



SKRIPSI

KLASTERISASI POLA PEMBELIAN KONSUMEN MENGUNAKAN *IMPROVED* K-MEDOIDS BERBASIS *CROW SEARCH ALGORITHM* DAN FP-GROWTH (STUDI KASUS: GERAJ MAKANAN CEPAT SAJI SURABAYA)

MELINDA PUTRI AZZAHRA

NPM 22083010035

DOSEN PEMBIMBING

Wahyu Syaifullah Jauharis Saputra, S.Kom., M.Kom.

Muhammad Nasrudin, S.Si., M.Stat.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2026**

LEMBAR PENGESAHAN

KLASTERISASI POLA PEMBELIAN KONSUMEN MENGGUNAKAN IMPROVED K-MEDOIDS BERBASIS CROW SEARCH ALGORITHM DAN FP-GROWTH (STUDI KASUS: GERAJ MAKANAN CEPAT SAJI SURABAYA)

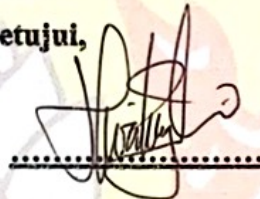
Oleh:
MELINDA PUTRI AZZAHRA
NPM. 22083010035

Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Sidang Skripsi Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur pada Tanggal 31 Agustus 2026:

Menyetujui,

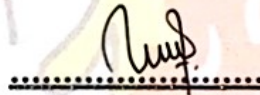
Wahvu Syaifullah J.S., S.Kom.,
M.Kom.

NIP/NPT 19860825 202121 1 003



(Pembimbing I)

Muhammad Nasrudin, S.Si., M.Stat.
NIP/NPT 19960909 202406 1 002



(Pembimbing II)

Amri Muhaimin, S.Stat., M.Stat.,
MS.

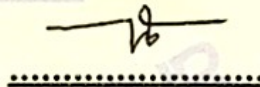
NIP/NPT 19950723 202406 1 002



(Ketua Penguji)

Dr. Ir. Mohammad Idhom, S.P.,
S.Kom., M.T.

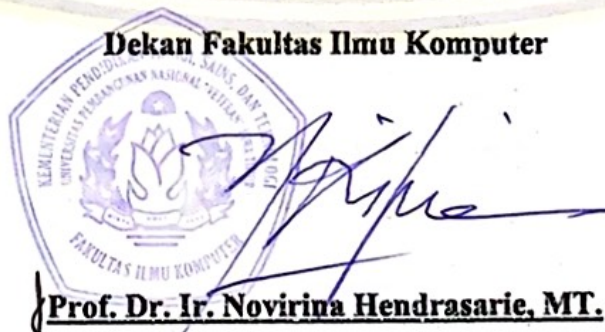
NIP/NPT 19830310 202121 1 006



(Penguji I)

Mengetahui,

Dekan Fakultas Ilmu Komputer



Prof. Dr. Ir. Novirina Hendrasarie, MT.

NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

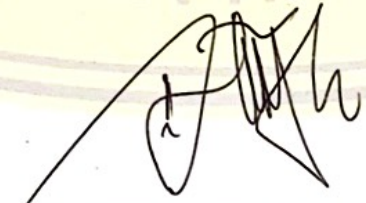
**KLASTERISASI POLA PEMBELIAN KONSUMEN MENGGUNAKAN
IMPROVED K-MEDOIDS BERBASIS CROW SEARCH ALGORITHM DAN
FP-GROWTH (STUDI KASUS: GERAJ MAKANAN CEPAT SAJI
SURABAYA)**

Oleh:
MELINDA PUTRI AZZAHRA
NPM. 22083010035

Telah disetujui untuk mengikuti Ujian Skripsi

Menyetujui,

**Koordinator Program Studi Sains Data
Fakultas Ilmu Komputer**



Amri Muhaimin, S.Stat., M.Stat., M.S.
NIP. 19950723 202406 1 002

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan di bawah ini:

Nama Mahasiswa : Melinda Putri Azzahra
NPM : 22083010035
Program : Sarjana (S1)
Program Studi : Sains Data
Fakultas : Fakultas Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila di kemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.

