



PENERAPAN METODE K-MEANS CLUSTERING UNTUK MENGELOMPOKKAN KETAHANAN PANGAN

Deki Setra Perdana^{1(*)}

¹Universitas Jambi, Jambi

Abstract

Food is a basic need that must be fulfilled and easily accessible by all people. After the end of the pandemic, it still caused several sectors to experience a decline, including the agricultural sector, which experienced a decline which resulted in crop yields also declining. The problem faced by several regions in Indonesia, one of which is the area of Indonesia, is the availability of food yields which has decreased and increased unstable due to lack of information about the grouping of harvest resilience each year. As a result, the food needs of the people in each region are not fulfilled. The purpose of this study was to classify regions with increasing and decreasing yields of food crops in Indonesia using the K-Medoids algorithm. The K-Means algorithm includes a clustering algorithm that is quite efficient in grouping small data and finding the most representative point and being able to overcome outliers. So that it can be used in grouping the effect of productivity and the level of food security.

Kata Kunci: Penerapan, Metode K-Means

Juli - Desember 2022, Vol 3 (2) : hlm 67-66
©2022 Institut Teknologi dan Bisnis Ahmad Dahlan.
All rights reserved.

(*) Korespondensi: dekisetra423@gmail.com (Deki Setra Perdana)

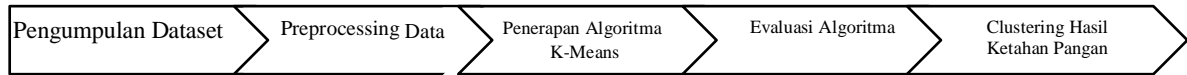
PENDAHULUAN

Pangan menjadi kebutuhan primer warga yang telah seharusnya terpenuhi, dan seiring bertambahnya populasi penduduk, kebutuhan akan pangan baik secara kualitas juga kuantitas harus semakin meningkat. Indonesia ialah salah satu negara dengan ketahanan pangan yang masih kurang stabil. Ketahanan pangan menjadi faktor yang sangat krusial pada menciptakan suatu negara, terutama di negara berkembang dikarenakan mempunyai dua peran penting, sehingga hal ini sebagai tujuan utama pada menciptakan ekonomi suatu negara. Kiprah pertama mempunyai fungsi pangan menjadi kondisi buat semua masyarakat mempunyai akses terhadap pangan menggunakan kualitas yang relatif baik buat kehidupan sehari-hari, sehat serta jua produktif. Peran kedua mempunyai arti fungsi ketahanan pangan yang esensial bagi pengembangan talenta kreatif serta produktif, penentu utama penemuan ilmiah, teknologi serta produktif di bidang ketenagakerjaan dan gizi. Selain itu ketahanan pangan pula terdiri berasal dari tiga subsistem: (1) Ketersediaan serta keamanan pangan, (2) Distribusi pasokan pangan bisa menjangkau semua wilayah, dan (3) konsumsi pangan yang relatif serta terpenuhinya gizi setiap warga sesuai dengan peraturan dan preferensi kesehatan yang terdapat.

Data mining menjadi teknik buat menjalankan proses manipulasi data dan menemukan data yang masih belum seksama, atau menggali berita berasal big data sehingga sebagai informasi data yang akurat serta sederhana. Metode data mining tak jarang ditemukan adalah clustering. Metode ini ditujukan buat mengelompokkan data berasal banyak sekali pola, titik, objek, serta sebagainya. Proses clustering bertujuan buat pengelompokan berita menjadi beberapa bagian sehingga berita pada satu perpaduan data memiliki banyak kemiripan serta perbedaan nyata menggunakan objek pada bagian yang tidak sama. Salah satu metode asal clustering yaitu K-Means, K-means merupakan salah satu algoritma clustering. Tujuan algoritma ini yaitu untuk membagi data menjadi beberapa kelompok. Algoritma ini menerima masukan berupa data tanpa label kelas. Hal ini berbeda dengan supervised learning yang menerima masukan berupa vektor $(-x-1, y_1), (-x-2, y_2), \dots, (-x-i, y_i)$, di mana x_i merupakan data dari suatu data pelatihan dan y_i merupakan label kelas untuk x_i .

