

# Implementasi Data Mining Pada *Clustering* Kualitas Bawang Merah Palu (*Alliumascalonium L.*)

Desy Lusiyanti<sup>a,\*</sup>, Selvy Musdalifah<sup>b</sup>, Agusman Sahari<sup>c</sup>,

<sup>a, b, c, d</sup> Universitas Tadulako, Palu, 94111, Indonesia

\* Alamat Surel: [desy.lusiyanti@untad.ac.id](mailto:desy.lusiyanti@untad.ac.id)

## Abstrak

Bawang merah Palu (*Allium Ascalonium L.*) merupakan salah satu sayuran umbi serbaguna yang memiliki nilai ekonomis tinggi. Di Sulawesi Tengah, ada ciri khas serikat yang disebut "Bawang Merah Palu", namun tidak semua Bawang Merah Palu memiliki kualitas yang sama. Oleh karena itu diperlukan suatu teknik untuk mendapatkan clustering kualitas bawang merah palu. Tujuan dari penelitian ini adalah untuk mendapatkan clustering kualitas bawang merah Palu dengan menggunakan algoritma K-Means. Proses clustering dilakukan dengan menggunakan 48 data. Atribut penelitian yang digunakan adalah ketuaan, kekerasan, busuk, kerusakan, kadar air dan diameter. Hasil yang diperoleh yaitu cluster pertama terdiri dari 28 data dan cluster kedua terdiri dari 20 data dengan parameter yang berpengaruh adalah kekerasan. Berdasarkan hasil penelitian, Bawang merah Palu untuk Desa Maranatha, Desa Sidera, dan Desa Watubula memiliki kualitas yang sama yakni berada pada klaster pertama sedangkan untuk Desa Waturalele, desa Maku, Desa Sidondo, Desa Oloboju, dan Desa Soulowe berada pada klaster kedua.

Kata kunci: Klastering, Kohonen, Bawang Merah Palu.

© 2023 Dipublikasikan oleh Jurusan Matematika, Universitas Negeri Semarang

## 1. Pendahuluan

Bawang Merah (*Allium Ascalonium L.*) merupakan salah satu sayuran umbi multiguna yang mempunyai nilai ekonomis yang cukup tinggi. Bawang merah banyak dimanfaatkan antara lain sebagai bumbu penyedap masakan maupun sebagai bahan baku obat-obatan, dan komoditi ini mempunyai peranan cukup penting di dunia perdagangan (Rukmana, 1995).

Di Sulawesi Tengah terdapat jenis bawang merah yang biasa disebut "Bawang Merah Palu". Bawang Merah Palu memiliki cita rasa yang khas dan cocok digunakan sebagai bawang goreng sehingga biasa juga disebut bawang goreng palu, dimana bawang goreng yang dihasilkan mempunyai sifat yang tidak mudah gosong apabila digoreng dan tetap garing apabila disimpan lama, serta aromanya yang tetap harum. Namun demikian tidak semua bawang merah Palu memiliki kualitas yang sama, oleh karena itu diperlukan suatu teknik untuk mendapatkan clustering kualitas bawang merah palu.

Clustering kualitas bawang merah Palu dapat dilakukan dengan menggunakan metode-metode data mining. Data mining sudah mulai banyak dimanfaatkan sebagai solusi terhadap berbagai macam kasus yang muncul dalam kehidupan sehari-hari. Penelitian mengenai Clustering kualitas bawang merah sebelumnya sudah pernah dilakukan. Desy Lusiyanti (2021) dalam penelitiannya telah mengaplikasikan metode data mining yaitu Kohonen's Self-Organizing Maps untuk mengclusterkan kualitas bawang merah Palu. Hasil penelitian diperoleh data bawang merah palu yang dibagi menjadi 2 cluster masing-masing Cluster1 terdiri dari 6 data (tidak ada yang memenuhi standar mutu) dan Cluster2 terdiri dari 42 data (26 data memenuhi syarat mutu II sesuai SNI 01-3159-1992 dan 16 data tidak memenuhi standar mutu). Daerah di Kabupaten Sigi yaitu desa sidondo desa oloboju dan desa soulowe memiliki kualitas bawang merah yang baik. Namun dari penelitian ini belum diperoleh akurasi dari hasil cluster yang dilakukan, sehingga penelitian ini akan dilakukan dengan menggunakan analisis cluster dengan memanfaatkan metode yang ada pada data mining. Salah satu metode yang bisa digunakan adalah K-Means.

