



Crawling Engine Pada Website Mann, Baldwin, Fleetguard Dan Pengelompokan Produk Menggunakan K-Means

Andi Rahman¹, Wowon Priatna^{2*}, Tyastuti Sri Lestari³, Agus Hidayat⁴

^{1,2,3,4}Informatika, Universitas Bhayangkara Jakarta Raya, Bekasi, Indonesia

E-mail: ¹andi.rahman18@mhs.ubharajaya.ac.id, ^{2*}wowon.priatna@dsn.ubharajaya.ac.id,
³tyas@ubharajaya.ac.id, ⁴agus.hidayat@dsn.ubharajaya.ac.id

Email Penulis Korespondensi: wowon.priatna@dsn.ubharajaya.ac.id

Abstrak

In today's highly competitive and fast-changing global economy, companies are required to leverage information technology to support their business processes. K-means is a data clustering algorithm that falls under machine learning algorithms. It is a popular method for partitioning data into one or more clusters or groups. The purpose of this study is to group products on several websites. The crawling engine will be very helpful in entering product data automatically from the product website, then entered into the Odoo application. The k-means clustering algorithm itself is an algorithm that groups observations into k groups, where k is the input parameter. Each data is then assigned to each cluster observation adjacent to the cluster mean value observation. This grouping will be very helpful in product classification based on cross reference. The result of this research is that the product is inputted automatically and the data is in accordance with the product's website and the products are grouped according to cross-references.

Informasi Artikel

Diterima: 6 Juli 2023

Direvisi: 7 Agustus 2024

Dipublikasikan: 30 September 2024

Keywords

Crawling, K-Means, Clustering, Data Mining, Odoo, API.

I. Pendahuluan

Di era ekonomi global saat ini yang sangat kompetitif dan cepat berubah, perusahaan dituntut untuk menggunakan teknologi informasi untuk mendukung kinerja proses bisnisnya. Manajemen teknologi informasi membutuhkan visi strategi yang dapat menyatukan antara teknologi informasi dan proses bisnis.

Strategi teknologi yang mengacu pada investasi spesifik dalam bentuk perbedaan jenis sistem, seperti sistem ERP (*Enterprise Resource Planning*). ERP adalah sebuah *software* yang memiliki fungsi untuk mengelola dan mengintegrasikan berbagai aktivitas operasional dalam bisnis [1].

Salah satu perusahaan PT. Rivindi Arta Mandiri yang sudah memiliki banyak produk menghindari melakukan penginputan data produk secara manual dikarenakan akan memakan waktu dan terjadinya kesalahan memasukan informasi data tentang produk, PT Rivindi juga ingin menghindari terjadinya kesalahan menawarkan produk yang akan bisa mengakibatkan kekecewaan kepada pembeli, maka akan dilakukannya pengelompokan produk filter secara agar lebih mudah dalam melakukan proses penawaran. metode yang cocok untuk melakukan pengelompokan produk adalah *K-Means Clustering* [2][3]. *K-mean* adalah algoritma pengolompokan data yang termasuk dalam algoritma *machine learning* [4], kmean merupakan algoritma pengelompokan populer berdasarkan partisi data [5] ke dalam bentuk satu atau lebih *cluster*/kelompok data lain [6]. Algoritma k-means adalah algoritma evolusioner, diturunkan dari bentuk operasinya [4], *K-Means Clustering* adalah, K dimaksudkan sebagai jumlah konstan *cluster* yang diinginkan sehingga dapat mengolompokan data [7], Algoritma pengelompokan mengeksploitasi struktur yang mendasari distribusi data dan menentukan aturan untuk mengelompokkan data dengan karakteristik yang serupa dan Proses ini menghasilkan partisi dari kumpulan data yang diberikan sesuai dengan kriteria pengelompokan tanpa pengetahuan sebelumnya tentang kumpulan data [8].

