian sebelumnya adalah Rapidminer sedangkan pada penelitian ini adalah Orange, serta yang dianalisis pada penelitian sebelumnya adalah data penjualan toko sedangkan penelitian ini adalah nominal pembelian. Berdasarkan hasil *clustering* diperoleh informasi pola pembelian sehingga lebih mudah untuk menentukan pola strategi promosi yang tepat dibandingkan dengan penelitian sebelumnya.

Penelitian ini bertujuan untuk mengidentifikasi penjualan saldo digital yang didasari pada data penjualan saldo digital seperti pulsa, kuota, *e-money* dan lain-lain dari bulan November 2021 sampai dengan bulan Oktober tahun 2022 agar dapat menjadi masukan kepada pemilik konter ZAR, khususnya strategi promosi konter ZAR. Dataset yang diolah berjumlah 52.388.

METODE

Penelitian ini diawali dengan melakukan studi literatur, yaitu mencari referensi-referensi dari jurnal ilmiah serta buku yang berkaitan dengan penelitian sehingga dapat mendukung kegiatan penelitian. Tahap selanjutnya adalah mengumpulkan data. Pada tahap ini, dataset yang digunakan adalah *dataset* yang diperoleh dari Konter ZAR yang menjual saldo digital yang berada di Tasikmalaya. Data yang diperoleh dari pemilik usaha adalah sebanyak 52.388 data. *File* yang didapatkan tidak bisa langsung diproses, melainkan harus dibersihkan terlebih dahulu. Data yang dihapus adalah atribut tanggal dan no hp. Selain itu, nominal yang kosong akan diisi dengan harga yang sesuai dengan transaksi lain. Setelah data di *cleaning*, selanjutnya data disimpan dan *dieksport* dalam format *csv* agar bisa diproses dengan aplikasi Orange. Setelah *dieksport* selanjutnya data *diimport* ke dalam aplikasi *Orange*. Proses klasterisasi dilakukan otomatis dengan aplikasi *Orange* sehingga tidak perlu mencari *cluster* secara

manual. Jumlah *cluster* ditentukan dengan fitur *k-means* pada aplikasi Orange dengan cara melihat nilai *silhouette* paling tinggi. Setelah fitur *k-means* diatur maka jumlah cluster yang dibuat sebanyak 2 buah dengan nilai *silhouette* 0,846 yaitu C1 dengan nominal rendah dan C2 dengan nominal tinggi. Setelah hasil klasterisasi muncul, tahap terakhir selanjutnya adalah melakukan analisis terhadap hasil klasterisasi tersebut agar menghasilkan informasi baru. Tabel 1 adalah dataset yang diperoleh dari pemilik usaha dan sebelum di *cleaning*. Tanggal merupakan waktu transaksi yang telah dilakukan sedangkan Trx adalah jenis transaksi seperti pulsa, transportasi online, voucher internet dan lain-lain. No. HP adalah nomor pelanggan yang melakukan transaksi serta Harga adalah nominal harga penjualan dari setiap transaksi.

Tabel 1. Dataset penjualan

Tanggal	Trx	No. HP	Harga
01/11/2021	PLS20	xxxxxxxxxx	22,000
	I10	xxxxxxxxxx	12,000
	S5	xxxxxxxxxx	7,000
	GFFP70	xxxxxxxxxx	10,000
	TZ2	xxxxxxxxxx	32,000
.	.	.	.
.	.	.	.
.	.	.	.
	m10	xxxxxxxxxx	10,000

