**Perbandingan Algoritma K-Means dan K-Medoids**

**Untuk Clustering Harga Beras di Provinsi Jawa Tengah**

**Galet Guntoro Setiaji1, Astrid Novita Putri2, Dinar Anggit Wicaksana3**

1Fakultas Teknologi Informasi dan Komunikasi, Universitas Semarang

Jl. Soekarno Hatta, Semarang, telp:024-6702757, e-mail: gallet@usm.ac.id

2Fakultas Teknologi Informasi dan Komunikasi, Universitas Semarang

Jl. Soekarno Hatta, Semarang, telp:024-6702757, e-mail: astrid@usm.ac.id

3Fakultas Teknologi Informasi dan Komunikasi, Universitas Semarang

Jl. Soekarno Hatta, Semarang, telp:024-6702757, e-mail: dinar\_anggit@usm.ac.id

**ARTICLE INFO**

*History of the article :*

Received 30 December 2010

Received in revised form 30 April 2011

Accepted 26 September 2012

Available online 8 October 2012

**Keywords:**

Clsutering; Beras; Jawa Tengah; K-Means;

K-Medoids

**\* Correspondece:**

Telepon:

+62 (024) 12345678

E-mail:

first\_author@afiliasi.xx.xx

***ABSTRACT***

Beras merupakan komoditas pangan utama bagi masyarakat di Indonesia khususnya di Jawa Tengah. Sebagai kebutuhan pokok, ketersediaan dan harga beras sangat mempengaruhi stabilitas ekonomi dan kesejahteraan masyarakat. Oleh karena itu, pemantauan dan analisis harga beras menjadi hal yang sangat penting. Salah satu hal dalam memahami harga beras adalah dengan melakukan clustering atau pengelompokan harga berdasarkan harga eceran beras medium. Clustering dapat memberikan gambaran lebih jelas variasi harga di jawa tengah metode yang digunakan yaitu *K-Means dan K-Medoid* dengan membandingkan metode tersebut diharapkan mendapatkan informasi detail dan pola harga beras di berbagai daerah. Hasil yang didapatkan bahwa validitas DBI K-Means dengan K=3 dan K=5 lebih unggul dari pada menggunakan K-Medoids.

1. **INTRODUCTION**

Beras merupakan komoditas pangan utama yang sangat vital bagi masyarakat Indonesia, termasuk di Provinsi Jawa Tengah. Sebagai kebutuhan pokok, ketersediaan dan harga beras sangat mempengaruhi stabilitas ekonomi dan kesejahteraan masyarakat. Oleh karena itu, pemantauan dan analisis harga beras menjadi hal yang sangat penting [1], [2].

Salah satu cara untuk memahami dinamika harga beras adalah dengan melakukan clustering atau pengelompokan harga beras berdasarkan harga eceran beras medium. Clustering harga beras dapat memberikan gambaran yang lebih jelas mengenai variasi harga di berbagai daerah, serta membantu dalam mengidentifikasi faktor-faktor yang mempengaruhi perbedaan harga tersebut [3],[4],[5].

Algoritma clustering seperti K-Means dan K-Medoids dapat digunakan untuk mengelompokkan data harga beras. K-Means menggunakan mean dari titik-titik dalam sebuah cluster sebagai pusat cluster, sedangkan K-Medoids menggunakan medoid atau titik yang paling representatif dari cluster tersebut[6],[7]. Meskipun kedua algoritma ini memiliki tujuan yang sama, yaitu mempartisi data ke dalam beberapa kelompok berdasarkan kemiripan, mereka memiliki perbedaan dalam cara kerja dan karakteristik saja [8],[9],[10].

