

**OPTIMASI *CLUSTER K-MEANS* MENGGUNAKAN
METODE *SILHOUETTE COEFFICIENT* UNTUK
POLA PENJUALAN MENU KOPI DI PEKPOK
COFFEE**

SKRIPSI

Oleh:

TIARA CAHYA FAUZIAH

201910225221



PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS BHAYANGKARA JAKARTA RAYA

2023

LEMBAR PERSETUJUAN PEMBIMBING

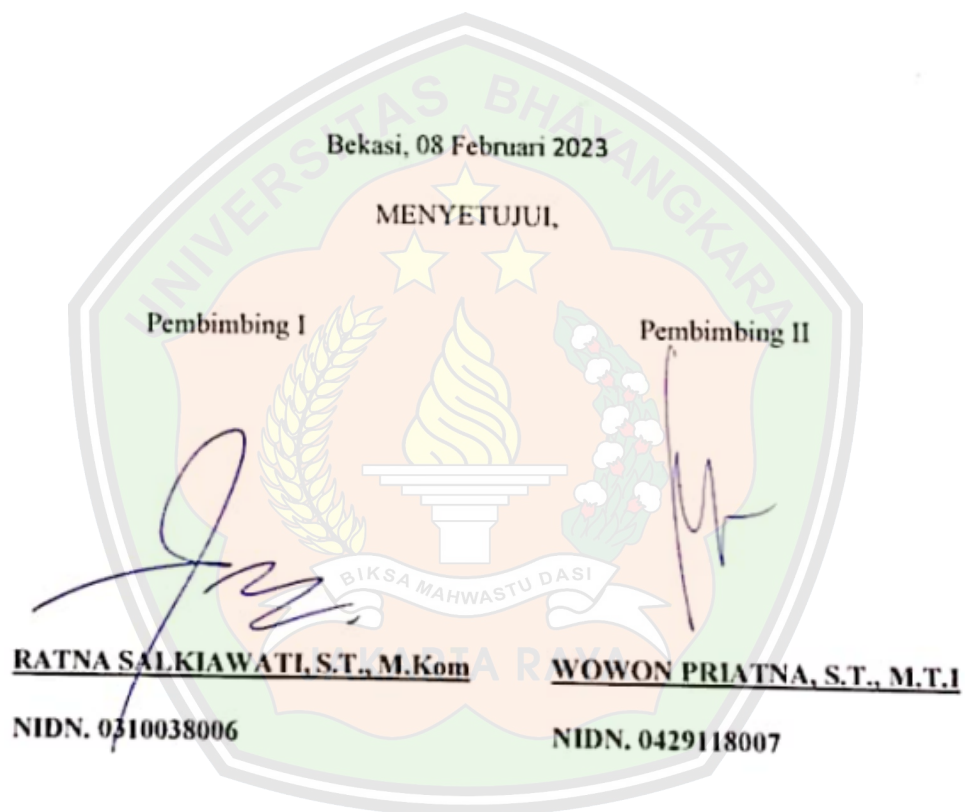
Judul Skripsi : Optimasi Cluster *K-Means* Menggunakan Metode
Silhouette Coefficient Untuk Pola Penjualan Menu
Kopi Di Pekpok Coffee

Nama Mahasiswa : Tiara Cahya Fauziah

Nomor Pokok Mahasiswa : 201910225221

Program Studi / Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 01 Februari 2023



LEMBAR PENGESAHAN

Judul Skripsi : Optimasi *Cluster K-Means* Menggunakan Metode
Silhouette Coefficient Untuk Pola Penjualan Menu
Kopi Di Pekpok Coffee

Nama Mahasiswa : Tiara Cahya Fauziah

Nomor Pokok Mahasiswa : 201910225221

Program Studi / Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 01 Februari 2023

Bekasi, 07 Februari 2023

MENGESAHKAN,

Ketua Tim Penguji : Prima Dina Atika, S.Kom., M.Kom.
NIDN. 0311037107

Penguji II : Ajif Yunizar Pratama Yusuf, S.Si., M.Eng.
NIDN. 0328068603

Penguji III : Ratna Salkiawati, S.T., M.Kom
NIDN. 0310038006

MENGETAHUI,

Ketua

Dekan

Program Studi Informatika

Fakultas Ilmu Komputer

Ahmad Fathurrozi, S.E., M.M.S.I.
NIP. 2012486

Dr. Dra. Tyastuti Sri Lestari, M.M.
NIP. 1408206



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan di bawah ini :

Nama : Tiara Cahya Fauziah
NPM : 201910225221
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Optimasi *Cluster K-Means* Menggunakan Metode *Silhouette Coefficient* Untuk Pola Penjualan Menu Kopi Di Pekpok Coffee

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggung jawabkan 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.