Surabaya, 11 September 2026
Yang Membuat Pernyataan,



MELINDA PUTRI AZZAHRA
NPM. 22083010035

ABSTRAK

Nama Mahasiswa / NPM : Melinda Putri Azzahra / 22083010035
Judul Skripsi : Klasterisasi Pola Pembelian Konsumen Menggunakan *Improved* K-Medoids Berbasis *Crow Search Algorithm* dan Fp-Growth (Studi Kasus: Gerai Makanan Cepat Saji Surabaya)
Dosen Pembimbing : 1. Wahyu Syaifullah Jauharis Saputra, S.Kom., M.Kom.
2. Muhammad Nasrudin, S.Si., M.Stat.

Industri *quick service restaurant* (QSR) di Indonesia terus berkembang seiring dengan perubahan gaya hidup masyarakat yang mengutamakan kecepatan dan kepraktisan. Persaingan antargera QSR mendorong perlunya pemanfaatan data transaksi untuk memahami karakteristik transaksi dan pola pembelian. Penelitian ini menerapkan *Improved* K-Medoids berbasis *Crow Search Algorithm* (CSA) untuk melakukan klasterisasi transaksi, kemudian menggunakan algoritma FP-Growth untuk menganalisis pola pembelian pada setiap klaster. Data penelitian diperoleh dari sistem *point of sale* (POS) salah satu gerai Hisana Fried Chicken di Kota Surabaya pada periode 1 Januari hingga 28 Maret 2025, terdiri atas 18.814 baris data *item-level* yang diagregasi menjadi 12.841 transaksi unik. Tahap pra-pemrosesan meliputi seleksi atribut, pemeriksaan dan perbaikan kualitas data, transformasi tipe data, agregasi, normalisasi *Min-Max Scaling*, serta eksplorasi data. Hasil klasterisasi menunjukkan empat klaster optimal dengan *Silhouette Score* sebesar 0,5594, *Davies-Bouldin Index* (DBI) sebesar 0,7703, dan *fitness* sebesar 15,1700. Dibandingkan K-Medoids konvensional, metode berbasis CSA menghasilkan kualitas klaster yang lebih baik berdasarkan ketiga metrik tersebut. Analisis sensitivitas mengidentifikasi 1.403 transaksi sebagai *outlier* dan menunjukkan bahwa keberadaannya memengaruhi struktur klaster. Penerapan FP-Growth tanpa klasterisasi menghasilkan 35 *frequent itemset* tanpa aturan asosiasi, sedangkan setelah klasterisasi aturan ditemukan pada Klaster 0, Klaster 2, dan Klaster 3. Aturan terbaik terdapat pada Klaster 0, yaitu Saos BBQ dan Sambal Samyang diikuti Ayam Bekakak dengan *confidence* 85% dan *lift* 22,8. Hasil penelitian disajikan melalui GUI untuk membantu pengelola gerai memahami hasil analisis sebagai bahan pertimbangan strategi pemasaran berbasis data.

Kata Kunci: *Quick Service Restaurant* (QSR), Klasterisasi Transaksi, *Improved* K-Medoids, *Crow Search Algorithm* (CSA), FP-Growth

ABSTRACT

Student Name / NPM : Melinda Putri Azzahra / 22083010035
Undergraduate thesis title : *Clustering of Consumer Purchase Patterns Using Improved K-Medoids Based On Crow Search Algorithm and FP-Growth (Case Study: Surabaya Fast Food Outlet)*
Advisors : 1. Wahyu Syaifullah Jauharis Saputra, S.Kom., M.Kom.
2. Muhammad Nasrudin, S.Si., M.Stat.

The quick service restaurant (QSR) industry in Indonesia continues to grow in line with changes in people's lifestyles that prioritize speed and convenience. Increasing competition among QSR outlets has encouraged the need to utilize transaction data to better understand transaction characteristics and purchasing patterns. This study applies an Improved K-Medoids method based on the Crow Search Algorithm (CSA) to cluster transactions, followed by the FP-Growth algorithm to analyze purchasing patterns within each cluster. The research data were obtained from the point-of-sale (POS) system of a Hisana Fried Chicken outlet in Surabaya for the period from January 1 to March 28, 2025. The dataset consisted of 18,814 item-level records, which were aggregated into 12,841 unique transactions. The preprocessing stages included attribute selection, data quality checking and correction, data type transformation, aggregation, Min-Max Scaling normalization, and exploratory data analysis. The clustering results identified four optimal clusters with a Silhouette Score of 0.5594, a Davies-Bouldin Index (DBI) of 0.7703, and a fitness value of 15.1700. Compared with conventional K-Medoids, the CSA-based method produced better clustering quality based on these three metrics. Sensitivity analysis identified 1,403 transactions as outliers and showed that their presence affected the cluster structure. FP-Growth without clustering produced 35 frequent itemsets but no association rules, whereas after clustering, association rules were found in Clusters 0, 2, and 3. The best rule was found in Cluster 0, where the purchase of Saos BBQ and Sambal Samyang was followed by the purchase of Ayam Bekakak, with 85% confidence and a lift value of 22.8. The research results were presented through a graphical user interface (GUI) to help outlet managers understand the analysis results as a basis for considering data-driven marketing strategies.