Pada algoritma ini, komputer mengelompokkan sendiri data-data yang menjadi masukannya tanpa mengetahui terlebih dulu target kelasnya. Pembelajaran ini termasuk dalam unsupervised learning. Masukan yang diterima adalah data atau objek dan k buah kelompok (cluster) yang diinginkan. Algoritma ini akan mengelompokkan data atau objek ke dalam k buah kelompok tersebut. Pada setiap cluster terdapat titik pusat (centroid) yang merepresentasikan cluster tersebut. Penelitian ini menerapkan metode algoritma K-Means untuk memilih di Indonesia bagian mana saja yang mengalami penurunan dan peningkatan jumlah pangan. menggunakan metode yang digunakan penulis berharap bisa membantu memprediksi ketersediaan pangan pada wilayah Indonesia serta bagian apa saja yang perlu diperhatikan pemerintah agar rakyat pada wilayah tersebut memiliki ketersediaan pangan yang relatif.

METODE PENELITIAN



Gambar 1. Prosedur Kerja

1. Pengumpulan Dataset

Pada Tahap awal dilakukan proses mengumpulkan dataset hasil tanaman pangan yang terdiri dari ketahanan pangan jagung, kacang tanah, kacang hijau, dan kacang kedelai pertahun 2013-2015 yang diperoleh dari Badan Pusat Statistik (BPS) Indonesia.

2. Pre-Processing Data

Setelah dataset dikumpulkan kemudian dilanjutkan dengan pre-processing data. Tahap pre-processing pada penelitian ini terdiri dari proses cleaning data untuk menghilangkan data yang kurang relevan atau kurang valid agar menjadi data yang relevan. Kemudian dilanjutkan dengan proses normalisasi dan transformasi data atau proses perubahan format data non-numerik menjadi numerik sebelum diterapkan algoritma *K-Medoid* dalam melakukan pengelompokan hasil tanaman pangan.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

3. Penerapan Algoritma K-Means

Setelah tahap Pre-processing data selanjutnya dilakukan pengolahan data dengan menerapkan algoritma *K-means*. Data clustering menggunakan metode K-Means ini secara umum dilakukan dengan algoritma dasar sebagai berikut (Yudi Agusta, 2007) :

- Tentukan jumlah cluster
- Alokasikan data ke dalam cluster secara random
- Hitung centroid/ rata-rata dari data yang ada di masing-masing cluster
- Alokasikan masing-masing data ke centroid/ rata-rata terdekat
- Kembali ke Step 3, apabila masih ada data yang berpindah cluster atau apabila perubahan nilai centroid, ada yang di atas nilai threshold yang ditentukan atau apabila perubahan nilai pada objective function yang digunakan di atas nilai threshold yang ditentukan.

HASIL DAN PEMBAHASAN

Penelitian ini menggunakan *dataset* dari Badan Pusat Statistik (BPS) Indonesia pada tahun 2013-2015. Tahap normalisasi data dilakukan dengan Microsoft Excel 2019 dan hasil pengelompokan dengan menerapkan algoritma *K-Means* menggunakan *software* R studio.

Tabel 1. Data Tanaman Pangan

	Padi	Jagung	Kedelai	Kacang tanah	Kacang hijau	Ubi Kayu	Ubi Jalar
ACEH	1956940.00	177842.00	45027.00	3861.00	955.00	34738.00	11602.00
SUMATERA UTARA	3727249.00	1183011.00	3229.00	11351.00	2344.00	1518221.00	116671.00
SUMATERA BARAT	2430384.00	547417.00	732.00	9093.00	753.00	218830.00	134453.00
RIAU	434144.00	28052.00	2211.00	1243.00	619.00	103070.00	8462.00
JAMBI	664535.00	25690.00	2372.00	1513.00	262.00	33291.00	68187.00
SUMATERA SELATAN	3676723.00	167457.00	5140.00	3475.00	1821.00	165250.00	15945.00
BENGKULU	622832.00	93988.00	3987.00	4679.00	1371.00	62193.00	31672.00
LAMPUNG	3207002.00	1760278.00	6156.00	10676.00	2643.00	8329201.00	45141.00
KEP. BANGKA BELITUNG	28480.00	783.00	N/A	357.00	N/A	14203.00	2863.00
KEP. RIAU	1370.00	790.00	18.00	168.00	N/A	8530.00	1891.00
DKI JAKARTA	10268.00	N/A	N/A	N/A	N/A	N/A	N/A
JAWA BARAT	12083162.00	1101998.00	51172.00	91573.00	11002.00	2138532.00	485065.00