Pada penelitian [9] menggunakan *K-mean clustering* untuk mengelompokan produk *e-commerce* berupa perawatan rambut dan wajah, mulut dan hewan peliharaan yang mendapatkan pengelompokan produk yang efektif, penelitian [10] menggunakan k-mean

untuk pengelompokan data Analisa pilpress dengan melakukan *crawling* data dari media *twitter* menghasilkan 3 *cluster* dalam pengelompokan. K-mean juga digunakan dalam pengelompokan data tidak terstruktur dengan memanfaatkan *crawling* data dari website [11] dan k-mean juga digunakan dalam mengelompokan customer potensial dalam menggunakan jasa pembagian produk [12], K-Mean juga dapat digunakan untuk pengelompokan data hasil MBKM [13][14] dan pengelompokan data hasil keuntungan sementara dari pembayaran sekolah [15].

Tujuan penelitian ini adalah untuk menyelesaikan permasalahan pengelompokan data dari berbagai website menggunakan algoritma K-mean dengan melakukan *crawling* data produk *Website Mann, Baldwin, Fleetguard*.

II. Metodologi

2.1. Desain Penelitian

Desain penelitian untuk melakukan sentiment analisis dapat dijabarkan pada gambar 1.



Gambar 1. Desain Penelitian

2.2. Identifikasi Masalah

Pada tahapan ini adalah identifikasi masalah mengenai produk yang belum mempunyai kelompok, selanjutnya menentukan pemecahan permasalahan yang akan digunakan dalam penelitian ini.

2.3. Pengumpulan Data

Fase ini merupakan fase pengambilan data. Data yang nantinya akan diolah berasal dari website masing masing produk dan akan di *crawling* dan

scraping di Odoo. Data-yang-telah-diambil-dari hasil *crawling* akan dilakukan-proses pemilihan untuk melanjutkan ke tahap pengolahan Data. Data yang akan di*crawling* merupakan data yang diberikan terhadap PT. Rivindi Arta Mandiri. Untuk melakukan proses *K-Means Clustering* tentu membutuhkan sebuah data yang cukup banyak sesuai dengan yang dibutuhkan, dalam penelitian ini memberikan gambaran data yang akan diolah pada tabel 1.

Tabel 1. Data Produk

No	Nama Produk	Merk	Cross Reference
1	18002	BALDWIN	[192] 51877,[191] 58877,[190] F2VY-7A098-A
2	18004	BALDWIN	[194] 4505048
3	18007	BALDWIN	[196] 8684221
4	18009	BALDWIN	[200] 51895,[199] 58895,[198] 8684953
5	18010	BALDWIN	[202] 24200796
6	18013	BALDWIN	[204] 35330-32022,[205] 51994,[206] 58994
7	18024	BALDWIN	[210] F9TZ-7A098-C,[211] ZZL2-21-500
8	18025	BALDWIN	[213] 85GT-7A098-AC,[211] ZZL2-21-500
9	C 1043/1	MANN	[1852] 1427 101 0140 00,[1854] 142710101400,[1853] 1427101014E11,[1843] 172.00002,[1862] 213146,[1849] 216 5021,[1850] 216 5024,[1845] 267 3698
10	BF 1018/1	MANN	[5] 000 477 34 15,[4] 1215 3199,[3] 2 C 2 H 4115,[7] 6.0541.14.2.0001,[8] 6.0541.14.2.0017,[9] 6.0541.14.2.0024,[10] 6.0541.29.2.0017,[11] 6.0541.29.4.0001

2.4. Preprocessing Data

Pra-pemrosesan data berfungsi sebagai dasar untuk analisis data yang valid [16]. agar data diatas dapat diolah dengan menggunakan metode K-Means

Clustering, maka data yang berjenis data nominal seperti nama *cross reference* harus diinisialisasikan berdasarkan ID yang telah termuat dalam Odoo seperti pada tabel 2.

Tabel 2. Data produk dengan *cross reference*

No	Nama Produk	Cross 1	Cross 2	Cross 3	Cross 4	Cross 5	Cross 6	Cross 7	Cross 8
1	18002	192	191	190	0	0	0	0	0
2	18004	194	0	0	0	0	0	0	0
3	18007	196	0	0	0	0	0	0	0

No	Nama Produk	Cross 1	Cross 2	Cross 3	Cross 4	Cross 5	Cross 6	Cross 7	Cross 8
4	18009	200	199	198	0	0	0	0	0
5	18010	202	0	0	0	0	0	0	0
6	18013	204	205	206	0	0	0	0	0
7	BF1018/1	5	4	3	7	8	9	10	11
8	18024	210	211	0	0	0	0	0	0
9	18025	213	211	0	0	0	0	0	0
10	C 1043/1	1852	1854	1853	1843	1862	1849	1850	1845