HASIL DAN PEMBAHASAN

Hasil

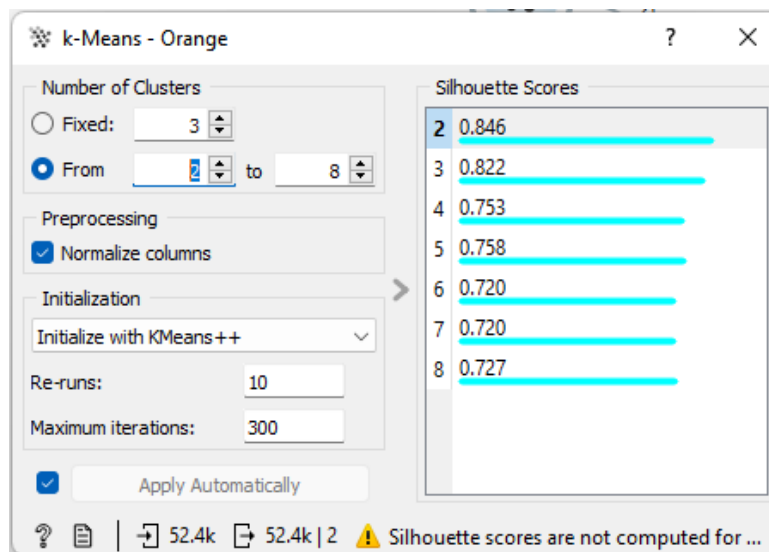
Tabel 1 merupakan dataset yang masih mentah, sehingga perlu dilakukan proses *cleaning* terlebih dahulu agar file siap diproses oleh aplikasi *Orange*. Tabel 2 adalah dataset yang telah di *cleaning*. Tanggal, Trx, dan no hp dihapus serta tanda baca pada kolom harga juga dihapus. Kolom No adalah transaksi yang telah dilakukan mulai dari transaksi ke 1 sampai dengan transaksi ke 52.388 sedangkan kolom Harga adalah nominal harga penjualan dari setiap transaksi.

Tabel 2. Dataset yang telah di *cleaning*

No	Harga
1	22000
2	12000
3	7000
4	10000
5	32000
.	.
.	.
.	.
52388	10000

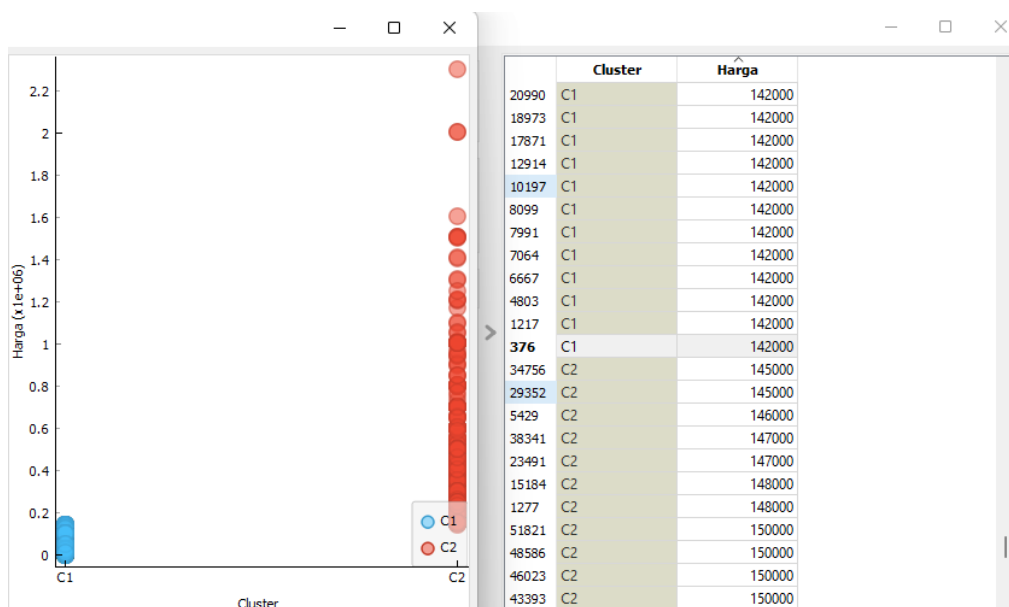
Setelah dataset di *cleaning* selanjutnya adalah diproses menggunakan aplikasi Orange. Fitur *K-Means* dalam aplikasi Orange perlu dilakukan pengaturan terlebih dahulu untuk mendapatkan jumlah cluster yang ideal. Pada gambar 1 merupakan jumlah *cluster* yang dihasilkan berdasarkan aplikasi *orange*, yang diatur dari angka 2 hingga 8 serta pengulangan algoritma sebanyak 10 kali. Berdasarkan pengaturan tersebut, *cluster* yang terbentuk sebanyak 2 buah dengan *silhouette score* sebesar 0,846. Hasil tersebut dibuat secara otomatis oleh

aplikasi sehingga tidak perlu dilakukan secara manual. Jika jumlah cluster dirubah menjadi 3 maka *silhouette score* lebih kecil dari 2 cluster yaitu 0,822 sehingga tingkat kemiripan data berkurang, oleh karena itu dipilih 2 cluster. Cluster yang terbentuk adalah C1 dengan nominal harga rendah dan C2 dengan nominal harga tinggi.



