

Penerapan K-Means pada Pengelompokan Penjualan Produk Smartphone

Fatimah Putri Arfani Hasibuan ^{1*}, Sumarno Sumarno ², Iin Parlina ³

^{1,2}Program Studi Teknik Informatika, STIKOM Tunas Bangsa, Pematangsiantar

³Program Studi Komputerisasi Akuntansi, AMIK Tunas Bangsa, Pematangsiantar

Email: ^{1*} fatimahputriarfanihasibuan@gmail.com

ABSTRACT

The development of smartphone technology companies that are increasingly rapidly today has more or less led to competition between these companies, thus requiring developers to find strategies or patterns that can increase product sales and marketing, as is the case at PT. Vivo Communication Indonesia Pematangsiantar branch. One of the strategies is to utilize transaction data by grouping them to see which products are more salable in the market and which are not, so that evaluations can be made in planning the next Vivo product promotion. The clustering method in this study uses the K-Means Clustering method. The K-Means method is a data mining method that is able to group some data into certain parts. In this paper the sales data that has been obtained will be divided into 3 groups, namely low sales, medium sales and high sales. Based on the results of testing using the K-Means method on Vivo smartphone sales data, it was found that the highest sales group had only 1 data, namely Vivo Y12 3+32GB. So it can be concluded that the K-Means method can be applied to group Vivo smartphone sales, because it is in accordance with the actual sales results.

Keywords: Clustering, Sales, Vivo, Data Mining, K-Means.

ABSTRAK

Perkembangan perusahaan teknologi Smartphone yang semakin pesat dewasa ini sedikit banyaknya menimbulkan persaingan diantara perusahaan tersebut, sehingga menuntut para pengembang untuk menemukan strategi atau suatu pola yang dapat meningkatkan penjualan dan pemasaran produk, seperti halnya pada PT. Vivo Communication Indonesia cabang Pematangsiantar. Salah satu strateginya adalah memanfaatkan data transaksi dengan melakukan pengelompokan untuk melihat produk mana yang lebih laku dipasaran dan mana yang tidak, agar dapat dilakukan evaluasi dalam perencanaan promosi produk Vivo selanjutnya. Metode pengelompokan pada penelitian ini menggunakan metode Clustering K-Means. Metode K-Means adalah metode data mining yang mampu mengelompokkan beberapa data menjadi bagian-bagian tertentu. Pada makalah ini data penjualan yang telah diperoleh akan dibagi menjadi 3 kelompok yaitu penjualan rendah, penjualan sedang dan penjualan tinggi. Berdasarkan hasil Pengujian menggunakan metode K-Means terhadap data penjualan smartphone Vivo, diperoleh bahwa penjualan kelompok tertinggi hanya ada 1 data, yakni Vivo Y12 3+32GB. Sehingga dapat disimpulkan bahwa metode K-Means dapat diterapkan untuk melakukan pengelompokan penjualan smartphone Vivo, karena sesuai dengan hasil penjualan yang sebenarnya..

Kata Kunci: Pengelompokan, Penjualan, Vivo, Data Mining, K-Means

1. Pendahuluan

Perkembangan pengguna smartphone dari tahun ke tahun semakin meningkat di Indonesia. Hal ini tidak lepas menjamurnya berbagai varian smartphone. Pada Era revolusi Industri 4.0 ini, smartphone sudah menjadi industri bisnis yang diperhitungkan karena cukup menguntungkan. Sehingga persaingan pasar smartphone semakin lama semakin ketat [1]. Sehingga tidak mengherankan apabila banyak sekali beredar merek smartphone. Banyak produsen baru yang muncul

menyajikan produk unggulannya. Hampir setiap waktu, para produsen smartphone ini selalu melakukan inovasi, seperti halnya PT. Vivo Communication Indonesia cabang Pematangsiantar. Jika tidak, tentu hal itu dapat menjadikan awal kehancuran dari bisnis tersebut. Memasuki pertengahan Juli 2021, sejumlah Lembaga riset mulai mengeluarkan laporan kuartal kedua pasar smartphone dunia, salah satunya Canalys. Berdasarkan laporan terbaru Canalys pada Quartile 2 tahun 2021, tercatat bahwa smartphone global