2.5. Penerapan K-Means

Setelah semua nama *cross reference* dan produk dimasukkan kedalam list, maka data data tersebut akan dikelompokkan dengan menggunakan algoritma *K-Means Clustering*. Untuk dapat melakukan pengelompokan data data tersebut dipecah menjadi beberapa *cluster*. Perlu dilakukan beberapa

langkah, yaitu:

1. Tentukan jumlah *cluster*, Dalam penelitian ini data yang akan dikelompokkan menjadi 3 *cluster*.
2. Tentukan titik pusat dari awal *cluster*, Dalam penelitian ini titik pusat ditentukan secara acak dan titik pusat dari setiap *cluster* dapat dilihat dari tabel 3.

Tabel 3. Contoh Data Produk Dengan *Cross Reference*

Titik Pusat awal	Nama Produk	Cross 1	Cross 2	Cross 3	Cross 4	Cross 5	Cross 6	Cross 7	Cross 8
C.1	C 1043/1	1852	1854	1853	1843	1862	1849	1850	1845
C.2	18025	213	211	0	0	0	0	0	0
C.3	BF 1018/1	5	4	3	7	8	9	10	11

2.6. Tahap Perhitungan K-Means

Setelah diketahui nilai *k* dan pusat *cluster*, selanjutnya adalah mengukur jarak antara pada pusat *cluster* menggunakan euclidian, kemudian akan didapatkan matriks jarak yaitu C1, C2, C3 sebagai berikut:

$$[(x, y), (a, b)] = \sqrt{(x - a)^2 + (y - b)^2} \quad (1)$$

perhitungan jarak data ke 9 dengan *cluster* pertama

$$d_{91} = \sqrt{(213-1852)^2 + (211-1854)^2 + (0-1853)^2 + (0-1843)^2 + (0-1862)^2 + (0-1849)^2 + (0-1850)^2 + (0-1845)^2} = 5091.993519$$

perhitungan jarak data ke 9 dengan *cluster* kedua

$$d_{92} = \sqrt{(213-213)^2 + (211-211)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2} = 0$$

perhitungan jarak data ke 9 dengan *cluster* ketiga

$$d_{92} = \sqrt{(213-5)^2 + (211-4)^2 + (0-3)^2 + (0-7)^2 + (0-8)^2 + (0-9)^2 + (0-10)^2 + (0-11)^2} = 294.1717186$$

Hasil dari perhitungan data dengan titik *cluster* dapat disimpulkan data ke 9 pada table 3.2 dekat dengan *cluster* ke 2, dan berikut hasil perhitungan dengan contoh data pada tabel 4

Tabel 4 Hasil perhitungan dengan C1, C2, C3

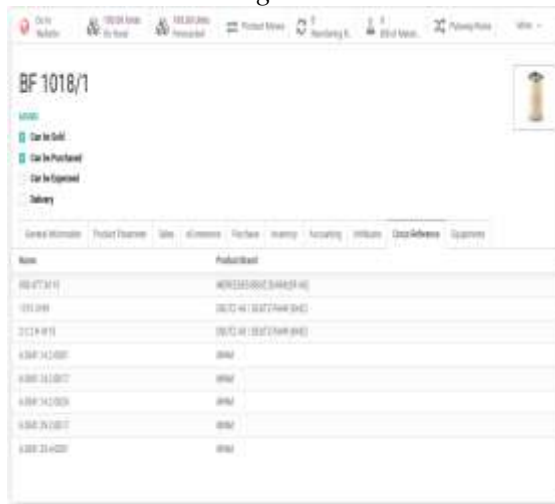
No	Nama Produk	C1	C2	C3	Jarak Terdekat
1	18002	5.039.420.304	1.922.004.162	3.245.335.114	1.922.004.162
2	18004	5.170.000.774	2.118.537.231	1.901.604.586	1.901.604.586
3	18007	5.169.359.728	2.116.837.263	1.921.483.802	1.921.483.802
4	18009	5.031.517.962	1.987.888.327	3.383.637.096	1.987.888.327
5	18010	5.167.440.759	2.112.865.353	1.981.136.038	1.981.136.038
6	18013	5.025.607.724	2.062.837.851	3.487.491.935	2.062.837.851
7	BF 1018/1	5.215.302.196	2.941.717.186	0	5.215.302.196
8	18024	5.092.959.945	3	2.920.582.134	5.092.959.945
9	18025	5.091.993.519	0	2.941.717.186	5.091.993.519
10	C 1043/1	0	5.091.993.519	5.215.302.196	0

