

**PERBANDINGAN PERFORMA GABUNGAN APRIORI K-MEANS
DENGAN APRIORI K-MEDOIDS DALAM MENEMUKAN POLA PADA
DATA HASIL PRODUKSI PADI**

SKRIPSI



Oleh:

AMAZIDA

NIM. 2010817120002

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS LAMBUNG MANGKURAT
BANJARMASIN, DESEMBER 2024**

**PERBANDINGAN PERFORMA GABUNGAN APRIORI K-MEANS
DENGAN APRIORI K-MEDOIDS DALAM MENEMUKAN POLA PADA
DATA HASIL PRODUKSI PADI**

SKRIPSI

Diajukan Untuk Memenuhi Salah Satu Syarat
Sarjana Strata-1 Teknologi Informasi



Oleh:

AMAZIDA

NIM. 2010817120002

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS LAMBUNG MANGKURAT
BANJARMASIN, DESEMBER 2024**


LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini

Nama : Amazida
NIM : 2010817120002
Fakultas : Teknik
Prodi : Teknologi Informasi
Judul Skripsi : Perbandingan Performa Gabungan *Apriori K-Means* dengan *Apriori K-Medoids* dalam Menemukan Pola pada Data Hasil Produksi Padi.
Pembimbing Utama : Dr. Ir. Yuslena Sari, S.Kom., M.Kom.

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar akademik di suatu perguruan tinggi, dan sepanjang pengetahuan saya, juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali tertulis diacu dalam naskah ini dan disebutkan dalam daftar rujukan.

Banjarmasin, Desember 2024



Amazida

NIM. 2010817120002

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

SKRIPSI PROGRAM STUDI S-1 TEKNOLOGI INFORMASI

**Perbandingan Performa Gabungan Apriori K-Means dengan Apriori K-Medoids dalam
Menemukan Pola pada Data Hasil Produksi Padi**

Oleh
Amazida (2010817120002)

Telah dipertahankan di depan Tim Penguji pada 9 Desember 2024 dan dinyatakan

LULUS

Komite Penguji :
Ketua : Andreyan Rizky Baskara, S.Kom., M.Kom
NIP. 199307032019031011
Anggota 1 : Helda Yunita, S.Kom., M.Kom
NIP. 199106192024062001
Anggota 2 : Nurul Fathanah Mustamin, S.Pd., M.T
NIP. 199110252019032018
Pembimbing : Dr. Ir. Yuslena Sari, S.Kom., M.Kom
Utama : NIP. 198411202015042002

.....
.....
.....
.....

18 DEC 2024

Banjarbaru,

Diketahui dan disahkan oleh:

Wakil Dekan Bidang Akademik
Fakultas Teknik ULM,



Dr. Mahmud, S.T., M.T.
NIP. 197401071998021001

Koordinator Program Studi
S-1 Teknologi Informasi,

.....

Andreyan Rizky Baskara, S.Kom., M.Kom.
NIP. 199307032019031011

LEMBAR PERSETUJUAN

PERBANDINGAN PERFORMA GABUNGAN APRIORI K-MEANS
DENGAN APRIORI K-MEDOIDS DALAM MENEMUKAN POLA
PADA DATA HASIL PRODUKSI PADI

OLEH
AMAZIDA
NIM. 2010817120002

Telah diperiksa dan terpenuhi semua persyaratan akademik, administrasi
dan disetujui untuk dipertahankan di hadapan dewan penguji

Banjarmasin, 22 November 2024

Pembimbing Utama,



Dr. Ir. Yuslena Sari, S.Kom., M.Kom.

