



IMPELEMENTASI ALGORITMA K-MEANS DALAM CLUSTERING BIBIT KAKAO UNGGUL (STUDI KASUS :KEBUNAN KAKAO PASAR 4 PATUMBAK)

Vini Kristin Septiani Situmorang¹, Efori Bu'ulolo², Rian Syahputra³

¹²Program Studi Informatika, ³Program Studi Sistem Informasi, Universitas Budi Darma, Medan
e-mail: ¹vinikssitumorang@gmail.com, ²buuloloefori21@gmail.com, ³ryansyah93@gmail.com

Abstrak

Kakao merupakan salah satu tanaman yang berjenis tanaman perkebunan yang berperan penting dalam perekonomian Indonesia dan terus mendapat perhatian untuk dikembangkan. Upaya pengembangan tanaman kakao disamping masih diarahkan pada peningkatan populasi (luas lahan) juga telah banyak diarahkan pada peningkatan jumlah produksi dan mutu hasil. Permasalahan yang sering terjadi pada Kebun kakao di daerah pasar 4 adalah petani yang sering salah dalam memilih bibit unggul kakao. Petani kakao tidak memanfaatkan data yang ada. Oleh karena itu, dibutuhkan suatu teknik yaitu data mining agar dapat memecahkan permasalahan yang ada. Pada penelitian ini penulis menggunakan metode K-Means agar memudahkan penulis dalam mengetahui bibit unggul. Berdasarkan penerapannya dengan menggunakan algoritma K-Means, clustering bibit unggul kakao mendapatkan hasil 2 clustering.

Kata kunci : K-Means, Clustering, bibit, kakao

Abstract

Cocoa is one of the plantation crops that plays an important role in the Indonesian economy and continues to receive attention for development. Efforts to develop cocoa plants in addition to still being directed at increasing the population (land area) have also been directed at increasing the amount of production and quality of results. The problem that often occurs in cocoa farms in market area 4 is that farmers are often wrong in choosing superior cocoa seeds. Cocoa farmers do not take advantage of existing data. Therefore, we need a technique that is data mining in order to solve the existing problems. In this study, the author uses the K-Means method to make it easier for the author to identify superior seeds. Based on its application using the K-Means algorithm, the clustering of superior cocoa seeds obtained 3 clustering results.

Keywords : K-Means, Clustering, seeds, cocoa

I. PENDAHULUAN

Kakao adalah Pohon budaya di perkebunan yang berasal dari Amerika Selatan, tetapi sekarang ditanam di berbagai kawasan tropika salah satunya yaitu Indonesia. Dari biji tumbuhan ini dihasilkan produk olahan yang dikenal sebagai coklat. Kandungan kakao (biji coklat) lebih dari 70% memiliki manfaat kesehatan, karena coklat kaya akan kandungan antioksidan yaitu fenol dan flavonoid dengan adanya antioksidan yang sangat dibutuhkan oleh tubuh untuk melawan radikal bebas dan besarnya antioksidan di dalam coklat 3 kali lebih besar dari teh hijau[1].

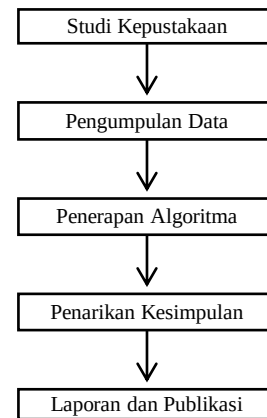
Pada pertumbuhan kakao itu cukup sangat mudah yaitu dengan melakukan pembibitan di awal tanaman dan dibiarkan tumbuh di kebun yang memiliki kadar air yang cukup, namun banyak petani yang kesulitan dalam memilih bibit kakao yang unggul sehingga petani sering salah memilih bibit yang unggul yang akhirnya berakibat pada tidak maksimalnya hasil perkebunan kakao dari bibit tersebut. Kesulitan para petani disebabkan karena bibit kakao memiliki bentuk, ukuran dan warna yang homogen [2][1]. Oleh karena itu sangat diperlukan teknik dalam clustering bibit kakao unggul.