pengirimannya meningkat 12%, ketika diseluruh dunia mulai menggunakan, dan mulai terbentuknya kebiasaan warga maupun ekonomi negara. Canalsy menjabarkan lima besar vendor smartphone global berturut-turut pada kuartal 2 tahun 2021 antara lain: Samsung (19%), Xiaomi (17%), Apple (14%), Oppo (10%), dan Vivo (10%) [2]. Berdasarkan data tersebut, maka PT. Vivo perlu menemukan strategi atau suatu pola yang mampu meningkatkan jumlah penjualan serta pemasaran produk mereka dimasa masa berikutnya. Salah satu strateginya adalah memanfaatkan data transaksi dengan melakukan pengelompokan untuk melihat produk Vivo mana yang banyak terjual dipasaran dan mana yang tidak, agar dapat dilakukan evaluasi dalam perencanaan promosi produk Vivo selanjutnya.

Banyak metode-metode yang menjadi alternatif untuk membantu penentu kebijakan, salah satunya dengan menggunakan Data Mining [3]–[10]. Pada sektor pemasaran/penjualan, data mining umumnya dimanfaatkan untuk target pemasaran (target penjualan), hubungan manajemen pelanggan (CRM), analisa pasar, *cross selling* hingga melakukan segmentasi pasar. Seperti menemukan kelompok/model pelanggan yang mempunyai karakteristik yang sama: minat, tingkat penghasilan, kebiasaan untuk berbelanja, dll. Selain itu untuk menentukan pelanggan yang memiliki pola pembelian dari waktu ke waktu. Menemukan hubungan / keterkaitan antar produk yang dijual, maupun prediksi berlandaskan asosiasi tersebut serta jenis pelanggan apa yang membeli produk apa, karena salah satu keunggulan data mining mampu melakukan pengelompokan [11]–[14] maupun klasifikasi [15]–[19], hingga identifikasi produk terbaik untuk berbagai kelompok pelanggan, memprediksi faktor apa yang akan menarik pelanggan baru, penyediaan informasi ringkasan, laporan ringkasan multidimensi, Informasi ringkasan statistik (kecenderungan dan variasi pusat data). Akan tetapi banyak nya metode Data Mining yang dapat digunakan untuk mengatasi masalah pengelompokan seperti K-Means [20], K-Medoids [21], X-Means [22], menimbulkan kebingungan tersendiri bagi pengambil kebijakan di PT. Vivo Communication Indonesia cabang Pematangsiantar. Sehingga pada makalah ini diusulkan pengelompokan data penjualan smartphone Vivo menggunakan metode K-Means.

Beberapa penelitian terdahulu yang berhubungan dengan topik ini diantaranya: Penelitian untuk mengelompokkan Laju Pertumbuhan PDRB Kota Surabaya menggunakan Metode K-Means. Data dibagi menjadi 3 cluster: tinggi, sedang dan rendah. Hasil yang diperoleh terdapat 9 kategori/sector dengan klaster tinggi, 5 kategori/sector dengan klaster sedang dan 3 kategori/sector dengan klaster rendah [23]. Selanjutnya penelitian yang dilakukan untuk mengelompokkan jumlah kasus terkonfirmasi Covid-19 dan jumlah kematian akibat virus ini di Asia Tenggara. Data yang digunakan adalah data statistik negara berdasarkan area kasus COVID-19 yang dikonfirmasi dan tercatat pada laboratorium, serta kematian pada bulan April 2020 dari WHO (Kesehatan Dunia Organisasi). Pada penelitian ini Data dibagi menjadi 3 cluster: tinggi (C1), sedang (C2) dan rendah (C3). Hasil yang diperoleh terdapat empat negara dengan cluster level tinggi (C1), satu negara dengan cluster level sedang (C2), dan 6 negara dengan cluster level rendah (C3) [24]. Berikutnya adalah penelitian menggunakan metode K-Means untuk mengelompokkan penjualan paket data telkomsel. Data penjualan dari penelitian ini dikelompokkan menjadi 3 bagian, antara lain: penjualan rendah, penjualan sedang dan penjualan tinggi. Pengujian clustering menggunakan K-Means terhadap data penjualan paket telkomsel, menghasilkan persentase kesesuaian 100% dibandingkan pengelompokan manual [25].