Penelitian ini bertujuan untuk membandingkan kinerja algoritma K-Means dan K-Medoids dalam clustering harga beras medium dengan studi kasus menggunakan data dari dinas pertahanan pangan Provinsi Jawa Tengah. Dengan melakukan clustering, diharapkan dapat diperoleh informasi yang lebih mendetail mengenai pola harga beras di berbagai daerah, yang nantinya dapat digunakan untuk pengambilan keputusan yang lebih baik dalam pengelolaan dan distribusi beras membandingan dua metode *K-Means* dan *K-Medoids* menggunakan *validitas cluster*. Salah satu validitas cluster yang akan kita gunakan adalah menggunakan *Davis Boulden Indeks* (*DBI*), dan nantinya akan bisa diketahui metode mana yang lebih bagus dalam pengelompokkan atau cluster. Sehingga menekankan pentingnya pemantauan harga beras dan bagaimana teknik clustering dapat digunakan untuk memahami pola harga beras medium. Melalui penelitian ini, diharapkan dapat diperoleh wawasan yang lebih mendalam mengenai efektivitas kedua algoritma clustering tersebut dalam konteks data harga beras, serta memberikan rekomendasi metode yang paling sesuai untuk analisis harga komoditas ini di Provinsi Jawa Tengah.

1. **RESEARCH METHODS**

Metode penelitian yang yang digunakan adalah membandingkan hasil validitas antara metode *K-Means* dengan *K-Medoids*, untuk perhitungan validitasnya kita menggunakan *Davis Boulden Index* (*DBI*).

**Gambar 1.** Alur metode penelitian

1. Analisa kebutuhan

Analisis kebutuhan penelitian ini dengan melakukan wawancara pada dinas pertahanan pangan Provinsi Jawa Tengah terdapat beberapa hal yang memerlukan pemantauan dan analisis harga dalam hal ini komoditas beras di Jawa Tengah. Sehingga dapat menghasilkan informasi yang lebih mendetail mengenai pola harga beras di berbagai daerah, yang nantinya dapat digunakan untuk pengambilan keputusan yang lebih baik dalam pengelolaan dan distribusi beras membandingan dua metode K-Means dan K-Medoids menggunakan validitas cluster. Data penelitian ini tahun 2021 sampai 2023 wilayah Jawa Tengah berupa data harga eceran beras medium.

1. Normalisasi Data

Normalisasi data adalah langkah penting dalam proses clustering, termasuk untuk analisis harga beras. Tujuan normalisasi adalah untuk mengubah skala data sehingga setiap fitur berkontribusi secara setara dalam pembentukan cluster. Hal ini penting karena fitur dengan skala yang lebih besar dapat mendominasi proses clustering dan menghasilkan cluster yang tidak akurat. Normalisasi datamenggunakan harga eceran beras medium, kemudian di normalisasikan dari data harga eceran 1 tahun dan mendapatkan nilai maksimal dalam 1 tahun pada setiap kota di Provinsi Jawa Tengah.

1. Pengujian Metode

Untuk pengujian data penelitian ini menggunakan metode *K-Means*, sedangkan untuk mengetahui metode *K-Means* lebih baik, maka diperlukan suatu pembanding menggunakan metode pengelompokan lainnya yaitu *K-Medoids*.

1. Validitas

Tahap Validitas digunakan sebagai perbandingan metode yang lebih unggul, sehingga diperlukan validitas metode clustering menggunakan *validitas Davis Boulde Index* (*DBI*). Dengan mempertimbangkan jumlah *K* pada tiap perhitungan masing-masing metode.

1. **HASIL DAN PEMBAHASAN**

Hasil dan pembahasan dari penelitian ini dengan tahapan perhitungan sebagai berikut.

1. Data Penelitian

Dari data yang di normalisasi didapatkan data sebanyak 35 kota dan kabupaten dari Provinsi Jawa Tengah, mulai tahun 2021 sampai dengan 2023.

