



**ANALISIS PENGELOMPOKKAN KEBIASAAN PELANGGAN PADA
TOKO *ONLINE* MENGGUNAKAN *K-MEANS* DENGAN METODE
JARAK *CHEBYSHEV* DAN *CANBERRA* SERTA EVALUASI *DAVIES-
BOULDIN INDEX***

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer

(S.Kom)

Disusun Oleh:

MUHAMMAD ZHILLAN ZHALILA

21107011

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS SAINS TEKNIK DAN DESAIN

UNIVERSITAS TRILOGI

JAKARTA

2026



LEMBAR PENGESAHAN SKRIPSI

Nama	:	Muhammad Zhillan Zhalila
NIM	:	21107011
Jenjang Pendidikan	:	Strata-1
Program Studi	:	Teknik Informatika
Judul	:	Analisis Pengelompokkan Kebiasaan Pelanggan Pada Toko Online Menggunakan <i>K-Means</i> Dengan Metode Jarak <i>Chebyshev</i> Dan <i>Canberra</i> Serta Evaluasi <i>Davies-Bouldin Index</i> .
Tanggal Sidang	:	9 Februari 2026

Disetujui Oleh :

(Gatot Tri Pranoto, S.Kom., M.Kom)

Dosen Pembimbing I

(Dr. Ketut Bayu Yogha Bintoro. Skom.,

M.CS)

Dosen Pembimbing II

Diketahui Oleh:

(Ir. Yaddarabullah, PhD, ASEAN Eng)

Ketua Program Studi Teknik Informatika



LEMBAR PENGESAHAN UJIAN SIDANG SKRIPSI/TUGAS AKHIR

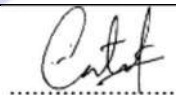
Pada hari Senin, 9 Februari 2026 telah diselenggarakan Ujian Sidang Tugas Akhir dengan judul: “ANALISIS PENGELOMPOKKAN KEBIASAAN PELANGGAN PADA TOKO *ONLINE* MENGGUNAKAN *K-MEANS* DENGAN METODE JARAK *CHEBYSHEV* DAN *CANBERRA* SERTA EVALUASI *DAVIES-BOULDIN INDEX*.” untuk memenuhi sebagai prasyarat akademik guna memperoleh gelar Sarjana/~~Magister~~*) Teknik Informatika Universitas Trilogi, bagi mahasiswa:

Nama : Muhammad Zhillan Zhalila

NIM : 21107011

Program Studi : Teknik Informatika

telah dinyatakan Lulus/~~Tidak Lulus~~*) oleh Dewan Penguji yang terdiri dari:

<u>Nama</u>	<u>Status</u>	<u>Tanda tangan</u>
Assoc. Prof. Silvester Dian Handy Permana, ST., MTI, Ph.D.	- Penguji I	
Assoc.Prof. Budi Arifitama, S.T., M.M.S.I.,Ph.D.	- Penguji II	
Gatot Tri Pranoto, S.Kom., M.Kom.	- Dosen Pembimbing I	
Dr. Ketut Bayu Yogha Bintoro, S.Kom., M.Cs.	- Dosen Pembimbing II	



SURAT PERNYATAAN KEASLIAN SKRIPSI

Saya, yang bertandatangan di bawah ini, menyatakan bahwa tugas akhir yang saya serahkan untuk maju ujian sidang adalah murni hasil karya saya sendiri, bukan merupakan hasil plagiarsime (penjiplakan).

Apabila pada saat ujian sidang, diketahui dan terbukti bahwa tugas akhir ini merupakan hasil plagiarisme (penjiplakan) baik sebagian atau seluruh isi tugas akhir saya, maka Universitas Trilogi mempunyai hak dan wewenang untuk membatalkan isi dari tugas akhir yang telah saya (sebagai peneliti) tulis dan saya bersedia menerima sanksi Drop Out atau tidak dapat melanjutkan studi dari Universitas Trilogi.

Demikianlah surat pernyataan ini saya buat dalam keadaan sehat dan sadar tanpa adanya campur tangan ataupun tekanan dari pihak manapun juga.