III. Hasil dan Pembahasan

3.1 Implementasi

Implementasi pada langkah ini dimulai dengan tahap pengumpulan data. Dengan cara melakukan *crawling*

pada website produk masing masing dan dalam tahap yang bsersamaan melakukan *crawling*, data produk akan dimasukan ke dalam apikasi Odoo.



Gambar 3. Hasil *crawling* dan tampilan pada aplikasi Odoo.

3.2 Tahap Pengumpulan Data

Tahap pengolahan data berguna untuk membersihkan kata agar mempermudah

dalam proses analisis. Tahap pengolahan data terdiri seperti yang sudah dijelaskan dalam tabel 5.

Tabel 5. Tahap *Preprocessing*

No	Nama Produk	Brand	Cross reference
1	101-30	BALDWIN	
2	101-W	BALDWIN	
3	1012	BALDWIN	
4	151	BALDWIN	

No	Nama Produk	Brand	Cross reference
5	151-30	BALDWIN	
6	151-W	BALDWIN	
7	18002	BALDWIN	[192] 51877,[191] 58877,[190] F2VY-7A098-A
8	18004	BALDWIN	[194] 4505048
9	18007	BALDWIN	[196] 8684221
10	18009	BALDWIN	[200] 51895,[199] 58895,[198] 8684953
11	18010	BALDWIN	[202] 24200796
12	18013	BALDWIN	[204] 35330-32022,[205] 51994,[206] 58994
13	18022	BALDWIN	[208] 24206433
14	18024	BALDWIN	[210] F9TZ-7A098-C,[211] ZZL2-21-500
15	18025	BALDWIN	[213] 85GT-7A098-AC,[211] ZZL2-21-500
16	18027	BALDWIN	[215] 24208847
17	18029	BALDWIN	[218] F4TZ-7A098-A,[217] F6AZ-7A098-A
18	18032	BALDWIN	[221] 3515966,[220] 3515996,[222] 732931-C2,[223] J8123042
19	18033	BALDWIN	[225] F5DZ-7A098-A
20	18034	BALDWIN	[227] F6DZ-7A098-A
21	116029S	FLEETGUARD	
22	142784S	FLEETGUARD	
23	151884S	FLEETGUARD	
24	153518S	FLEETGUARD	
25	153520S	FLEETGUARD	
26	153528S	FLEETGUARD	
27	154711S	FLEETGUARD	
28	161917S	FLEETGUARD	
29	164159S	FLEETGUARD	
30	164159SB	FLEETGUARD	
31	173368S	FLEETGUARD	
32	183342S	FLEETGUARD	
33	193302S	FLEETGUARD	
34	204163S	FLEETGUARD	
35	212013S	FLEETGUARD	
36	215617S	FLEETGUARD	
37	250825S	FLEETGUARD	
38	250856S	FLEETGUARD	
39	251081S	FLEETGUARD	
40	251152S	FLEETGUARD	
41	BF 1018/1	MANN	[5] 000 477 34 15,[4] 1215 3199,[3] 2 C 2 H 4115,[7] 6.0541.14.2.0001,[8] 6.0541.14.2.0017,[9] 6.0541.14.2.0024,[10] 6.0541.29.2.0017,[11] 6.0541.29.4.0001

