

SKRIPSI
ANALISIS PRODUK E-COMMERCE UNTUK PENGELOMPOKAN
DATA MENGGUNAKAN K-MEANS CLUSTERING

UTILIZING CLUSTERING METHODS FOR CATEGORIZING
DELIVERY REQUIREMENTS BASED ON ANALYSIS OF E-COMMERCE
PRODUCT DATA

Diajukan untuk memenuhi salah satu syarat
memperoleh Gelar Sarjana Komputer



Disusun Oleh:

Jumat

311710240

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS PELITA BANGSA
BEKASI
2024

LEMBAR PERSETUJUAN SKRIPSI

**UTILIZING CLUSTERING METHODS FOR CATEGORIZING DELIVERY
REQUIREMENTS BASED ON ANALYSIS OF E-COMMERCE PRODUCT DATA**

Disusun oleh:

Jum'at

311710240

Telah diperiksa dan disahkan
pada tanggal : 14 Januari 2024

Dosen Pembimbing 1



**Suprpto, S.T., M.Kom.
NIDN.0414078104**

Dosen Pembimbing 2



**Muhamad Fatchan, S.Kom., M.Kom.
NIDN.0403107605**

Mengetahui :

Ketua Program Studi Teknik Informatika



**Wahyu Hadikristanto, S.Kom., M.Kom.
NIDN.0415088207**

LEMBAR PENGESAHAN

**UTILIZING CLUSTERING METHODS FOR CATEGORIZING DELIVERY
REQUIREMENTS BASED ON ANALYSIS OF E-COMMERCE PRODUCT DATA**

Disusun oleh:

Jum'at

311710240

Telah dipertahankan pada tanggal : 14 Januari 2024

Dosen Pembimbing 1



**Suprpto, S.T., M.Kom.
NIDN.0414078104**

Dosen Pembimbing 2



**Muhamad Fatchan, S.Kom., M.Kom.
NIDN.0403107605**

Mengetahui :

Ketua Program Studi Teknik Informatika



**Wahyu Hadikristanto, S.Kom., M.Kom.
NIDN.0415088207**

Dekan Fakultas Teknik



**Putri Anggun Sari, S.Pt., M.Si.
NIDN.0424088403**

PERNYATAAN KEASLIAN SKRIPSI

Sebagai mahasiswa Universitas Pelita Bangsa, yang bertanda-tangan dibawah ini, saya

Nama : Jumat

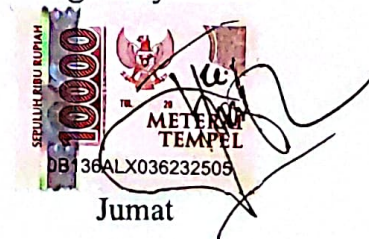
NIM : 311710240

Menyatakan bahwa karya ilmiah saya yang berjudul : “Analisis Produk E-Commerce Untuk Pengelompokan Data Menggunakan K-Means Clustering” . Merupakan karya asli saya (kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya dan perangkat pendukung seperti web cam dll). Apabila di kemudian hari, karya saya disinyalir bukan merupakan karya asli saya, yang disertai dengan bukti-bukti yang cukup, maka saya bersedia untuk dibatalkan gelar saya beserta hak dan kewajiban yang melekat pada gelar tersebut. Demikian surat pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Kabupaten Bekasi

Pada tanggal : 27 November 2023

Yang Menyatakan



SEPULEH RIBU RUPIAH
1000
METERAN
TEMPEL
DB136ALX036232505

Jumat

**PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Sebagai mahasiswa Universitas Pelita Bangsa, yang bertanda-tangan dibawah ini,

Nama : Jumat

NIM : 311710240

Demi mengembangkan Ilmu Pengetahuan, menyetujui untuk memberikan kepada Universitas Pelita Bangsa Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah yang berjudul :

**“ANALISIS PRODUK E-COMMERCE UNTUK PENGELOMPOKAN
DATA MENGGUNAKAN K-MEANS CLUSTERING ”**

Beserta perangkat yang di perlukan (bila ada), dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Pelita Bangsa berhak untuk menyimpan, mengcopy ulang (memperbanyak), menggunakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan / mempublikasikannya di internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis /pencipta.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Universitas Pelita Bangsa, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini. Demikian surat pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Bekasi,

Pada Tanggal : 27 November 2023

Yang Menyatakan,


Jumat

UCAPAN TERIMAKASIH

Dengan memanjatkan puji syukur kehadirat Allah Subhanahu Wata'ala, Tuhan Yang Maha Pengasih dan Maha Penyayang yang telah melimpahkan segala rahmat, hidayat dan inayah-Nya kepada Penulis, sehingga dapat menyelesaikan laporan tugas akhir skripsi yang berjudul "**ANALISIS PRODUK E-COMMERCE UNTUK PENGELOMPOKAN DATA MENGGUNAKAN K-MEANS CLUSTERING**" laporan ini disusun untuk memenuhi salah satu syarat guna menyelesaikan pendidikan S1 (Strata Satu) Program Studi Teknik Informatika di Universitas Pelita Bangsa Cikarang.

