



**ANALISIS POLA KARIER ALUMNI MENGGUNAKAN
ALGORITMA K-MEANS, HIERARCHICAL CLUSTERING,
DAN DBSCAN BERDASARKAN DATA TRACER STUDY
UNIVERSITAS MERCU BUANA**

**TUGAS AKHIR
SKRIPSI**

**ANTONIO PARLINDUNGAN SIMATUPANG
41522010245**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026**



**ANALISIS POLA KARIER ALUMNI MENGGUNAKAN
ALGORITMA K-MEANS, HIERARCHICAL CLUSTERING,
DAN DBSCAN BERDASARKAN DATA TRACER STUDY
UNIVERSITAS MERCU BUANA**

**TUGAS AKHIR
SKRIPSI**

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana

**ANTONIO PARLINDUNGAN SIMATUPANG
41522010245**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : ANTONIO PARLINDUNGAN SIMATUPANG
NIM : 41522010245
Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika

Menyatakan dengan sebenar-benarnya bahwa Tugas Akhir berjudul:

“Analisis Pola Karier Alumni Menggunakan Algoritma K-Means, Hierarchical Clustering, dan DBSCAN Berdasarkan Data Tracer Study Universitas Mercu Buana” adalah hasil karya saya sendiri, tidak mengandung unsur plagiarisme, pelanggaran hak cipta, atau konten ilegal dalam bentuk apapun dan tidak melanggar hukum atau hak pihak manapun.

Apabila di kemudian hari ditemukan pelanggaran terhadap pernyataan ini, saya bersedia menanggung seluruh konsekuensi hukum dan membebaskan Universitas Mercu Buana dari segala bentuk tuntutan hukum dan saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Demikian pernyataan ini saya buat dengan sebenar-benarnya untuk digunakan sebagaimana mestinya.

Jakarta, 28 Juli 2026

UNIVERSITAS
MERCU BUANA



Antonio Parlindungan Simatupang

HALAMAN SURAT KETERANGAN HASIL UJI TURNITIN



072.423.4.03.01

PERNYATAAN SIMILARITY CHECK /SIMILARITY CHECK STATEMENT

Saya yang bertanda tangan dibawah ini menyatakan, bahwa karya ilmiah yang ditulis oleh
The undersigned, hereby declare that the scientific work written by

Nama /Name : Antonio Parlindungan Simatupang
NIM /Student ID Number : 41522010245
Program Studi /Study Program : Teknik Informatika

Dengan Judul Tugas Akhir

The title:

“ANALISIS POLA KARIER ALUMNI MENGGUNAKAN ALGORITMA K-MEANS, HIERARCHICAL CLUSTERING, DAN DBSCAN BERDASARKAN DATA TRACER STUDY UNIVERSITAS MERCU BUANA”

telah dilakukan pengecekan *similarity* dengan sistem Turnitin pada tanggal:

/Similarity checks have been carried out with the Turnitin system on the date:

21 Juli 2026

dengan nilai persentase sebesar :

/with a percentage value of:

20%

dinyatakan memenuhi standar sesuai dengan ketentuan berlaku di Fakultas Ilmu Komputer Universitas Mercu Buana. */declared to meet standards in accordance with applicable regulations at the Faculty of Computer Science, Universitas Mercu Buana.*

File hasil cek similarity turnitin:

/Turnitin similarity report file

https://drive.google.com/file/d/1pQEvvQQHmsk-WZ_LdwXyCA92wZUWmWx/view?usp=drive_link



Jakarta, 21 Juli 2026
Admin Turnitin Fasilkom UMB

Agung Prawoto, S.Kom., B.Sc
NIK : 322970503

Fakultas Ilmu Komputer

KAMPUS MENARA BHAKTI

Jl. Raya Meruya Selatan No. 1 Kembangan, Jakarta Barat 11650

Telp. 021-5840816 (Hunting), Psw : 5700 Fax. 021-5840813

<http://www.mercubuana.ac.id>, e-mail : fasilkom@mercubuana.ac.id

HALAMAN PENGESAHAN

Tugas Akhir ini diajukan oleh:

Nama : ANTONIO PARLINDUNGAN SIMATUPANG
NIM : 41522010245
Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika
Judul Tugas Akhir : Analisis Pola Karier Alumni Menggunakan Algoritma K-Means, Hierarchical Clustering, dan DBSCAN Berdasarkan Data Tracer Study Universitas Mercu Buana

Telah berhasil dipertahankan pada sidang tanggal 28 Juli 2026 dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing

Dr. Ir. Eliyani, M.Kom.

NIDN/NUPTK: 0321026901

Jakarta, 28 Juli 2026

Mengetahui,

Dekan Fakultas Ilmu Komputer



Dr. Bambang Jokonowo, S.Si., MTI
NIDN/NUPTK: 0320037002

Ketua Program Studi
Teknik Informatika



Dr. Hadi Santoso, S.Kom., M.Kom
NIDN/NUPTK: 0225067701

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Tugas Akhir ini. Penulisan Tugas Akhir ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Teknik Informatika pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan Tugas Akhir ini, sangatlah sulit bagi saya untuk menyelesaikan Tugas Akhir ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercu Buana.
4. Ibu Dr. Ir. Eliyani, M.Kom. selaku dosen pembimbing tugas akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan tugas akhir ini terjadwal dengan baik.
5. Kedua Orang Tua saya yang selalu mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercu Buana.

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Tugas Akhir ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 28 Juli 2026



Antonio Parlindungan Simatupang

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR DI REPOSITORI UMB**

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : ANTONIO PARLINDUNGAN SIMATUPANG
NIM : 41522010245
Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika
Judul Laporan Skripsi : Analisis Pola Karier Alumni Menggunakan Algoritma K-Means, Hierarchical Clustering, dan DBSCAN Berdasarkan Data Tracer Study Universitas Mercu Buana

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif** (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 28 Juli 2026
Yang menyatakan,

UNIVERSITAS
MERCU BUANA



Antonio Parlindungan Simatupang

ANALISIS POLA KARIER ALUMNI MENGGUNAKAN ALGORITMA K-MEANS, HIERARCHICAL CLUSTERING, DAN DBSCAN BERDASARKAN DATA TRACER STUDY UNIVERSITAS MERCU BUANA

ANTONIO PARLINDUNGAN SIMATUPANG

ABSTRAK

Tracer study merupakan salah satu instrumen yang digunakan perguruan tinggi untuk mengevaluasi keberhasilan lulusan dalam memasuki dunia kerja. Namun, hasil tracer study umumnya masih dimanfaatkan dalam bentuk statistik deskriptif sehingga belum mampu mengidentifikasi pola karier alumni secara lebih mendalam. Penelitian ini bertujuan untuk mengidentifikasi pola karier alumni berdasarkan data tracer study Universitas Mercu Buana menggunakan algoritma K-Means, Hierarchical Clustering, dan DBSCAN, membandingkan performa ketiga algoritma tersebut, serta mengimplementasikan model terbaik ke dalam sebuah aplikasi berbasis web. Dataset yang digunakan berasal dari data tracer study alumni Universitas Mercu Buana periode 2020–2024. Dataset awal terdiri atas 4.261 data alumni dengan 82 atribut. Setelah melalui proses seleksi, pembersihan data, *feature engineering*, *ordinal encoding*, transformasi logaritmik pada variabel pendapatan, dan standarisasi menggunakan *StandardScaler*, diperoleh 2.191 data alumni yang siap digunakan pada proses *clustering*. Evaluasi kualitas cluster dilakukan menggunakan *Silhouette Score* dan *Davies–Bouldin Index (DBI)*. Hasil penelitian menunjukkan bahwa algoritma Hierarchical Clustering memperoleh nilai evaluasi sedikit lebih baik dibandingkan K-Means. Namun, K-Means dipilih sebagai model akhir karena menghasilkan pembagian cluster yang lebih seimbang, lebih mudah diinterpretasikan, serta lebih sesuai untuk kebutuhan implementasi pada aplikasi. Model K-Means berhasil mengidentifikasi empat pola karier alumni, yaitu Karier Adaptif, Karier Selaras, Karier Tertunda, dan Karier Progresif, yang dibedakan berdasarkan karakteristik masa tunggu memperoleh pekerjaan, pendapatan, keselarasan pekerjaan dengan bidang studi, kesesuaian tingkat pendidikan terhadap pekerjaan, dan skala perusahaan tempat alumni bekerja. Model *clustering* yang dihasilkan kemudian diimplementasikan ke dalam aplikasi Career Pattern Identification System (CPIS) berbasis Streamlit yang menyediakan fitur prediksi individu, prediksi batch, visualisasi distribusi hasil, profil pola karier, dan informasi model. Hasil penelitian menunjukkan bahwa metode *clustering* mampu mengidentifikasi pola karier alumni secara efektif serta berpotensi mendukung pengelolaan data tracer study sebagai dasar evaluasi lulusan dan pengambilan keputusan di lingkungan perguruan tinggi.