NIP. 198411202015042002

ABSTRAK

Indonesia merupakan negara agraris yang menjadikan pertanian sebagai salah satu sektor ekonomi yang berpengaruh, salah satu komoditas pertanian itu adalah padi. Efisiensi produk untuk meningkatkan hasil pertanian sangat diperlukan untuk mengoptimalkan produksi, sehingga diperlukan analisis yang mendalam untuk memahami pola yang mempengaruhi hasil produksi. Untuk mengatasi hal tersebut, *Apriori* dapat digunakan untuk menemukan pola yang mempengaruhi antara data. Algoritma tersebut banyak digunakan penelitian sebelumnya untuk menemukan aturan asosiasi atau hubungan secara signifikan. Selain itu untuk membantu pengelompokan data berdasarkan kesamaan tertentu dapat menggunakan algoritma *clustering K-Means* dan *K-Medoids*. Kedua metode ini pada penelitian sebelumnya menunjukkan bahwa masing-masing mempunyai keunggulan tergantung kebutuhan analisis data yang diperlukan. Maka, penelitian ini dilakukan untuk membandingkan performa gabungan dari *Apriori K-Means* dengan *Apriori K-Medoids* dilihat dari nilai *Silhouette Score* masing-masing *clustering*. Data yang digunakan pada penelitian ini adalah *dataset* hasil produksi padi di pulau Sumatera yang diambil dari *website* kaggle.com, data ini berjumlah 224 data. Tahapan penelitian yang dilakukan terdiri dari *pre-processing data*, penerapan algoritma *Apriori* untuk mendapatkan pola asosiasi dari masing-masing fitur, dan penerapan *K-Means* dan *K-Medoids* untuk pengelompokan data berdasarkan fitur asosiasi yang didapatkan oleh *Apriori*. Hasil penelitian menunjukkan bahwa hasil gabungan analisis dari *Apriori K-Means* mendapatkan nilai *Silhouette Score* yang lebih baik dari pada *Apriori K-Medoids*, yaitu *Apriori K-Means* dengan nilai 0.59 sedangkan *Apriori K-Medoids* 0.58. Berdasarkan hasil tersebut, maka model *Apriori K-Means* diimplementasikan pada sebuah *website* yang dapat digunakan untuk menganalisis pola pada data hasil produksi padi dengan beberapa parameter yang dapat disesuaikan tergantung kebutuhan analisis data.

Kata Kunci: Analisis Pola, *Apriori*, *Data Mining*, *K-Means*, *K-Medoids*, Produksi Padi

ABSTRACT

Indonesia is an agrarian country where agriculture plays a significant role in the economy. One of its key commodities is rice. Enhancing product efficiency is essential to boost agricultural output. Therefore, a deeper analysis is required to understand the patterns influencing production. To address this issue, the Apriori algorithm can be utilized to identify patterns and relationships within the data. This algorithm is frequently used by researchers to uncover association rules or identify significant relationships in datasets. In addition, K-Means and K-Medoids clustering algorithms are applied to assist in data clustering based on specific similarities. Both methods, as shown in previous research, have distinct advantages depending on the specific requirements of the data analysis. This research aims to compare the performance of combining Apriori with K-Means and Apriori with K-Medoids, using the silhouette score to evaluate the effectiveness of each clustering approach. The data used in this research is a dataset of rice production in Sumatra, sourced from Kaggle.com, containing a total of 224 entries. This research begins with data preprocessing, followed by the implementation of the Apriori algorithm to identify association patterns among features. Subsequently, K-Means and K-Medoids clustering algorithms are applied to group the data based on the association features derived from Apriori. The results of this research indicate that the combination of Apriori and K-Means outperforms Apriori and K-Medoids, achieving a silhouette score of 0.59 compared to 0.58. Based on these findings, the Apriori–K-Means approach was implemented in a website designed to analyze patterns in rice production datasets, offering customizable parameters depending on the data analysis requirements.

Keywords: Rice Analysis, Apriori, Data Mining, K-Means, K-Medoids, Rice Production

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur kepada Allah SWT, skripsi ini saya persembahkan kepada mereka yang telah menjadi kekuatan, motivasi, dan inspirasi dalam setiap langkah perjalanan hidup penulis:

1. Kepada Allah SWT, Tuhan yang Maha Esa, yang telah memberikan penulis kekuatan, kesempatan, dan kesehatan untuk terus belajar, dan menjalani hidup. Tanpa rahmat dan kasih sayang-Nya penulis tidak mampu melewati segala rintangan yang telah penulis hadapi hingga saat ini.
2. Kepada Ibu dan Ayah tercinta, yang selalu memberikan penulis semangat, nasihat, dan doa-doa yang tidak pernah putus diucapkan. Terima kasih atas setiap kerja keras dan tetes keringat yang telah kalian curahkan demi pendidikan dan masa depan anak ini. Segala pencapaian ini tidak sebanding dengan yang telah kalian berikan, semoga skripsi ini menjadi langkah awal yang membuat kalian bangga.
3. Kepada adik tersayang, yang selalu menjadi penghibur disaat lelah melanda. Terima kasih atas canda tawa, dan dukungan yang telah kamu berikan untuk kakak.
4. Kepada Ibu Dr. Ir. Yuslena Sari, S.Kom., M.Kom, selaku Dosen Pembimbing Utama yang telah membimbing penulis dalam menyusun skripsi. Terima kasih atas bimbingan, ilmu, dan kesabaran yang telah ibu berikan kepada penulis, semoga ilmu yang telah ibu berikan menjadi amal jariyah yang terus mengalir.
5. Kepada seluruh Dosen, yang telah memberikan banyak ilmu dan membantu penulis dalam perjalanan akademik selama masa perkuliahan. Terima kasih atas kesabaran dan kegigihan Bapak Ibu dalam memberikan ilmu, semoga ilmu yang telah kalian berikan menjadi amal jariyah yang terus mengalir.