Tabel 1.Data harga beras medium periode tahun 2021 s/d 2023

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No** | **Nama Kota/ Kabupaten** | **Tahun** | | |
| **2021** | **2022** | **2023** | |
| 1 | Kab. Banjanegara | 12.000 | 12.000 | 13.500 | |
| 2 | Kab. Banyumas | 9.800 | 11.000 | 13.500 | |
| 3 | Kab. Batang | 11.000 | 12.000 | 13.500 | |
| 4 | Kab. Blora | 10.300 | 10.500 | 12.700 | |
| 5 | Kab. Boyolali | 11.500 | 11.400 | 13.500 | |
| … | … | … | … | … | |
| … | … | … | … | … | |
| 34 | Kota Surakarta | 11.000 | 11.000 | 14.500 | |
| 35 | Kota Tegal | 10.170 | 11.500 | 13.500 | |

1. Perhitungan K-Means
2. Menentukan *centroid* baru dengan *k=2*

Tabel 2.*Centroid* awal *k=2* perhitungan *K-Means*

|  |  |  |  |
| --- | --- | --- | --- |
| **Kota/Kabupaten** | **2021** | **2022** | **2023** |
| Kab. Klaten | 10.200 | 11.510 | 22.500 |
| Kab. Kudus | 10.500 | 12.500 | 23.000 |

1. Menghitung jarak dengan *ecludian distance* pada interasi ke 1

Tabel 3.Perhitungan *Ecludian Distance i=1* pada *K-Means*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Kota** | **Beras Medium Pertahun** | | | **C1** | | **C2** | | **Centroid** | |
| **2021** | **2022** | **2023** | |  | |  |  |
| 1 | Kab. Banjanegara | 12.000 | 12.000 | 13.500 | | 9191 | | 9630 | C1 |
| 2 | Kab. Banyumas | 9.800 | 11.000 | 13.500 | | 9023 | | 9643 | C1 |
| 3 | Kab. Batang | 11.000 | 12.000 | 13.500 | | 9048 | | 9526 | C1 |
| 4 | Kab. Blora | 10.300 | 10.500 | 12.700 | | 9852 | | 10494 | C1 |
| 5 | Kab. Boyolali | 11.500 | 11.400 | 13.500 | | 9094 | | 9615 | C1 |
| ... | ... | ... | ... | ... | | ... | | ... | ... |
| ... | ... | ... | ... | ... | | ... | | ... | ... |
| 15 | Kab. Kudus | 10.500 | 12.500 | 23.000 | | 1148 | | 0 | C2 |
| ... | ... | ... | ... | ... | | ... | | ... | ... |
| ... | ... | ... | ... | ... | | ... | | ... | ... |
| 34 | Kota Surakarta | 11.000 | 11.000 | 14.500 | | 8056 | | 8645 | C1 |
| 35 | Kota Tegal | 10.170 | 11.500 | 13.500 | | 9000 | | 9558 | C1 |

Di *interasi* ke 1 didapatkan *cluster* 1 dengan jumlah 34, sedangkan *cluster* 2 didapatkan jumlah 1.

1. Perhitungan *K-Means* berhenti pada interasi ke 4 diakrenakan centroid baru sama dengan interasi ke 3, didapatkan centroid akhir dengan hasil sebagai berikut.

Tabel 4.Centroid akhir *K-Means* tidak berubah pada *i=4*

|  |  |  |  |
| --- | --- | --- | --- |
| **Centroid Baru** | **2021** | **2022** | **2023** |
| C1 | 10508 | 11298 | 13621 |
| C2 | 10350 | 12005 | 22750 |

1. Dari hasil centroid akhir didapatkan cluster akhir sebagai berikut.

Tabel 5.Hasil cluster akhir *K-Means* pada *i=4*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Kota** | **Beras Medium Pertahun** | | | **C1** | **C2** | **Centroid** |
| **2021** | **2022** | **2023** |
| 1 | Kab. Banjanegara | 12.000 | 12.000 | 13.500 | 1653 | 9396 | C1 |
| 2 | Kab. Banyumas | 9.800 | 11.000 | 13.500 | 777 | 9320 | C1 |
| 3 | Kab. Batang | 11.000 | 12.000 | 13.500 | 865 | 9272 | C1 |
| 4 | Kab. Blora | 10.300 | 10.500 | 12.700 | 1236 | 10162 | C1 |
| 5 | Kab. Boyolali | 11.500 | 11.400 | 13.500 | 1004 | 9340 | C1 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | .... | ... | ... | ... |
| 14 | Kab. Klaten | 10.200 | 11.510 | 22.500 | 8886 | 574 | C2 |
| 15 | Kab. Kudus | 10.500 | 12.500 | 23.000 | 9455 | 574 | C2 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 34 | Kota Surakarta | 11.000 | 11.000 | 14.500 | 1050 | 8336 | C1 |
| 35 | Kota Tegal | 10.170 | 11.500 | 13.500 | 411 | 9265 | C1 |

Dengan *cluster* akhir didapatkan C1 dengan jumlah 33 dan C2 didapatkan dengan jumlah 2.