Kata kunci: *Tracer Study*, Clustering, K-Means, Hierarchical Clustering, DBSCAN, Pola Karier Alumni.

**ANALISIS POLA KARIER ALUMNI MENGGUNAKAN ALGORITMA K-MEANS, HIERARCHICAL CLUSTERING, DAN DBSCAN
BERDASARKAN DATA TRACER STUDY UNIVERSITAS MERCU
BUANA**

ANTONIO PARLINDUNGAN SIMATUPANG

ABSTRACT

Graduate tracer studies play an important role in evaluating higher education outcomes by providing information on graduates' transition into the labor market. However, tracer study data are commonly analyzed using descriptive statistics, which limits the identification of underlying graduate career patterns. This study aims to identify graduate career patterns using tracer study data from Universitas Mercu Buana by applying K-Means, Hierarchical Clustering, and DBSCAN algorithms, comparing their clustering performance, and implementing the selected model in a web-based application. The dataset was obtained from the graduate tracer study conducted at Universitas Mercu Buana covering the period 2020–2024. The initial dataset consisted of 4,261 graduate records with 82 attributes. After data selection, data cleaning, feature engineering, ordinal encoding, logarithmic transformation of the income variable, and feature standardization using StandardScaler, a total of 2,191 records were retained for clustering analysis. Cluster quality was evaluated using the Silhouette Score and the Davies–Bouldin Index (DBI). The experimental results indicate that Hierarchical Clustering achieved slightly better internal evaluation scores than K-Means. Nevertheless, K-Means was selected as the final model because it produced more balanced clusters, offered better interpretability, and was more suitable for deployment in the proposed application. The selected model successfully identified four graduate career patterns, namely Adaptive Career, Aligned Career, Delayed Career, and Progressive Career, which were differentiated based on job waiting time, income level, horizontal job–education alignment, vertical education–job alignment, and company scale. The selected clustering model was subsequently implemented in a Streamlit-based Career Pattern Identification System (CPIS) featuring individual prediction, batch prediction, career pattern profiling, result visualization, and model information. The findings demonstrate that clustering techniques can effectively identify graduate career patterns and provide meaningful insights to support tracer study analysis, graduate outcome evaluation, and data-driven decision-making in higher education institutions.