Pada Tabel 1. dataset yang digunakan antara lain, data tanaman pangan padi, jagung, kacang hijau, kacang kedelai, kacang tanah, ubi kayu, dan ubi jalar dari 35 daerah di Indonesia pada tahun 2013 – 2015 sedangkan variable yang digunakan dalam melakukan pengelompokan berdasarkan rata-rata produksi bahan pangan, jumlah produksi bahan pangan, dan luas panen bahan pangan. Proses normalisasi data untuk menghilangkan data yang kurang akurat sebelum diterapkan algoritma *K-Means* dapat dilihat pada Tabel 2.

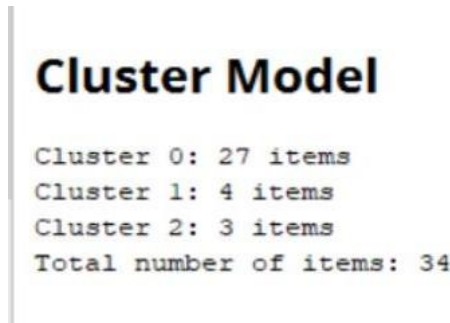
Tabel 2. Hasil Standarisasi Data Tanaman Pangan

	Padi	Jagung	Kedelai	Kacang tanah	Kacang hijau	Ubi Kayu	Ubi Jalar
[1.]	-0.1746	-0.2742	0.0032	-0.2960	-0.2932	-0.3136	[1.] -0.3058
[2.]	-0.0285	0.0372	-0.2937	-0.2350	-0.2552	0.0351	[2.] -0.0490
[3.]	-0.1356	-0.1597	-0.3114	-0.2534	-0.2987	-0.2703	[3.] -0.0055
[4.]	-0.3003	-0.3206	-0.3009	-0.3173	-0.3023	-0.2975	[4.] -0.3135
[5.]	-0.2813	-0.3213	-0.2998	-0.3151	-0.3121	-0.3139	[5.] -0.1675
[6.]	-0.0327	-0.2774	-0.2801	-0.2991	-0.2695	-0.2829	[6.] -0.2952
[7.]	-0.2847	-0.3001	-0.2883	-0.2893	-0.2818	-0.3071	[7.] -0.2567
[8.]	-0.0715	0.2160	-0.2729	-0.2405	-0.2471	1.6360	[8.] -0.2238
[9.]	-0.3338	-0.3290	-0.3166	-0.3245	-0.3192	-0.3184	[9.] -0.3271
[10.]	-0.3360	-0.3290	-0.3165	-0.3261	-0.3192	-0.3197	[10.] -0.3295
[11.]	-0.3353	-0.2742	-0.3166	-0.2960	-0.3192	-0.3136	[11.] -0.3058
[12.]	0.6610	0.0121	0.0469	0.4183	-0.0189	0.1809	[12.] 0.8514
[13.]	0.5176	0.5787	0.3889	0.7152	1.4352	0.6395	[13.] 0.1148
[14.]	-0.2601	-0.2396	-0.0916	0.2494	-0.3105	-0.0835	[14.] -0.3220
[15.]	0.6582	1.4553	2.0237	1.3663	1.2553	0.5247	[15.] 0.6269
[16.]	-0.1642	-0.3255	-0.2433	-0.2231	-0.3009	-0.2987	[16.] -0.2658
[17.]	-0.2633	-0.3114	-0.2638	-0.2377	-0.2869	-0.2848	[17.] -0.1857
[18.]	-0.1551	-0.1329	0.3303	0.0137	0.2834	-0.3078	[18.] -0.3064
[19.]	-0.2759	-0.1100	-0.3047	-0.1967	-0.0425	-0.1311	[19.] -0.1412
[20.]	-0.2171	-0.2797	-0.3047	-0.3167	-0.3041	-0.2821	[20.] -0.2968
[21.]	-0.2691	-0.3273	-0.3046	-0.3223	-0.3164	-0.3121	[21.] -0.3117
[22.]	-0.1685	-0.2961	-0.2877	-0.2359	-0.2986	-0.3012	[22.] -0.2937
[23.]	-0.2999	-0.3278	-0.3066	-0.3156	-0.3090	-0.3087	[23.] -0.3024
[24.]	-0.3258	-0.3290	-0.3160	-0.3255	-0.3173	-0.3140	[24.] -0.3265
[25.]	-0.2834	-0.1905	-0.2755	-0.2557	-0.2772	-0.3087	[25.] -0.3265
[26.]	-0.2510	-0.2861	-0.2267	-0.2680	-0.2963	-0.2980	[26.] -0.2369
[27.]	0.0794	0.0580	0.0080	-0.0961	0.1814	-0.2199	[27.] -0.2815
[28.]	-0.2898	-0.3083	-0.2911	-0.2872	-0.2897	-0.2793	[28.] -0.1612
[29.]	-0.3117	-0.1220	-0.2853	-0.3170	-0.3143	-0.3207	[29.] -0.2752
[30.]	-0.2994	-0.2895	-0.3082	-0.3226	-0.3024	-0.3093	[30.] -0.3292
[31.]	-0.3277	-0.3256	-0.3148	-0.3158	-0.2950	-0.2987	[31.] -0.3061
[32.]	-0.3301	-0.3201	-0.3079	-0.2887	-0.3104	-0.2936	[32.] -0.2862
[33.]	-0.3337	-0.3286	-0.3119	-0.3222	-0.3141	-0.3189	[33.] -0.2437
[34.]	-0.3221	-0.3271	-0.2839	-0.3108	-0.3006	-0.3126	[34.] -0.2977
[35.]	5.5461	5.4052	5.2239	5.3870	5.2672	5.3044	[35.] 0.6570

