

CLUSTERING KUALITAS BERAS BERDASARKAN CIRI FISIK MENGUNAKAN METODE K-MEANS

Silvi Agustina¹⁾, Dhimas Yhudo²⁾, Hadi Santoso³⁾, Nofiadi Marnasusanto⁴⁾, Arif Tirtana⁵⁾,
Fakhris Khusnu^{6*)}

Program Studi Ilmu Komputer, Program Teknologi Informasi dan Ilmu Komputer
Universitas Brawijaya Malang
Jalan Veteran Malang 65145, Indonesia

Email : silviagustina07@gmail.com¹⁾, dhimas.tomz00@gmail.com²⁾, hadi.santoso@gmail.com³⁾,
nofiadi.m@gmail.com⁴⁾, asgether@gmail.com⁵⁾, fakhris.cr7@gmail.com^{6*)}

Abstrak

Saat ini pemeriksaan kualitas beras telah dilakukan secara manual oleh inspektur yang telah berpengalaman. Dengan cara ini memiliki kelemahan seperti: (1) adanya subjektivitas penilaian mutu antara pengamat yang satu dengan yang lain; (2) adanya kelelahan fisik jika pengamat bekerja terlalu lama, sehingga menyebabkan hasil pengamatan tidak konsisten, dan (3) waktu yang dibutuhkan untuk pengamatan mutu lebih lama. Sehubungan dengan permasalahan diatas, maka diperlukan cara untuk menentukan klasifikasi mutu beras dengan cepat, akurat dan mudah untuk dioperasikan, sehingga dapat meningkatkan efisiensi pengkelasan mutu fisik beras. K-Means merupakan metode dengan tipe klasifikasi unsupervised dimana mengelompokkan data menjadi satu atau lebih kluster. K-Means memodelkan dataset menjadi kluster-kluster dimana data pada satu kluster memiliki karakteristik yang sama dan memiliki karakteristik yang berbeda dari kluster lain. Penelitian ini bertujuan untuk membuat sistem penunjang keputusan untuk menentukan klasifikasi mutu beras.

Kata kunci : Beras, kualitas beras, k-means, unsupervised, kluster.

Abstract

Currently rice quality inspection is usually done manually by inspectors who are experienced. This method has weaknesses such as: (1) there is a different quality assessment objectivity among the observers; (2) there is physical exhaustion if the observers worked too long which is causing inconsistent observations, and (3) the observation take a long time. In connection with the problems above, it's needed a way to determine the classification of rice quality which can be operated fastly, accurately and easily, so the classifying of rice physical quality can be increased. K-Means is a type of unsupervised classification method which partitions data items into one or more clusters. K-Means tries to model a dataset into clusters so that data items in a cluster have similar characteristic and have different characteristics from the other clusters. This research proposed to create a decision support system for determining the classification of rice quality.

Keywords : Rice, rice quality, k-means, unsupervised, cluster.

PENDAHULUAN

Beras merupakan makanan pokok sebagian besar penduduk dunia, termasuk penduduk Indonesia. Tingginya konsumsi beras tersebut menuntut pemerintah untuk selalu mengembangkan varietas padi yang lebih unggul dengan produktivitas tinggi. Konsumsi beras yang tinggi juga memicu terjadinya perdagangan bebas pada produk beras di Indonesia, sehingga pemerintah menerbitkan standar mutu beras giling agar beras yang diperdagangkan memenuhi standar. Pemilihan beras merupakan ungkapan selera pribadi konsumen, ditentukan oleh faktor subjektif dan dipengaruhi oleh lokasi, suku bangsa atau etnis, lingkungan, pendidikan, status sosial ekonomi, jenis pekerjaan, dan tingkat pendapatan.

Mutu beras secara umum dipengaruhi oleh empat faktor utama, yaitu: (1) sifat genetik, (2) lingkungan dan kegiatan pra-panen, (3) perlakuan pemanenan dan (4) perlakuan pasca panen. Rangkaian kegiatan pasca panen di tingkat petani meliputi pemanenan, perontokan, pembersihan, pengeringan, pengemasan, penyimpanan, dan penggilingan. Faktor-faktor ini sangat mempengaruhi terjadinya butir patah [1].

Di Indonesia, mutu beras lebih dikenal berdasarkan cara pengolahan, seperti beras tumbuk atau beras giling,

berdasarkan derajat sosoh seperti beras slip, berdasarkan asal daerah seperti beras Cianjur, dan berdasarkan jenis atau kelompok varietas seperti beras IR [2].