No	Nama Produk	Brand	Cross reference
42	BFU 700 X	MANN	[31] 00 213 15108,[19] 000 090 11 51,[20] 000 092 27 05,[21] 000 092 54 05,[14] 0133 537 0,[27] 110722,[18] 202608,[22] 352 470 00 92
43	BFU 707	MANN	[273] 00 0099 104 1,[352] 00 213 1510,[309] 000 092 15 05,[310] 000 092 19 05,[311] 000 470 01 92,[312] 000 477 38 15,[313] 000 477 42 15,[314] 000 477 54 15
44	BFU 715	MANN	[404] 0.4001.062.0,[407] 003.1407.0,[397] 116 8402,[403] 116 8402,[398] 1215 3114,[399] 5574 980,[402] 5574980,[405] 6.0540.19.1.0002
45	BFU 811	MANN	[1414] 00 24 152 004,[1415] 00 24 866 064,[1431] 00 Z62 546,[1400] 000 477 31 15,[1365] 000 477 3115,[1401] 000 477 32 15,[1420] 001.0780.0,[1425] 5.402.08.001.BO2
46	BFU 900 X	MANN	[1575] 00 0133 602 1,[1583] 000 090 14 51,[1573] 0133 602 0,[1574] 0133 602 1,[1576] 118 1060,[1595] 33023,[1572] 38136,[1579] 400 02 049
47	C 10 005	MANN	[1617] 909548
48	C 10 006	MANN	[1644] 1E51967G01,[1643] 32/925894,[1645] P030047120
49	C 10 050	MANN	[1694] 1000141558,[1695] 16546-LC50A,[1691] 16546LC50A,[1690] 242541,[1697] 270764,[1689] 3-69225,[1692] 32/925742,[1696] 74 85 138 330
50	C 1036/1	MANN	[1723] 000 1024 V 001,[1725] 000 3123 V 007,[1724] 000 3124 V 001
51	C 1036/2	MANN	[1916] 000 9996 V 001,[1915] 000 9997 V 001,[1913] 160 094 01 04,[1914] A 160 094 01 04
52	C 1041	MANN	[1753] 000 4591 V 001
53	C 1043/1	MANN	[1852] 1427 101 0140 00,[1854] 142710101400,[1853] 1427101014E11,[1843] 172.00002,[1862] 213146,[1849] 216 5021,[1850] 216 5024,[1845] 267 3698
54	C 1049	MANN	[1885] 730 512,[1884] 90957400000
55	C 105	MANN	[1905] 3515175
56	C 11 003	MANN	[1936] 10400511200,[1935] 10400511200,[1918] 72280475,[1923] 72285-432,[1924] 72285-433,[1920] 72285431,[1921] 84221831,[1919] 87643356
57	C 11 004	MANN	
58	C 11 005	MANN	
59	C 11 100	MANN	[8233] 000 983 9002,[8223] 131 9141,[8234] 16546-FA01A,[8243] 2127067,[8228] 23487929,[8218] 3-69217,[8235] 41748,[8238] 5.501.660.947
60	C 11 103/2	MANN	[8253] 057 301 33,[8254] 057 301 36,[8247] 1-952583,[8313] 100 000 4566,[8314] 100 023

No	Nama Produk	Brand	Cross reference
			1364,[8291] 1000004566,[8278] 1000052549,[8283] 104.2175.1240

3.3 Preprocessing Data

Pada tahap ini akan dilakukan pembersihan data, data yang tidak memiliki cross reference tidak akan lanjut ketahap berikutnya, jika setelah pembersihan data yang tidak layak maka selanjutnya adalah transtformasi data.

Data yang dianggap layak yaitu data produk yang memiliki nilai *cross reference* akan melakukan transformasi data, dikarenakan data *cross reference* harus diubah ke data nomina, data yang memiliki kelayakan dan sudah ditransformasi seperti pada tabel 6.