KATA PENGANTAR

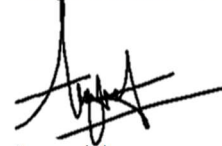
Rasa puji dan syukur yang selalu penulis panjatkan ke hadirat Allah SWT dan Nabi Muhammad SAW, yang telah memberikan limpahan rahmat, hidayah dan karunia-Nya sehingga penulis mampu menyelesaikan skripsi dengan judul “Perbandingan Performa Gabungan *Apriori K-Means* dengan *Apriori K-Medoids* dalam Menemukan Pola pada Data Hasil Produksi Padi”. Penulis sadari bahwa proses penyusunan skripsi ini bukanlah perjalanan yang singkat dan mudah. Banyak perjuangan yang penulis lalui untuk menyelesaikan skripsi. Namun, berkat doa-doa yang telah dipanjatkan, semangat, arahan, dan dukungan dari orang-orang tercinta lah penulis bisa melalui dan menyelesaikan skripsi ini dengan sebaik-baiknya. Pada kesempatan ini, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Ibu, Ayah, dan Adik tercinta. Setiap kerja keras yang kalian berikan merupakan alasan penulis untuk selalu berjuang. Ibu dan Ayah, orang tua terhebat yang selalu memberikan kasih sayang, doa, dan mengusahakan yang terbaik untuk anaknya. Adik, saudara terkasih yang selalu menjadi obat lelah penulis dalam menjalani hidup.
2. Ibu Dr. Ir. Yuslena Sari, S.Kom., M.Kom. selaku Dosen Pembimbing Utama yang telah bersedia meluangkan waktu menjadi pembimbing penulis, yang penuh kesabaran dan keikhlasan dalam memberikan arahan, bimbingan, dan panduan hingga skripsi ini selesai.
3. Bapak Andreyan Rizky Baskara, S.Kom., M.Kom. selaku Koordinator Program Studi Teknologi Informasi Universitas Lambung Mangkurat yang telah memberikan dukungan dan motivasi kepada mahasiswa, terutama penulis. Penulis sangat mengapresiasi dedikasi bapak selama masa jabatan dalam memajukan program studi ini.
4. Seluruh Dosen dan Staff Administrasi Program Studi Teknologi Informasi. Dosen yang telah memberikan ilmu yang berlimpah kepada penulis, dan membantu penulis dalam menjalani perkuliahan.

5. Oling Anugrah Fajar Pratama, yang selalu berada disisi penulis dalam berbagi suka duka, memberikan semangat, perhatian, bantuan, dan motivasi hingga membuat perjalanan penulis lebih bermakna.
6. Seluruh sahabat dan teman seperjuangan. Laily, Cane, Fajra, Ghinaa, Fariz, Pram, dan seluruh teman, kakak tingkat, adik tingkat yang tidak bisa penulis sebutkan satu persatu yang pernah berkontribusi dalam membantu penulis, menghibur, dan menemani dalam melewati masa-masa sulit selama perkuliahan dan penyusunan skripsi.

Terima kasih atas semua pihak yang pernah berkontribusi dalam perjalanan penulis menyusun skripsi. Penulis sadari bahwa skripsi ini masih jauh dari kata sempurna. Oleh karena itu, penulis dengan rendah hati menerima kritik maupun saran yang membangun demi perbaikan di masa yang akan datang. Penulis berharap semoga skripsi ini dapat bermanfaat bagi pembaca, dan dengan segala kerendahan hati penulis memohon doa agar langkah penulis kedepan mendapatkan ridho dan berkah dari Allah SWT.

