



SKRIPSI

PENGELOMPOKAN PELANGGAN DAN REKOMENDASI PAKET PRODUK BERDASARKAN POLA PEMBELIAN KONSUMEN MENGGUNAKAN *K-PROTOTYPE* DAN *FP-GROWTH*

DENDY ARIZKI KUSWARDANA
NPM 21083010006

DOSEN PEMBIMBING

Dr. Eng. Ir. Dwi Arman Prasetya, ST., MT., IPU., Asean, Eng.
Trimono, S.Si., M.Si.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**



SKRIPSI

PENGELOMPOKAN PELANGGAN DAN REKOMENDASI PAKET PRODUK BERDASARKAN POLA PEMBELIAN KONSUMEN MENGGUNAKAN *K-PROTOTYPE* DAN *FP-GROWTH*

DENDY ARIZKI KUSWARDANA
NPM 21083010006

DOSEN PEMBIMBING

Dr. Eng. Ir. Dwi Arman Prasetya, ST., MT., IPU., Asean, Eng.
Trimono, S.Si., M.Si.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**



SKRIPSI

PENGELOMPOKAN PELANGGAN DAN REKOMENDASI PAKET PRODUK BERDASARKAN POLA PEMBELIAN KONSUMEN MENGGUNAKAN *K-PROTOTYPE* DAN *FP-GROWTH*

DENDY ARIZKI KUSWARDANA
NPM 21083010006

DOSEN PEMBIMBING

Dr. Eng. Ir. Dwi Arman Prasetya, ST., MT., IPU., Asean, Eng.
Trimono, S.Si., M.Si.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**

LEMBAR PENGESAHAN

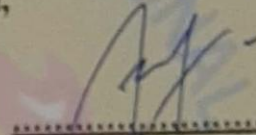
PENGELOMPOKAN PELANGGAN DAN REKOMENDASI PAKET PRODUK BERDASARKAN POLA PEMBELIAN KONSUMEN MENGUNAKAN *K-PROTOTYPE* DAN *FP-GROWTH*

Oleh :
DENDY ARIZKI KUSWARDANA
NPM. 21083010006

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 4 Juni 2025:

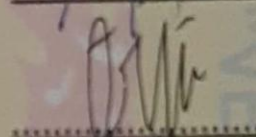
Menyetujui,

Dr. Eng. Ir. Dwi Arman Prasetya, ST., MT.,
IPU., Asean, Eng.
NIP. 19801205 200501 1 002



(Pembimbing I)

Trimono, S.Si., M.Si.
NIP. 19950908 202203 1 003



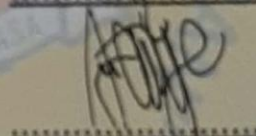
(Pembimbing II)

Kartika Maulida Hindrayani, S.Kom., M.Kom.
NIP. 19920909 202203 2 009



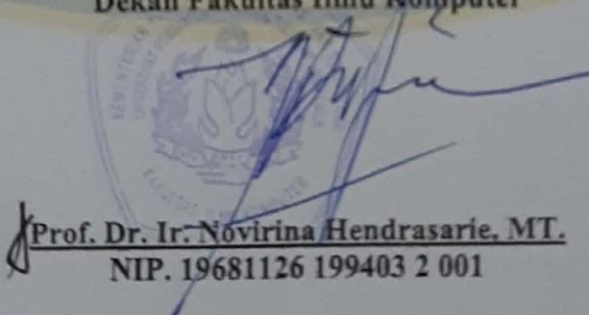
(Ketua Penguji)

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



(Penguji I)

Mengetahui,
Dekan Fakultas Ilmu Komputer



Prof. Dr. Ir. Novirina Hendrasarie, MT.
NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

PENGELOMPOKAN PELANGGAN DAN REKOMENDASI PAKET PRODUK BERDASARKAN POLA PEMBELIAN KONSUMEN MENGUNAKAN *K-PROTOTYPE* DAN *FP-GROWTH*

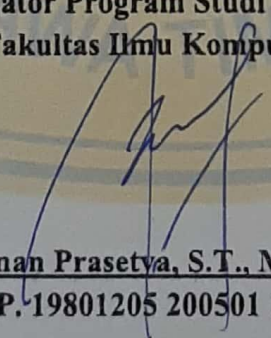
Oleh:
DENDY ARIZKI KUSWARDANA
NPM. 21083010006

Telah disetujui untuk mengikuti Ujian Skripsi

The logo of Universitas Pembangunan Nasional Veteran is a large, light-colored watermark in the background. It is a shield-shaped emblem with a yellow field. In the center is a white lotus flower with five petals. Above the lotus is a blue circle containing a white star. On either side of the lotus are red flames. Below the lotus is a white banner with the text 'WIDYA MWAT YASA' in black. The words 'UNIVERSITAS PEMBANGUNAN NASIONAL' are written in a semi-circle along the top left, and 'VETERAN' is written in a semi-circle along the top right.