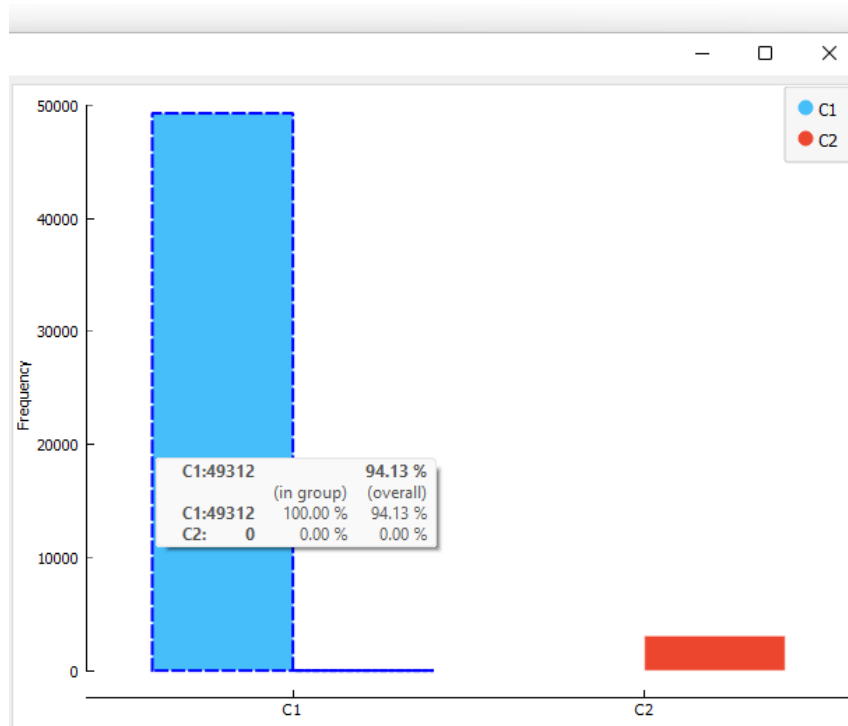
Gambar 1. Jumlah cluster yang dihasilkan

Setelah cluster dibentuk maka langkah selanjutnya adalah menampilkan sebaran data. Gambar 2 menampilkan sebaran data cluster dengan diagram *scatter plot* dengan nominal tertinggi pada cluster C1 adalah 142.000. Pada gambar 2, nilai *silhouette* tidak ditampilkan karena jumlah data lebih dari 5.000, oleh karena itu grafik yang dihasilkan berbentuk vertikal.

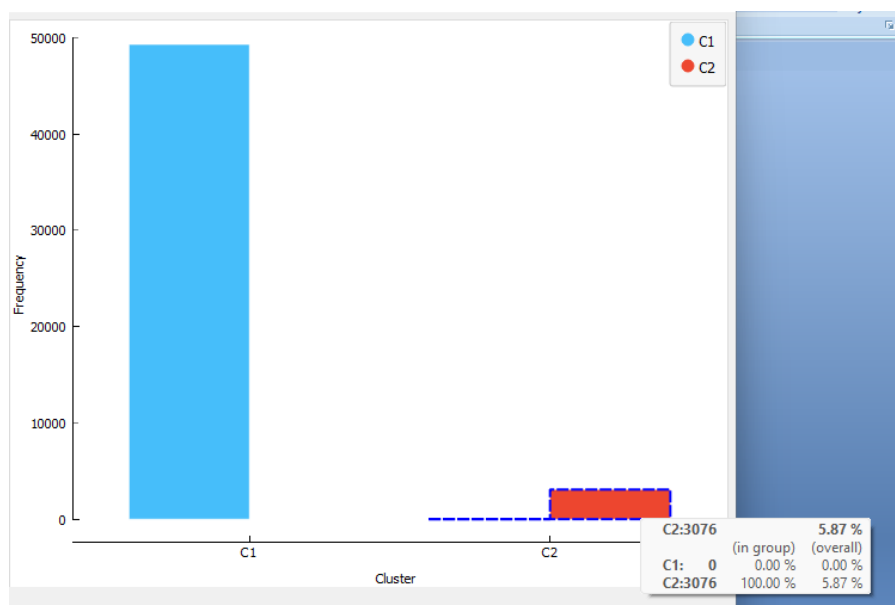


Gambar 2. Sebaran data berdasarkan cluster

Jumlah data pada cluster C1 sebanyak 49.312 data dengan persentase 94,13%. Gambar 3 menampilkan jumlah dan persentase data pada cluster C1. Sedangkan jumlah data pada cluster C2 sebanyak 3.076 dengan persentase sebesar 5,87% seperti terlihat pada gambar 4.