Data mining merupakan suatu proses menemukan korelasi yang berarti, pola serta kecenderungan dengan memeriksa pada sekumpulan data besar yang tersimpan dalam penyimpanan, menggunakan teknik pengenalan pola seperti teknik statistik dan matematika[3][4]. Algoritma clustering dapat digunakan untuk menemukan jumlah cluster yang tepat pada data bibit unggul kakao[5][5].

Adapun tujuan dari penelitian ini adalah memilih bibit kakao unggul dan mengimplementasikan[6]. algoritma clustering K-Means dalam clustering bibit kakao unggul, serta manfaat dari penelitian ini adalah memberikan kemudahan cluster biji kakao unggul dan dengan bibit kakao yang terpilih unggul akan memberikan hasil yang maksimal pada hasil panen[7].

II. METODOLOGI PENELITIAN

Adapun metodologi penelitian yang dilakukan pada penelitian ini adalah sebagai berikut:



Gambar 1 Kerangka metodologi penelitian

Metode penelitian diawali dengan studi pustaka, dimana penulis mencari, mengumpulkan[8] dan memahami referensi yang berkaitan dengan judul penelitian.[9] Kemudian penulis mengumpulkan data untuk, data yang dikumpulkan adalah data kakao serta kriterianya masing-masing. Selanjutnya data tersebut dianalisis dengan algoritma K-Means untuk mengolompokkan bibit[10] kakao yang unggul. Dari hasil analisis tersebut ditarik sebuah kesimpulan yang menjadi informasi / pengetahuan yang diperoleh dari data. Tahap terakhir adalah pembuatan laporan penelitian dan publikasi[8].

III. HASIL DAN PEMBAHASAN

Untuk Menentukan jarak masing *centroid* berdasarkan jarak terpendek, kriteria, kondisi atau karakteristik yang sama atau hampir sama, menggunakan model *Euclidean* yang dirumuskan sebagai berikut:

$$d_{ij} = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2}$$

Ket :

d_{ij} = Jarak dari data ke i ke pusat *cluster* j

x_{ki} = Data dari ke-i pada *attribute* data ke-k



x_{kj} = Data dari ke-j pada *attribute* data ke-k

Lonjong	1	Coklat	1
Bulat	2	Hitam	2

Berikut ini adalah data biji kakao yang akan dicluster:

Tabel 1. Kriteria biji kakao

No	Nama	Bentuk	Warna	Berat
1	K1	Bulat	Coklat	0,06
2	K2	Bulat	Hitam	0,09
3	K3	Lonjong	Coklat	0,02
4	K4	Bulat	Coklat	0,04
5	K5	Lonjong	Hitam	0,07
6	K6	Lonjong	Coklat	0,05

Data dalam bentuk string dirubah dalam bentuk angka seperti tabel berikut:

Tabel 2. Konversi nilai string ke angka

Bentuk	Angka Bentuk	Warna	Angka Warna
--------	--------------	-------	-------------

Tabel 3. Kriteria biji kakao setelah dikonversi

No	Nama	Bentuk	Warna	Berat
1	K1	2	1	0,06
2	K2	2	2	0,09
3	K3	1	1	0,02
4	K4	2	1	0,04
5	K5	1	2	0,07
6	K6	1	1	0,05

Tabel 4 Nilai Centroid Awal

K1	2	1	0,06
K6	1	1	0,05

Rekapitulasi hasil perhitungan cluster pada iterasi ke-1 sebagai berikut.