Menyetujui,

Koordinator Program Studi Sains Data
Fakultas Ilmu Komputer

A blue ink signature is written over the text of the coordinator's name and NIP.

Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU., Asean, Eng.
NIP. 19801205 200501 1 002

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan di bawah ini:

Nama : Dendy Arizki Kuswardana
NPM : 2108301010006
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, 4 Juni 2025

Yang Membuat Pernyataan



Dendy Arizki Kuswardana
NPM. 2108301010006

ABSTRAK

Nama Mahasiswa / NPM : Dendy Arizki Kuswardana / 21083010006
Judul Skripsi : Pengelompokan Pelanggan dan Rekomendasi Paket Produk Berdasarkan Pola Pembelian Konsumen Menggunakan *K-Prototype* dan *FP-Growth*
Dosen Pembimbing : 1. Dr. Eng. Ir. Dwi Arman Prasetya, ST., MT., IPU., Asean, Eng.
2. Trimono, S.Si., M.Si.

Persaingan bisnis kuliner di Kota Surabaya semakin ketat, termasuk di sektor UMKM yang menyajikan menu khas Jepang. Di tengah tantangan pencapaian target penjualan, terutama pada akhir pekan, diperlukan strategi berbasis data untuk memahami perilaku pelanggan dan meningkatkan efektivitas pemasaran. Penelitian ini menggabungkan dua pendekatan data mining, algoritma *K-Prototype* untuk pengelompokan pelanggan dan algoritma *FP-Growth* untuk analisis asosiasi pola pembelian menu. Hasil klusterisasi menunjukkan dua kelompok pelanggan, yaitu Klaster 0 (75%) yang didominasi pelanggan dengan frekuensi tinggi dan loyalitas stabil, serta Klaster 1 (25%) yang memiliki pola belanja lebih spesifik dan terbiasa dengan layanan digital. Nilai *silhouette score* sebesar 0,6168 mengindikasikan kualitas kluster yang baik. Sementara itu, analisis asosiasi mengidentifikasi kombinasi menu populer seperti *Ocha Reffil* → *Gohan* (Nasi Putih) dengan nilai *confidence* sebesar 64,17% dan *lift* sebesar 1,21. Temuan ini digunakan untuk merancang paketmenu strategis seperti “Paket Hemat Jepang” untuk Klaster 0, serta “Paket *Mentai Lovers*” dengan promosi digital untuk Klaster 1. Pendekatan gabungan ini tidak hanya mendorong strategi *cross-selling* dan personalisasi layanan, tetapi juga membantu dalam optimalisasi stok dan peningkatan loyalitas pelanggan.

Kata kunci: Pola Pembelian Konsumen, Pengelompokan Pelanggan, Rekomendasi Menu, *K-Prototype*, *FP-Growth*, UMKM

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Dendy Arizki Kuswardana / 21083010006
Thesis Title : *Customer Clustering and Product Package Recommendations Based on Consumer Purchasing Patterns Using K-Prototype and FP-Growth*
Advisor : 1. Dr. Eng. Ir. Dwi Arman Prasetya, ST., MT., IPU., Asean, Eng.
2. Trimono, S.Si., M.Si.