Keywords: *Quick Service Restaurant (QSR), Transaction Clustering, Improved K-Medoids, Crow Search Algorithm (CSA), FP-Growth*

KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas segala rahmat, hidayah dan karunia-Nya kepada penulis sehingga skripsi dengan judul **“Klasterisasi Pola Pembelian Konsumen Menggunakan *Improved K-Medoids* Berbasis *Crow Search Algorithm* dan *FP-Growth* (Studi Kasus: Gerai Makanan Cepat Saji Surabaya)”** dapat terselesaikan dengan baik.

Penulis mengucapkan terima kasih kepada Bapak Wahyu Syaifullah Jauharis Saputra, S.Kom., M.Kom. selaku Dosen Pembimbing I dan Bapak Muhammad Nasrudin, S.Si., M.Stat. selaku Dosen Pembimbing 2 yang telah meluangkan waktunya untuk memberikan bimbingan, nasehat serta motivasi kepada penulis. Selain itu, selama penyusunan skripsi penulis juga banyak menerima bantuan dari berbagai pihak. Untuk itu penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, M.MT., selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Dr. Ir. Novirina Henrasarie, M.T., selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Bapak Amri Muhaimin, S.Stat., M.Stat., M.S., selaku Koordinator Program Studi Sains Data UPN “Veteran” Jawa Timur.
4. Bapak/Ibu dosen yang telah memberi pengetahuan kepada penulis selama menempuh pendidikan di program studi S1 Sains Data, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur.
5. Alm. Bapak yang telah berpulang 10 tahun yang lalu saat penulis masih duduk di bangku SMP. Meskipun beliau tidak lagi mendampingi penulis secara langsung dalam menempuh pendidikan hingga jenjang perguruan tinggi, kasih sayang, didikan, dan nilai-nilai kehidupan yang telah beliau tanamkan sejak kecil tetap menjadi bekal dan sumber kekuatan bagi penulis dalam menjalani setiap proses, termasuk dalam menyelesaikan skripsi ini.
6. Ibu, Mas Heru, Mbak Intani, Mas Handi, Azam, Alisha, dan Alila yang senantiasa memberikan doa, kasih sayang, dukungan moral, dan dukungan

finansial kepada penulis selama menempuh perkuliahan hingga penyelesaian penelitian ini.

7. Komploten (Lidya, Fira, Sela, Chelin, Aden, Ajeng, Serlinda, Bitu, Ayak) yang menemani penulis dari awal perkuliahan hingga akhir menyelesaikan masa studi ini.
8. Keluarga Keciyl (Fanya, Nia, Rahma, Balya, Muplik) sahabat terdekat dan terlama penulis yang selalu membersamai penulis diberbagai fase kehidupan.
9. Teman terdekat penulis yang senantiasa mendengarkan keluhan kesah, memberikan dukungan, dan selalu menemani selama proses penyusunan penelitian ini.
10. Semua pihak yang memberikan dukungan kepada penulis selama proses penelitian ini berlangsung.

Penulis menyadari bahwa di dalam penyusunan skripsi ini banyak terdapat kekurangan. Untuk itu kritik dan saran yang membangun dari semua pihak sangat diharapkan demi kesempurnaan penulisan skripsi ini. Akhirnya, dengan segala keterbatasan yang penulis miliki semoga laporan ini dapat bermanfaat bagi semua pihak umumnya dan penulis pada khususnya.