Pada kesempatan kali ini, penulis ingin menyampaikan rasa terima kasih kepada pihak terkait. Karena penulis menyadari tanpa adanya uliran dan bantuan mereka belum tentu laporan ini dapat diselesaikan, pihak-pihak tersebut adalah :

1. Bapak Hamzah Muhammad Mardiputra, S.K.M., MM sebagai Ketua Universitas Pelita Bangsa
2. Ibu Putri Anggun Sari, S.Pt., M.Si sebagai Dekan Teknik Universitas Pelita Bangsa
3. Bapak Wahyu Hadikristanto, S. Kom, M.Kom sebagai Kaprodi Teknik Informatika - S1
4. Bapak Suprpto, S.T., M.Kom selaku Dosen Pembimbing I dan Bapak M. Fatchan, S.Kom., M.Kom, MTCNA selaku Dosen Pembimbing II yang memberikan ide penelitian, serta informasi referensi yang penulis butuhkan.
5. Orang Tua, Kerabat dan Keluarga yang telah memberi motivasi dan dukungan sehingga kami dapat menyelesaikan studi di Universitas Pelita Bangsa.
6. Seluruh Seluruh sahabat dan rekan-rekan yang telah banyak membantu dan atas dukungannya sehingga kami dapat menyelesaikan studi di Universitas Pelita Bangsa.

Semoga Allah Subhanahu Wata'ala memberikan balasan yang lebih besar kepada beliau-beliau.

Penulis menyadari bahwa penulisan skripsi ini masih jauh dari kata sempurna, maka dari itu bahwasanya penulis memohon saran positif yang bersifat membangun untuk kemajuan penulis di masa mendatang. Akhir kata semoga laporan tugas akhir skripsi ini dapat dipergunakan sebagaimana mestinya bagi penulis serta berguna bagi pembaca pada umumnya, Aamiin Ya Robbal ‘Alamin.

Bekasi, 27 November 2023

Jumat

ABSTRAK

Model algoritma K-Means yang diterapkan memiliki hasil yang menunjukkan sebuah wawasan baru, yaitu pengelompokan tingkat risiko dalam proses pengiriman 2 klaster, klaster 1 (C0) merupakan tingkat risiko tinggi yang terdiri dari 53 data dari 360 dataset yang diuji dan klaster 2 (C1) adalah merupakan tingkat risiko rendah yang terdiri dari 307 data dari 360 dataset yang diuji. Pengujian menggunakan aplikasi RapidMiner Studio juga dapat menghasilkan wawasan yang serupa yaitu masing-masing klaster memiliki anggota kelompok yang terbagi menjadi 2 klaster, yaitu C0 dengan anggota kelompok klaster sebanyak 53 data, dan C1 dengan anggota klaster sebanyak 307 data. Masing-masing klaster memiliki nilai centroid optimal yakni 131,717 & 385,075 untuk C0 serta 119,932 & 111,414 untuk C1, dengan nilai evaluasi Davies-Bouldin Index sebesar 0,626.

Kata Kunci : Data Mining, *K-Means*, *Klaster*, *E-Commerce*, *Produk*

ABSTRACT

The K-Means algorithm model applied has results that show a new insight, namely the grouping of risk levels in the delivery process into 2 clusters, cluster 1 (C0) is a high risk level consisting of 53 data from the 360 datasets tested and cluster 2 (C1) is a low risk level consisting of 307 data from the 360 datasets tested. Testing using the RapidMiner Studio application can also produce similar insights, namely that each cluster has group members who are divided into 2 clusters, namely C0 with 53 data cluster group members, and C1 with 307 data cluster members. Each cluster has an optimal centroid value, namely 131,717 & 385,075 for C0 and 119,932 & 111,414 for C1, with a Davies-Bouldin Index evaluation value of 0.626.

Keywords: Data Mining, K-Means, Klastrer, E-Commerce, Products

DAFTAR ISI

LEMBAR PERSETUJUAN SKRIPSI.....	i
LEMBAR PENGESAHAN	ii
PERNYATAAN KEASLIAN SKRIPSI.....	iii
PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	iv
UCAPAN TERIMAKASIH.....	v
ABSTRAK.....	vii
ABSTRACT.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xiii
BAB I PENDAHULUAN.....	1
1.2 Identifikasi Masalah.....	2
1.3 Rumusan Masalah.....	3
1.4 Batasan Masalah	3
1.5 Tujuan Penelitian	3
1.6 Manfaat	4
1.6.1 Penulis.....	4
1.6.2 Teknik Informatika	4
1.6.3 Umum	4
BAB II TINJUAN PUSTAKA DAN LANDASAN TEORI	6
2.1 Penelitian Terdahulu	6
2.2 Landasan Teori.....	8
2.2.1 Pengelompokan Produk	8
2.2.2 Data Mining	8
2.2.3 Fungsi Data Mining	10
2.2.4 Proses Tahapan Data mining.....	11
2.2.5 Algoritma K-Means	13
BAB III METODOLOGI PENELITIAN	16
3.1 Objek Penelitian.....	16
3.2 Tahapan Penelitian.....	16
3.3 Pengumpulan Data.....	16