The culinary business competition in Surabaya City is getting tighter, including in the MSMEs sector that serves Japanese specialties. Amidst the challenges of achieving sales targets, especially on weekends, a data-based strategy is needed to understand customer behavior and increase marketing effectiveness. This study combines two data mining approaches, the K-Prototype algorithm for customer clustering and the FP-Growth algorithm for menu purchasing pattern association analysis. The clustering results show two customer groups, namely Cluster 0 (75%) which is dominated by customers with high frequency and stable loyalty, and Cluster 1 (25%) which has a more specific shopping pattern and is accustomed to digital services. The silhouette score value of 0,6168 indicates good cluster quality. Meanwhile, the association analysis identified popular menu combinations such as Ocha Reffil → Gohan (White Rice) with a confidence value of 64,7% and a lift of 1,21. These findings were used to design strategic menu packages such as the “Japanese Value Package” for Cluster 0, and the “Mentai Lovers Package” with digital promotions for Cluster 1. This combined approach not only drives cross-selling and service personalization strategies, but also helps in stock optimization and increasing customer loyalty.

Keywords: *Consumer Purchasing Patterns, Customer Grouping, Menu Recommendations, K-Prototype, FP-Growth, MSME*

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas segala rahmat, hidayah dan karunia-Nya kepada penulis sehingga skripsi dengan judul **“Pengelompokan Pelanggan dan Rekomendasi Paket Produk Berdasarkan Pola Pembelian Konsumen Menggunakan *K-Prototype* dan *FP-Growth*”** dapat terselesaikan dengan baik.

Dalam proses penyusunan skripsi ini, tidak lepas dari tantangan, namun berkat bantuan moril, spiritual, dan materil, semua hambatan berhasil diatasi. Oleh karena itu, dengan penuh rasa hormat dan terima kasih, penyusun ingin menyampaikan apresiasi kepada:

1. Prof. Dr. Ir. Akhmad Fauzi, M.MT., IPU, selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Bapak Dr. Eng. Ir. Dwi Arman Prasetya., ST., MT., IPU. selaku Ketua Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur sekaligus dosen pembimbing skripsi pertama yang telah meluangkan waktu serta memberikan bimbingan selama pengerjaan skripsi.
4. Bapak Trimono, S.Si., M.Si. selaku dosen pembimbing skripsi kedua yang selalu memberikan bimbingan, arahan, dan motivasi kepada penulis.
5. Bapak Dr. Ir. I Gede Susrama Mas Diyasa, ST., MT., selaku dosen wali yang telah mengarahkan penyusunan penulisan.
6. Dosen-dosen Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
7. Kedua orang tua, kakak Dyndi dan seluruh keluarga besar yang terus memberikan semangat dan dorongan positif pada penulis.
8. Rahma, Rangga, Calvin yang juga turut memberikan dukungan intelektual, emosional, dan teknis yang tak terukur pada penulis.
9. Teman-teman Program Studi Sains Data yang juga turut memberikan motivasi selama penyusunan penulisan pada penulis.
10. Pihak-pihak lain yang tidak dapat penulis tuliskan satu per satu

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, 4 Juni 2025

Penulis

DAFTAR ISI

HALAMAN SAMPUL	i
LEMBAR PENGESAHAN	ii
LEMBAR PERSETUJUAN	iii
SURAT PERNYATAAN BEBAS PLAGIASI	iv
ABSTRAK.....	v
<i>ABSTRACT</i>	vii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvii
DAFTAR LAMPIRAN	xix
DAFTAR NOTASI	xxi
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah.....	6
1.3. Batasan Masalah	6
1.4. Tujuan Penelitian	6
1.5. Manfaat Penelitian	7
BAB II TINJAUAN PUSTAKA	9
2.1. Penelitian Terdahulu	9
2.2. Dasar Teori	18
2.1.1. UMKM	18
2.1.2. Pola Pembelian Konsumen.....	20
2.1.3 <i>Data Mining</i>	20
2.1.4. Klasterisasi	22
2.1.5. <i>Similarity Measure</i>	24
2.1.6. Metode <i>Elbow</i>	25
2.1.7. Algoritma <i>K-Prototype</i>	27
2.1.10. <i>Association Rules</i>	29