Jakarta, 6 Februari 2026

Yang Menyatakan,



(**Muhammad Zhillan Zhalila**)

NIM : 2110701

ABSTRAK

Perkembangan teknologi digital dan *e-commerce* meningkatkan aktivitas transaksi daring yang menghasilkan *big data* perilaku pelanggan. Data transaksi tersebut menyimpan informasi penting mengenai kebiasaan belanja pelanggan yang dapat dimanfaatkan untuk melakukan segmentasi dan mendukung penyusunan strategi bisnis yang lebih tepat sasaran. Namun, pemanfaatan data pelanggan pada toko daring masih belum optimal sehingga diperlukan analisis berbasis data (*data-driven analysis*) untuk memahami karakteristik pelanggan secara lebih sistematis. Penelitian ini menggunakan pendekatan *data mining* dengan teknik *clustering* untuk mengelompokkan pelanggan berdasarkan metode *Recency, Frequency, Monetary (RFM)*. Proses pengelompokan dilakukan menggunakan algoritma K-Means dengan perbandingan dua metode jarak, yaitu *Chebyshev* dan *Canberra*, guna memperoleh metode dengan kinerja paling optimal. *Dataset* yang digunakan berjumlah 541.908 transaksi dari *Kaggle* dan *UCI Machine Learning Repository*. Untuk meningkatkan efisiensi komputasi, data direduksi menggunakan teknik *sampling* menjadi 4.339 data yang tetap merepresentasikan pola perilaku pelanggan. Evaluasi kinerja dilakukan menggunakan waktu eksekusi dan *Davies-Bouldin Index (DBI)*. Hasil penelitian menunjukkan terbentuk tiga *cluster*, yaitu *Low Value Customer* sebanyak 790 pelanggan, *Medium Value Customer* sebanyak 1.150 pelanggan, dan *High Value Customer* sebanyak 2.350 pelanggan. Metode *Canberra* lebih efisien dalam waktu eksekusi, sedangkan *Chebyshev* menghasilkan kualitas *cluster* yang lebih baik berdasarkan nilai DBI yang lebih rendah. Hasil segmentasi ini dapat dimanfaatkan sebagai dasar pengambilan keputusan bisnis, seperti penentuan promosi, program loyalitas, strategi retensi pelanggan, serta perencanaan penjualan secara lebih efektif.

Kata Kunci: Perilaku Pelanggan, *RFM (Recency, Frequency, Monetary)*, *Clustering*, *K-Means*, *Chebyshev*, *Canberra*, *Davies-Bouldin Index*.

KATA PENGANTAR

Segala puji dan syukur atas kehadiran Tuhan Yang Maha Esa yang telah memberikan rahmatnya dan juga nikmat akal serta pikiran yang saya gunakan dalam penyusunan proposal ini dengan judul “Analisis Pengelompokan Kebiasaan Pelanggan Pada Toko *Online* Menggunakan *K-Means* Dengan Metode Jarak *Chebyshev* Dan *Canberra* Serta Evaluasi *Davies-Bouldin Index*.”. Penulis mengucapkan terima kasih sebesar-besarnya atas pihak yang telah membantu, mendukung dan mendoakan agar cepat terselesaikannya proposal ini, kepada:

1. Bapak Ir. Yaddarabullah, Phd., IPM., ASEAN.Eng selaku Ketua Program Studi Teknik Informatika Universitas Trilogi, terima kasih atas usaha dan bantuannya dalam membantu mahasiswa untuk aktif dan kreatif dalam bidang akademik maupun non akademik.
2. Bapak Gatot Tri Pranoto, S.Kom., M.Kom, selaku dosen dari Program Studi Sistem Informasi sekaligus dosen pembimbing 1 yang telah memberikan, informasi, motivasi dan masukan dalam pembuatan proposal ini agar hasil dan harapan dari penelitian ini lebih terarah dan bermanfaat.
3. Bapak Dr. Ketut Bayu Yogha Bintoro, S.Kom., M.Cs. selaku dosen dari Program Studi Teknik Informatika sekaligus dosen pembimbing 2 saya dalam memberikan arahan dan masukan dalam proses penulisan proposal ini agar penggunaan format kata maupun bahasa lebih baik dan benar.
4. Kepada kedua orang tua serta seluruh anggota keluarga, saya mengucapkan terima kasih atas doa, dukungan, dan semangat yang telah diberikan sehingga saya dapat menyelesaikan penyusunan proposal ini dengan penuh semangat.
5. Seluruh teman-teman kampus yang selalu menemani, memberikan motivasi dan semangat, sehingga penulis dapat menyelesaikan proposal ini sampai selesai.
6. Seluruh sahabat saya yang selalu ada ketika saya membutuhkan semangat, dan membutuhkan tempat untuk bercerita, ataupun butuh sesuatu yang membuat saya beristirahat dari diri saya sendiri.