1. Hasil Clustering Ketahanan Pangan dengan Algoritma *K-Means*

Penelitian ini menerapkan *Indeks Davies Bouldin (DBI)* dalam penentuan Jumlah Cluster yang optimum untuk menggolongkan hasil pangan setiap tahun. Adapun tahapan-tahapan adalah sebagai berikut :

- a. Membangun model clustering memanfaatkan algoritma *K-Means* dengan jumlah *k-cluster* yang akan diuji adalah 5 sampai 10 *cluster*.



Gambar 1. Anggota Cluster

- b. Centroid Table Centroid table merupakan titik pusat dalam menentukan banyaknya jumlah data yang ada pada setiap cluster. Tampilan centroid table dapat dilihat pada gambar berikut ini:

Attribute	cluster_0	cluster_1	cluster_2
PROVINSI	18.963	10.750	13.333
JAGUNG	394724.889	1903812.250	3927600.667
KACANG TANAH	7415.519	6976.750	94690.667
KEDELAI	11851.444	38895.750	169022
PADI	921734.630	5313275	12299071.667
UBI KAYU	133153.333	2072951.250	2484782.667

Gambar 2. Tampilan Centroid Table

KESIMPULAN

Berdasarkan hasil analisis terhadap data produksi bahan pangan menurut provinsi menggunakan metode K- Means, maka dapat disimpulkan hal-hal sebagai berikut:

1. Penerapan metode K-Means Clustering untuk mengelompokkan provinsi menurut jumlah produksi bahan pangan dimulai dari analisa data yang terdiri dari produksi jagung, produksi kacang tanah, produksi ubi kayu, produksi kedelai, dan produksi padi.
2. Diperoleh 3 provinsi tingkat tinggi dalam produksi bahan pangan yaitu Jawa Barat, Jawa Tengah, dan Jawa Timur, 4 provinsi dengan tingkat sedang, dan 27 provinsi dengan tingkat produksi bahan pangan rendah.
3. Clustering yang dihasilkan dapat menjadi masukan kepada pemerintah dalam menganalisa provinsi mana saja yang memiliki jumlah produksi bahan pangan yang dominan terhadap jagung, kacang tanah, ubi kayu, kedelai, dan padi.
4. Berdasarkan hasil perbandingan cara manual dengan software menampilkan hasil akhir yang sama yang dapat dilihat pada cluster model, folder view, dan centroid data.

DAFTAR PUSTAKA

<http://www.ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/4939>

<https://www.bps.go.id/indicator/53/23/1/produksi.html>

https://www.academia.edu/download/60095974/jurnal_1411420190723-61818-3emuou.pdf

<https://www.ejurnal.stmik-budidarma.ac.id/index.php/komik/article/view/1681>