



**PERBANDINGAN ALGORITMA *K-MEANS* DAN *K-MEDOIDS* UNTUK
ANALISIS POLA BELANJA PELANGGAN MENGGUNAKAN
ALGORITMA *FP- GROWTH***

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

Oleh

MAULANA RAHMAN

1611016310013

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

JULI 2023

SKRIPSI

PERBANDINGAN ALGORITMA *K-MEANS* DAN *K-MEDOIDS* UNTUK ANALISIS POLA BELANJA PELANGGAN MENGGUNAKAN ALGORITMA *FP-GROWTH*

Oleh :

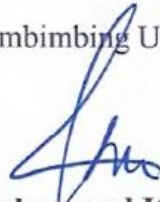
MAULANA RAHMAN

NIM 1611016310013

Telah dipertahankan di depan Dosen Penguji pada tanggal 20 Juli 2023

Susunan Penguji:

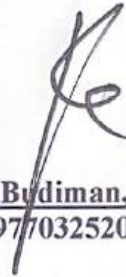
Pembimbing Utama



Muhammad Itqan Mazdadi,
S.Kom., M.Kom.
NIP. 199006122019031013

Penguji

1.



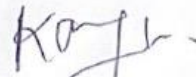
Irwan Budiman, S.T., M.Kom.
NIP. 197703252008121001

Pembimbing Pendamping



Friska Abadi, S.Kom., M.Kom.
NIP.19880913201612110001

2.



Dwi Kartini, S.Kom., M.Kom.
NIP. 198704212012122003



Banjarbaru, Juli 2023
Koordinator PS Ilmu Komputer

Irwan Budiman, S.T., M.Kom.
NIP. 19770325200812100

PERNYATAAN

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

Banjarbaru, 2023

Yang Menyatakan,



Maulana Rahman

NIM. 1611016310013

ABSTRAK

PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS UNTUK ANALISIS POLA BELANJA PELANGGAN MENGGUNAKAN ALGORITMA FP- GROWTH

(Oleh: Maulana Rahman; Pembimbing: Muhammad Itqan Mazdadi, S.Kom, M.Kom dan Friska Abadi, S.Kom, M.Kom; 2023; 52 halaman)

Toko ATK adalah bisnis yang menyediakan berbagai perlengkapan kantor dan memiliki potensi data yang berharga untuk dianalisis melalui teknik *data mining*. Pembeli sering mengalami kesulitan dalam menemukan barang yang mereka cari karena tata letak barang yang masih dilakukan secara acak dan tidak sesuai dengan pola belanja pembeli. Asosiasi merupakan salah satu tipe *data mining* yang digunakan untuk analisis keranjang pasar. Fungsi ini mengidentifikasi item-item barang yang mungkin dibeli pembeli bersamaan dengan barang lain. Metode FP-Growth dapat menghasilkan lebih banyak rule dengan akurasi yang lebih baik, namun proses asosiasi cenderung mengabaikan *itemset* yang besar pada ukuran *dataset* yang besar sehingga untuk *dataset* yang besar hasil rekomendasi kurang akurat. Pada penelitian ini pengelompokan *dataset* dilakukan dengan menggunakan metode *clustering*, yaitu *K-Means* dan *K-Medoids*. Tahapan penelitian terdiri dari pengambilan data, *preprocessing data*, *clustering* dengan menggunakan *K-Means* dan *K-Medoids*, asosiasi menggunakan FP-Growth dan evaluasi model. Hasil dari penelitian ini menunjukkan bahwa *clustering* dengan metode *K-Means* lebih banyak menghasilkan aturan asosiasi dibandingkan *K-Medoids*. *Clustering* dengan nilai DBI terbaik menggunakan *K-Means* berada pada $K=6$ sedangkan *K-Medoids* berada pada $K=3$. Model asosiasi menggunakan metode *K-Means* mampu menghasilkan 33.242 aturan asosiasi yang sesuai dengan nilai *Support* dan *Confidence* dan 18.405 di antaranya memenuhi nilai *lift ratio*.