1. Perhitungan *K-Medoids*
2. Menentukan *Medoids* awal dengan *k=2* pada Tabel 6.

Tabel 6.Menentukan *Medoids* awal dengan k=2

|  |  |  |  |
| --- | --- | --- | --- |
| **Kota/Kabupaten** | **2021** | **2022** | **2023** |
| Kota Tegal | 10.170 | 11.500 | 13.500 |
| Kab. Kudus | 10.500 | 12.500 | 23.000 |

1. Menghitung *cluster medoids* dengan *ecludian distance.*

Tabel 7.Perhitungan *Medoids* interasi ke 1

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Kota** | **2021** | **2022** | **2023** | **C1** | **C2** | **Kedekatan** | **centroid** |
| Kab. Banjanegara | 12000 | 12000 | 13500 | 1,897.08 | 9,630.68 | 1,897.08 | C1 |
| Kab. Banyumas | 9800 | 11000 | 13500 | 622.01 | 9,643.13 | 622.01 | C1 |
| Kab. Batang | 11000 | 12000 | 13500 | 968.97 | 9,526.28 | 968.97 | C1 |
| Kab. Blora | 10300 | 10500 | 12700 | 1,287.21 | 10,494.28 | 1,287.21 | C1 |
| Kab. Boyolali | 11500 | 11400 | 13500 | 1,333.75 | 9,615.61 | 1,333.75 | C1 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| Kab. Klaten | 10200 | 11510 | 22500 | 9,000.06 | 1,148.96 | 1,148.96 | C2 |
| Kab. Kudus | 10500 | 12500 | 23000 | 9,558.18 | 0.00 | 0.00 | C2 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| Kota Surakarta | 11000 | 11000 | 14500 | 1,392.44 | 8,645.81 | 1,392.44 | C1 |
| Kota Tegal | 10170 | 11500 | 13500 | 0.00 | 9,558.18 | 0.00 | C1 |
| **Total Kedekatan** | | | | | | **35364.8838** |  |

1. Menentukan *Medoids* baru lagi untuk menghitung *cluster medoids.*

Tabel 8.*Medoids* baru untuk menentukan nilai kedekatan pada *cluster i=2*

|  |  |  |  |
| --- | --- | --- | --- |
| **Kota/Kab** | **2021** | **2022** | **2023** |
| Kab. Banyumas | 9800 | 11000 | 13500 |
| Kab. Pati | 9500 | 11000 | 13500 |

1. Menghitung cluster medoids kembali

Disini menghitung cluster medoids baru pada Tabel 8, dengan data Tabel 1. Kemudian menghitung selisih kedekatan. Apabila selisih kedekatan > 0, maka berhenti.

Selisih kedekatan = *35364.88 - 57359.44*

Selisih kedekatan = *21994.56*

Dihasilkan selisih kegiatan ternyata lebih dari 0 yaitu *21994.56*, maka perhitungan berhenti. Dan menghasilkan cluster dengan C1 dengan jumlah 33 sedangkan C2 sebanyak 2.

1. Validitas *Davis Boulden Indeks*

Setelah data penelitian diolah mengunakan *K-Means* dengan k=2 dihasilkan perhitungan *DBI* sebesar *0.1711*, sedangkan perhitungan *K-Medoids* dengan *k=2* didapatkan *DBI* sebesar *0.1686*. untuk perbandingan pada *k=2*, *k=3* dan *k=5*, dengan hasil tabel sebagai berikut.