Hal-hal terkait tersebut yang melatar belakangi dilakukannya penelitian ini dengan menggunakan metode K-Means, karena Metode K-Means mudah untuk diimplementasikan dan dijalankan, waktu yang diperlukan untuk pembelajaran relatif cepat serta mudah untuk diadaptasi.

2. Metode Penelitian

2.1. Lokasi dan Waktu Penelitian

Data Penelitian diperoleh dari semua Store yang bekerjasama dengan PT. Vivo yang ada dikota Pematangsiantar. Waktu pengumpulan data dilakukan selama bulan April sampai dengan bulan Juni 2020.

Tabel 1. Data Penjualan Vivo

No	Type Smartphone Vivo	Bulan April 2020	Bulan Mei 2020	Bulan Juni 2020
1	S1 Pro 8+128	23	7	2
2	V17 Pro	4	3	6
3	V19 8+128GB	113	95	88
4	V19 8+256GB	9	6	9
5	Y11 2+32GB	71	117	98
6	Y12 3+32GB	996	1154	1133
7	Y12 3+64GB	5	1	25
8	Y154+64 GB	159	194	271

DOI: xxxx

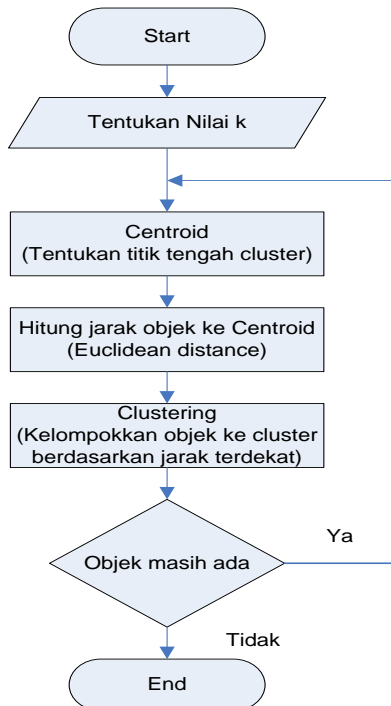
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

9	Y17 4+128GB	162	88	12
10	Y19 6+128GB	106	72	34
11	Y30 4+128GB	37	237	330
12	Y50 8+128GB	526	91	188
13	Y91C 2+32GB	3	213	522

Tabel 1 adalah data tabel penjualan Smartphone Vivo yang dikumpulkan dari semua Gerai HP di Pematangsiantar yang merupakan mitra PT. Vivo Communication Indonesia.

2.2. Flowchart Penelitian

Berikut ini akan disajikan Flowchart penelitian menggunakan metode K-Means.



Gambar 1. Diagram alir dari metode Clustering K-Means [26]

Tahapan demi tahapan metode *K-Means* sebagai berikut [8], [27]–[30]:

1. Atur jumlah *cluster* (*k*) pada data set.
2. Atur nilai pusat (*Centroid*).

Pada tahap awal Pengaturan nilai *Centroid* dilakukan secara acak. Pada tahap iterasi digunakan rumus persamaan (1) seperti berikut.

