

**ANALISIS PENYEBAB KETERLAMBATAN *INVOICE*
MENGUNAKAN ALGORITMA *K-MEANS* DAN DBSCAN DI PT
SEBASTIAN JAYA METAL**

SKRIPSI

**Oleh:
Daffa Rayhan Fadhillah
202110715243**



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2025**

LEMBAR PERSETUJUAN PEMBIMBING

Judul Skripsi : Analisis Penyebab Keterlambatan
Invoice Menggunakan Algoritma K-
Means dan DBSCAN di PT Sebastian
Jaya Metal

Nama Mahasiswa : Daffa Rayhan Fadhillah

Nomor Pokok Mahasiswa : 202110715243

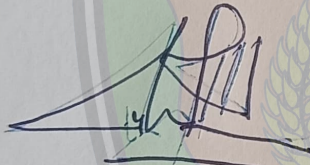
Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 29 Juli 2025

Jakarta, 05 Agustus 2025

MENYETUJUI,

Pembimbing I



Prof. Dr. Mujiono Sadikin, S.T., M.T.

NIDN : 0406127002

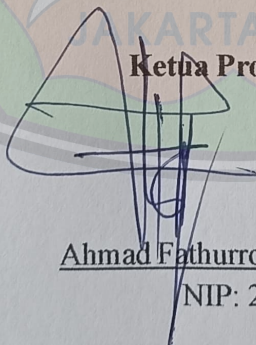
Pembimbing II



Sugiyatno, S.Kom., M.Kom.

NIDN : 0313077206

Ketua Program Studi



Ahmad Fathurrozi, S.E., M.M.S.I.

NIP: 2012486

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS BHAYANGKARA JAKARTA RAYA

2025

LEMBAR PENGESAHAN

Judul Skripsi : Analisis Penyebab Keterlambatan *Invoice*
Menggunakan Algoritma K-Means dan
DBSCAN di PT Sebastian Jaya Metal

Nama Mahasiswa : Daffa Rayhan Fadhillah

Nomor Pokok Mahasiswa : 202110715243

Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 29 Juli 2025

Jakarta, 05 Agustus 2025

MENGESAHKAN,

Ketua Tim Penguji : Adi Muhajirin, S.Kom., M.Kom., M.M
NIDN: 0318038501

Penguji I : Rafika Sari, S.Si., M.Si.
NIDN: 0329098902

Penguji II : Prof. Dr. Mujiono Sadikin, S.T., M.T.
NIDN: 0406127002

MENGETAHUI,

Ketua

Program Studi Informatika

Ahmad Fathurrozi, S.E., M.M.S.I

NIP: 2012486

Dekan

Fakultas Ilmu Komputer

Dr. Robertus Suraji, S.S., M.A

NIP: 1904406



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini:

Nama : Daffa Rayhan Fadhillah
NPM : 202110715243
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Analisis Penyebab Keterlambatan *Invoice* Menggunakan
Algoritma K-Means dan DBSCAN di PT Sebastian Jaya
Metal

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila di kemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Jakarta, 05 Agustus 2025

Penulis



Daffa Rayhan Fadhillah

ABSTRAK

Daffa Rayhan Fadhillah. 202110715243. Analisis keterlambatan *invoice* menggunakan metode clustering k-means dan dbscan di pt sebastian jaya metal. Bekasi: Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya. 2025.

Keterlambatan penerbitan *invoice* pada PT Sebastian Jaya Metal menyebabkan hambatan pada arus kas dan mengganggu efisiensi operasional perusahaan manufaktur, sehingga diperlukan analisis berbasis data untuk mengidentifikasi akar masalah dan memberikan solusi yang tepat. Penelitian ini bertujuan membandingkan efektivitas dua metode klasterisasi, yaitu K-Means dan DBSCAN, dalam memetakan pola keterlambatan *invoice* berdasarkan fitur transaksi yang relevan. Proses penelitian mengikuti tahapan *Knowledge Discovery in Databases* (KDD), mulai dari seleksi data, pra-pemrosesan untuk mengatasi duplikasi dan outlier, transformasi dengan *label encoder* untuk variabel kategori serta *MinMax Scaler* untuk variabel numerik, hingga modeling menggunakan kedua algoritma tersebut. Data yang dianalisis meliputi nama pelanggan, nama part, total transaksi, dan bulan transaksi dari sistem internal perusahaan periode bulan April – Agustus 2024. Keunikan penelitian ini terletak pada pendekatan perbandingan algoritma serta interpretasi hasil dengan mengembalikan fitur asli agar segmentasi yang dihasilkan mudah dipahami dan diterapkan secara bisnis. Hasil analisis menunjukkan bahwa metode K-Means lebih efektif dalam menghasilkan segmentasi pelanggan yang terstruktur dan mudah diinterpretasikan, dengan skor evaluasi *Silhouette Score* 0.583 dan *Davies-Bouldin Index* 0.523 yang lebih baik dibandingkan DBSCAN dengan *Silhouette Score* 0.31 dan *Davies-Bouldin Index* 1.20. Sebaliknya, DBSCAN unggul dalam mengidentifikasi kelompok anomali. Penelitian ini diharapkan dapat menjadi rujukan untuk pengembangan strategi penagihan dan pengelolaan pelanggan berbasis segmentasi data, serta menjadi dasar optimalisasi proses bisnis di sektor manufaktur.