Kata kunci: *Graduate Tracer Study, Career Pattern Analysis, Clustering, K-Means, Hierarchical Clustering, DBSCAN.*

DAFTAR ISI

HALAMAN SAMPUL.....	0
HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN SURAT KETERANGAN HASIL UJI TURNITIN	iii
HALAMAN PENGESAHAN.....	iv
KATA PENGANTAR.....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB.....	vi
ABSTRAK	vii
ABSTRACT.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
1.5 Batasan Masalah	4
BAB II TINJAUAN PUSTAKA	6
2.1 Penelitian Terdahulu	6
2.2 Teori Utama	14
2.2.1 <i>Data Mining</i>	14
2.2.2 <i>Clustering</i>	15
2.2.3 K-Means Clustering	16
2.2.4 Hierarchical Clustering.....	16
2.2.5 DBSCAN.....	17
2.2.6 <i>Evaluasi Clustering</i>	17
2.3 Teori Pendukung.....	18
2.3.1 <i>Tracer Study</i>	18
2.3.2 <i>Graduate Employability</i> dan <i>Career Outcome</i>	19
2.3.3 <i>Horizontal Mismatch</i> dan <i>Vertical Mismatch</i>	19

2.3.4	Prapemrosesan (<i>Preprocessing</i>) Data.....	20
2.4	Gap Penelitian	21
BAB III METODE PENELITIAN		23
3.1	Pendekatan Penelitian	23
3.2	Subjek Penelitian	24
3.3	Instrumen Penelitian	25
3.3.1.	Perangkat Keras	25
3.3.2.	Perangkat Lunak	26
3.4	Teknik Pengumpulan Data.....	27
3.5	Analisis Data	27
3.5.1	Identifikasi Variabel Penelitian	28
3.5.2	Teknik Preprocessing Data.....	29
3.5.3	Teknik <i>Clustering</i>	30
3.5.4	Teknik Evaluasi Clustering	30
3.6	Prosedur Penelitian	30
3.6.1	Identifikasi Permasalahan.....	32
3.6.2	Studi Literatur.....	33
3.6.3	Pengumpulan Dataset	33
3.6.4	Seleksi dan Pembersihan Data	34
3.6.5	<i>Feature Engineering</i> dan Transformasi Data	35
3.6.6	Implementasi Algoritma <i>Clustering</i>	38
3.6.7	Evaluasi Model Clustering	38
3.6.8	Interpretasi Pola Karier Alumni	39
3.7	Evaluasi Hasil Penelitian	39
BAB IV PEMBAHASAN		41
4.1	Persiapan data penelitian.....	41
4.1.1	Dataset Awal	41
4.1.2	<i>Exploratory Data Analysis</i> (EDA)	42
4.1.3	Proses Seleksi dan Pembersihan Dataset.....	46
4.1.4	<i>Feature Engineering</i> dan Transformasi Data	47
4.1.5	Dataset Siap Clustering	49
4.2	Implementasi Clustering	50
4.2.1	Hasil Clustering K-Means	50
4.2.2	Hasil Clustering Hierarchical Clustering	53
4.2.3	Hasil Clustering DBSCAN.....	55
4.3	Evaluasi dan Analisis Komparatif Model Clustering	59
4.3.1	Hasil Evaluasi Model <i>Clustering</i>	60
4.3.2	Analisis Komparatif Antar Algoritma	61
4.3.3	Pemilihan Model Clustering.....	62
4.4	Interpretasi Pola Karier Alumni.....	62
4.4.1	Cluster 0 – Karier Adaptif	64

4.4.2 Cluster 1 – Karier Selaras.....	65
4.4.3 Cluster 2 – Karier Tertunda.....	65
4.4.4 Cluster 3 – Karier Progresif.....	66
4.5 Analisis Hasil Penelitian	66
4.5.1 Perbandingan dengan Penelitian Terdahulu	67
4.5.2 Analisis Hasil Berdasarkan Landasan Teori.....	69
4.5.3 Implikasi dan Kontribusi Penelitian	70
4.6 Implementasi Model Clustering pada Aplikasi Streamlit.....	72
4.6.1 Halaman Beranda (Home).....	73
4.6.2 Halaman Prediksi Individu	75
4.6.3 Halaman Prediksi Batch	77
4.6.4 Halaman Profil Pola Karier	80
4.6.5 Halaman Tentang Model	81
4.6.6 Pembahasan Implementasi	83
BAB V KESIMPULAN DAN SARAN	86
5.1 Kesimpulan	86
5.2 Saran	87
DAFTAR PUSTAKA.....	89
LAMPIRAN.....	92