Bekasi, 07 Februari 2023
Penulis

Tiara Cahya Fauziah

ABSTRAK

Tiara Cahya Fauziah 201910225221, Optimasi *Cluster K-Means* Menggunakan Metode *Silhouette Coefficient* Untuk Pola Penjualan Menu Kopi Di Pekpok Coffee. Bekasi: Fakultas Ilmu Komputer. Universitas Bhayangkara Jakarta Raya. 2023.

Pekpok Coffee merupakan salah satu kedai kopi yang menjual berbagai macam menu kopi dengan domisili di Kecamatan Cibitung Kabupaten Bekasi. Menu Kopi yang dijual di Pekpok Coffee adalah *Americano*, *Cappuccino*, *Affogato* dan masih ada 11 menu kopi lainnya. Terdapat permasalahan yang seiring berjalannya waktu muncul akibat dari banyaknya persaingan seperti, adanya menu kopi yang tidak laku dan sulitnya untuk mempertahankan pelanggan. Tentunya permasalahan ini akan berdampak pada turunnya pendapatan dari penjualan kopi, kerugian bahkan yang lebih parahnya mengalami kebangkrutan. Tujuan dari penelitian ini untuk mengetahui pola penjualan berdasarkan informasi yang ada pada hasil pengelompokkan dengan menggunakan *algoritma K-Means* dengan optimasi *cluster* menggunakan metode *Silhouette Coefficient*. *Algoritma K-Means* termasuk kedalam *algoritma* penggolongan yang *iterative*, *algoritma* ini bekerja dengan memecah data set menjadi bagian – bagian kelompok berdasarkan jumlah kelompok yang sudah ditentukan pada awal tahapan. Tahapan dalam penelitian ini menggunakan metode *CRISP-DM* yang di dalamnya terdapat 6 proses untuk memecahkan sebuah permasalahan yaitu mulai dari *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation*, dan *Deployment*. Hasil Pengujian metode *algoritma K-Means* dengan optimasi *cluster* menggunakan metode *Silhouette Coefficient* mendapatkan kelompok optimal $k=7$ dengan nilai *silhouette score* mendekati 1 yaitu 0,58591

Kata Kunci : *Clustering*, *Pola Penjualan*, *K-Means*, *CRISP-DM*, *Silhouette Coefficient*.

ABSTRACT

Tiara Cahya Fauziah 201910225221, *K-Means Cluster Optimization Using the Silhoutte Coefficient Method for Coffee Menu Sales Patterns at Pekkop Coffee. Bekasi. Faculty of Computer Science. Bhayangkara University, Greater Jakarta. 2023.*

Pekkop Coffee is a coffee shop that sells various kinds of coffee menus with domiciles in Cibitung District, Bekasi Regency. Coffee menus sold at Pekkop Coffee are Americano, Cappuccino, Affogato and there are 11 other coffee menus. There are problems that over time arise as a result of lots of competition, such as coffee menus that don't sell well and it's difficult to retain customers. Of course, this problem will have an impact on decreasing income from coffee sales, losses and even worse bankruptcy. The purpose of this study is to determine sales patterns based on the information contained in the grouping results using the K-Means algorithm with cluster optimization using the Silhoutte Coefficient method. The K-Means algorithm is included in an iterative classification algorithm, this algorithm works by breaking the data set into groups based on the number of groups that have been determined at the beginning of the stage. The stages in this research use the CRISP-DM method in which there are 6 processes to solve a problem, starting from Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment. Test results of the K-Means algorithm method with cluster optimization using the Silhoutte Coefficient to get the optimal group $k = 7$ with a Silhoutte score close to 1, namely 0.5859.