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \dots\dots\dots(1)$$

Keterangan :

V_{ij} = *Centroid* rata-rata *cluster* ke-I variabel ke-j

N_i = Jumlah anggota *cluster* ke-i

i, k = Indeks *cluster*

j = Indeks variabel

X_{kj} = nilai ke-k variabel ke-j untuk *cluster*

3. Berdasarkan masing-masing *record*, jarak terdekat dihitung dengan *Centroid*.

Beberapa cara yang biasa digunakan untuk melakukan pengukuran jarak data ke pusat *cluster*,

diantaranya *Euclidean*, *Manhattan/City Block*, dan *Minkowsky*. Setiap cara memiliki kelebihan dan kekurangan masing-masing. Untuk penulisan pada bab ini, jarak *Centroid* yang digunakan adalah *Euclidean Distance*, dengan rumus seperti dibawah ini:

$$De = \sqrt{(xi - si)^2 + (yi - ti)^2} \dots\dots\dots(2)$$

Keterangan :

De = *Distance Euclidean*

i = Jumlah objek ²

(x, y) = Koordinat objek

(s, t) = Koordinat *Centroid*

4. Kelompokkan objek berdasarkan jarak ke *Centroid* terdekat

5. Ulangi langkah ke-3 hingga langkah ke-4, lakukan iterasi hingga *Centroid* bernilai optimal

3. Hasil dan Pembahasan

3.1. Penentuan Centroid

Penentuan pusat awal Cluster (*Centroid*) ditentukan secara manual atau random yang diambil dari data yang ada dalam range. Nilai Cluster 0 diambil dari data paling rendah (C1), Nilai Cluster 1 diambil dari data acak atau yang paling tengah/sedang (C2) dan Nilai Cluster 2 diambil dari data yang paling tertinggi (max).

Tabel 2. Centroid Awal

	C1 (Min)	4	3	6
<i>Centroid</i>	C2 (Ave)	159	194	271
	C3 (Max)	996	1154	1133

3.2. Mengitung Jarak Centroid

Perhitungan jarak antara titik *Centroid* dengan titik tiap objek menggunakan rumus Euclidian Distance. (Persamaan 1). Sehingga menghasilkan Jarak dari *Centroid* yang disajikan pada Tabel 3.

Tabel 3. Hasil Perhitungan Iterasi 1

C1	C2	C3	Jarak Terpendek
19,82423	354,7196	1881,887	19,8242276
0	361,5674	1891,823	0
164,5266	213,0868	1730,085	164,5265936
6,557439	355,6515	1885,59	6,557438524
161,0869	208,8109	1732,691	161,0869331
1891,823	1537,925	0	0
19,13113	348,5412	1881,264	19,13112647
361,5674	0	1537,925	0
179,5132	279,8678	1757,428	179,5132307
126,2894	271,7756	1780,625	126,2893503
401,0249	142,1759	1550,928	142,1759473
559,7785	390,1115	1497,963	390,1115225

C1	C2	C3	Jarak Terpendek
557,0969	296,1385	1498,283	296,1384811

C1	C2	C3	Jarak Terpendek
351,510	153,781	1550,928	153,781
490,998	383,386	1497,963	383,386
518,031	265,258	1498,283	265,258

3.3. Menentukan Pengelompokan (*Cluster*)

Penentuan kelompok dilakukan dengan mencari nilai kelompok berdasarkan nilai minimal dari nilai kelompok dan diletakkan pada kelompok yang sesuai dengan nilai minimal pada Iterasi 1.

Berdasarkan tabel jarak Centroid diatas, maka Cluster atau pengelompokan Iterasi ke-2 dapat dilihat pada tabel 7.

Tabel 4. Cluster Iterasi 1

No	Type	C1	C2	C3
1	S1 Pro 8+128	1		
2	V17 Pro	1		
3	V19 8+128GB	1		
4	V19 8+256GB	1		
5	Y11 2+32GB	1		
6	Y12 3+32GB			1
7	Y12 3+64GB	1		
8	Y154+64 GB		1	
9	Y17 4+128GB	1		
10	Y19 6+128GB	1		
11	Y30 4+128GB		1	
12	Y50 8+128GB		1	
13	Y91C 2+32GB		1	

Tabel 7. Cluster Iterasi 1

No	Type	C1	C2	C3
1	S1 Pro 8+128	1		
2	V17 Pro	1		
3	V19 8+128GB	1		
4	V19 8+256GB	1		
5	Y11 2+32GB	1		
6	Y12 3+32GB			1
7	Y12 3+64GB	1		
8	Y154+64 GB		1	
9	Y17 4+128GB	1		
10	Y19 6+128GB	1		
11	Y30 4+128GB		1	
12	Y50 8+128GB		1	
13	Y91C 2+32GB		1	