Tabel 6. Data *Preprocessing*

No	Nama Produk	Cross 1	Cross 2	Cross 3	Cross 4	Cross 5	Cross 6	Cross 7	Cross 8
1	18002	192	191	190	0	0	0	0	0
2	18004	194	0	0	0	0	0	0	0
3	18007	196	0	0	0	0	0	0	0
4	18009	200	199	198	0	0	0	0	0
5	18010	202	0	0	0	0	0	0	0
6	18013	204	205	206	0	0	0	0	0
7	18022	208	0	0	0	0	0	0	0
8	18024	210	211	0	0	0	0	0	0
9	18025	213	211	0	0	0	0	0	0
10	18027	215	0	0	0	0	0	0	0
11	18029	218	217	0	0	0	0	0	0
12	18032	221	222	223	0	0	0	0	0
13	18033	225	0	0	0	0	0	0	0
14	18034	227	0	0	0	0	0	0	0
15	BF 1018/1	5	4	3	7	8	9	10	11
16	BFU 700 X	31	19	20	21	14	27	18	22
17	BFU 707	273	352	309	310	311	312	313	314
18	BFU 715	404	407	397	403	398	399	402	405
19	BFU 811	1414	1431	1400	1365	1401	1420	1425	0
20	BFU 900 X	1575	1583	1573	1574	1576	1595	1572	1579
21	C 10 005	1617	0	0	0	0	0	0	0
22	C 10 006	1644	1643	1645	0	0	0	0	0
23	C 10 050	1694	1695	1691	1690	1697	1689	1692	1696
24	C 1036/1	1723	1725	1724	0	0	0	0	0
25	C 1036/2	1916	1915	1913	1914	0	0	0	0
26	C 1041	1753	0	0	0	0	0	0	0
27	C 1043/1	1852	1854	1853	1843	1862	1849	1850	1845
28	C 1049	1885	1884	0	0	0	0	0	0
29	C 105	1905	0	0	0	0	0	0	0

No	Nama Produk	Cross 1	Cross 2	Cross 3	Cross 4	Cross 5	Cross 6	Cross 7	Cross 8
30	C 11 003	1936	1935	1918	1923	1924	1920	1921	1919
31	C 11 100	8233	8223	8234	8243	8228	8218	8235	8238
32	C 11 103/2	8253	8254	8247	8313	8314	8291	8278	8283

3.4 Uji Data dengan K-Means

Data yang memenuhi kriteria akan dianalisis dan dikelompokkan menggunakan metode k-means clustering. Tahap clustering antara lain sebagai berikut.

1. Tentukan jumlah *cluster*, Dalam

penelitian ini data yang akan dikelompokkan menjadi 3 *cluster*.

2. Tentukan titik pusat dari awal *cluster*, Dalam penelitian ini titik pusat ditentukan secara acak dan titik pusat dari setiap *cluster* dapat dilihat dari tabel 3.3.

Tabel 7. Contoh Data Produk Dengan *Cross Reference*

Titik Pusat awal	Nama Produk	Cross 1	Cross 2	Cross 3	Cross 4	Cross 5	Cross 6	Cross 7	Cross 8
C.1	C 1043/1	1852	1854	1853	1843	1862	1849	1850	1845
C.2	18025	213	211	0	0	0	0	0	0
C.3	BF 1018/1	5	4	3	7	8	9	10	11

Setelah diketahui nilai k dan pusat *cluster*, selanjutnya adalah mengukur jarak antara pada pusat *cluster* menggunakan euclidian, kemudian akan didapatkan matriks jarak yaitu C1, C2, C3 sebagai berikut:

$$[x, y), (a, b)] = \sqrt{(x - a)^2 + (y - b)^2} \quad (2)$$

perhitungan jarak data ke 9 dengan *cluster* pertama

$$d_{91} = \sqrt{(213-1852)^2 + (211-1854)^2 + (0-1853)^2 + (0-1843)^2 + (0-1862)^2 + (0-1849)^2 + (0-1850)^2 + (0-1845)^2} = 5091.993519$$

perhitungan jarak data ke 9 dengan *cluster* kedua

$$d_{92} = \sqrt{(213-213)^2 + (211-211)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2} = 0$$