2.1.11. Analisis Keranjang Pasar	33
2.1.12. Algoritma <i>FP-Growth</i>	35
BAB III DESAIN DAN IMPLEMENTASI SISTEM	37
3.1. Variabel Penelitian dan Sumber Data	37
3.2. Langkah Analisis	38
3.2.1. Pengumpulan Data	40
3.2.2. <i>Pre-Processing</i> Data	40
3.2.3. Pengelompokan Pelanggan	41
3.2.4. Rekomendasi Menu	44
3.3. Desain Sistem	46
BAB IV HASIL DAN PEMBAHASAN	49
4.1. Analisis Deskriptif	49
4.1.1. Persiapan <i>Dataset</i>	49
4.1.2. Data <i>Pre-Processing</i>	58
4.2. Pengelompokan Pelanggan	61
4.2.1. <i>Feature Engineering</i>	61
4.2.2. Menghitung Jarak	64
4.2.3. Menghitung Jumlah Nilai Kluster Optimal	67
4.2.4. Pemodelan klasterisasi <i>K-Prototype</i>	69
4.2.5. Evaluasi Model	73
4.3. Rekomendasi Menu	86
4.3.1. <i>Feature Engineering</i>	86
4.3.2. Pemodelan Rekomendasi Menu	89
4.3.3. Evaluasi Model	92
4.3.4. Hasil <i>Output</i> Pemodelan	93
4.4. Implementasi <i>User Interface</i> (UI) Berbasis Streamlit	95
4.4.1. Tampilan Fitur “ <i>Home</i> ”	95
4.4.3. Rekomendasi Menu	97
BAB V PENUTUP	101
5.1. Kesimpulan	101
5.2. Saran Pengembangan	102

DAFTAR PUSTAKA	103
LAMPIRAN	109

Halaman ini sengaja dikosongkan

DAFTAR GAMBAR

Gambar 2.1 Tahapan <i>Data Mining</i>	21
Gambar 2.2 Ilustrasi Data Dikelompokkan Menjadi Klaster.....	24
Gambar 2.3 Contoh Grafik Metode <i>Elbow</i>	26
Gambar 4.1 Distribusi Total Penjualan Dari Data Mentah.....	52
Gambar 4.2 Distribusi Jenis Order Dari Data Mentah.....	54
Gambar 4.3 Distribusi Metode Pembayaran Dari Data Mentah	55
Gambar 4.4 Distribusi Order Perjam Dari Data Mentah	57
Gambar 4.5 Distribusi 10 Produk Terpopuler Dari Data Mentah.....	58
Gambar 4.6 Klaster Optimal Dengan <i>Elbow</i>	69
Gambar 4.7 Distribusi Lama Pembayaran	77
Gambar 4.8 Distribusi Jumlah Produk.....	79
Gambar 4.9 Distribusi Jenis Order.....	81
Gambar 4.10 Distribusi Metode Pembayaran.....	82
Gambar 4.11 Distribusi Kategori Hari.....	83
Gambar 4.12 Distribusi Kategori Waktu	84
Gambar 4.13 Distribusi Bulan	85
Gambar 4.14 Tampilan Fitur “ <i>Home</i> ”.....	95
Gambar 4.15 Memilih Fitur Pengelompokan Pelanggan.....	96
Gambar 4.16 Tampilan Tahap <i>Preprocessing</i> Pengelompokan Pelanggan.....	96
Gambar 4.17 Tampilan Metode <i>Elbow</i>	97
Gambar 4.18 Tampilan Pemodelan <i>K-Prototype</i>	97
Gambar 4.19 Tampilan Fitur Rekomendasi Menu.....	98
Gambar 4.20 Tampilan Tahap <i>Preprocessing</i> Rekomendasi Menu	98
Gambar 4.21 Tampilan Pemodelan <i>FP-Growth</i>	99
Gambar 4.22 Tampilan Rekomendasi Menu	99

Halaman ini sengaja dikosongkan

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	9
Tabel 2.2 Pengukuran <i>Silhouette Coefficient</i>	29
Tabel 3.1 Variabel Dataset Penelitian.....	37
Tabel 4.1 Dataset Mentah	49
Tabel 4.2 Statistika Deskriptif Total Penjualan	51
Tabel 4.3 Data Yang Hilang	59
Tabel 4.4 Tipe Data	60
Tabel 4.5 Data Yang Duplikat	61
Tabel 4.6 Sebelum dan Sesudah Standarisasi Data	63
Tabel 4.7 <i>Dataset</i> Pemodelan	64
Tabel 4.8 Nilai <i>Euclidean Distances</i>	65
Tabel 4.9 Nilai <i>Hamming Distances</i>	66
Tabel 4.10 Total Jumlah Tiap Klaster.....	72
Tabel 4.11 Nilai <i>Centroid</i> Tiap Klaster	72
Tabel 4.12 Hasil Evaluasi <i>Silhouette Score</i>	75
Tabel 4.13 <i>Dataset</i> Pemodelan <i>FP-Growth</i>	88
Tabel 4.14 Perbandingan <i>Threshold</i> Aturan Asosiasi.....	89
Tabel 4.15 Hasil Nilai <i>Support</i> dan <i>Confidence</i>	91
Tabel 4.16 Hasil Nilai <i>Lift</i>	93
Tabel 4.17 Rekomendasi Paket Menu	94