Tujuan dari pelaksanaan kegiatan ini adalah mempelajari karakteristik mutu fisik beras berdasarkan panjang dan lebar beras menggunakan metode K-means Clustering. Hasil dari proyek akhir ini berupa sistem pendukung keputusan yang dapat melakukan klasterisasi mutu beras.

K-Means Clustering

Algoritma *K-Means* merupakan metode *nonheirarchial* yang pada awalnya mengambil sebagian dari banyaknya komponen dari populasi untuk dijadikan pusat *cluster* awal. Pada step ini pusat *cluster* dipilih secara acak dari sekumpulan populasi data. Berikutnya Kmeans menguji masing-masing komponen di dalam populasi data dan menandai komponen tersebut ke salah satu pusat cluster yang telah didefinisikan tergantung dari jarak minimum antar komponen dengan tiap-tiap pusat *cluster*. Posisi pusat cluster akan dihitung kembali sampai semua komponen data digolongkan ke dalam tiap-tiap pusat cluster dan terakhir akan terbentuk posisi pusat cluster baru. Beberapa alternatif penerapan *K-Means* dengan beberapa pengembangan teori-teori penghitungan terkait telah diusulkan. Hal ini termasuk pemilihan [3] :

1. *Distance space* untuk menghitung jarak di antara suatu data dan *centroid*
2. Metode pengalokasian data kembali ke dalam setiap *cluster*
3. *Objective function* yang digunakan.

Distance Space

Beberapa *distance space* telah diimplementasikan dalam menghitung jarak (*distance*) antara data dan *centroid* termasuk di antaranya L_1 , L_2 , dan L_p . Jarak antara dua titik x_1 dan x_2 pada *Manhattan/City Block distance space* dihitung dengan menggunakan rumus pada persamaan 1 [3]:

$$D_{L_1}(x_2, x_1) = |x_2 - x_1|_1$$

$$= \sqrt{\sum_{j=1}^p (x_{2j} - x_{1j})^2} \dots\dots\dots(1)$$

Secara umum algoritma dasar dari K-Means Clustering adalah sebagai berikut [5]:

1. Tentukan jumlah cluster
2. Alokasikan data ke dalam cluster secara random
3. Hitung centroid/rata-rata dari data yang ada di masing-masing cluster
4. Alokasikan masing-masing data ke centroid/rata-rata terdekat
5. 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

METODOLOGI

Metodologi penelitian yang dilakukan dengan langkah-langkah sebagai berikut.

1. **Studi Literatur**
Studi literatur merupakan langkah awal dalam penelitian ini, studi literatur dilakukan untuk melengkapi pengetahuan dasar dan teori-teori yang berasal dari buku-buku, jurnal, maupun media internet.
2. **Pengumpulan Data**
Data yang digunakan pada penelitian ini adalah beberapa sampel beras yang didapatkan dari pedagang beras secara acak.
3. **Identifikasi Masalah**
Tahap ini dilakukan setelah mendapatkan *dataset* yang sesuai untuk dilakukan proses *clustering*.
4. **Pre Processing**
Tahap pre processing merupakan tahap seleksi data yang bertujuan untuk mendapatkan data yang bersih. Tahap ini ditunjukkan pada proses binarisasi gambar hasil *capture image*.
5. **Proses Clustering dengan Metode K-Means**
Tahap ini merupakan proses dimana data-data yang sudah bersih dan tidak ada noise, di *cluster*-kan sesuai dengan

ciri fisiknya yaitu panjang dan lebar beras.

6. Hasil

Hasil yang didapat setelah proses *clustering* adalah sejumlah *cluster* data yang dapat digunakan sebagai estimasi terhadap kualitas mutu beras.

7. Evaluasi

Tahap evaluasi yaitu tahap dimana dilakukan evaluasi terhadap hasil yang diperoleh. Jika evaluasi akurasi belum mendapatkan hasil yang sesuai maka dikerjakan kembali proses *K-Means*.

| | | |
|----|------|------|
| 11 | 6,96 | 2,89 |
| 12 | 7,1 | 3,07 |
| 13 | 6,98 | 3,06 |
| 14 | 7,01 | 2,94 |
| 15 | 6,22 | 2,7 |
| 16 | 7,25 | 2,06 |
| 17 | 6,62 | 2,98 |
| 18 | 7,25 | 2,74 |
| 19 | 6,62 | 2,99 |
| 20 | 6,66 | 2,76 |