Banjarmasin, Desember 2024



Amazida

NIM. 2010817120002

DAFTAR ISI

SAMPUL DALAM.....	i
LEMBAR PERNYATAAN.....	ii
LEMBAR PENGESAHAN.....	iii
LEMBAR PERSETUJUAN.....	iv
ABSTRAK.....	v
<i>ABSTRACT</i>	vi
HALAMAN PERSEMBAHAN.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL.....	xiv
DAFTAR LAMPIRAN.....	xvi
DAFTAR RIWAYAT HIDUP.....	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Penelitian Terkait.....	6
2.1.1 <i>Analysis and Research on Library User Behavior</i>	6
2.1.2 <i>Apriori and K-Means Algorithms of Machine Learning</i>	7
2.1.3 <i>Apriori Association Rule and K-Means Clustering Algorithms</i>	7
2.1.4 <i>Research on Improved Full-Factor Deep Information Mining</i>	8
2.1.5 <i>Comparison Performance of K-Medoids and K-Means Algorithms</i> . 8	
2.1.6 <i>Implementation of K-Means, K-Medoid and DBSCAN Algorithms</i> . 9	
2.1.7 <i>Comparison of K-Means and K-Medoids Algorithms</i>	10
2.1.8 <i>Application of Apriori Algorithm in Data Mining</i>	10
2.2 Landasan Teori.....	14
2.2.1 Analisis Pola Hasil Produksi Padi.....	14
2.2.2 <i>Data Preprocessing</i>	14

2.2.3	<i>Association Analysis</i>	15
2.2.4	<i>Cluster Analysis</i>	15
2.2.5	<i>Performance Analysis</i>	17
2.3	Kerangka Pemikiran	18
BAB III METODOLOGI PENELITIAN.....		19
3.1	Alat dan Bahan Penelitian	19
3.1.1	Alat Penelitian.....	19
3.1.2	Bahan Penelitian.....	20
3.2	Alur Penelitian.....	20
3.2.1	Identifikasi Masalah	20
3.2.2	Studi Literatur	21
3.2.3	Pengumpulan Data	21
3.2.4	<i>Preprocessing data</i>	21
3.2.5	Implementasi <i>Apriori</i>	22
3.2.6	Implementasi <i>K-Means</i> dan <i>K-Medoids</i>	24
3.2.7	<i>Silhouette Score</i>	28
3.2.8	Implementasi gabungan <i>Apriori K-Means</i> dan <i>Apriori K-Medoids</i>	29
3.2.9	Evaluasi Performa	30
BAB IV HASIL DAN PEMBAHASAN		31
4.1	Pengumpulan Data	31
4.2	<i>Preprocessing Data</i>	31
4.2.1	<i>Data Cleaning</i>	31
4.2.2	<i>Encoding Data</i>	32
4.2.3	<i>Data preparation</i>	33
4.3	Implementasi <i>Apriori</i>	34
4.3.1	Eksperimen <i>Apriori</i>	36
4.4	Fitur <i>Clustering</i>	42
4.5	Implementasi <i>K-Means</i>	43
4.6	Implementasi <i>K-Medoids</i>	44
4.7	Analisis dan Pembahasan	45
4.7.1.	Evaluasi Performa	46
4.8	Implementasi <i>Apriori K-Means</i>	50
BAB V KESIMPULAN DAN SARAN.....		57
5.1	Kesimpulan.....	57

5.2	Saran.....	57
	DAFTAR PUSTAKA	58
	LAMPIRAN.....	64

DAFTAR GAMBAR

Gambar 1. 1 Perkembangan produksi padi di Indonesia tahun 2020-2021	1
Gambar 2.1 Kerangka Pemikiran.....	18
Gambar 3.1 Alur Penelitian.....	20
Gambar 3. 2 <i>Apriori</i>	22
Gambar 3. 3 <i>Clustering</i> sebelum iterasi	26
Gambar 3. 4 <i>Clustering</i> dengan <i>centroid</i> baru	26
Gambar 3.5 Tahapan implementasi.....	29
Gambar 4. 1 <i>Dataset</i> dari <i>kaggle</i>	31
Gambar 4. 2 Visualisasi <i>Frequent Itemsets</i>	45
Gambar 4. 3 Aturan Asosiasi	46
Gambar 4. 6 Tampilan awal <i>website</i>	50
Gambar 4. 7 Tampilan <i>preview data</i>	51
Gambar 4. 8 Tampilan <i>dataset summary</i>	51
Gambar 4. 9 Tampilan <i>data cleaning</i>	52
Gambar 4. 10 Tampilan <i>Numerical data binning options</i>	53
Gambar 4. 11 Tampilan <i>preview</i> kategori data	53
Gambar 4. 12 Tampilan <i>encode data</i> kategori	53
Gambar 4. 13 Tampilan <i>Generate Association Rules</i>	54
Gambar 4. 14 Tampilan bagian <i>clustering</i>	55
Gambar 4. 15 Visualisasi <i>clustering</i>	55

