



SKRIPSI

**KLASTERISASI TRACER STUDY ALUMNI
UNIVERSITAS PEMBANGUNAN NASIONAL
"VETERAN" JAWA TIMUR MENGGUNAKAN
ALGORITMA K-MEANS**

FABIYAN ATHA FERNALDY
NPM 20082010077

DOSEN PEMBIMBING
Amalia Anjani Arifiyanti, S.Kom., M.Kom.
Dhian Satria Yudha Kartika, S.Kom, M.Kom.

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
2024**



SKRIPSI

**KLASTERISASI TRACER STUDY ALUMNI
UNIVERSITAS PEMBANGUNAN NASIONAL
"VETERAN" JAWA TIMUR MENGGUNAKAN
ALGORITMA K-MEANS**

FABIYAN ATHA FERNALDY
NPM 20082010077

DOSEN PEMBIMBING

Amalia Anjani Arifiyanti, S.Kom., M.Kom.
Dhian Satria Yudha Kartika, S.Kom, M.Kom.

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
2024**

Halaman ini sengaja dikosongkan

LEMBAR PENGESAHAN

KLASTERISASI TRACER STUDY ALUMNI UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAWA TIMUR MENGUNAKAN ALGORITMA K-MEANS

Oleh :

FABIYAN ATHA FERNALDY

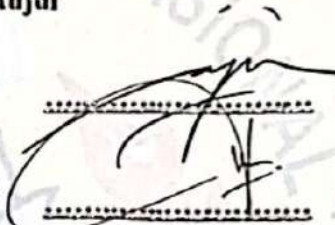
NPM. 20032010077

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 12 Desember 2024.

Menyetujui

Amalia Anjani Arifiyanti, S.Kom., M.Kom.

NIP. 19920812 2018032 001


..... (Pembimbing I)

Dhian Satria Yudha Kartika, S.Kom., M.Kom.

NPT. 2 0119 86 052224 9


..... (Pembimbing II)

Agung Brastama Putra, S.Kom., M.Kom.

NIP. 19851124 2021211 003


..... (Penguji I)

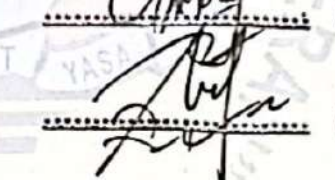
Anita Wulansari, S.Kom., M.Kom.

NIP. 19871015 2022032 005


..... (Penguji II)


Anindo Saka Fitri, S.Kom., M.Kom.

NIP. 19930325 2024062 001


..... (Penguji III)

Mengetahui,

Dekan Fakultas Ilmu Komputer


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

Halaman ini sengaja dikosongkan

LEMBAR PERSETUJUAN

**KLASTERISASI TRACER STUDY ALUMNI UNIVERSITAS
PEMBANGUNAN NASIONAL "VETERAN" JAWA TIMUR
MENGUNAKAN ALGORITMA K-MEANS**

Oleh :

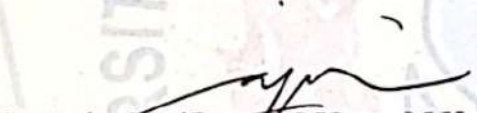
FABIYAN ATHA FERNALDY
NPM. 20082010077

Telah disetujui untuk mengikuti Ujian Skripsi

Menyetujui,

Dosen Pembimbing I

Dosen Pembimbing II



Amalia Anjani Arifiyanti, S.Kom., M.Kom.

NIP. 19920812/2018032 001


Dhian Satria Yudha Kartika, S.Kom., M.Kom.

NPT. 2 0119 86 052224 9

Koordinator Skripsi
Program Studi Sistem Informasi
Fakultas Ilmu Komputer


Eristva Maya Safitri, S.Kom., M.Kom.

NIP. 19930316 2019032 020

Halaman ini sengaja dikosongkan

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini :

Nama Mahasiswa / NPM : Fabiyan Atha Fernaldy / 20082010077
Program Studi : Sistem Informasi
Dosen Pembimbing : 1. Amalia Anjani Arifiyanti, S.Kom., M.Kom.
2. Dhian Satria Yudha Kartika, S.Kom., M.Kom.

dengan ini menyatakan bahwa isi sebagian maupun keseluruhan disertasi dengan judul:

Klasterisasi Tracer Study Alumni Universitas Pembangunan Nasional "Veteran" Jawa Timur Menggunakan Algoritma K-Means

adalah benar-benar hasil karya intelektual mandiri, diselesaikan tanpa menggunakan bahan-bahan yang tidak diizinkan dan bukan merupakan karya pihak lain yang saya akui sebagai karya sendiri. Semua referensi yang dikutip maupun dirujuk telah ditulis secara lengkap pada daftar pustaka. Apabila ternyata pernyataan ini tidak benar, saya bersedia menerima sanksi sesuai peraturan yang berlaku.