Selanjutnya dalam metode K-Means, perhitungan berhenti apabila Cluster pada iterasi yang dihasilkan sama pada iterasi sebelumnya. Maka pencarian kelompok pada iterasi selanjutnya hingga sama nilai iterasinya. Mencari nilai Centroid berikutnya dengan menggunakan Centroid baru berdasarkan Iterasi ke-1 dengan menjumlahkan nilai sesuai yang tertera pada kelompok di tabel 3. Adapun Centroid baru untuk mencari Cluster selanjutnya adalah dengan menjumlahkan nilai yang terpilih pada Cluster tersebut kemudian membagikannya sebanyak jumlah nilai.

Tabel 5. Centroid Baru Iterasi 2

	C1 (Min)	61,625	48,625	34,25
<i>Centroid</i>	C2 (Ave)	181,25	183,75	327,75
	C3 (Max)	996	1154	1133

Berikutnya adalah perhitungan jarak antara titik Centroid dengan titik tiap objek, juga menggunakan rumus Euclidian Distance (Persamaan 1). Sehingga menghasilkan Jarak dari Centroid yang disajikan pada Tabel 6.

Tabel 6. Hasil Perhitungan Iterasi 2

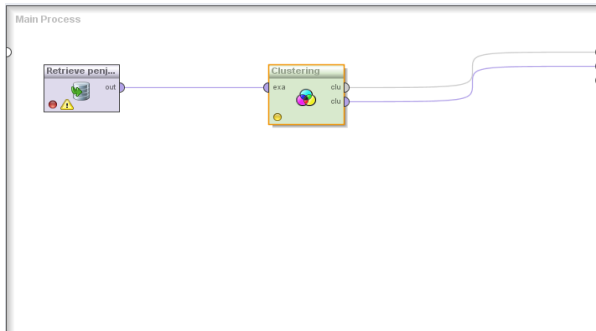
C1	C2	C3	Jarak Terpendek
65,304	402,985	1881,887	65,304
78,742	409,403	1891,823	78,742
87,630	264,603	1730,085	87,630
72,276	403,567	1885,590	72,276
93,953	263,431	1732,691	93,953
1817,185	1501,210	0,000	0,000
74,566	395,119	1881,264	74,566
294,391	61,812	1537,925	61,812
110,094	330,510	1757,428	110,094
50,156	323,171	1780,625	50,156

Tabel Cluster Iterasi ke-1 dan tabel Cluster Iterasi ke-2 memiliki nilai Cluster yang sama atau tidak berubah pada Cluster Iterasi ke-2 maka perhitungan dihentikan dan hasil yang diperoleh yaitu :

1. Cluster 1 (C1) memiliki 8 data yang diartikan bahwa kelompok pertama adalah kategori penjualan terendah pada bulan April, Mei, Juni Tahun 2020 yaitu Handphone Vivo dengan merk sebagai berikut : S1 Pro 8+128, V17 Pro, V19 8+128GB, V19 8+256GB, Y11 2+32GB Y11 2+32GB, Y12 3+64GB, Y17 4+128GB, Y19 6+128GB.
2. Cluster 2 (C2) memiliki 4 data yang diartikan bahwa kelompok kedua adalah kategori penjualan sedang pada bulan April, Mei, Juni Tahun 2020 yaitu Handphone Vivo dengan merk sebagai berikut: Y154+64 GB, Y30 4+128GB, Y50 8+128GB, Y91C 2+32GB.
3. Cluster 3 (C3) memiliki 1 data yang diartikan bahwa kelompok ketiga adalah kategori penjualan tertinggi pada bulan April, Mei, Juni Tahun 2020 yaitu Handphone Vivo dengan merk Y12 3+32GB

3.4. Implementasi dengan Rapidminer

Berikut adalah proses pengelompokan dan hasil dari metode K-Means yang dilakukan dengan Rapidminer..