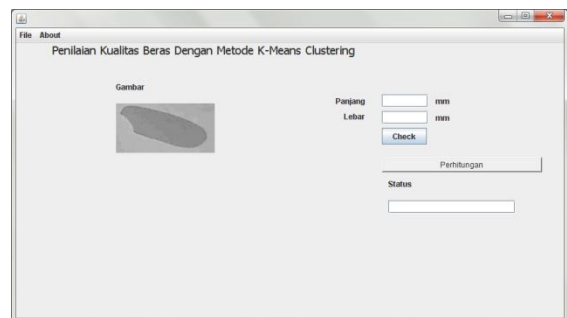
HASIL DAN PEMBAHASAN

Data sampel beras yang digunakan pada penelitian ini sebanyak 20 dengan fitur X1 yaitu panjang beras (mm) dan X2 yaitu lebar beras (mm) dengan rincian pada Tabel 1.

Tabel 1 Data Penelitian

| Sampel | Fitur | |
|--------|-------|------|
| | X1 | X2 |
| 1 | 5,52 | 2,04 |
| 2 | 6,51 | 2,43 |
| 3 | 6,22 | 2,24 |
| 4 | 7,24 | 3,15 |
| 5 | 6,45 | 2,82 |
| 6 | 6,01 | 2,54 |
| 7 | 6,19 | 2,01 |
| 8 | 7,14 | 3,11 |
| 9 | 5,97 | 2,1 |
| 10 | 6,76 | 3,24 |

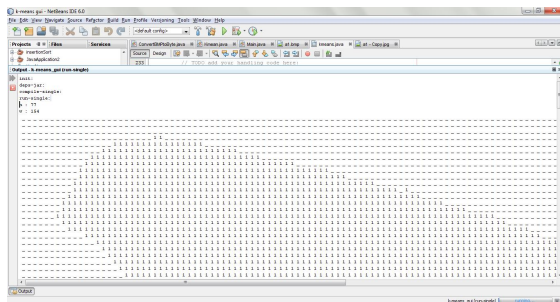
Tampilan interface program dapat dilihat pada gambar 1.



Gambar 1 Tampilan Menu Awal

Dari Gambar 1 di atas, input yang dibutuhkan berupa image beras dengan format bitmap dengan latar belakang warna putih. Image beras diperoleh dengan *capture* objek dengan jarak 15 cm dengan kamera. Setelah image diinputkan, maka dilakukan proses binarisasi, yaitu proses mengkonversi citra grayscale ke dalam bentuk citra biner. Tiap-tiap piksel dalam citra levelnya dirubah melalui suatu range warna tertentu untuk memudahkan dalam mengenali objek dengan *background*

dengan keterangan bahwa angka 1 adalah objek, dan $_$ adalah *background*. Hasil dari proses binarisasi dapat dilihat pada Gambar 2.



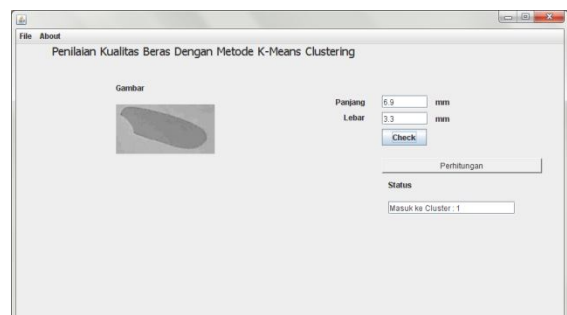
Gambar 2 Hasil Binarisasi

Proses K-Means Clustering

Perhitungan pada program Klasterisasi Kualitas Beras dengan K-Means diawali dengan menentukan jumlah *cluster*. Penelitian ini menggunakan 3 *cluster* untuk menentukan kualitas beras, dengan keterangan bahwa *cluster* 1 adalah beras kualitas buruk, *cluster* 2 adalah beras kualitas sedang, dan *cluster* 3 adalah beras kualitas baik. Setelah menentukan jumlah *cluster*, kemudian menentukan 3 pusat *cluster* awal, yaitu pusat *cluster* 1 (5,5 ; 2), pusat *cluster* 2 (6,2 ; 2,45), dan pusat *cluster* 3 (7 ; 3,2).