DAFTAR TABEL

Tabel 2.1 Penelitian Terkait	6
Tabel 3.1 Karakteristik Dataset.....	25
Tabel 3.2 Perangkat Keras	26
Tabel 3.3 Perangkat Lunak	26
Tabel 3.4 Variabel Penelitian.....	29
Tabel 3.5 Mapping Ordinal.....	37
Tabel 4.1 Hasil Seleksi Dataset	46
Tabel 4.2 Dataset Akhir Hasil Feature Engineering	49
Tabel 4.3 Jumlah Anggota Setiap Cluster K-Means.....	52
Tabel 4.4 Jumlah Anggota Setiap Cluster Hierarchical Clustering	55
Tabel 4.5 Hasil Eksperimen Parameter DBSCAN.....	57
Tabel 4.6 Distribusi Anggota Cluster DBSCAN	58
Tabel 4.7 Perbandingan Hasil Evaluasi Algoritma Clustering	60
Tabel 4.8 Karakteristik Cluster Hasil Clustering K-Means	63



DAFTAR GAMBAR

Gambar 3.1 Alur Pendekatan Penelitian	24
Gambar 3.2 Flowchart Prosedur Penelitian	31
Gambar 3.3 Proses Seleksi dan Pembersihan Dataset	35
Gambar 3.4 Tahapan Feature Engineering dan Transformasi Data.....	36
Gambar 3.5 Implementasi Algoritma Clustering.....	38
Gambar 4.1 Tampilan Dataset Awal Tracer Study Alumni.....	42
Gambar 4.2 Komposisi Jenis Atribut Dataset	43
Gambar 4.3 Distribusi Status Alumni	44
Gambar 4.4 Visualisasi 10 Atribut dengan Missing Value Terbesar.....	45
Gambar 4.5 Distribusi Pendapatan Sebelum dan Sesudah Transformasi Logaritmik.....	48
Gambar 4.6 Grafik Elbow Method untuk Penentuan Jumlah Cluster K-Means...	51
Gambar 4.7 Visualisasi Hasil Clustering K-Means	52
Gambar 4.8 Dendrogram Hierarchical Clustering Menggunakan Ward Linkage	53
Gambar 4.9 Visualisasi Hasil Hierarchical Clustering	54
Gambar 4.10 K-Distance Graph untuk Penentuan Nilai Parameter eps pada Algoritma DBSCAN	56
Gambar 4.11 Visualisasi Hasil Clustering DBSCAN.....	58
Gambar 4.12 Perbandingan Nilai Silhouette Score dan Davies–Bouldin Index ..	61
Gambar 4.13 Profil Karakteristik Cluster Hasil K-Means Berdasarkan Variabel Penelitian.....	64
Gambar 4.14 Arsitektur Implementasi Model Clustering pada Aplikasi Streamlit	73
Gambar 4.15 Halaman Beranda Aplikasi Streamlit.....	74
Gambar 4.16 Halaman Prediksi Individu Sebelum Proses Prediksi	75
Gambar 4.17 Hasil Prediksi Individu.....	76
Gambar 4.18 Halaman Prediksi Batch.....	77
Gambar 4.19 Preview Dataset Batch	78
Gambar 4.20 Hasil Prediksi Batch.....	79
Gambar 4.21 Visualisasi Distribusi Hasil Prediksi Batch.....	79
Gambar 4.22 Halaman Profil Pola Karier Alumni.....	80
Gambar 4.23 Halaman Tentang Model.....	82
Gambar 4.24 Pipeline dan Evaluasi Model.....	83

DAFTAR LAMPIRAN

Lampiran 1. Kartu Asistensi	92
Lampiran 2. Curriculum Vitae	93
Lampiran 3. Surat Pernyataan & Pengalihan HAKI.....	94
Lampiran 4. Sertifikat BNSP	96
Lampiran 5. Surat Ijin Riset Perusahaan.....	97