Gambar 2. Proses K-Means dengan Rapidminer (Nilai K = 3)

Gambar 2 menjelaskan proses pengelompokan atau pengklusteran metode K-means dengan menggunakan Rapidminer yang diawali dengan mengimport data excel Penjualan Smartphone di Pematangsiantar, kemudian dilanjutkan dengan pemilihan operator metode K-means untuk pengklusterannya. Nilai k = 3, measure types yang digunakan adalah MixedMeasures. Setelah itu dihubungkan ke operator Apply Model untuk menerapkan model yang sudah dipelajari atau dilatih. Tujuannya adalah untuk mendapatkan prediksi pada unlabeled data (data testing) yang belum memiliki label. Tahap berikutnya adalah menghubungkan ke operator Performance untuk mengevaluasi kinerja model yang memberikan daftar nilai kriteria kinerja secara otomatis sesuai dengan tugas yang diberikan. Hasil nya dapat dilihat pada gambar 3, gambar 4, gambar 5, gambar 6 dan gambar 7.

Gambar 3. Hasil Cluster

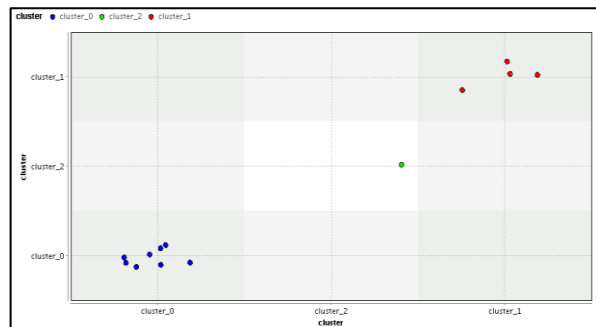
Berdasarkan gambar 3 dapat dijelaskan bahwa dari tiga cluster yang dihasilkan terdapat 8 items untuk Cluster_0 (C1), 4 items untuk Cluster_1 (C2) dan 1 items untuk Cluster_2 (C3). Untuk hasil akhir Centroid table dapat dilihat pada gambar 4. Sedangkan hasil plot view cluster penjualan smartphone Vivo disajikan pada gambar 5.

Gambar 5 berikut menjelaskan bahwa titik warna biru merupakan kelompok Cluster_0 (Penjualan rendah) yang terdiri dari 4 items. Titik warna merah merupakan kelompok Cluster_1 (Penjualan sedang) yang terdiri dari 4 items data dan titik warna merah merupakan

kelompok Cluster_2 (Penjualan tertinggi) yang terdiri dari 1 item.

id	attribute	attribute	attribute
S1 Pro 8+12	23	7	2
V17 Pro	4	3	6
V19 8+128G	113	95	88
V19 8+256G	9	6	9
Y11 2+32GB	71	117	98
Y12 3+32GB	996	1154	1133
Y12 3+64GB	5	1	25
Y154+64 GB	159	194	271
Y17 4+128G	162	88	12
Y19 6+128G	106	72	34
Y30 4+128G	37	237	330
Y50 8+128G	526	91	188
Y91C 2+32G	3	213	522

Gambar 4. Tabel Centroid



Gambar 5. Plot View Penjualan Smartphone Vivo di Pematangsiantar

4. Kesimpulan

Berdasarkan hasil penelitian dapat dilihat bahwa Ade Putri Mekaria Laila terpilih sebagai alternatif terbaik dengan nilai Optimasi 2,22692. Sehingga dapat disimpulkan Metode Moora merupakan metode yang cocok untuk mendapatkan hasil yang baik di dalam menentukan siswa yang layak menerima bantuan siswa miskin karena dapat mengolah data secara cepat dan tepat sesuai dengan yang diharapkan berdasarkan kriteria kriteria yang telah disajikan.

Referensi

[1] J. Martins, C. Costa, T. Oliveira, R. Gonçalves, and F. Branco, "How smartphone advertising influences consumers' purchase intention," *Journal of Business Research*, vol. 94, no. August 2017, pp. 378–387, 2019.