Surabaya, 12 Desember 2024

Mahasiswa



(Fabiyan Atha Fernaldy)

NPM. 20082010077

Halaman ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa / NPM : Fabiyan Atha Fernaldy / 20082010077
Judul Skripsi : Klasterisasi Tracer Study Alumni Universitas
Pembangunan Nasional “Veteran” Jawa Timur
Menggunakan Algoritma K-Means
Dosen Pembimbing : 1. Amalia Anjani Arifiyanti, S.Kom., M.Kom.
2. Dhian Satria Yudha Kartika, S.Kom., M.Kom.

Skripsi ini bertujuan untuk melakukan klasterisasi terhadap data tracer study alumni Universitas Pembangunan Nasional "Veteran" Jawa Timur menggunakan algoritma K-Means. Klasterisasi ini dilakukan untuk mengelompokkan alumni berdasarkan karakteristik karir mereka, yang diukur melalui Indeks Prestasi Kumulatif (IPK), waktu tunggu untuk mendapatkan pekerjaan, dan hubungan antara pekerjaan dengan program studi. Data yang digunakan dalam skripsi ini adalah hasil pengisian kuesioner tracer study dari tahun 2021 hingga 2022, dengan total 5.313 data alumni. Metode yang digunakan dalam skripsi ini meliputi preprocessing data, normalisasi, dan visualisasi menggunakan teknik Improved Visual Assessment for Tendency (iVAT) serta metode elbow untuk menentukan jumlah klaster optimal. Hasil klasterisasi menunjukkan bahwa terdapat tiga klaster berdasarkan IPK dan waktu tunggu, serta dua klaster berdasarkan hubungan pekerjaan dengan program studi dan waktu tunggu. Dalam penelitian dilakukan 2 klasterisasi dengan menggunakan dimensi yang berbeda, klasterisasi pertama yaitu ipk dengan masa tunggu alumni mendapatkan pekerjaan menghasilkan sejumlah 3 klaster dengan nilai evaluasi *silhouette coefficient* sebesar 0.499046 (memiliki struktur cluster cenderung lemah). Sebagai output, skripsi ini menghasilkan sebuah website yang menyajikan visualisasi dari hasil klasterisasi, sehingga dapat diakses dan digunakan oleh pihak universitas untuk mengevaluasi dan meningkatkan kualitas pendidikan serta mendukung pengembangan karir alumni.

Kata Kunci: Klasterisasi, Tracer Study, K-Means

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Fabiyan Atha Fernaldy / 20082010077
Thesis Title : Clustering Tracer Study Alumni of Universitas
Pembangunan Nasional “Veteran” Jawa Timur Using
K-Means Algorithm
Advisors : 1. Amalia Anjani Arifiyanti, S.Kom., M.Kom.
2. Dhian Satria Yudha Kartika, S.Kom., M.Kom.

This thesis aims to cluster the alumni tracer study data of Universitas Pembangunan Nasional “Veteran” East Java using K-Means algorithm. This clustering is done to group alumni based on their career characteristics, which are measured through Grade Point Average (GPA), waiting time to get a job, and the relationship between work and study program. The data used in this thesis are the results of filling out the tracer study questionnaire from 2021 to 2022, with a total of 5,313 alumni data. The methods used in this thesis include data preprocessing, normalization, and visualization using the Improved Visual Assessment for Tendency (iVAT) technique and the elbow method to determine the optimal number of clusters. The clustering results show that there are three clusters based on GPA and waiting time, and two clusters based on the relationship between work and study program and waiting time. In the study, 2 clusterizations were conducted using different dimensions, the first clusterization, namely the GPA with the waiting period for alumni to get a job, resulted in a number of 3 clusters with a Silhouette coefficient evaluation value of 0.499046 (having a cluster structure that tends to be weak). As an output, this thesis produces a website that presents a visualization of the clustering results, so that it can be accessed and used by the university to evaluate and improve the quality of education and support alumni career development.

