



Eksplorasi Data UMKM di Jawa Barat menggunakan Metode Clustering dengan algoritma K-Means dan Fuzzy C-Means untuk Rekomendasi Bisnis

LAPORAN TUGAS AKHIR



DAVID DARMAWAN

41520010066

**UNIVERSITAS
MERCU BUANA**

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2024



Eksplorasi Data UMKM di Jawa Barat menggunakan Metode Clustering dengan algoritma K-Means dan Fuzzy C-Means untuk Rekomendasi Bisnis

LAPORAN TUGAS AKHIR



DAVID DARMAWAN

41520010066

**UNIVERSITAS
MERCU BUANA**

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2024

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : David Darmawan
NIM : 41520010066
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Eksplorasi Data UMKM di Jawa Barat menggunakan Metode Clustering dengan algoritma K-Means dan Fuzzy C-Means untuk Rekomendasi Bisnis

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.



Jakarta, 22 Juli 2024



David Darmawan

UNIVERSITAS
MERCU BUANA

HALAMAN PENGESAHAN



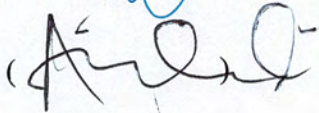

Laporan Skripsi ini diajukan oleh:

Nama : David Darmawan
NIM : 41520010066
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Eksplorasi Data UMKM di Jawa Barat menggunakan Metode Clustering dengan algoritma K-Means dan Fuzzy C-Means untuk Rekomendasi Bisnis

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Saruni Dwiasnati, ST, MM, M.Kom
NIDN : 0325128802
Ketua Penguji : Prastika Indriyanti, S. Kom, MCS
NIDN : 0312089401
Penguji 1 : Anis Cherid, S.E, M.TI
NIDN : 0328127203
Penguji 2 : Mohamad Yusuf, S. Kom, MCS
NIDN : 0307097606

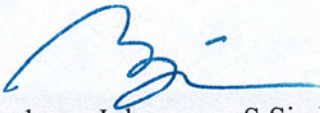
()
( ng/au
/or
)
()
()

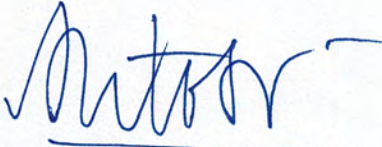
Jakarta, 22 Juli 2024

Mengetahui,

Dekan

Ketua Program Studi


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


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

KATA PENGANTAR

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

1. Prof. Dr. Ir. Andi Adriansyah, M.Eng selaku Rektor Universitas Mercu Buana
2. Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu komputer
3. Dr. Hadi Santoso S.Kom, M.Kom selaku Ketua Program Studi Teknik Informatika
4. Saruni Dwiasnati ST, MM, M.Kom selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan Skripsi ini;
5. Anis Cherid SE, M.TI dan Mohamad Yusuf S. Kom, MCS selaku Dosen Penguji Tugas Akhir atas koreksi dan arahan serta masukannya..
6. Teman – Teman Angkatan 20 Teknik Informatika selaku teman berdiskusi selama penelitian ini berlangsung.
7. Keluarga saya yang memberikan support selama perkuliahan berlangsung

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

Jakarta 22 Juli 2024



David Darmawan

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

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

Nama : David Darmawan
NIM : 41520010066
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Eksplorasi Data UMKM di Jawa Barat menggunakan Metode Clustering dengan algoritma K-Means dan Fuzzy C-Means untuk Rekomendasi Bisnis

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, mengalih media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 22 Juli 2024

UNIVERSITAS Yang menyatakan,

MERCU BUANA



David Darmawan

ABSTRAK

Nama : David Darmawan
NIM : 41520010066
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Eksplorasi Data UMKM di Jawa Barat menggunakan Metode Clustering dengan algoritma K-Means dan Fuzzy C-Means untuk Rekomendasi Bisnis

Pembimbing : Saruni Dwiasnati ST, MM, M.Kom

Indonesia, sebagai negara yang kaya akan sumber daya alam dan potensi sumber daya manusia yang besar, menghadapi tantangan berupa tingginya tingkat pengangguran karena keterbatasan lapangan pekerjaan. Usaha mikro, kecil, dan menengah merupakan kelompok pelaku ekonomi terbesar dalam perekonomian Indonesia dan terbukti menjadi katup pengaman perekonomian nasional pada masa kritis, serta mendorong pertumbuhan ekonomi pasca krisis ekonomi. Untuk itu penulis membuat sebuah rekomendasi untuk membantu calon UMKM usaha apa yang sebaiknya dibuka dengan mengelompokkan kategori usaha menggunakan algoritma K-Means dan Fuzzy C-Means dengan menggunakan python dan menggunakan library numpy, pandas, matplotlib, skfuzzy dan scikit-learn. K-Means merupakan suatu teknik pengelompokan data non-hierarki yang digabungkan ke dalam satu atau beberapa klaster. Algoritma Fuzzy C-Means adalah metode grup data di mana tingkat keanggotaan menentukan keberadaan setiap titik data dalam kelompok. Hasil dari pengujian kedua algoritma tersebut menunjukkan algoritma K-Means dengan nilai silhouette score 0,7781 dan Fuzzy C-Means dengan 0,630 yang mana algoritma K-Means lebih unggul dibanding Fuzzy C-Means, dan hasil rekomendasinya berupa Bisnis Top Brand dengan kategori usaha kuliner, makanan, dan jasa/lainnya.