- [2] Canals, "Global smartphone market share Q2 2021," *Canalys estimates (sell-in shipments)*, *Smartphone Analysis*, July 2021, 2021. [Online]. Available: <https://www.canalys.com/newsroom/worldwide-smartphone-market-q2-2021>. [Accessed: 10-Aug-2021].
- [3] A. Wanto *et al.*, *Data Mining : Algoritma dan Implementasi*. Yayasan Kita Menulis, 2020.
- [4] T. H. Sinaga, A. Wanto, I. Gunawan, S. Sumarno, and Z. M. Nasution, "Implementation of Data Mining Using C4.5 Algorithm on Customer Satisfaction in Tirta Lihou PDAM," *Journal of Computer Networks, Architecture, and High-Performance Computing*, vol. 3, no. 1, pp. 9–20, 2021.
- [5] D. Hartama, A. Perdana Windarto, and A. Wanto, "The Application of Data Mining in Determining Patterns of Interest of High School Graduates," *Journal of Physics: Conference Series*, vol. 1339, no. 1, pp. 1–6, 2019.
- [6] I. S. Damanik, A. P. Windarto, A. Wanto, Poningsih, S. R. Andani, and W. Saputra, "Decision Tree Optimization in C4.5 Algorithm Using Genetic Algorithm," in *Journal of Physics: Conference Series*, 2019, vol. 1255, no. 1, pp. 1–6.
- [7] M. G. Sadewo, A. P. Windarto, and A. Wanto, "Penerapan Algoritma Clustering dalam Mengelompokkan Banyaknya Desa/Kelurahan Menurut Upaya Antisipasi/ Mitigasi Bencana Alam Menurut Provinsi dengan K-Means," *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. 2, no. 1, pp. 311–319, 2018.
- [8] I. Parlina, A. P. Windarto, A. Wanto, and M. R. Lubis, "Memanfaatkan Algoritma K-Means dalam Menentukan Pegawai yang Layak Mengikuti Assessment Center untuk Clustering Program SDP," *CESS (Journal of Computer Engineering System and Science)*, vol. 3, no. 1, pp. 87–93, 2018.
- [9] W. Katrina, H. J. Damanik, F. Parhusip, D. Hartama, A. P. Windarto, and A. Wanto, "C.45 Classification Rules Model for Determining Students Level of Understanding of the Subject," *Journal of Physics: Conference Series*, vol. 1255, no. 1, pp. 1–6, Aug. 2019.
- [10] H. Siahaan, H. Mawengkang, S. Efendi, A. Wanto, and A. Perdana Windarto, "Application of Classification Method C4.5 on Selection of Exemplary Teachers," *Journal of Physics: Conference Series*, vol. 1235, no. 1, pp. 1–7, Jun. 2019.
- [11] C. Astria, A. P. Windarto, A. Wanto, and E. Irawan, "Metode K-Means pada Pengelompokan Wilayah Pendistribusian Listrik," *Seminar Nasional Sains & Teknologi Informasi (SENSASI)*, pp. 306–312, 2019.
- [12] M. A. Hanafiah, A. Wanto, and P. B. Indonesia, "Implementation of Data Mining Algorithms for Grouping Poverty Lines by District/City in North Sumatra," *International Journal of Information System & Technology*, vol. 3, no. 2, pp. 315–322, 2020.
- [13] R. W. Sari, A. Wanto, and A. P. Windarto, "Implementasi Rapidminer Gengan Metode K-Means (Study Kasus: Imunisasi Campak Pada Balita Berdasarkan Provinsi)," *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. 2, no. 1, pp. 224–230, 2018.
- [14] S. R. Ningsih, I. S. Damanik, A. P. Windarto, H. S. Tambunan, J. Jalaluddin, and A. Wanto, "Analisis K-Medoids Dalam Pengelompokan Penduduk Buta Huruf Menurut Provinsi," *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, vol. 1, pp. 721–730, Sep. 2019.