Surabaya, September 2026

Melinda Putri Azzahra

DAFTAR ISI

HALAMAN SAMPUL.....	i
LEMBAR PENGESAHAN	ii
LEMBAR PERSETUJUAN	iii
SURAT PERNYATAAN BEBAS PLAGIASI.....	iv
ABSTRAK	v
<i>ABSTRACT</i>	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xiv
DAFTAR NOTASI.....	xv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	6
1.3. Batasan Masalah	7
1.4. Tujuan Penelitian	7
1.5. Manfaat Penelitian	8
BAB II TINJAUAN PUSTAKA	9
2.1. Penelitian Terdahulu	9
2.2. Kerangka Teori	16
2.2.1. <i>Quick Service Restaurant</i>	16
2.2.2. Klasterisasi	17
2.2.3. Algoritma K-Medoids	18
2.2.4. <i>Crow Search Algorithm (CSA)</i>	21
2.2.5. <i>Improved K-Medoids berbasis Crow Search Algorithm (CSA)</i> ...	24

2.2.6.	<i>Silhouette Score</i>	27
2.2.7.	<i>Davies Bouldin-Index</i>	29
2.2.8.	<i>Association Rules Mining</i>	30
2.2.9.	Algoritma FP-Growth.....	32
BAB III METODOLOGI PENELITIAN.....		37
3.1.	Variabel Penelitian dan Sumber Data.....	37
3.2.	Langkah Analisis	38
3.2.1.	Pengumpulan Data.....	39
3.2.2.	Pra-pemrosesan Data	39
3.2.3.	<i>Improved K-Medoids berbasis Crow Search Algorithm (CSA)</i>	41
3.2.4.	Pembentukan Dataset Transaksi Berdasarkan Hasil Klasterisasi..	43
3.2.5.	Analisis Pola Pembelian Menggunakan Algoritma FP-Growth....	44
3.3.	Desain Sistem	47
BAB IV HASIL DAN PEMBAHASAN		51
4.1.	Hasil Penelitian.....	51
4.1.1	Pengumpulan Data.....	51
4.1.2	Pra-pemrosesan Data	53
4.1.3	<i>Improved K-Medoids berbasis Crow Search Algorithm (CSA)</i>	67
4.1.4	Pembentukan Dataset Transaksi Berdasarkan Hasil Klasterisasi..	94
4.1.5	Analisis Pola Pembelian Menggunakan Algoritma FP-Growth....	97
4.2.	Tampilan Antar Muka (Graphical User Interface)	115
BAB V PENUTUP		121
5.1.	Kesimpulan.....	121
5.2.	Saran Pengembangan.....	123
DAFTAR PUSTAKA		125
LAMPIRAN		133

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Klasterisasi.....	17
Gambar 2.2 Ilustrasi Cara Kerja K-Medoids [21]	18
Gambar 2.3 Alur Algoritma K-Medoids	20
Gambar 2.4 Ilustrasi Cara Kerja CSA [31]	22
Gambar 2.5 Alur <i>Crow Search Algorithm</i> (CSA)	23
Gambar 2.6 Alur <i>Improved</i> K-Medoids berbasis CSA [13].....	25
Gambar 2.7 Pembentukan <i>Frequent</i> Itemset-1 [43]	33
Gambar 2.8 Pengurutan Item secara Menurun (<i>Descending</i>) [43].....	34
Gambar 2.9 Pembentukan <i>FP-Tree</i> T1 [43].....	34
Gambar 2.10 Hasil <i>FP-Tree</i> Seluruh Transaksi [43]	34
Gambar 2.11 Pembentukan CPB [43]	35
Gambar 2.12 <i>Conditional</i> <i>FP-Tree</i> untuk I5 [43].....	35
Gambar 2.13 Hasil <i>Conditional Pattern Base</i> dan <i>Conditional FP-Tree</i> [43]....	35
Gambar 2.14 Hasil Penentuan <i>Frequent Patterns</i> [43]	36
Gambar 2.15 Hasil Pembentukan Aturan Asosiasi [44].....	36
Gambar 3.1 Alur Penelitian.....	38
Gambar 3.2 Pra-pemrosesan Data	39
Gambar 3.3 <i>Improved</i> K-Medoids berbasis <i>Crow Search Algorithm</i> (CSA).....	41
Gambar 3.4 Implementasi <i>FP-Growth</i>	44
Gambar 3.5 Desain Sistem <i>Website</i>	47
Gambar 4.1 Visualisasi Distribusi 10 Produk Terlaris	61
Gambar 4.2 Visualisasi Distribusi Nilai Transaksi	63
Gambar 4.3 Visualisasi Distribusi Jumlah Item per Transaksi	66
Gambar 4.4 Perbandingan Hasil Evaluasi Parameter CSA	78
Gambar 4.5 Perbandingan Hasil Evaluasi <i>Random State</i>	82
Gambar 4.6 Tampilan Halaman <i>Home</i>	116
Gambar 4.7 Tampilan Setelah <i>Upload</i> Dataset	117
Gambar 4.8 Tampilan EDA.....	117
Gambar 4.9 Tampilan Halaman <i>Clustering</i> K-Medoids CSA	118