Tabel 5 Rakapitulasi Hasil Perhitungan Cluster Ke-1

No	Nama	Bentuk	Warna	Bert	dc1	dc2	Cluster
1	K1	2	1	0,06	0	1,00	1
2	K2	2	2	0,09	1,00	1,41	1
3	K3	1	1	0,02	1,00	0,03	2
4	K4	2	1	0,04	0,00	1,00	1
5	K5	1	2	0,07	1,41	1,00	2
6	K6	1	1	0,05	1,00	0	2

Aturan yang di gunakan untuk mengelompokan/cluster data berdasarkan hasil perhitungan jarak adalah

- Jika $dc1 < dc2$ dan $dc1 < d3$ maka cluster = 1
- Jika $dc2 < dc1$ dan $dc2 < d2$ maka cluster = 2

Iterasi Ke-2 :

Pertama, pencarian nilai centroid / titik pusat awal pada iterasi ke-2 diperoleh dari nilai rata-rata penjumlahan anggota masing – masing cluster, setelah itu dibagi dengan jumlah anggota masing – masing cluster.

Perolehan nilai C1 :

Tabel 6 Perolehan Nilai C1 Untuk Iterasi Ke-2

No	Nama	Bentuk	Warna	Berat	Cluster
1	K1	2	1	0,06	1
2	K2	1	2	0,09	1
3	K4	2	1	0,04	1
C1		$(2+1+2)/3 = 1,6$	$(1+2+1)/3 = 1,3$	$(0,06+0,09+0,04)/3 = 0,063$	

Perolehan Nilai C2 :



Tabel 7 Perolehan Nilai C2 Untuk Iterasi Ke-2

No	Nama	Bentuk	Warna	Berat	Cluster
1	K3	1	1	0,02	2
2	K5	1	2	0,07	2
3	K6	1	1	0,05	2
	C2	$(1+1+1)/3 = 1$	$(1+2+1)/3 = 1,3$	$(0,02+0,07+0,05)/3 = 0,046$	

Sehingga nilai centroid / awal/titik pusat awal adalah sebagai berikut.

Tabel 8 Nilai Centroid Awal Untuk Iterasi Ke-2

C1	1,6	1,3	0,063
C2	1	1,3	0,046

Kedua : Menghitung jarak dengan menggunakan model Euclidean pada iterasi ke- 2 . Rekapitulasi hasil perhitungan cluster pada iterasi ke-2 sebagai berikut :

Tabel 9 Rekapitulasi Hasil Perhitungan Pada iterasi Ke- 2

No	Nama	Bentuk	Berat	Warna	dc1	dc2	Cluster
1	K1	2	1	0,06	0,50	1,04	1
2	K2	2	2	0,09	0,80	1,22	1
3	K3	1	1	0,02	0,67	0,30	2
4	K4	2	1	0,04	0,50	1,04	1
5	K5	1	2	0,07	0,92	0,70	2
6	K6	1	1	0,05	0,67	0,30	2

Aturan yang di gunakan untuk mengelompokan / cluster data berdasarkan hasil perhitungan jarak adalah

- Jika $dc1 < dc2$ dan $dc1 < d3$ maka cluster = 1
- Jika $dc2 < dc1$ dan $dc2 < d2$ maka cluster = 2

Hasil cluster pada iterasi ke-1 dan iterasi ke-2 sama, maka proses pembentukan cluster berhenti sehingga hasil cluster adalah sebagai berikut:

Tabel 10. Hasil Akhir:

Cluster 1	Cluster 2
K1	K3
K2	K5
K4	K6

Maka berdasarkan Tabel 10. Hasil akhir, maka clustering bibit kakao unggul berada pada cluster 1 karena data cluster 1 cenderung ke kriteria bibit kakao

unggul baik dari segi warna, bentuk maupun berat.



IV. KESIMPULAN

Setelah dilakukan tahap dalam pengujian data bibit kakao unggul, maka dapat diambil kesimpulan bahwa dengan menggunakan Algoritma K-Means clustering ini dapat membagi menjadi 2 clustering bibit kakao unggul berdasarkan kriteria warna, bentuk dan berat. Berdasarkan hasil perbandingan dari kedua cluster tersebut maka bibit yang layak untuk digunakan oleh petani adalah berada pada cluster 1.