Keywords: *Clustering, Sales Patterns, K-Means, CRISP-DM, Silhoutte Coefficient.*

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Tiara Cahya Fauziah
NPM : 201910225221
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 :

OPTIMASI CLUSTER *K-MEANS* MENGGUNAKAN METODE *SILHOUTTE COEFFICIENT* UNTUK POLA PENJUALAN MENU KOPI DI PEKPOK COFFEE

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-ekslusif 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 : Bekasi
Pada tanggal : 07 Februari 2023
Yang Menyatakan



Tiara Cahya Fauziah

KATA PENGANTAR

Puji dan Syukur kepada Allah SWT yang telah memberikan Rahmat, Hidayah beserta Karunia-Nya sehingga penulis dapat menyelesaikan penyusunan Skripsi yang berjudul “Optimasi *Cluster K-Means* Menggunakan Metode *Silhouette Coefficient* Untuk Pola Penjualan Menu Kopi Di Pekpok Coffee”. Dalam penyusunan Skripsi penulis ingin mengungkapkan rasa berterima kasih pada pihak – pihak yang telah banyak membantu dalam penyusunan Skripsi ini, yaitu:

1. Bapak Irjen Pol. (Purn) Dr. Drs. H. Bambang Karsono, SH, MM. selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Ibu Dr. Dra. Tyastuti Sri Lestari, M.M selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Bapak Ahmad Fathurrozi, S.E., MMSI selaku Ketua Program Studi Informatika.
4. Ibu Ratna Salkiawati, S.T., M.Kom selaku Dosen Pembimbing I yang telah memberikan arahan dan motivasi dalam kelancaran pelaksanaan skripsi ini.
5. Bapak Wowon Priatna, S.T., M.T.I selaku Dosen Pembimbing II yang memberikan saran dan motivasi demi kelancaran penyusunan skripsi ini.
6. Bapak Tjahyo Wulan Purnomo, Ibu Trijuni Merdianingsih dan kedua adik selaku keluarga yang selalu memberikan dukungan dan do'a.
7. Seluruh rekan Discord dan teman-teman Program Studi Informatika. Dengan atas bantuan, kerja samanya dalam hal penyelesaian laporan skripsi serta tugas-tugas semasa masih berkuliah.

Semoga penelitian ini bermanfaat untuk mahasiswa yang sedang menyelesaikan skripsi sebagai referensi penulisan. Penulis ucapkan terimakasih, semoga mendapatkan balasan dengan baik dari Allah SWT.

Bekasi, 11 Januari 2023

Peneliti,



Tiara Cahya Fauziah

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iii
ABSTRAK	iv
ABSTRACT	v
LEMBAR PERNYATAAN	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah	3
1.3 Rumusan Masalah	3
1.4 Batasan Masalah.....	4
1.5 Tujuan Penelitian.....	4
1.6 Manfaat Penelitian.....	4
1.7 Metode Pengumpulan Data	5
1.8 Sistematika Penulisan.....	5
BAB II LANDASAN TEORI	7
2.1 Tinjauan Pustaka	7
2.2 Data Mining	9
2.2.1 CRISP – DM	9
2.2.2 Clustering.....	11