Tabel 9.Hasil perbandingan metode *K-Means* dan *K-Medoids*

|  |  |  |  |
| --- | --- | --- | --- |
| **Metode** | ***K=2*** | ***K=3*** | ***K=5*** |
| ***K-Means*** | *DBI = 0.171*  C1= 32  C2= 2 | *DBI* = 0.717  C1= 18  C2= 2  C3= 15 | *DBI* = 0.926  C1= 6  C2= 2  C3= 9  C4= 15  C5= 3 |
| ***K-Medoids*** | *DBI* = 0,168  C1= 32  C2= 2 | *DBI* = 0,926  C1= 9  C2= 24  C3= 2 | *DBI* = 1,242  C1= 9  C2= 12  C3= 5  C4= 7  C5= 2 |

1. **KESIMPULAN**

Dari hasil penelitian dapat disimpulkan *Clustering* dapat memberikan gambaran lebih jelas variasi harga di jawa tengah metode yang digunakan yaitu *K-Means* dan *K-Medoid* dengan membandingkan metode tersebut diharapkan mendapatkan informasi detail dan pola harga beras di berbagai daerah. Hasil yang didapatkan bahwa validitas *DBI K-Means* dengan *K=3* dan *K=5* lebih unggul dari pada menggunakan *K-Medoids*.

**REFERENCES**

|  |  |
| --- | --- |
| [1] | N. I. Sarumaha. Implementasi Algoritma K-Means Clustering Pada Analisa Impor Beras. *Jurnal Sains Dan Teknologi Informasi.* 2021; 1(1): 19-27. |
| [2] | E. R. Arini. Penerapan K-Means Cluster Di Provinsi Jawa Timur Berdasarkan Ketahanan Pangan. *Journal of Science Nusantara.* 2023; 3(1): 32-36. |
| [3] | S. Wijayanto and M. Y. Fathoni. Pengelompokkan Produktivitas Tanaman Padi di Jawa Tengah Menggunakan Metode Clustering K-Means. *Jurnal Penelitian Ilmu dan Teknologi Komputer.* 2021; 13(2): 212-219. |
| [4] | F. Marisa, A. Zahma, A. M. Bau, E. Noviansa, A. S. Neno and A. L. Maukar. Digitasi Produktivitas Panen Padi Berbasis K-Means Clustering. *SMARTICS Journal.* 2021; 7(1): 21-26. |
| [5] | N. Mirantika, T. S. Syamfithriani and R. Trisudarmo. Implementasi Algoritma K-Medoids Clustering untuk Menentukan Segmentasi Pelanggan. *Jurnal Nuansa Informatika.* 2023; 17(1): 196-204. |
| [6] | S. Abdullah, B. . D. Nuryanto, R. Sahara, M. I. Saputra and C. R. Hassolthine. Clustering Pasar Tradisional di Wilayah Kabupaten Tangerang Berdasarkan Harga Bahan Pokok dengan Metode K-means. *Jurnal Riset Teknik Informatika dan Komputer.* 2023; 5(3): 334-339. |
| [7] | N. Novitasari, N. D. Nuris and R. Herdiana. Penerapan Algoritma K-Means Untuk Clustering Data Jumlah Penduduk Miskin Berdasarkan Kota/Kabupaten Di Jawa Barat Menggunakan Rapidminer. *Jurnal Informatika Terpadu*. 2023; 9(1): 68-73. |
| [8] | I. Febriani, M. Safii and O. Alfina. Implementasi Data Mining Peningkatan Produksi Beras Menggunakan Metode K-Means Clustering. *Majalah Ilmiah METHODA.* 2022; 12(3): 258-268. |
| [9] | D. Lestari, N. Fadillah and A. Ihsan. Clustering Kualitas Beras Berdasarkan Ciri Fisik Menggunakan Metode K-Means. *Jurnal Nasional Informatika dan Teknologi Jaringan.* 2019; 3(2): 136-142. |
| [10] | Y. Febriani, Y. P. Sari and D. Octaria. Metode K-Means Cluster Untuk Mengelompokkan Kota/Kabupaten di Sumatera Selatan Berdasarkan Produksi Ikan Air Tawar. *Jurnal Ilmiah Matematika dan Ilmu Pengetahuan Alam.* 2021; 18(2): 175-182. |