3.4	Pemodelan.....	19
3.5	Interpretasi dan Evaluasi.....	20
3.6	Instrumen Penelitian	20
3.6.1	Kebutuhan Perangkat Lunak (<i>Software</i>)	20
3.6.2	Kebutuhan Perangkat Keras (<i>Hardware</i>).....	21
3.7	Waktu Penelitian	21
BAB IV HASIL PENGUJIAN DAN PEMBAHASAN		22
4.1	Penerapan Algoritma K-Mean	22
4.1.1	Analisis Data	22
4.1.2	Tahapan dan Contoh Perhitungan Algoritma K-Means	23
4.2	Pembentukan Kelompok Klaster	59
4.3	Hasil Pengujian	60
4.3.1	Evaluasi Pengujian pada Aplikasi RapidMiner Studio	60
4.3.2	Analisa Hasil Pengujian	61
BAB V KESIMPULAN DAN SARAN.....		64
5.1	Kesimpulan	64
5.2	Saran	64
DAFTAR PUSTAKA		66

DAFTAR TABEL

Tabel 3. 1 Data Selection	17
Tabel 3. 2 Data Selection	18
Tabel 3. 3 Dataset Klasterisasi K-Means	18
Tabel 3. 4 Perangkat Lunak (Software)	20
Tabel 3. 5 Perangkat Keras (Hardware).....	21
Tabel 3. 6 Waktu Penelitian.....	21
Tabel 4. 1 Dataset Barang	22
Tabel 4. 2 Euclidean Distance Iterasi 1.....	25
Tabel 4. 3 Matriks Kelompok Data Iterasi ke-1.....	26
Tabel 4. 4 Titik Pusat Awal Klaster Baru ke-2	27
Tabel 4. 5 <i>Euclidean Distance</i> Iterasi 2	29
Tabel 4. 6 Matriks Kelompok Data Iterasi ke-2.....	30
Tabel 4. 7 Titik Pusat Awal Klaster Baru ke-3	31
Tabel 4. 8 <i>Euclidean Distance</i> Iterasi 3	33
Tabel 4. 9 Matriks Kelompok Data Iterasi ke-3.....	34
Tabel 4. 10 Titik Pusat Awal Klaster Baru ke-4	35
Tabel 4. 11 <i>Euclidean Distance</i> Iterasi 4	37
Tabel 4. 12 Matriks Kelompok Data Iterasi ke-4.....	38
Tabel 4. 13 Titik Pusat Awal Klaster Baru ke-5	39
Tabel 4. 14 <i>Euclidean Distance</i> Iterasi 5	41
Tabel 4. 15 Matriks Kelompok Data Iterasi ke-5.....	42
Tabel 4. 16 Titik Pusat Awal Klaster Baru ke-6	43
Tabel 4. 17 <i>Euclidean Distance</i> Iterasi 6	45
Tabel 4. 18 Matriks Kelompok Data Iterasi ke-6.....	46
Tabel 4. 19 Titik Pusat Awal Klaster Baru ke-7	47
Tabel 4. 20 <i>Euclidean Distance</i> Iterasi 7	49
Tabel 4. 21 Matriks Kelompok Data Iterasi ke-7.....	50
Tabel 4. 22 Titik Pusat Awal Klaster Baru ke-8	51
Tabel 4. 23 <i>Euclidean Distance</i> Iterasi 8	53

Tabel 4. 24 Matriks Kelompok Data Iterasi ke-8.....	54
Tabel 4. 25 Titik Pusat Awal Klaster Baru ke-9	55
Tabel 4. 26 <i>Euclidean Distance</i> Iterasi 9	57
Tabel 4. 27 Matriks Kelompok Data Iterasi ke-9.....	58
Tabel 4. 28 Perbandingan Jumlah Objek Data dari Iterasi ke 7 & ke 8.....	59
Tabel 4. 29 Kelompok Klaster Data.....	60

DAFTAR GAMBAR

Gambar 2. 1 Tahapan proses KDD	12
Gambar 3. 1 Tahap Penelitian.....	16
Gambar 4. 1 Proses Rapid Miner Klasterisasi K-Means.....	61
Gambar 4. 2 Hasil Cluster Model pada aplikasi RapidMiner Studio.....	62
Gambar 4. 3 Hasil Optimal Cluster pada aplikasi RapidMiner Studio	62
Gambar 4. 4 Hasil Davies-Bouldin Index pada aplikasi RapidMiner Studi	63