Keywords: Clustering, Tracer Study, K-Means

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala rahmat, hidayah dan karunia-Nya kepada penulis sehingga skripsi dengan judul **“Klasterisasi Tracer Study Alumni Universitas Pembangunan Nasional “Veteran” Jawa Timur Menggunakan Algoritma K-Means”** dapat terselesaikan dengan baik.

Penulis mengucapkan terima kasih kepada Ibu Amalia Anjani Arifiyanti, S.Kom., M.Kom. selaku Dosen Pembimbing pertama dan Bapak Dhian Satria Yudha Kartika, S.Kom, M.Kom., sebagai dosen pembimbing ke-dua yang telah meluangkan waktunya untuk memberikan bimbingan, nasehat serta motivasi kepada penulis. Dan penulis juga banyak menerima bantuan dari berbagai pihak, baik itu berupa moril, spiritual maupun materiil. Untuk itu penulis mengucapkan terima kasih kepada:

1. Orang tua dan keluarga penulis yang selalu menyertai penulis dalam doa. Tidak lupa memberikan dukungan moral dan moril kepada penulis sehingga dapat menyelesaikan skripsi dengan baik.
2. Bapak Agung Brastama Putra, S.Kom., M.Kom., selaku Koordinator Program Studi Sistem Informasi di Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Bapak Arista Pratama, S.Kom., M.Kom, sebagai pihak UPA-PKK yang telah mendukung, menyediakan data, dan fasilitas yang diperlukan untuk pelaksanaan penelitian ini.
4. Ibu Siti Mukaromah, S.Kom, M.Kom., selaku dosen wali yang meluangkan waktunya untuk memberikan nasihat dalam menyelesaikan skripsi.
5. Seluruh Bapak dan Ibu dosen Sistem Informasi yang telah memberikan tenaga, waktu dan ilmunya selama proses belajar mengajar pada 8 semester terakhir.
6. Teman-teman terdekat saya yaitu nana, sopi, hur, dinka, fairuz, fira, anggi, takul, fikri, ddn, wsn yang selalu mendoakan, membantu, menghibur, memberikan motivasi dan semangat kepada penulisan dalam menyelesaikan skripsi ini.
7. Ivana Elfirdaus yang telah menemani, memotivasi dan memberikan dukungan penuh dalam pengerjaan skripsi.

8. Teman-teman Sistem Informasi Angkatan 2020 yang memberikan semangat kepada penulis sehingga dapat menyelesaikan skripsi.
9. Diri Saya sendiri yang sudah bertahan dan mampu menyelesaikan skripsi.
10. Seluruh pihak yang tidak bisa penulis sebutkan satu per satu yang telah memberikan dukungan dalam penyusunan skripsi.

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, November 2024

Penulis

DAFTAR ISI

LEMBAR JUDUL SKRIPSI	i
LEMBAR PENGESAHAN	iii
LEMBAR PERSETUJUAN	v
SURAT PERNYATAAN ORISINALITAS	vii
ABSTRAK	ix
ABSTRACT	xi
KATA PENGANTAR	xiii
DAFTAR ISI	xv
DAFTAR GAMBAR	xviii
DAFTAR TABEL	xxi
DAFTAR LAMPIRAN	xxii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian.....	3
BAB II TINJAUAN PUSTAKA	4
2.1 Dasar Teori	4
2.1.1 Tracer Study	4
2.1.2 <i>Data Mining</i>	4
2.1.3 <i>Clustering</i>	4
2.1.4 <i>K-Means Algorithm</i>	5
2.1.5 Python	5
2.1.6 iVAT	6

2.1.7	<i>Elbow Method</i>	6
2.1.8	<i>Silhouette Coefficient</i>	6
2.2	Penelitian Terdahulu	7
BAB III METODOLOGI PENELITIAN		10
3.1	Identifikasi Masalah	10
3.2	Studi Literatur.....	10
3.3	Pengumpulan Data	10
3.4	<i>Preprocessing Data</i>	18
3.5	Visualisasi Tendency Cluster (iVAT).....	19
3.6	<i>Clustering Data</i>	19
3.7	Evaluasi	20
3.8	Visualisasi Berbasis Website	20
BAB IV HASIL DAN PEMBAHASAN.....		21
4.1	Pengumpulan Data	21
4.2	<i>Preprocessing Data</i>	21
4.2.1	<i>Import Library</i>	21
4.2.2	Membaca Data CSV.....	22
4.2.3	Pemilihan Fitur.....	22
4.2.4	Data Cleansing	23
4.2.5	Normalisasi Data.....	26
4.3	Visualisasi Tendency Cluster (IVAT)	27
4.4	Clustering Data.....	28
4.4.1	Menentukan Nilai K yang Optimal	28
4.4.2	Assign Hasil Klaster.....	31
4.5	Evaluasi	35
4.6	Visualisasi Berbasis Website	36