Kata kunci: *clustering, FP-Growth, data mining, analisis keranjang pasar, k-means, k-medoids*

ABSTRACT

THE COMPARISON OF K-MEANS AND K-MEDOIDS ALGORITHMS FOR THE ANALYSIS OF CUSTOMER PATTERN USING THE FP-GROWTH ALGORITHM

(By: Maulana Rahman; Advisor : Muhammad Itqan Mazdadi, S.Kom, M.Kom and Friska Abadi, S.Kom, M.Kom; 2023; 52 pages)

A stationery store is a business that provides various office supplies and has potentially valuable data to be analyzed through data mining techniques. Buyers often experience difficulty finding the items they want because the layout of the items is unorganized and incompatible with the buyer's shopping patterns. The association is a type of data mining used for market basket analysis. This function identifies items that a buyer might buy concurrently with other things. The FP-Growth method can produce more rules with better accuracy. Still, the association process disregards large item sets at large dataset sizes, making the recommendations for large datasets less accurate. In this study, datasets were grouped using the clustering method, i.e., K-Means, and K-Medoids. The stages of the research consisted of data collection, data preprocessing, clustering using K-Means and K-Medoids, Association using FP-Growth, and model evaluation. The results of this study indicate that clustering with the K-Means method produces more association rules than K-Medoids. Clustering with the best DBI value using K-Means is at $K=6$ while K-Medoids is at $K=3$. Using the K-Means method, the association model can produce 33.242 association rules corresponding to the Support and Confidence values, and 18.405 rules fulfill the lift ratio value.

Keywords: *clustering, FP-Growth, data mining, market basket analysis, k-means, k-medoids*

PRAKATA

Puji dan syukur kepada Allah SWT karena atas berkat rahmat-Nya penulis dapat menyelesaikan skripsi yang berjudul “PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS UNTUK ANALISIS POLA BELANJA PELANGGAN MENGGUNAKAN ALGORITMA FP- GROWTH” untuk memenuhi syarat dalam menyelesaikan pendidikan program S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.

Tidak lupa penulis menyampaikan ucapan terimakasih kepada pihak-pihak terkait yang sangat mendukung dan membantu penulis dalam penyusunan skripsi ini, adapun yang dimaksud adalah sebagai berikut :

1. Kedua orang tua serta keluarga yang senantiasa memberikan doa, semangat, dukungan, hingga kepercayaan yang membuat penulis selalu bekerja keras menyelesaikan skripsi.
2. Bapak Muhammad Itqan Mazdadi, S.Kom, M.Kom selaku dosen pembimbing utama yang senantiasa membimbing, membantu, dan meluangkan waktu dalam proses penyelesaian skripsi ini.
3. Bapak Friska Abadi, S.Kom, M.Kom., selaku dosen pembimbing pendamping yang turut serta memberi arahan, membantu dan meluangkan waktu dalam proses penyelesaian skripsi ini.
4. Teman-teman keluarga Ilmu Komputer angkatan 2016.
5. Semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu dalam menyelesaikan skripsi ini.

Akhir kata penulis menyadari bahwa skripsi ini jauh dari sempurna, kritik dan saran yang membangun sangat diperlukan. Semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT

Banjarbaru, Juni 2023



Maulana Rahman

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
PERNYATAAN.....	iii
ABSTRAK	iv
ABSTRACT	v
PRAKATA.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN.....	xii
BAB I PENDAHULUAN	
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah.....	2
1.4. Tujuan Penelitian.....	3
1.5. Manfaat Penelitian.....	3
BAB II TINJAUAN PUSTAKA	
2.1 Kajian Terdahulu	4
2.2 Keaslian Penelitian	6
2.3 <i>Knowledge Discovery in Database</i>	7
2.4 Data Mining.....	8
2.5 Association Rules	11
2.6 Algoritma <i>Fp-Growth</i>	13
2.7 <i>FP- Tree</i>	14
2.8 <i>Lift Ratio</i>	15
2.9 <i>Clustering</i>	16
2.10 <i>K-Means</i>	18
2.11 <i>K-Medoids</i>	20