Langkah selanjutnya adalah mengalokasikan data ke dalam *cluster*, kemudian menghitung jarak setiap data terhadap setiap pusat *cluster* dengan menggunakan Persamaan 1. Suatu data akan menjadi anggota dari suatu *cluster* yang memiliki jarak terkecil dari pusat *cluster*-nya. Setelah itu menghitung pusat

cluster baru sebagai acuan untuk iterasi berikutnya. Pada iterasi 1 dihasilkan 3 pusat *cluster* baru, pusat *cluster* 1 (5,89333;2,05), pusat *cluster* 2 (6,28199;2,546), dan pusat *cluster* 3 (6,96583;2,99917). Perhitungan pada iterasi selanjutnya sama dengan iterasi 1, hingga posisi data tidak mengalami perubahan. Pusat *cluster* yang dihasilkan pada iterasi 2 yaitu pusat *cluster* 1 (5,74499;2,07), pusat *cluster* 2 (6,26667;2,45667), dan pusat *cluster* 3 (6,96583;2,99917). Pusat *cluster* yang dihasilkan pada iterasi 3 yaitu pusat *cluster* 1 (6,28199;2,546), pusat *cluster* 2 (5,89333;2,05), dan pusat *cluster* 3 (6,96583;2,99917). Pusat *cluster* yang dihasilkan pada iterasi 4 yaitu pusat *cluster* 1 (5,89333;2,05), pusat *cluster* 2 (6,28199;2,546), dan pusat *cluster* 3 (6,96583;2,99917). karena pada iterasi selanjutnya posisi data tidak berubah, maka iterasi dihentikan dan hasil akhir yang diperoleh adalah 3 pusat *cluster* pada iterasi ke-4.



Gambar 3 Hasil Clustering

Setelah klik tombol Chek pada program seperti pada Gambar 3, maka akan keluar output berupa panjang beras yaitu 6,9 mm dan lebar 3,3 mm. Kemudian klik tombol Perhitungan, maka akan dihasilkan output berupa status kualitas beras. Berdasarkan perhitungan pada program, image beras yang telah diinputkan masuk ke dalam *Cluster 3* yaitu kualitas baik.

AKURASI

Validasi dilakukan untuk memastikan ketepatan hasil *clustering*. Proses validasi tersebut dengan membandingkan rata-rata pusat *cluster* akhir dengan data validasi yang diperoleh dari pedagang beras yaitu data validasi panjang 6,7 mm dan validasi lebar 2,8 mm. Sedangkan rata-rata pusat *cluster* akhir yaitu 6,38038 mm dan 2,53172 mm. Hasil validasi dapat dilihat pada Tabel 2.

Tabel 2 Validasi

| | Panjang | Lebar |
|---------------------------|---------------|---------|
| Pusat <i>Cluster</i> | 6,38038 | 2,53172 |
| Data validasi | 6,7 | 2,8 |
| Validasi | 95,23% | 90,42% |
| Rata-rata validasi | 92,82% | |

KESIMPULAN

K-Means Clustering merupakan metode klasterisasi berdasarkan persamaan karakteristik, dan merupakan metode yang sangat berguna karena mampu mentranslasi ukuran persamaan yang intuitif menjadi ukuran yang kuantitatif. Penelitian ini menggunakan 20 data uji, dimana ke-20 data tersebut dibagi menjadi 3 *cluster* dengan keterangan *Cluster 1* merupakan beras kualitas buruk, *Cluster 2* beras kualitas sedang, dan *Cluster 3* beras kualitas baik. Dari hasil penelitian, didapatkan 3 pusat cluster akhir yaitu pusat *cluster 1* (5,89333;2,05), pusat *cluster 2* (6,28199;2,546), dan pusat *cluster 3* (6,96583;2,999167) serta dihasilkan validasi sebesar 92,82% yang menunjukkan bahwa program ini dapat dijadikan sebagai acuan dalam klasterisasi kualitas beras.

DAFTAR PUSTAKA

- [1] Damardjati, D.S., dan E.Y.Purwani, (1991). Padi Buku 3. Penyunting Edi Soenarjo, D.S. dan Mahyudin Syam. Pusat Penelitian dan Pengembangan Tanaman Pangan Bogor.
- [2] Damardjati, D.S, (1987), Prospek Peningkatan Mutu Beras di Indonesia. Jurnal Penelitian dan Pengembangan Pertanian. Pusat Penelitian dan Pengembangan Pertanian. Bogor.

4:85-92. Di akses pada tanggal 19 Desember 2012.

- [3] McLachlan, G. J. and Peel, D. (2000). *Finite Mixture Models*. New York : John Wiley and Sons.
- [4] Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York : Plenum Press.
- [5] Agusta, Y. (2007). *K-Means, Penerapan, Permasalahan dan Metode Terkait*. *Jurnal Sistem Informatika*. 3:47-60. Di akses pada tanggal 18 Desember 2012.