DAFTAR TABEL

Tabel 2.1 Ringkasan penelitian terkait.....	11
Tabel 3.1 Alat Penelitian.....	19
Tabel 3. 2 <i>Encode data</i>	22
Tabel 3. 3 contoh transaksi 1	23
Tabel 3. 4 Contoh <i>itemset</i>	23
Tabel 3. 5 Contoh data	24
Tabel 3. 6 Perhitungan <i>Euclidean Distance</i>	25
Tabel 3. 7 Pengelompokkan <i>cluster</i> awal	25
Tabel 3. 8 Perhitungan rata-rata <i>cluster</i>	25
Tabel 3. 9 Iterasi perhitungan <i>centroid</i> baru	25
Tabel 3. 10 Pembagian <i>cluster</i> dari <i>centroid</i> baru	26
Tabel 3. 11 Pengelompokan <i>cluster medoid</i> awal.....	27
Tabel 3. 12 Inisiasi <i>medoid</i> baru	27
Tabel 3. 13 Hasil perhitungan <i>euclidean distance</i>	28
Tabel 3. 14 Hasil akhir <i>clustering K-Medoids</i>	28
Tabel 3. 15 <i>Silhouette Score</i>	28
Tabel 4. 1 Sebelum <i>data cleaning</i>	32
Tabel 4. 2 Sesudah <i>data cleaning</i>	32
Tabel 4. 3 Sebelum diskritisasi	32
Tabel 4. 4 Sesudah diskritisasi	33
Tabel 4. 5 Sesudah <i>ordinal encoding</i>	33
Tabel 4. 6 Sebelum <i>data preparation</i>	33
Tabel 4. 7 Sesudah <i>data preparation</i>	34
Tabel 4. 9 <i>Association rules</i>	35
Tabel 4. 10 <i>Frequent itemsets</i>	35
Tabel 4. 11 <i>Association rules</i>	36
Tabel 4. 12 <i>Association rules</i>	36
Tabel 4. 13 Kategori.....	36
Tabel 4. 14 <i>Frequent itemsets</i> dengan 2 kategori	37
Tabel 4. 15 <i>Association rules</i> 2 kategori.....	37
Tabel 4. 16 <i>Frequent itemset</i> 2 kategori.....	39

Tabel 4. 17 <i>Association rules</i> 2 kategori.....	40
Tabel 4. 18 <i>Frequent itemsets</i> 4 kategori.....	40
Tabel 4. 19 <i>Association rule</i> 4 kategori	40
Tabel 4. 20 <i>Frequent itemsets</i> 5 kategori.....	40
Tabel 4. 21 <i>Association rule</i> 5 kategori	41
Tabel 4. 22 <i>Dataframe clustering</i> pertama	42
Tabel 4. 23 <i>Dataframe clustering</i> kedua.....	42
Tabel 4. 24 Eksperimen distribusi <i>K-Means</i> dengan 2 kategori	43
Tabel 4. 25 Eksperimen distribusi <i>K-Means</i> dengan 3 kategori	43
Tabel 4. 26 Eksperimen distribusi <i>K-Medoids</i> dengan 2 kategori	44
Tabel 4. 27 Eksperimen distribusi <i>K-Medoids</i> dengan 3 kategori	44
Tabel 4. 28 <i>Silhouette score K-Means</i> dengan 2 kategori.....	47
Tabel 4. 29 <i>Silhouette score K-Medoids</i> dengan 2 kategori	47
Tabel 4. 30 <i>Silhouette score K-Means</i> dengan 3 kategori.....	47
Tabel 4. 31 <i>Silhouette score K-Medoids</i> dengan 3 kategori	48
Tabel 4. 32 Perbandingan <i>silhouette score</i>	48
Tabel 4. 33 <i>Silhouette score clustering</i> tanpa penggabungan fitur	50

DAFTAR LAMPIRAN

Lampiran 1 Manual Book	64
Lampiran 2 Baris Kode Implementasi	70
Lampiran 3 Baris Kode Perbandingan Model.....	79
Lampiran 4 Surat Pernyataan Kesiapan Membimbing Skripsi	82
Lampiran 5 Lembar Konsultasi.....	83