Halaman ini sengaja dikosongkan

DAFTAR LAMPIRAN

Lampiran 1. LoA Jurnal.....	109
Lampiran 2. Kode Program UI	110
Lampiran 3. Kode Program <i>Machine Learning</i>	110

Halaman ini sengaja dikosongkan

DAFTAR NOTASI

Similarity Measure

Euclidean Distance

$d_{i,j}$	:	Perbedaan antara objek i dan objek
x_{il}	:	Merujuk pada besarnya objek i dalam variabel ke- l
x_{jl}	:	Merujuk pada besarnya objek j dalam variabel ke- l
p	:	Merepresentasikan jumlah variabel
$d(x,y)$	:	Jarak
X	:	Vektor koordinat lokasi pertama, misal $(x_1, x_2, \dots, x_n)$
Y	:	Vektor koordinat lokasi kedua, misal $(y_1, y_2, \dots, y_n)$
x_i	:	Objek pengamatan ke- i
$\bar{x}$	:	Rata-rata variabel
Σ^{-1}	:	Matriks kovarian dari variabel
$d(i, j)$	:	Jarak antara objek ke- i dengan objek ke- j (<i>mixed variable</i>)
$\sum_{r=1}^p (x_{ir} - x_{jr})^2$	:	Ukuran jarak untuk variabel data bertipe numerik
X_{ir}	:	Nilai objek ke- i pada variabel ke- r
X_{jr}	:	Nilai objek ke- j pada variabel ke- r
P	:	Jumlah variabel dengan tipe data numerik

Hamming Distance

$H_1(x, y)$	:	Total nilai <i>Hamming distance</i> berbobot antara dua objek x dan y .
$\sum_i^n = 1$	:	Simbol penjumlahan dari $i = 1$ hingga n , yaitu menjumlahkan semua perhitungan untuk setiap atribut
n	:	Jumlah total atribut atau fitur dari vektor x dan y
w_i	:	Bobot untuk atribut ke- i , yang menunjukkan seberapa penting atribut tersebut dalam perhitungan jarak
$H(x_i, y_i)$	:	Fungsi <i>Hamming distance</i> $\begin{cases} 0, & \text{jika } x_i = y_i \\ 1, & \text{jika } x_i \neq y_i \end{cases}$
(x_i, y_i)	:	Nilai dari atribut ke- i pada vektor x dan y .

Metode Elbow

x_i	:	Nilai dari atribut pada data ke- i
u_k	:	Nilai atribut dari pusat titik klaster i

Silhouette Coefficient

$ A $	:	Jumlah objek dalam klaster A
i, j	:	Indeks yang berkaitan dengan objek
$d_{i,j}$	:	Perbedaan antara objek ke- i dan objek ke- j
$d_{i,A}'$	:	Jarak rata-rata antara objek i dan semua objek yang ada di klaster lain
A'	:	Jumlah objek dalam klaster yang berbeda dari A
$a(i)$	:	Rata-rata jarak antara objek i dan seluruh objek lain dalam klaster yang sama
$b(i)$	:	Rata-rata jarak antara objek i dan objek-objek dalam kluster terdekat yang berbeda
$S(i)$	:	<i>Silhouette Coefficient</i> untuk objek i , dengan nilai yang bervariasi antara -1 hingga 1.

Support

$P(A \cap B)$	:	Kemungkinan transaksi yang mengandung item A dan item B secara bersamaan
---------------	---	--

Confidence

$P(A B)$:	:	Kemungkinan <i>item</i> B dibeli jika item A telah dibeli
$Support(A \cap B)$	:	Kemungkinan transaksi yang mengandung A dan B
$P(A)$	:	Kemungkinan <i>item</i> A muncul dalam transaksi

Lift

<i>Confidence</i>	:	Nilai <i>confidence</i> dari aturan $A \rightarrow B$
<i>Expected Confidence</i>	:	Kemungkinan <i>item</i> B muncul