To cite this article:

Lusiyanti, D., Musdalifah, S. & Sahari, A. (2023). Implementasi Data Mining Pada Clustering Kualitas Bawang Merah Palu (*Alliumascalonium L.*). *PRISMA, Prosiding Seminar Nasional Matematika 6*, 724-728

K-Means merupakan salah satu metode data clustering non-hierarki yang mengelompokkan data dalam bentuk satu atau lebih cluster atau kelompok (Prasetyo, 2014). Algoritma K-Means merupakan pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah K cluster yang sudah ditetapkan di awal. Algoritma K-Means sederhana untuk diimplementasikan dan dijalankan, relatif cepat, mudah beradaptasi, umum penggunaannya dalam praktek (Wu dan Kumar, 2009). Penelitian ini akan mengclusterkan kualitas bawang merah Palu berdasarkan parameter-parameter standar mutu bawang merah (SNI 01-3159-1992) yaitu ketuaan, kekerasan, diameter, kerusakan, kebusukan dan kadar air. Diharapkan hasil penelitian ini dapat memberikan kontribusi dibidang pertanian dan masyarakat petani bawang merah Palu pada khususnya.

## 2. Metode

Penelitian ini dilakukan dalam beberapa tahapan sebagai berikut :

### Tahap 1 Identifikasi Permasalahan

Pada tahap pertama ini akan dilakukan identifikasi permasalahan yang akan dibahas yaitu mulai dari mencari topik, literatur-literatur yang berhubungan dengan bahan penelitian dan membuat proposal penelitian.

### Tahap 2 Pengambilan dan Pengolahan Data

Pada tahap ini dilakukan penentuan data input dan target. Variabel yang digunakan dalam penelitian ini adalah ketuaan ( $x_1$ ), kekerasan ( $x_2$ ), kebusukan ( $x_3$ ), kerusakan ( $x_4$ ), kadar air ( $x_5$ ), diameter ( $x_6$ ). Metode pengambilan sampel yang digunakan adalah teknik quota sampling, yaitu teknik pengambilan sampel dengan cara menetapkan jumlah tertentu sebagai target yang harus dipenuhi dalam pengambilan sampel dari populasi. Sebelum data diolah menggunakan metode K-Means, data tersebut perlu ditransformasi. Adapun ketentuan transformasi dari masing-masing variabel penelitian adalah sebagai berikut :

**Tabel 1.** Variabel dan Transformasinya

Variabel	Deskripsi Variabel	Transformasi
$x_1$	Ketuaan	Nilai 1 untuk tua (60-70 HST) Nilai 0 untuk cukup tua (< 60 HST)
$x_2$	Kekerasan	Nilai 1 untuk keras Nilai 0 untuk cukup keras
$x_3$	Busuk	Nilai 1 untuk busuk Nilai 0 untuk tidak busuk
$x_4$	Kerusakan	Nilai 1 untuk rusak Nilai 0 untuk tidak rusak
$x_5$	Kadar Air	Nilai 1 untuk 80-85% Nilai 0 untuk 75-80%
$x_6$	Diameter	Nilai 1 untuk > 2,5 cm Nilai 0,5 untuk 1,5-2,5 cm Nilai 0 untuk < 1,5 cm

### Tahap 3 Implementasi Metode

Data yang telah ditransformasi, selanjutnya akan diolah dengan menggunakan metode K-Means.

### Tahap 4 Analisa dan Pembahasan

Setelah pengolahan data dilakukan, kemudian data tersebut akan dianalisa dengan memperoleh hasil akhir berupa cluster wilayah terhadap kualitas bawang merah Palu.