perhitungan jarak data ke 9 dengan *cluster* ketiga

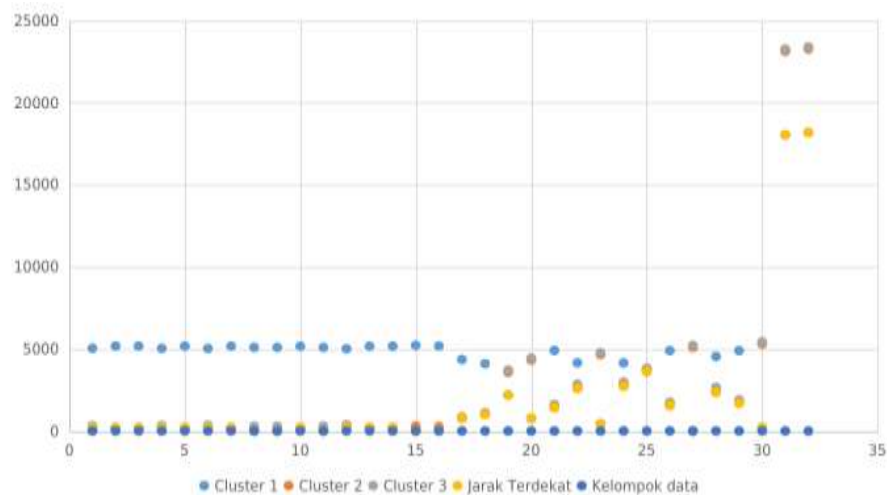
$$d_{93} = \sqrt{(213-5)^2 + (211-4)^2 + (0-3)^2 + (0-7)^2 + (0-8)^2 + (0-9)^2 + (0-10)^2 + (0-11)^2} = 294.1717186$$

Hasil dari perhitungan data dengan titik *cluster* dapat disimpulkan data ke 9 pada tabel 3.2 dekat dengan *cluster* ke 2, dan berikut hasil perhitungan dengan contoh data pada tabel 3.2.

Tabel 8. Hasil perhitungan dengan C1, C2, C3

No	Nama Produk	C1	C2	C3	Jarak Terdekat
1	18002	5.039.420.304	1.922.004.162	3.245.335.114	1.922.004.162
2	18004	5.170.000.774	2.118.537.231	1.901.604.586	1.901.604.586
3	18007	5.169.359.728	2.116.837.263	1.921.483.802	1.921.483.802
4	18009	5.031.517.962	1.987.888.327	3.383.637.096	1.987.888.327

No	Nama Produk	C1	C2	C3	Jarak Terdekat
5	18010	5.167.440.759	2.112.865.353	1.981.136.038	1.981.136.038
6	18013	5.025.607.724	2.062.837.851	3.487.491.935	2.062.837.851
7	BF 1018/1	5.215.302.196	2.941.717.186	0	5.215.302.196
8	18024	5.092.959.945	3	2.920.582.134	5.092.959.945
9	18025	5.091.993.519	0	2.941.717.186	5.091.993.519
10	C 1043/1	0	5.091.993.519	5.215.302.196	0



Gambar 4. Tabel charter bersarkan hasil *clustering*

Pada gambar 4 merupakan *table charter* hasil dari tabel 4.6 yang dapat disimpulkan, sebagai berikut:

1. Dalam *cluster* C1 dengan pusat produk terdapat 7 data produk yang akan dikelompokan.
2. Dalam *cluster* C2 terdapat 18 data produk yang akan dikelompokan.
3. Dalam *cluster* C3 terdapat 7 data produk yang akan dikelompokan.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang sudah dilakukan mengenai *crawling* engine dan pengelompokan produk dengan algoritma k-means *clustering* di PT. Rivindi Arta Mandiri, maka dapat diambil kesimpulan sebagai berikut :

1. Dengan dibuatnya *crawling* engine pihak PT. Rivindi Arta Mandiri hanya tinggal mengecek data produk

apakah sesuai dengan yang ada pada website produk tersebut.

2. Pengelompokan produk yang telah ditetapkan bisa digunakan sebagai produk referensi untuk diberikan kepada pembeli.

Daftar Pustaka

- [1] Wicaksono, H. H. Mulyo, and I. E. Riantono, "TERHADAP KINERJA PENGGUNA Aries Wicaksono ; Hery Harjono Mulyo ; Ignatius Edward Riantono PENDAHULUAN Latar Belakang," *Account. Finance*, no. 9, pp. 25–34, 2015.
- [2] N. A. Khairani and E. Sutoyo, "Application of K-Means *Clustering* Algorithm for Determination of Fire-Prone Areas Utilizing Hotspots in West Kalimantan Province," *Int. J. Adv. Data Inf. Syst.*, vol. 1, no. 1,