4.6.1	Halaman Dashboard	41
4.6.2	Halaman Persebaran Alumni (Klaster IPK dengan Masa Tunggu Alumni Mendapatkan Pekerjaan)	42
4.6.3	Halaman Persebaran Alumni (Klaster Hubungan erat pekerjaan dan Program Studi dengan Masa Tunggu Alumni Mendapatkan Pekerjaan)	51
4.6.4	Halaman Tabel Data Persebaran Alumni (Klaster IPK dengan Masa Tunggu Alumni Mendapatkan Pekerjaan)	59
4.6.5	Halaman Tabel Data Persebaran Alumni (Klaster Hubungan erat pekerjaan dan Program Studi dengan Masa Tunggu Alumni Mendapatkan Pekerjaan)	60
4.6.6	Halaman Upload Data	61
BAB V PENUTUP		62
5.1	Kesimpulan.....	62
5.2	Saran	63
DAFTAR PUSTAKA.....		64
LAMPIRAN.....		68

DAFTAR GAMBAR

Gambar 3. 1 Diagram Alir Penelitian.....	10
Gambar 3. 2 Data Tracer study Universitas Pembangunan Nasional Veteran Jawa Timur tahun 2021-2022.....	11
Gambar 4. 1 Source Code Tahap Import Library.....	21
Gambar 4. 2 Source Code Tahap Membaca Data CSV	22
Gambar 4. 3 Source Code Tahap Pemilihan Fitur.....	22
Gambar 4. 4 Hasil Source Code Tahap Pemilihan Fitur	23
Gambar 4. 5 Source Code Tahap Pertama Data Cleansing	23
Gambar 4. 6 Hasil Source Code Tahap Pertama Data Cleansing.....	24
Gambar 4. 7 Source Code Tahap Penanganan Nilai Yang Dan Duplikat Data Frame 1.....	25
Gambar 4. 8 Source Code Tahap Penanganan Nilai Yang Dan Duplikat Data Frame 2.....	25
Gambar 4. 9 Hasil Source Code Tahap Penanganan Nilai Yang Dan Duplikat DataFrame1 dan DataFrame2	25
Gambar 4. 10 Source Code Tahap Normalisasi Data (Data Frame 1)	26
Gambar 4. 11 Source Code Tahap Normalisasi Data (Data Frame 2)	26
Gambar 4. 12 Hasil Source Code Tahap Normalisasi Data (DataFrame1 dan DataFrame2).....	26
Gambar 4. 13 Source Code Tahap Visualisasi Tendency Cluster Data Frame 1....	27
Gambar 4. 14 Source Code Tahap Visualisasi Tendency Cluster Data Frame 2....	27
Gambar 4. 15 Hasil Source Code Tahap Visualisasi Tendency Cluster Data Frame 1 dan Data Frame 2	27
Gambar 4. 16 Source Code Tahap Perhitungan SSE untuk K-means Clustering ..	28
Gambar 4. 17 Source Code Tahap Membuat Tabel SSE.....	28
Gambar 4. 18 Source Code Tahap Menghitung Penurunan SSE	29
Gambar 4. 19 Source Code Tahap Visualisasi Grafik Metode Elbow	29
Gambar 4. 20 Source Code Tahap Penerapan Fungsi elbow_method	30
Gambar 4. 21 Hasil Source Code Tahap Penerapan Fungsi elbow_method.....	30
Gambar 4. 22 Source Code Tahap Persiapan dan Definisi Fungsi	31