## 3. Hasil dan Pembahasan

Clustering menggunakan Algoritma K-Means terdiri dari beberapa tahapan. Tahapan yang paling awal dari proses ini adalah penentuan jumlah cluster. Cluster menentukan berapa banyak kelompok data yang ingin

kita buat. Dalam penelitian ini digunakan 2 cluster. Setelah menentukan jumlah cluster, langkah selanjutnya yaitu menentukan pusat cluster (centroid). Jumlah centroid yang dibuat disesuaikan dengan jumlah cluster yang di tentukan sebelumnya yaitu dua cluster. Centroid berupa angka yang di pilih secara acak.

Langkah berikutnya yaitu melakukan perhitungan jarak antara titik centroid dengan data yang ada. Perhitungan ini dilakukan dengan menggunakan persamaan Euclidean Distance. Metode Euclidean Distance yaitu metode dengan menghitung jarak antara dua buah objek. Metode ini biasa disebut jarak Euclidean. Berikut merupakan formula yang dipakai menggunakan persamaan Euclidean Distance :

$$d_2(x, y) = \|x - y\|_2 = \sqrt{\sum_{i=1}^r (x_i - y_i)^2}$$

dimana  $x, y \in X$  dan  $x_i, y_i$  adalah nilai fitur ke- $i$  dari  $x$  dan  $y$ , sedangkan  $r$  adalah jumlah fitur dalam vektor. Setelah melakukan perhitungan jarak, dilanjutkan dengan membentuk nilai centroid baru dengan cara yaitu mengelompokkan data berdasar dimensi, pengelompokan diambil dengan memilih jarak terpendek terhadap centroid. Setelah menemukan kelompok cluster berikutnya menghitung nilai rata-rata. Nilai rata-rata ini akan menentukan nilai dari centroid berikutnya. Setelah mendapatkan nilai centroid baru, maka dilakukan tahapan yang sama dengan proses sebelumnya, yaitu dengan melakukan perhitungan jarak terdekat menggunakan persamaan Euclidean Distance untuk cluster 1 dan cluster 2, namun menggunakan titik centroid yang baru. Langkah ini akan dilakukan sampai kondisi berhenti terpenuhi. Hasil akhir yang diperoleh adalah nilai centroid yang bisa digunakan untuk mengclusterkan data baru. Berikut hasil pengelompokan atau clustering bawang merah Palu.

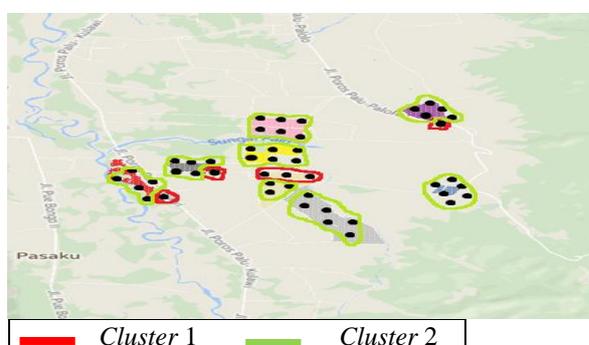
**Tabel 2.** Hasil clustering pembagian 2 cluster

Sampel	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	Cluster	
							1	2
1	0	0	0	0	0	0,5		*
2	0	0	0	0	0	0,5		*
3	0	0	0	0	0	0,5		*
4	0	0	0	0	0	0,5		*
5	0	0	0	0	0	0,5		*
6	0	1	0	1	1	0,5	*	
7	0	0	0	0	0	0,5		*
8	0	0	0	0	0	0,5		*
9	0	0	1	0	0	0		*
10	0	0	1	0	0	0,5		*
11	0	0	0	0	0	0		*
12	0	1	0	0	0	1		*
13	1	1	0	0	0	1	*	
14	1	1	0	0	1	0,5	*	
15	1	1	0	0	1	0,5	*	
16	1	0	0	0	0	0,5	*	
17	1	0	0	0	0	0,5	*	
18	1	0	0	0	0	0,5	*	
19	0	0	0	0	0	0,5		*
20	0	0	0	0	0	0,5		*
21	0	0	1	0	0	0,5		*
22	0	0	0	0	0	0,5		*
23	0	0	0	0	0	0,5		*
24	0	0	0	0	0	0,5		*
25	1	0	0	0	0	0,5	*	
26	1	0	0	0	0	0,5	*	
27	1	0	0	0	0	0,5	*	
28	1	1	0	1	1	0	*	
29	1	0	0	0	0	0,5	*	