- pp. 9–16, 2020, doi: 10.25008/ijadis.v1i1.13.
- [3] K. Kusriani, “Grouping of Retail Items by Using K-Means Clustering,” *Procedia Comput. Sci.*, vol. 72, pp. 495–502, 2015, doi: 10.1016/j.procs.2015.12.131.
- [4] R. Vankayalapati, K. B. Ghutugade, R. Vannapuram, and B. P. S. Prasanna, “K-means algorithm for clustering of learners performance levels using machine learning techniques,” *Rev. d’Intelligence Artif.*, vol. 35, no. 1, pp. 99–104, 2021, doi: 10.18280/ria.350112.
- [5] J. Xie, S. Jiang, W. Xie, and X. Gao, “An efficient global K-means clustering algorithm,” *J. Comput.*, vol. 6, no. 2, pp. 271–279, 2011, doi: 10.4304/jcp.6.2.271-279.
- [6] F. Nasari and C. J. M. Sianturi, “Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Diare Di Kabupaten Langkat,” *CogITO Smart J.*, vol. 2, no. 2, p. 108, 2016, doi: 10.31154/cogito.v2i2.19.108-119.
- [7] N. Buslim, R. P. Iswara, and F. Agustian, “the Modeling of ‘Mustahiq’ Data Using K-Means Clustering Algorithm and Big Data Analysis (Case Study: Laz),” *J. Tek. Inform.*, vol. 13, no. 2, pp. 213–230, 2021, doi: 10.15408/jti.v13i2.19610.
- [8] M. Ahmed, R. Seraj, and S. M. S. Islam, “The k-means algorithm: A comprehensive survey and performance evaluation,” *Electron.*, vol. 9, no. 8, pp. 1–12, 2020, doi: 10.3390/electronics9081295.
- [9] N. M. N. Mathivanan, N. A. M. Ghani, and R. M. Janor, “Analysis of K-Means Clustering Algorithm: A Case Study Using Large Scale E-Commerce Products,” *2019 IEEE Conf. Big Data Anal. ICBDA 2019*, pp. 41–44, 2019, doi: 10.1109/ICBDA47563.2019.898714
- [10] F. M. Evanita, I. Cholissodin, and S. Adinugroho, “Pengelompokan Toko E-commerce Shopee berdasarkan Reputasi Toko menggunakan Metode Clustering K-Medoids,” vol. 5, no. 3, pp. 1230–1236, 2021.
- [11] N. Jain and P. Mangal, “Available Online at www.jgrcs.info AN APPROACH TO BUILD A WEB CRAWLER USING CLUSTERING BASED K-MEANS ALGORITHM,” vol. 4, no. 12, 2013.
- [12] R. Nainggolan and E. Purba, “Cluster Analysis of Online Shop Product Reviews Using K-Means Clustering,” *IJEED (International J. Entrep. Bus. Dev.)*, vol. 3, no. 02, pp. 142–151, 2020, doi: 10.29138/ijeed.v3i02.977.
- [13] Mayadi, S. Setiawati, and W. Priatna, “Pengelompokan Hasil Survei MBKM Menggunakan K-Mean dan K-Medoids Clustering,” *J. Media Inform. Budidarmajurnal Media Inform. Budidarma*, vol. 7, pp. 426–435, 2023, doi: 10.30865/mib.v7i1.5003.
- [14] Mayadi, S. Setiawati, and W. Priatna, “Pengelompokan Hasil Survei Merdeka Belajar Kampus Merdeka Di Universitas Bhayangkara Jakarta Raya Menggunakan Kmean Dan K-Medoids Clustering,” *Pelita Teknol.*, vol. 17, no. 2, pp. 1–11, 2022.
- [15] O. Hermadi, W. Priatna, and A. D. Alexander, “Implementasi Algoritma K-Medoids Clustering Untuk Mencari Keuntungan Sementara Dalam Laporan Keuangan,” vol. 6, no. April, pp. 6–

11, 2023.

- [16] Fan, M. Chen, X. Wang, J. Wang, and B. Huang, "A Review on Data Preprocessing Techniques Toward Efficient and Reliable Knowledge Discovery From Building Operational Data," *Front. Energy Res.*, vol. 9, no. March, pp. 1–17, 2021, doi: 10.3389/fenrg.2021.652801.