V. SARAN

Berdasarkan hasil penelitian di atas untuk mengetahui atau untuk mendapatkan bibit kakao yang unggul perlu memperhatikan lagi di bagian cara merawat dan cara pemberian pupuk dan juga jenis biji kakao yang ingin dijadikan sebagai bibit yang ingin di tanam.



VI. DAFTAR PUSTAKA

- [1] M. Trisanti Saragih, H. Harianto, and H. Kuswanti, "Pengaruh Penerapan Bea Keluar Biji Kakao Terhadap Daya Saing Serta Ekspor Produk Kakao Indonesia," *Forum Agribisnis*, vol. 11, no. 2, pp. 133–152, 2021, doi: 10.29244/fagb.11.2.133-152.
- [2] S. Rosniawaty, R. Sudirja, and H. Afrianto, "Pemanfaatan urin kelinci dan urin sapi sebagai alternatif pupuk organik cair pada pembibitan kakao (*Theobroma cacao* L.)," *Kultivasi*, vol. 14, no. 1, pp. 32–36, 2015, doi: 10.24198/kultivasi.v14i1.12094.
- [3] Y. Prayoga, H. S. Tambunan, and I. Parlina, "Penerapan Clustering Pada Laju Inflasi Kota Di Indonesia Dengan Algoritma K-Means," *BRAHMANA J. Penerapan Kecerdasan Buatan*, vol. 1, no. 1, pp. 24–30, 2019, doi: 10.30645/brahmana.v1i1.4.
- [4] A. Lia Hananto *et al.*, "Analysis of Drug Data Mining with Clustering Technique Using K-Means Algorithm," *J. Phys. Conf. Ser.*, vol. 1908, no. 1, 2021, doi: 10.1088/1742-6596/1908/1/012024.
- [5] mohamad jajuli nurul rohmawati, sofi defiyanti, "Implementasi Algoritma K-Means Dalam Pengklasteran Mahasiswa Pelamar Beasiswa," *Jitter 2015*, vol. I, no. 2, pp. 62–68, 2015.
- [6] Z. Syahara, R. N. Adiha, and A. P. Windarto, "Implementasi Data Mining Algoritma Apriori Pada Sistem," vol. 2, no. 2, pp. 107–115, 2021.
- [7] A. Azis, A. Rosmana, and V. Dewi, "Pengendalian Penyakit Hawar Daun Phytophthora pada Bibit Kakao dengan *Trichoderma asperellum*," *J. Fitopatol. Indones.*, vol. 9, no. 1, pp. 15–20, 2013, doi: 10.14692/jfi.9.1.15.
- [8] F. Nurdiyansyah and I. Akbar, "Implementasi Algoritma K-Means untuk Menentukan Persediaan Barang pada Poultry Shop," *J. Teknol. dan Manaj. Inform.*, vol. 7, no. 2, pp. 86–94, 2021, doi: 10.26905/jtmi.v7i2.6377.
- [9] A. Nurzahputra, M. A. Muslim, and M. Khusniati, "Penerapan Algoritma K-Means Untuk Clustering Penilaian Dosen Berdasarkan Indeks Kepuasan Mahasiswa," *Techno.Com*, vol. 16, no. 1, pp. 17–24, 2017, doi: 10.33633/tc.v16i1.1284.
- [10] C. Desiana, I. S. Banuwa, R. Evizal, and S. Yusnaini, "PENGARUH PUPUK ORGANIK CAIR URIN SAPI DAN LIMBAH TAHU TERHADAP PERTUMBUHAN BIBIT KAKAO (*Theobroma cacao* L.)," *J. Agrotek Trop.*, vol. 1, no. 1, pp. 113–119, 2013, doi: 10.23960/jat.v1i1.1927.