Kata Kunci : UMKM, cluster, k means, fuzzy c-means, tableau

ABSTRACT

Nama : David Darmawan

NIM : 41520010066

Program Studi : Teknik Informatika

Judul Laporan Skripsi : Eksplorasi Data UMKM di Jawa Barat menggunakan Metode Clustering dengan algoritma K-Means dan Fuzzy C-Means untuk Rekomendasi Bisnis

Pembimbing : Saruni Dwiasnati ST, MM, M.Kom

Indonesia, a country rich in natural resources and possessing significant human resource potential, faces the challenge of high unemployment due to limited job opportunities. Micro, small, and medium enterprises (MSMEs) constitute the largest group of economic actors in Indonesia's economy and have proven to be a safety valve for the national economy during critical times, while also driving economic growth post-economic crisis. Therefore, the author has made a recommendation to help prospective MSMEs determine which type of business to start by categorizing business types using the K-Means and Fuzzy C-Means algorithms, utilizing Python and libraries such as numpy, pandas, matplotlib, skfuzzy, and scikit-learn. K-Means is a non-hierarchical data clustering technique that groups data into one or more clusters. The Fuzzy C-Means algorithm is a data grouping method where the membership degree determines the existence of each data point within the groups. The results from testing these two algorithms show that the K-Means algorithm achieved a silhouette score of 0.7781, while the Fuzzy C-Means algorithm scored 0.630, indicating that the K-Means algorithm is superior to the Fuzzy C-Means. The recommendation results include Top Brand Businesses in categories such as culinary, food, and services/others.

Keyword : UMKM, cluster, k means, fuzzy c-means, tableau

UNIVERSITAS
MERCU BUANA

DAFTAR ISI

HALAMAN JUSUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN	iii
KATA PENGANTAR	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
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	3
BAB II TINJAUAN PUSTAKA	5
2.1 Penelitian Terdahulu	5
2.2 Teori Pendukung	11
BAB III METODE PENELITIAN	15
3.1 Jenis Penelitian	15
3.2 Tahapan Penelitian	15
BAB IV HASIL DAN PEMBAHASAN	18
4.1 Business Understanding	18
4.2 Data Understanding	18
4.3 Data Preparation	18
4.4 Modeling	19
4.4.1 Label Encoder	20
4.4.2 Elbow Method	21

4.4.3 K-Means.....	24
4.4.4 Fuzzy C-Means	26
4.5 Evaluation.....	28
4.5.1 Silhouette Score	28
4.6 Deployment	29
4.6.1 Visualization	29
BAB V KESIMPULAN DAN SARAN	33
5.1 Kesimpulan.....	33
5.2 Saran.....	33
DAFTAR PUSTAKA.....	34
LAMPIRAN.....	36



UNIVERSITAS
MERCU BUANA

DAFTAR TABEL

Tabel 2.1. 1 Penelitian Terdahulu.....	5
Tabel 4.4.3. 1 Hasil clustering algoritma K-Means cluster 0.....	24
Tabel 4.4.3. 2 Hasil clustering algoritma K-Means cluster 1.....	25
Tabel 4.4.4. 1 Hasil algoritma Fuzzy C-Means cluster 0.....	26
Tabel 4.4.4. 2 Hasil algoritma Fuzzy C-Means cluster 1.....	27
Tabel 4.4.4. 3 Hasil algoritma Fuzzy C-Means cluster 2.....	28



DAFTAR GAMBAR

Gambar 2.2.5. 1 CRISP-DM.....	12
Gambar 2.2.6. 1 Logo Google Colab	13
Gambar 2.2.7 1 Logo Tableau.....	13
Gambar 3.2. 1 CRISP-DM prosedur	15
Gambar 4.2. 1 Dataset.....	18
Gambar 4.3. 1 Kolom kategori usaha dan jumlah umkm	19
Gambar 4.4. 1 Import library dan data pada algoritma K-Means	19
Gambar 4.4. 2 Import library dan data pada algoritma Fuzzy C-Means.....	20
Gambar 4.4.1. 1 Label encoder.....	20
Gambar 4.4.1. 2 Hasil label encoder pada fuzzy c-means	21
Gambar 4.4.1. 3 Hasil label encoder pada k-means.....	21
Gambar 4.4.2. 1 Codingan elbow method K-means	22
Gambar 4.4.2. 2 Codingan elbow method Fuzzy C-Means	22
Gambar 4.4.2. 3 Elbow method K-Means	23
Gambar 4.4.2. 4 Elbow method Fuzzy C-Means.....	23
Gambar 4.4.3. 1 Codingan algoritma K-Means	24
Gambar 4.4.4. 1 Proses clustering menggunakan Fuzzy C-Means.....	26
Gambar 4.5.1. 1 Hasil silhouette score pada K-Means	29
Gambar 4.5.1. 2 Hasil silhouette score pada Fuzzy C-means.....	29
Gambar 4.6.1. 1 Visualisasi 2 cluster pada K-means.....	30
Gambar 4.6.1. 2 Visualisasi cluster 0 top by 5 daerah	30
Gambar 4.6.1. 3 Visualisasi cluster 1 top by 5 daerah	31
Gambar 4.6.1. 4 Centroid K-Means.....	31

DAFTAR LAMPIRAN

Lampiran 1 : Kartu Asistensi Bimbingan.....	36
Lampiran 2 : Curriculum Vitae	37
Lampiran 3 : Lampiran Surat Pernyataan HKI.....	38
Lampiran 4 : Sertifikat BNSP.....	40
Lampiran 5 : Bukti Submit Jurnal ke Publisher	41
Lampiran 6 : Lembar Persetujuan Maju Sidang.....	42
Lampiran 7 : Form Revisi Dosen Penguji	43
Lampiran 8 : Hasil Cek Turnitin.....	45