Kata Kunci: *Clustering*, K-Means, DBSCAN, Keterlambatan *Invoice*, Segmentasi, Operasional Bisnis

ABSTRACT

Daffa Rayhan Fadhillah. 202110715243. *Analysis of Invoice Delays Using K-Means and DBSCAN Clustering Methods at PT Sebastian Jaya Metal.* Bekasi: Department of Informatics, Faculty of Computer Science, Bhayangkara University Jakarta Raya. 2025.

Delayed invoice issuance at PT Sebastian Jaya Metal disrupts cash flow and reduces operational efficiency, highlighting the need for data-driven analysis to pinpoint root causes and guide effective solutions. This research aims to compare the performance of two clustering algorithms, K-Means and DBSCAN, in revealing patterns behind invoice delays using relevant transaction features. The methodology adopts the Knowledge Discovery in Databases (KDD) framework, encompassing data selection, preprocessing to address duplicates and outliers, transformation through one-hot encoding for categorical variables and robust scaling for numerical data, and modeling with both algorithms. The analyzed dataset includes customer names, part names, total transaction amounts, and transaction months, sourced from the company's internal system for 2023–2024. The novelty of this research lies in its comparative approach to clustering methods and the re-mapping of clusters to original features, making the resulting segments both interpretable and actionable for business use. Results show that K-Means produces more coherent and interpretable customer segments, with higher Silhouette Score and better Davies-Bouldin Index values than DBSCAN, while DBSCAN proves useful for detecting anomalous transaction groups. The outcomes are expected to inform the development of targeted collection strategies and customer management based on data segmentation, offering practical recommendations for optimizing business processes in the manufacturing sector.

Keywords: Clustering, K-Means, DBSCAN, Invoice Delay, Segmentation, Business Operations

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik universitas bhayangkara jakarta raya, saya yang bertanda tangan dibawah ini:

Nama : Daffa Rayhan Fadhillah
NPM : 202110715243
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (Non-Exclusive Royalty-Free Right)**, atas karya ilmiah saya yang berjudul :

Analisis Penyebab Keterlambatan *Invoice* Menggunakan Algoritma K-Means dan DBSCAN di PT Sebastian Jaya Metal

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 05 Agustus 2025
Yang Menyatakan



Daffa Rayhan Fadhillah

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas limpahan rahmat, kasih, dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Keterlambatan Invoice Menggunakan Metode Clustering K-Means dan DBSCAN” sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya.

Penyusunan skripsi ini tentu tidak terlepas dari peran, dukungan, serta bantuan dari berbagai pihak. Untuk itu, pada kesempatan ini penulis ingin menyampaikan rasa terima kasih dan penghargaan yang setinggi-tingginya kepada:

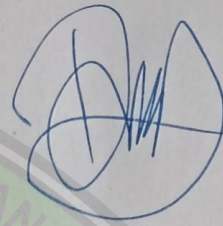
1. Bapak Irjen Pol. (Purn) Prof. Dr., Drs. H. Bambang Karsono, S.H., M.M., Ph.D., D.Crim (HC) selaku Rektor Universitas Bhayangkara Jakarta Raya, atas kepemimpinan dan dukungan beliau dalam menciptakan lingkungan akademik yang kondusif selama penulis menempuh studi.
2. Bapak Dr. Robertus Suraji, S.S., M.A. selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya, atas arahan dan dukungan yang senantiasa diberikan dalam proses akademik penulis.
3. Bapak Ahmad Fathurrozi, S.E., M.M.S.I selaku Ketua Program Studi Informatika Universitas Bhayangkara Jakarta Raya, atas bimbingan dan fasilitasi selama penulis menjalani perkuliahan hingga tahap penyusunan skripsi.

4. Bapak Prof. Dr. Mujiono Sadikin, S.T., M.T. selaku Dosen Pembimbing I. ucapan terima kasih yang sebesar-besarnya, atas kesabaran, dedikasi, dan ketelitian beliau dalam membimbing penulis hingga skripsi ini terselesaikan dengan baik
5. Bapak Sugiyatno S.Kom., M.kom. selaku Dosen Pembimbing Akademik dan Dosen Pembimbing II. Program Studi Informatika Universitas Bhayangkara Jakarta Raya, penulis mengucapkan terima kasih atas segala masukan dan pendampingan yang sangat membantu dalam menyelesaikan tugas akhir ini.
6. Kedua orang tua penulis, Bapak Sutrisno dan Ibu Leny atas segala cinta, doa, serta dukungan moril dan materil yang tiada henti diberikan, sehingga penulis mampu menyelesaikan skripsi ini dengan penuh semangat dan keyakinan
7. Seluruh tim dan staff Departemen IT di PT Sebastian Jaya Metal, atas bantuan, bimbingan, serta kerjasama yang diberikan selama proses kerja praktik dan pengumpulan data.
8. Ucapan terima kasih penulis kepada teman-teman yang telah memberikan dukungan, dan bantuan, baik secara langsung maupun tidak, dalam proses penyusunan skripsi ini
9. Untuk diri penulis sendiri, atas keberanian dalam menghadapi segala tantangan, komitmen untuk terus belajar, dan keteguhan hati dalam menyelesaikan tanggung jawab hingga akhir.