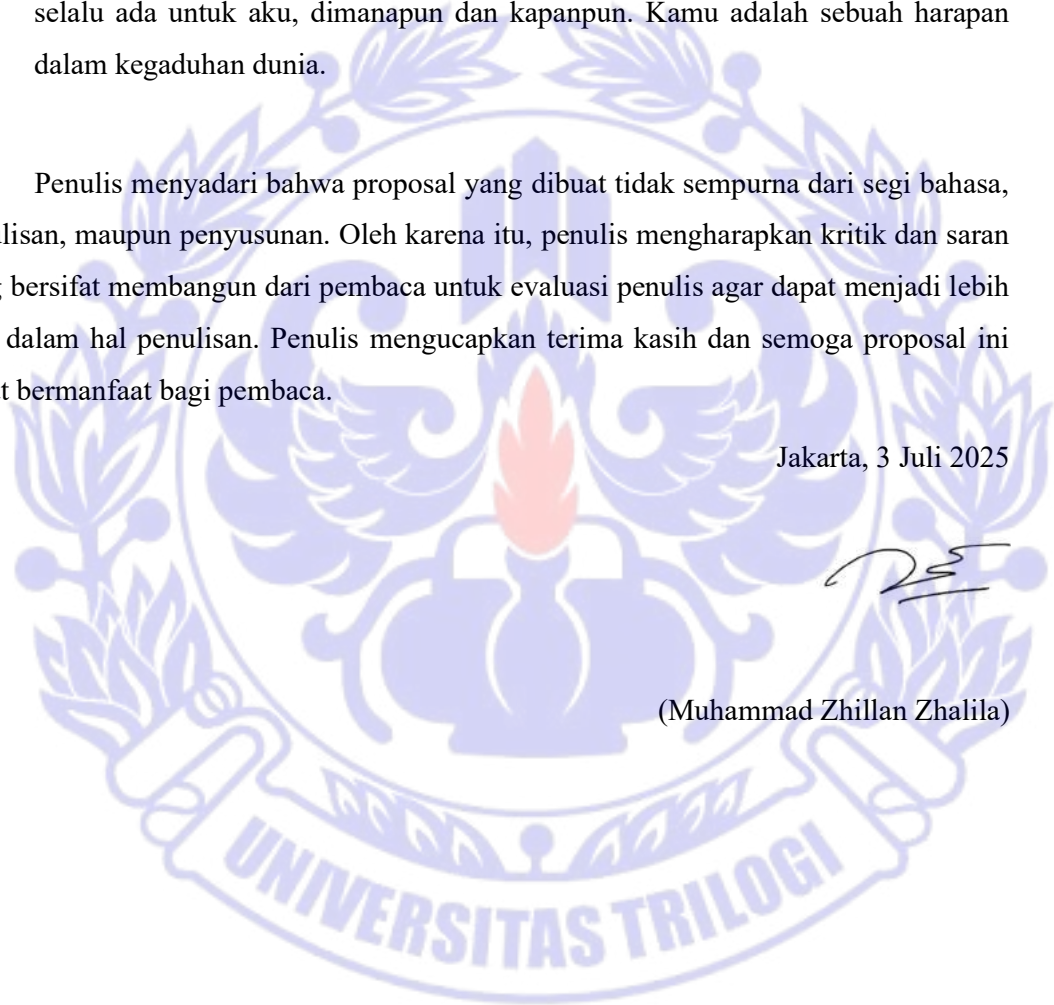
7. Kepada Tarissa Salma Andira, meskipun kita kini berpisah, kau tetap sebuah cerita yang akan selalu kubawa. Terima kasih, sedalam-dalamnya, untuk segalanya, pelajaran, momen, dan jejak yang kau tinggalkan dalam hidupku, kau adalah sesuatu sebelum semuanya hilang.
8. Kepada Sarah Corniva, terima kasih atas segala hal baik yang menyertaimu, kamu selalu ada untuk aku, dimanapun dan kapanpun. Kamu adalah sebuah harapan dalam kekaduan dunia.