Gambar 3. Jumlah dan persentase data pada *cluster* C1



Gambar 4. Jumlah dan persentase data pada *cluster* C1

Tabel 3. *Cluster* dan jumlah data yang terbentuk

No	Cluster	Jumlah	Persentase (%)
1	C1	49.312	94,13
2	C2	3.076	5,87
Jumlah		52.388	100

Jumlah dataset yang digunakan pada penelitian ini sebanyak 52.388 data. Setelah data diproses dengan algoritma *k-means* dan metode *silhouette* maka menghasilkan 2 *cluster* yaitu *cluster* C1 dengan nominal harga rendah sebanyak 49.312 data dan persentase sebesar 94,13% dan

cluster C2 dengan nominal harga tinggi sebanyak 3.076 data dan persentase sebesar 5,87% berdasarkan hasil yang disajikan pada tabel 3.

Pembahasan

Data penjualan saldo digital konter ZAR dari bulan November 2021 sampai dengan bulan Oktober 2022 sebanyak 52.388 data. Sebelum dataset tersebut diproses dengan aplikasi Orange, terlebih dahulu dilakukan proses *cleaning* yaitu menghapus kolom tanggal, Trx dan No HP. Setelah dataset di *cleaning* selanjutnya adalah di *export* menjadi file *csv* agar bisa diproses dengan aplikasi Orange. Jumlah *cluster* ditentukan oleh aplikasi dengan melakukan pengaturan pada fitur *K-Means* terlebih dahulu yaitu *number of clusters* diatur dari angka 2 sampai 8. Penelitian ini menggunakan metode *silhouette* pada saat penentuan jumlah cluster karena dilakukan secara otomatis oleh aplikasi *Orange* pada fitur *k-means* sehingga lebih cepat dan efisien, sedangkan metode *elbow* hasilnya tidak bisa langsung diketahui tetapi harus membandingkan nilai *centroid* secara manual sehingga memerlukan lebih banyak waktu dan kurang efisien. *Cluster* yang terbentuk adalah 2 yaitu C1 dengan nominal harga rendah dan C2 dengan nominal harga tinggi. Nominal tertinggi *cluster* C1 adalah 142.000 Setelah jumlah *cluster* ditentukan maka langkah selanjutnya adalah menampilkan hasil data dengan diagram *Scatter Plot*. Jumlah data pada *cluster* nominal harga rendah sebanyak 49.312 dengan persentase 94,13% dan jumlah data pada *cluster* nominal harga tinggi sebanyak 3.076 dengan persentase 5,87%.

Hasil pada tabel 3 menjelaskan bahwa penjualan saldo digital konter ZAR cukup baik dengan jumlah data nominal harga rendah lebih banyak daripada transaksi dengan jumlah data nominal harga tinggi, namun ada 3.076 data transaksi dengan harga tinggi. Menurut pemilik konter ZAR, semakin kecil harga transaksi maka semakin tinggi margin keuntungan yang didapat karena jumlah transaksi yang dilakukan akan lebih banyak. Oleh karena itu salah satu masukan strategi promosi yang bisa dilakukan adalah dengan membuat potongan harga dengan persentase tertentu pada harga dibawah Rp. 142.000 sehingga pelanggan yang biasanya membeli dengan nominal harga tinggi pindah ke nominal harga rendah.

Penelitian sebelumnya yang dilakukan oleh (Dana et al., 2018) tentang *data mining* dan menggunakan algoritma *k-means* menggunakan metode *elbow* untuk mencari jumlah *cluster* yang optimal dengan menghasilkan 3 *cluster*, sementara penelitian ini menggunakan metode *silhouette* untuk mencari jumlah *cluster* yang optimal dengan menghasilkan 2 *cluster*. Untuk mencari jumlah cluster, metode *elbow* hasilnya tidak bisa langsung diketahui tetapi harus membandingkan nilai *centroid* secara manual sedangkan metode *silhouette* pada penelitian ini dilakukan dengan otomatis oleh aplikasi. Penelitian sebelumnya menggunakan data mahasiswa berdasarkan asal kota untuk menentukan strategi marketing sedangkan penelitian ini menggunakan data penjualan saldo digital konter untuk menentukan strategi promosi. Hal ini membuktikan relevansi penelitian sebelumnya dengan penelitian ini.

SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, diketahui bahwa algoritma *k-means* dan metode *silhouette* bisa digunakan untuk proses klasterisasi penjualan saldo digital pada konter ZAR. Analisa serta pengujian dengan algoritma *k-means* dalam aplikasi Orange dapat diketahui klasterisasi data berdasarkan besarnya nominal pembelian yang menghasilkan 2 *cluster* yaitu *cluster* nominal harga rendah dan *cluster* nominal harga tinggi. Dari hasil penerapan algoritma *k-means* didapatkan hasil nominal harga rendah lebih banyak daripada nominal harga tinggi. Penjualan saldo digital konter ZAR cukup baik namun diharapkan kedepannya penjualan saldo digital konter ZAR bisa lebih maksimal.

REFERENSI

- Anggarwal, C. (2015). Data Mining. In *IBM T.J. Watson Research Center* (Vol. 14, Issue 3). Springer.
- Aqila, N., & Bijaksana, M. arif. (2020). Developing Set of Word Senses of Vocabulary in Al-Qur'an. *Edumatic : Jurnal Pendidikan Informatika*, 4(1), 83–90. <https://doi.org/10.29408/edumatic.v4i1.2119>
- Ashari, I. A., Negara, I. S. M., & Sumantri, R. B. B. (2022). Evaluasi Pembayaran Keuangan Siswa berdasarkan Penghasilan Wali Siswa menggunakan Metode Clustering K-Means. *Edumatic: Jurnal Pendidikan Informatika*, 6(2), 324–333. <https://doi.org/10.29408/edumatic.v6i2.6395>
- Astuti, N., Utamajaya, J. N., & Pratama, A. (2022). Penerapan Data Mining Pada Penjualan Produk Digital Konter Leppangeng Cell Menggunakan Metode K-Means Clustering. *Jurnal Riset Komputer*, 9(3), 2407–389. <https://doi.org/10.30865/jurikom.v9i3.4351>
- Dana, R. D., Rohmat, C. L., & Rinaldi, A. R. (2018). Strategi Marketing Penerimaan Mahasiswa Baru Menggunakan Machine Learning dengan Teknik Clustering. *Jurnal Pengembangan IT (JPIT)*, 04(2), 201–204. <https://doi.org/10.30591/jpit.v4i2-2.1879>
- Hadikristanto, W., & Setyaningsih, A. (2020). Penerapan Algoritma Naive Bayes Untuk Menentukan Klasifikasi Produk Terlaris Pada Penjualan Pulsa. *Sigma*, 11(2), 49–58.
- Haryatmi, E., & Pramita Hervianti, S. (2021). Penerapan Algoritma Support Vector Machine Untuk Model Prediksi Kelulusan Mahasiswa Tepat Waktu. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(2), 386–392. <https://doi.org/10.29207/resti.v5i2.3007>
- Hermawan, A., Hartati, T., & Wijaya, Y. A. (2022). Analisa Keamanan Data melalui Website Zahra Software Menggunakan Metode Keamanan Informasi CIA Triad. *JPIT*, 7(3), 125–130.
- Indriyani, F., & Irfiani, E. (2019). Clustering Data Penjualan pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means. *JUITA : Jurnal Informatika*, 7(2), 109–114. <https://doi.org/10.30595/juita.v7i2.5529>
- Larose, D. T., & Larose, C. D. (2005). *Discovering Knowledge in Data: an Introduction to Data Mining* (Vol. 100, Issue 472). New Jersey: John Wiley & Sons. <https://doi.org/10.1002/0471687545>
- Mardalius, M. (2018). Pemanfaatan Rapid Miner Studio 8.2 Untuk Pengelompokan Data Penjualan Aksesoris Menggunakan Algoritma K-Means. *Jurteks*, 4(2), 123–132. <https://doi.org/10.33330/jurteks.v4i2.36>
- Nishom, M., & Fathoni, M. Y. (2018). Implementasi Pendekatan Rule-Of-Thumb untuk Optimasi Algoritma K-Means Clustering. *Jurnal Informatika: Jurnal Pengembangan IT*, 3(2), 237–241. <https://doi.org/10.30591/jpit.v3i2.909>
- North, M. (2012). *Data Mining For the Masses*. Georgia: A Global Text Project.
- Nuiaa, R. R., Alsaeedi, A. H., Manickam, S., & Al-Shammary, D. E. J. (2022). Evolving Dynamic Fuzzy Clustering (EDFC) to Enhance DRDoS DNS Attacks Detection Mechnism. *Intelligent Engineering and Systems*, 15(1), 509–519. <https://doi.org/10.22266/IJIES2022.0228.46>
- Nurdiyansyah, F., Arifin, S., & Marisa, F. (2018). Penerapan Clustering Algorithm Untuk Mendukung Promosi Server Pulsa Reload. *J I M P - Jurnal Informatika Merdeka Pasuruan*, 3(2), 73–78. <https://doi.org/10.37438/jimp.v3i2.174>
- Palembang, C. F., & Palembang, S. P. (2021). Pengelompokan Tingkat Harga Cabai Rawit Berdasarkan Provinsi di Indonesia Menggunakan Algoritma K-Means Clustering. *Variance*, 3(2), 48–60. <https://doi.org/10.30598/variancevol3iss2page48-60>
- Pramadhana, D. (2021). Klasifikasi Penyakit Diabetes Menggunakan Metode CFS Dan ROS dengan Algoritma J48 Berbasis Adaboost. *Edumatic: Jurnal Pendidikan Informatika*, 5(1), 89–98. <https://doi.org/10.29408/edumatic.v5i1.3336>