Gambar 4. 23 Source Code Tahap Pemilihan Jumlah Cluster dan Penerapan K-Means	31
Gambar 4. 24 Source Code Tahap Analisis Pusat Cluster.....	32
Gambar 4. 25 Source Code Tahap Visualisasi Hasil Clustering	32
Gambar 4. 26 Hasil Source Code Tahap Analisis Pusat Cluster dan Tahap Visualisasi Hasil Clustering	33
Gambar 4. 27 Source Code Tahap Menampilkan Distribusi Data Tiap Cluster	34
Gambar 4. 28 Hasil Source Code Tahap Menampilkan Distribusi Data Tiap Cluster	35
Gambar 4. 29 Source Code Tahap Evaluasi Cluster	35
Gambar 4. 30 Hasil Source Code Tahap Evaluasi Cluster	36
Gambar 4. 31 Source Code Tahap Koneksi Database dan Fetch Data	37
Gambar 4. 32 Source Code Tahap Eksekusi Query SQL.....	37
Gambar 4. 33 Source Code Tahap Transformasi Label pada Dataset.....	38
Gambar 4. 34 Source Code Tahap Visualisasi Berdasarkan Jenis Pekerjaan dan Program Studi	39
Gambar 4. 35 Source Code Tahap Inisialisasi Visual Grafik.....	40
Gambar 4. 36 Tampilan Halaman Dashboard Utama	41
Gambar 4. 37 Tampilan Halaman Persebaran Alumni (Klaster IPK dengan Masa Tunggu Alumni Mendapatkan Pekerjaan).....	42
Gambar 4. 38 Hasil Scatter Chart Cluster masa tunggu alumni mendapatkan pekerjaan dengan IPK.....	43
Gambar 4. 39 Tampilan Halaman Persebaran Alumni (Klaster IPK dengan Masa Tunggu Alumni Mendapatkan Pekerjaan) berdasarkan Jenis Pekerjaan	44
Gambar 4. 40 Visualisasi Grafik Batang Berdasarkan Jenis Pekerjaan Menggunakan Data Klaster 1.....	45
Gambar 4. 41 Tampilan Halaman Persebaran Alumni (Klaster IPK dengan Masa Tunggu Alumni Mendapatkan Pekerjaan) berdasarkan Program studi.....	47
Gambar 4. 42 Visualisasi Grafik Batang Berdasarkan Program Studi Menggunakan Data Klaster 1.....	48

Gambar 4. 43 Tampilan Halaman Persebaran Alumni (Klaster Hubungan erat pekerjaan dan Program Studi dengan Masa Tunggu Alumni Mendapatkan Pekerjaan).....	51
Gambar 4. 44 Hasil Scatter Chart Cluster masa tunggu alumni mendapatkan pekerjaan dengan hubungan erat pekerjaan - prgoram studi.....	52
Gambar 4. 45 Tampilan Halaman Persebaran Alumni (Klaster Hubungan erat pekerjaan dan Program Studi dengan Masa Tunggu Alumni Mendapatkan Pekerjaan) berdasarkan Jenis Pekerjaan.....	53
Gambar 4. 46 Visualisasi Grafik Batang Berdasarkan Jenis Pekerjaan Menggunakan Data Klaster 2.....	54
Gambar 4. 47 Tampilan Halaman Persebaran Alumni (Klaster Hubungan erat pekerjaan dan Program Studi dengan Masa Tunggu Alumni Mendapatkan Pekerjaan) berdasarkan Prgram studi.....	56
Gambar 4. 48 Visualisasi Grafik Batang Berdasarkan Program Studi Menggunakan Data Klaster 2.....	57
Gambar 4. 49 Tampilan Halaman Tabel Data Klaster IPK dengan Masa Tunggu Alumni Mendapatkan Pekerjaan	59
Gambar 4. 50 Tampilan Halaman Tabel Data Klaster Hubungan erat pekerjaan dan Program Studi dengan Masa Tunggu Alumni Mendapatkan Pekerjaan.....	60
Gambar 4. 51 Tampilan Halaman Upload Data	61

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu.....	7
Tabel 3. 1 Keterangan Atribut Kuesioner Tracer Study	11
Tabel 4. 1 Distribusi Data Berdasarkan Jenis Pekerjaan Menggunakan Data Klaster 1.....	46
Tabel 4. 2 Distribusi Data Berdasarkan Program Studi Menggunakan Data Klaster 1.....	48
Tabel 4. 3 Distribusi Data Berdasarkan Jenis Pekerjaan Menggunakan Data Klaster 2.....	54
Tabel 4. 4 Distribusi Data Berdasarkan Program Studi Menggunakan Data Klaster 2.....	57

DAFTAR LAMPIRAN

Lampiran 1 Surat Izin Penelitian.....	68
Lampiran 2 Dataset alumni Universitas Pembangunan Nasional “Veteran” Jawa Timur.....	69
Lampiran 3 Wawancara Pengumpulan Data	70