Gambar 4.10 Tampilan Ketika Halaman <i>Clustering</i> K-Medoids CSA Setelah Dijalankan.....	118
Gambar 4.11 Tampilan Halaman <i>FP-Growth Analysis</i>	119
Gambar 4.12 Tampilan Ketika Halaman <i>FP-Growth Analysis</i> Setelah Dijalankan	119
Gambar 4.13 Tampilan Halaman Hasil dan Report.....	120

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	9
Tabel 2.2 Penjelasan Nilai <i>Silhouette Score</i>	28
Tabel 3.1 Metadata	37
Tabel 4.1 Hasil Seleksi Atribut	53
Tabel 4.2 Hasil Identifikasi Nilai Hilang	54
Tabel 4.3 Hasil Pemeriksaan Nilai Nol	55
Tabel 4.4 Hasil Transformasi Tipe Data	57
Tabel 4.5 Hasil Agregasi Data	58
Tabel 4.6 Hasil Normalisasi Data	60
Tabel 4.7 Hasil Penentuan Jumlah Klaster Awal	72
Tabel 4.8 Hasil Evaluasi Parameter <i>Crow Search Algorithm</i> (CSA)	76
Tabel 4.9 Hasil Evaluasi <i>Random State</i>	81
Tabel 4.10 Hasil Pengujian Jumlah Klaster dengan Parameter Final	83
Tabel 4.11 Parameter Klasterisasi Akhir	86
Tabel 4.12 Hasil Perhitungan Jarak <i>Euclidean</i>	86
Tabel 4.13 Ringkasan Karakteristik Klaster	88
Tabel 4.14 Hasil Penggabungan Label Klaster ke Data Transaksi	95
Tabel 4.15 Hasil Segmentasi Transaksi Berdasarkan Klaster	96
Tabel 4.16 Nilai <i>Minimum Support</i> pada Setiap Klaster	101
Tabel 4.17 Jumlah <i>Frequent Itemset</i> pada Setiap Klaster	102
Tabel 4.18 Hasil <i>Frequent Itemset</i> Klaster 0	103
Tabel 4.19 Hasil <i>Frequent Itemset</i> Klaster 1	103
Tabel 4.20 Hasil <i>Frequent Itemset</i> Klaster 2	103
Tabel 4.21 Hasil <i>Frequent Itemset</i> Klaster 3	104
Tabel 4.22 Nilai Minimum <i>Confidence</i> pada Setiap Klaster	106
Tabel 4.23 Aturan Asosiasi Terbaik per Klaster	106
Tabel 4.24 Saran Strategi Tiap Klaster	108

DAFTAR LAMPIRAN

Lampiran 1. Data	133
Lampiran 2. Kode <i>Script</i>	134

DAFTAR NOTASI

K	:	Jumlah klaster
n	:	Jumlah data dalam dataset
X	:	Dataset atau kumpulan data transaksi
x_i	:	Data transaksi ke- i
m_k	:	<i>Medoid</i> pada klaster ke- k
$d(x_i, m_i)$	:	Jarak <i>Euclidean</i> antara data ke- i dengan <i>medoid</i> klaster ke- k
F	:	Nilai <i>fitness</i> atau fungsi objektif K-Medoids
P	:	Ukuran populasi (<i>population size</i>)
AP	:	<i>Awareness probability</i> , yaitu probabilitas burung gagak menyadari adanya burung gagak lain yang mengikutinya
fl	:	<i>Flight length</i> , yaitu panjang perpindahan posisi burung gagak
t	:	Iterasi ke- t
M_i	:	Posisi memori terbaik yang pernah dicapai burung gagak ke- i
$X_i(t)$	:	Posisi burung gagak ke- i pada iterasi ke- t
r	:	Bilangan acak yang digunakan dalam proses pembaruan posisi CSA
$S(i)$	:	Nilai <i>Silhouette Score</i> untuk data ke- i
$a(i)$	:	Rata-rata jarak data ke- i terhadap data lain dalam klaster yang sama
$b(i)$	:	Rata-rata jarak minimum data ke- i terhadap data pada klaster lain