- Purwani, F., Wahyudi, R. T., & Jaya, I. D. (2022). Penerapan Algoritma K-Nearest Neighbor dengan Euclidean Distance untuk Menentukan Kelompok Uang Kuliah Tunggal Mahasiswa. *Edumatic: Jurnal Pendidikan Informatika*, 6(2), 344–353. <https://doi.org/10.29408/edumatic.v6i2.6547>
- Rahayu, S., Yumarlin, M. Z., Bororing, J. E., & Hadiyat, R. (2022). Implementasi Metode K-Nearest Neighbor (K-NN) untuk Analisis Sentimen Kepuasan Pengguna Aplikasi Teknologi Finansial FLIP. *Edumatic: Jurnal Pendidikan Informatika*, 6(1), 98–106. <https://doi.org/10.29408/edumatic.v6i1.5433>
- Siburian, T., Safii, M., & Parlina, I. (2019). Penerapan Algoritma K-Means Clustering untuk Pengelompokan Harga Eceran Beras di Pasar Tradisional Berdasarkan Wilayah Kota. *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, 1(100), 927–936. <https://doi.org/10.30645/senaris.v1i0.101>
- Sihombing, L. O., Hannie, H., & Dermawan, B. A. (2021). Sentimen Analisis Customer Review Produk Shopee Indonesia Menggunakan Algoritma Naïve Bayes Classifier. *Edumatic: Jurnal Pendidikan Informatika*, 5(2), 233–242. <https://doi.org/10.29408/edumatic.v5i2.4089>
- Syahputra, T., Halim, J., & Sintho, E. P. (2018). Penerapan Data Mining Dalam Menentukan Pilihan Jurusan Bidang Studi SMA Menggunakan Metode. *Jurteks*, 4(2), 1–4.
- Takdirillah, R. (2020). Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Penjualan Bisnis Ritel. *Edumatic: Jurnal Pendidikan Informatika*, 4(1), 37–46. <https://doi.org/10.29408/edumatic.v4i1.2081>
- Uska, M., Wirasasmita, R., Usuluddin, U., & Arianti, B. (2020). Evaluation of Rapidminer-Application in Data Mining Learning using PeRSIVA Model. *Edumatic: Jurnal Pendidikan Informatika*, 4(2), 164–171. <https://doi.org/10.29408/edumatic.v4i2.2688>
- Widiawati, C. R. A. dkk. (2021). K-Means Clustering Berdasarkan Otsu Thresholding untuk Segmentasi Inti. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 1(5), 907–914. <https://doi.org/https://doi.org/10.29207/resti.v4i5.2309>
- Wiguna, R. A. raffaidy, & Rifai, A. I. (2021). Analisis Text Clustering Masyarakat Di Twitter Mengenai Omnibus Law Menggunakan Orange Data Mining. *Journal of Information Systems and Informatics*, 3(1), 1–12. <https://doi.org/10.33557/journalisi.v3i1.78>
- Wulandari, N., Farida, I. N., & Mahdiyah, U. (2022). Implementasi Metode K-Means Clustering Dalam Pengadaan Barang Di Toko N-Case. *Prosiding SEMNAS Inovasi Teknologi*, 6(1), 308–313.