Penulis menyadari bahwa proposal yang dibuat tidak sempurna dari segi bahasa, penulisan, maupun penyusunan. Oleh karena itu, penulis mengharapkan kritik dan saran yang bersifat membangun dari pembaca untuk evaluasi penulis agar dapat menjadi lebih baik dalam hal penulisan. Penulis mengucapkan terima kasih dan semoga proposal ini dapat bermanfaat bagi pembaca.

Jakarta, 3 Juli 2025



(Muhammad Zhillan Zhalila)



DAFTAR ISI

HALAMAN JUDUL.....	i
LEMBAR PENGESAHAN SKRIPSI.....	ii
LEMBAR PENGESAHAN UJIAN SIDANG SKRIPSI/TUGAS AKHIR	iii
SURAT PERNYATAAN KEASLIAN SKRIPSI.....	iv
SURAT PERSETUJUAN UJIAN SIDANG TUGAS AKHIR.....	v
ABSTRAK	vi
KATA PENGANTAR.....	vii
DAFTAR ISI	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xiii
DAFTAR PERSAMAAN.....	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
1.6 Metode Penelitian	6
1.7 Sistematika Penulisan	7
BAB II TINJAUAN PUSTAKA	10
BAB III METODOLOGI PENELITIAN.....	23
3.1 Pengumpulan Data.....	24
3.1.1 Sumber Data	25
3.1.2 Format Data	27
3.1.3 Attribute Data	27
3.1.4 Tujuan Pengumpulan Data	28
3.2 Preprocessing Data	29
3.2.1 Pembersihan Data Awal.....	29
3.2.2 Transformasi Data Mentah menjadi RFM.....	29
3.2.3 Sampel Data.....	32

3.2.4	Validasi Sampel	33
3.3	K-Means Clustering.....	34
3.3.1	Algoritma K-Means.....	34
3.3.2	Metode Perhitungan Jarak	38
3.4	Evaluasi Davies Bouldin-Index	40
3.5	Interpretasi Hasil.....	42
BAB IV HASIL DAN PEMBAHASAN.....		43
4.1	Gambaran Umum <i>Dataset</i>	43
4.2	Hasil Processing Data	44
4.2.1	Pembersihan Data (Data Cleaning)	45
4.2.2	Transformasi Data ke Model RFM.....	51
4.2.3	Skoring RFM.....	57
4.3	Implementasi K-Means.....	62
4.3.1	Analisa Perhitungan.....	62
4.3.2	K-Means dengan Chebyshev Distance.....	63
4.3.3	K-Means dengan Canberra Distance	66
4.3.4	Centroid yang dihasilkan.....	69
4.4	Pengujian Metode yang diusulkan.....	72
4.4.1	Persiapan Data dan Parameter	72
4.4.2	Proses Clustering dengan Metode Jarak Chebyshev	74
4.4.3	Proses Clustering dengan Metode Jarak Canberra	77
4.4.4	Perbandingan Hasil Kedua Metode	79
4.4.5	Evaluasi Kualitas Cluster Menggunakan Davies-Bouldin Index	80
4.4.6	Perbandingan Waktu Eksekusi Jarak	84
4.5	Analisis dan Pembahasan.....	85
4.5.1	Analisis Pengelompokan Pelanggan Berdasarkan RFM	85
4.5.2	Analisis Efisiensi Waktu Eksekusi.....	87
4.5.3	Pembahasan Perbandingan Metode Jarak <i>Chebyshev</i> dan <i>Canberra</i>	89
BAB V PENUTUP.....		91
5.1	Kesimpulan.....	91
5.2	Saran	91
DAFTAR PUSTAKA		93
LAMPIRAN.....		95