- [15] M. Widyastuti, A. G. Fepdiani Simanjuntak, D. Hartama, A. P. Windarto, and A. Wanto, "Classification Model C.45 on Determining the Quality of Customer Service in Bank BTN Pematangsiantar Branch," *Journal of Physics: Conference Series*, vol. 1255, no. 1, pp. 1–5, Aug. 2019.
- [16] A. Pradipta, D. Hartama, A. Wanto, S. Saifullah, and J. Jalaluddin, "The Application of Data Mining in Determining Timely Graduation Using the C45 Algorithm," *IJISTECH (International Journal of Information System & Technology)*, vol. 3, no. 1, pp. 31–36, 2019.
- [17] T. Imandasari, E. Irawan, A. P. Windarto, and A. Wanto, "Algoritma Naive Bayes Dalam Klasifikasi Lokasi Pembangunan Sumber Air," *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, vol. 1, pp. 750–761, Sep. 2019.
- [18] I. Parlina *et al.*, "Naive Bayes Algorithm Analysis to Determine the Percentage Level of visitors the Most Dominant Zoo Visit by Age Category," *Journal of Physics: Conference Series*, vol. 1255, no. 1, pp. 1–5, Aug. 2019.
- [19] H. Siahaan, H. Mawengkang, S. Efendi, A. Wanto, and A. Perdana Windarto, "Application of Classification Method C4.5 on Selection of Exemplary Teachers," in *Journal of Physics: Conference Series*, 2019, vol. 1235, no. 1, pp. 1–7.
- [20] S. Sudirman, A. P. Windarto, and A. Wanto, "Data Mining Tools | RapidMiner : K-Means Method on Clustering of Rice Crops by Province as Efforts to Stabilize Food Crops In Indonesia," *IOP Conference Series: Materials Science and Engineering*, vol. 420, no. 012089, pp. 1–8, 2018.
- [21] S. Sundari, I. S. Damanik, A. P. Windarto, H. S. Tambunan, J. Jalaluddin, and A. Wanto, "Analisis K-Medoids Clustering Dalam Pengelompokan Data Imunisasi Campak Balita di Indonesia," *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, vol. 1, pp. 687–696, Sep. 2019.
- [22] N. Arminarahmah, A. D. GS, G. W. Bhawika, M. P. Dewi, and A. Wanto, "Mapping the Spread of Covid-19 in Asia Using Data Mining X-Means Algorithms," *IOP Conf. Series: Materials Science and Engineering*, vol. 1071, no. 012018, pp. 1–7, 2021.
- [23] N. A. Febriyati, A. D. GS, and A. Wanto, "GRDP Growth Rate Clustering in Surabaya City uses the K- Means Algorithm," *International Journal of Information System & Technology*, vol. 3, no. 2, pp. 276–283, 2020.
- [24] J. Hutagalung, N. L. W. S. R. Ginantra, G. W. Bhawika, W. G. S. Parwita, A. Wanto, and P. D. Panjaitan, "COVID-19 Cases and Deaths in Southeast Asia Clustering using K-Means Algorithm," *Journal of Physics: Conference Series*, vol. 1783, no. 1, p. 012027, 2021.
- [25] S. Handoko, F. Fauziah, and E. T. E. Handayani, "Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering," *Jurnal Ilmiah Teknologi dan Rekayasa*, vol. 25, no. 1, pp. 76–88, 2020.
- [26] Z. S. Younus *et al.*, "Content-based image retrieval using PSO and k-means clustering algorithm," *Arabian Journal of Geosciences*, vol. 8, no. 8, pp. 6211–6224, 2015.
- [27] E. Prasetyo, *Data Mining: Konsep dan Aplikasi menggunakan Matlab*. Yogyakarta: Andi Offset, 2012.
- [28] D. T. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining: Second Edition*. New Jersey: John Wiley & Sons, 2005.
- [29] R. Primartha, *Belajar Machine Learning Teori dan Praktik*. Bandung: Informatika Bandung, 2018.
- [30] T. Khotimah, "Pengelompokan Surat dalam Al Qur'an menggunakan Algoritma K-Means," *Jurnal Simetris*, vol. 5, no. 1, pp. 83–88, 2014.