2.12	<i>Davies Bouldin Index</i>	21
2.13	RapidMiner	22
2.14	Bahasa Pemrograman <i>Python</i>	23
BAB III METODE PENELITIAN		
3.1.	Alat Penelitian	26
3.2.	Bahan Penelitian.....	26
3.3.	Prosedur Penelitian.....	26
BAB IV HASIL DAN PEMBAHASAN		
4.1.	Hasil.....	30
4.1.1.	Pengambilan Data.	30
4.1.2.	Preprocessing.	31
4.1.3.	<i>Clustering K-Means</i>	33
4.1.4.	<i>Clustering K-Medoids</i>	40
4.1.5.	Evaluasi.....	46
4.2.	Pembahasan	47
BAB V PENUTUP		
5.1.	Kesimpulan.....	52
5.2.	Saran.....	52
DAFTAR PUSTAKA		
LAMPIRAN		

DAFTAR TABEL

Tabel 1.Keaslian Penelitian.....	6
Tabel 2. Tabel Rancangan Penelitian.....	7
Tabel 3. List FP-Tree	14
Tabel 4. Contoh data yang akan di gunakan	32
Tabel 5. Perhitungan Jarak K-Means	33
Tabel 6. SourceCode K-Means	34
Tabel 7. Hasil K-Means	34
Tabel 8. Hasil Cluster K-Means.....	35
Tabel 9.Frekuensi Item K-Means.....	36
Tabel 10. Conditional FP-Tree K-means	37
Tabel 11. Pembangkitan Frequent Pattern	37
Tabel 12. Perhitungan Nilai Support.....	37
Tabel 13. Hasil Perhitungan Nilai Support	38
Tabel 14. Perhitungan Nilai Confidence	38
Tabel 15. Hasil Perhitungan Nilai Confidence	38
Tabel 16. Hasil Aturan Asosiasi.....	39
Tabel 17. Hasil Asosiasi K-Means.....	40
Tabel 18.Perhitungan Jarak K-medoids	40
Tabel 19. SourceCode K-Medoids.....	41
Tabel 20. Hasil K-Medoids	41
Tabel 21. Hasil Cluster K-Medoids	42
Tabel 22. Frekuensi Item K-medoids.....	43
Tabel 23. Conditional FP-Tree.....	43
Tabel 24. Frequent Pattern	44
Tabel 25. Perhitungan Nilai Support.....	44
Tabel 26. Hasil Perhitungan Nilai Support	44
Tabel 27. Perhitungan Nilai Confidence	45
Tabel 28.Hasil Perhitungan Nilai Confidence	45
Tabel 29.Hasil Aturan Asosiasi.....	45

Tabel 30. Hasil Asosiasi K-Medoids 46
Tabel 31. Hasil Asosiasi Fp-Growth..... 47

DAFTAR GAMBAR

Gambar 1. Tahapan Proses KDD	7
Gambar 2. Tahapan Data Mining.....	9
Gambar 3. Contoh FP-Tree	15
Gambar 4. Contoh Pembentukan Cluster.....	16
Gambar 5. Alur penelitian.....	27
Gambar 6. Data Transaksi.....	31
Gambar 7. Data Selection	31
Gambar 8. Data yang telah di transformasi.....	33
Gambar 10. Fp-Tree K-Means	36
Gambar 11. Rangkaian Fp-Growth dan K-Means	39
Gambar 12. Fp-Tree K-Medoids.....	43
Gambar 13. Algoritma FP-Growth dengan K-Medoids.....	46
Gambar 14. Grafik Perbandingan Hasil Asosiasi	50
Gambar 15. Nilai Lift Ratio	51

DAFTAR LAMPIRAN

Lampiran 1. Data Transaksi	57
Lampiran 2. SourceCode Clustering K-Means	57
Lampiran 3. SourceCode Clustering K-Medoids.....	58