30	1	0	0	0	0	0,5	*
31	0	0	0	0	0	0,5	*
32	0	0	0	0	0	0,5	*
33	0	0	0	0	0	0,5	*
34	0	0	0	0	0	0,5	*
35	0	0	0	0	0	0,5	*
36	0	0	0	0	0	0,5	*
37	0	0	0	0	0	0,5	*
38	0	0	0	0	0	0,5	*
39	0	0	0	0	0	0,5	*
40	0	0	0	0	0	0,5	*
41	0	0	0	0	0	0,5	*
42	0	0	0	0	0	0,5	*
43	1	0	0	0	0	0	*
44	1	0	0	0	1	0,5	*
45	1	0	0	0	0	0,5	*
46	1	0	0	0	0	0,5	*
47	1	0	0	0	0	0,5	*
48	1	0	0	0	1	0,5	*

Keterangan : $x_1$  = ketuaan,  $x_2$  = kekerasan,  $x_3$  = kebusukan,  $x_4$  = kerusakan,  $x_5$  = kadar air,  $x_6$  = diameter

Hasil clustering data yang dibagi menjadi 2 cluster masing-masing cluster 1 terdiri dari 28 data dan cluster 2 terdiri dari 20 data. Berdasarkan hasil penelitian, Bawang merah Palu untuk Desa Maranatha, Desa Sidera, dan Desa Watubula memiliki kualitas yang sama yakni berada pada klaster pertama sedangkan untuk Desa Waturalele, desa Maku, Desa Sidondo, Desa Oloboju, dan Desa Soulowe berada pada klaster kedua.



**Gambar 1.** Hasil *clustering* pembagian 2 cluster

Berikut adalah nilai centroid akhir yang diperoleh. Nilai centroid ini dapat digunakan untuk mengclusterkan data baru.

$$C1 = [0.9474 \quad 0.2632 \quad 0 \quad 0.1053 \quad 0.3158 \quad 0.4737]$$

$$C2 = [ \quad 0 \quad 0.0345 \quad 0.1034 \quad 0 \quad 0 \quad 0.4828]$$

#### 4. Simpulan

Hasil clustering dari 48 data bawang merah palu yang dibagi menjadi 2 cluster masing-masing Cluster1 terdiri dari 28 data dan Cluster2 terdiri dari 20 data . Daerah di Kabupaten Sigi yaitu Desa Maranatha, Desa Sidera, dan Desa Watubula memiliki kualitas bawang merah yang sama yakni berada pada klaster pertama

---

sedangkan untuk Desa Waturalele, desa Maku, Desa Sidondo, Desa Oloboju, dan Desa Soulowe berada pada klaster kedua.

---

### **Daftar Pustaka**

- Fausett, L.V. (1993). *Fundamental of Neural Network : Architectures, Algorithm, and Application*. Prentice Hall, 1st edition. ISBN-13 : 978-0133341867.
- Irawan, M. I. (2008). *Exploratory Data Analysis dengan JST-Kohonen SOM : Struktur Tingkat Kesejahteraan Darah Tk II se-Jawa Timur*. ITS, Surabaya.
- Prasetyo, E. (2014). *Data Mining Mengolah Data Menjadi informasi menggunakan Matlab*. Andi, Yogyakarta.
- Rukmana, R. (1995). *Bawang Merah Budidaya dan Pengolahan Pasca Panen*. Kanisius, Yogyakarta.
- Stergiou, C. and Siganos, D. (1989). *Neural Networks*. Pasific Northwest National Laboratory : New York (diakses pada 13 April 2017 pada pukul 13.07 WITA).