Penulis menyadari bahwa laporan Tugas Akhir ini masih jauh dari kesempurnaan, oleh sebab itu, peneliti berharap saran dan masukan dari berbagai pihak

agar laporan ini menjadi lebih baik agar, serta memberikan manfaat bagi semua pihak yang membutuhkannya.

Jakarta, 05 Agustus 2025



Daffa Rayhan Fadhilla



DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iv
ABSTRAK	v
ABSTRACT	vi
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK	vii
KATA PENGANTAR	viii
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	3
1.3 Rumusan Masalah	4
1.4 Tujuan	4
1.5 Manfaat	5
1.6 Batasan Masalah	6
1.7 Sistematika Tugas Akhir	6
BAB II TINJAUAN PUSTAKA	7
2.1 Penelitian Terkait	7
2.2 Data Mining	11
2.3 Clustering	11
2.4 Algoritma K-Means	12
2.5 Algoritma DBSCAN	13
2.6 Knowledge Discovery in Databases (KDD)	14
2.7 Davies Bouldin Index	16
BAB III METODOLOGI PENELITIAN	18

3.1 Jenis Penelitian.....	18
3.2 Objek Penelitian.....	18
3.3 Kerangka Penelitian	19
3.4 Metode Pengumpulan Data.....	22
3.4.1 Observasi Lapangan.....	22
3.4.2 Studi Literatur	22
3.5 Data Penelitian	23
3.6 Metode Analisis Data.....	24
3.6.1 Data Preparation.....	24
3.6.2 Modeling	24
3.6.3 Evaluasi.....	26
3.6.4 Analisis dan Interpretasi Hasil	26
BAB IV HASIL DAN PEMBAHASAN	28
4.1 Persiapan Data	28
4.2 Preprocessing Data.....	29
4.3 Pembuatan Model K-Means dan DBSCAN.....	32
4.3.1 Algoritma K-Means	33
4.3.2 Algoritma DBSCAN.....	35
4.4 Evaluasi Model	38
4.5 Interpretasi Hasil Klasterisasi	40
4.5.1 Hasil Klasterisasi dengan K-Means	41
4.5.2 Hasil Klasterisasi dengan DBSCAN.....	42
4.5.3 Hasil Perbandingan Algoritma.....	45
BAB V PENUTUP.....	49
5.1 Kesimpulan	49
5.2 Saran	50
DAFTAR PUSTAKA	52
LAMPIRAN.....	57

DAFTAR TABEL

Tabel 4. 1	Preview Nama Customer Encoded.....	30
Tabel 4. 2	Preview Partname Encoded.....	31
Tabel 4. 3	Preview Hasil Cluster.....	36
Tabel 4. 4	Hasil Perbandingan Algoritma	46



DAFTAR GAMBAR

Gambar 1. 1 Persentase Keterlambatan Invoice	2
Gambar 2. 3 Proses Knowledge Discovery In Database	15
Gambar 3. 1 Kerangka Penelitian.....	20
Gambar 3. 2 Data Penelitian.....	24
Gambar 4. 1 Preview Dataset Invoice	28
Gambar 4. 2 Checking Missing Values	29
Gambar 4. 3 Checking Duplicates.....	30
Gambar 4. 4 Feature Encoding.....	30
Gambar 4. 5 Hasil Normalisasi	31
Gambar 4. 6 Hasil Transformasi Data	32
Gambar 4. 7 Elbow Method for Optimal k pada pemodelan K-Means.....	33
Gambar 4. 8 Penerapan Algoritma K-Means	34
Gambar 4. 9 Visualisasi Hasil Klastering K-Means.....	35
Gambar 4. 10 Visualisasi Nilai eps untuk DBSCAN	36
Gambar 4. 11 Penerapan Algoritma DBSCAN	36
Gambar 4. 12 Hasil Davies-Bouldin Index dan Silhouette Score K-Means	39
Gambar 4. 13 Hasil Silhouette Score DBSCAN	39
Gambar 4. 14 Klasterisasi K-Means dan DBSCAN.....	40

DAFTAR LAMPIRAN

Lampiran 1 Hasil Cek Plagiarisme

Lampiran 2 Biodata Mahasiswa

Lampiran 3 Kartu Konsultasi Bimbingan Pembimbing I

Lampiran 4 Kartu Konsultasi Bimbingan Pembimbing II

Lampiran 5 Rekomendasi Pembimbing I

Lampiran 6 Rekomendasi Pembimbing II

Lampiran 7 Surat Permohonan Izin Penelitian

Lampiran 8 Surat Balasan Penelitian