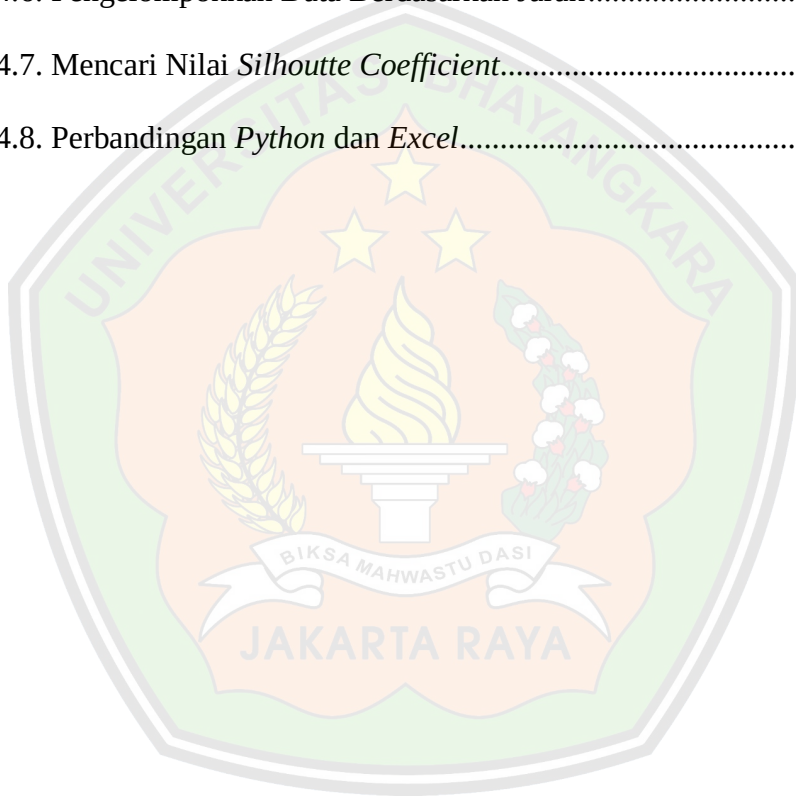
2.3	<i>Algortima K-Means</i>	12
2.3.1	Langkah <i>Clustering K-Means</i>	13
2.4	Metode <i>Silhoutte Coefficient</i>	14
2.5	<i>Python</i>	15
BAB III METODOLOGI PENELITIAN		17
3.1	Tempat dan Waktu Penelitian	17
3.1.1	Tempat Penelitian.....	17
3.1.2	Waktu Penelitian	18
3.2	Struktur Organisasi.....	19
3.3	Kerangka Penelitian.....	20
3.3.1	Identifikasi Masalah.....	20
3.3.2	Rumusan Masalah	21
3.3.3	Tahapan CRISP-DM	21
3.3.4	Metode Pengumpulan Data.....	23
3.3.5	Implementasi	25
3.4	Analisis Sistem Berjalan.....	25
3.5	Analisis Permasalahan.....	26
3.6	Analisis Sistem yang Diusulkan.....	27
BAB IV HASIL DAN IMPLEMENTASI		28
4.1	Perancangan Sistem.....	28
4.1.1	CRISP-DM Tahap 1 (<i>Business Undestanding</i>).....	28
4.1.2	CRISP-DM Tahap 2 (<i>Data Undestanding</i>)	28
4.1.3	CRISP-DM Tahap 3 (<i>Data Preparation</i>)	34
4.1.4	CRISP-DM Tahap 4 (<i>Modelling</i>).....	36
4.1.5	CRISP-DM Tahap 5 (<i>Evaluation</i>).....	38
4.2	Pengujian <i>Algoritma K-Means</i>	43

BAB V PENUTUP.....	47
5.1 Kesimpulan.....	47
5.2 Saran.....	48
DAFTAR PUSTAKA.....	49



DAFTAR TABEL

Tabel 4.1. Nilai <i>silhoutte</i>	39
Tabel 4.2. Penjualan Menu Kopi.....	42
Tabel 4.3. Sampel 10 Data Penjualan Menu Kopi Pekpok Coffee	43
Tabel 4.4. Penentuan <i>Centroid</i>	44
Tabel 4.5. Perhitungan <i>Cluster</i> dan Jarak Pengelompokkan.....	44
Tabel 4.6. Pengelompokkan Data Berdasarkan Jarak.....	45
Tabel 4.7. Mencari Nilai <i>Silhoutte Coefficient</i>	46
Tabel 4.8. Perbandingan <i>Python</i> dan <i>Excel</i>	46



DAFTAR GAMBAR

Gambar 2.1. Tahapan CRISP-DM	9
Gambar 2.2. <i>Developer Survey</i>	15
Gambar 3.1. Logo Pekkop Coffee.....	17
Gambar 3.2. Menu Pekkop Coffee.....	17
Gambar 3.3. Keadaan Saat Ramai Pelanggan	18
Gambar 3.4. Waktu Penelitian	18
Gambar 3.5. Struktur Organisasi.....	19
Gambar 3.6. Kerangka Penelitian	20
Gambar 3.7. <i>Flowchart</i> CRISP-DM	21
Gambar 3.8. <i>Flowchart</i> Fase <i>Modelling</i>	23
Gambar 3.9. <i>Flowmap</i> Sistem Berjalan Dalam Menentukan Strategi Bisnis	26
Gambar 3.10. <i>Flowmap</i> Sistem Yang Diusulkan Untuk Pola Penjualan.....	27
Gambar 4.1. Melihat Jumlah Baris dan Kolom Data	28
Gambar 4.2. Melihat 5 Data terakhir	28
Gambar 4.3. Melihat Informasi Mengenai Data	29
Gambar 4.4. Merubah Tipe Data dan Melihat 5 Data Pertama.....	29
Gambar 4.5. Melihat Tipe Data.....	29
Gambar 4.6. Melihat <i>Missing Value</i>	30
Gambar 4.7. Melihat <i>Outlier</i>	30
Gambar 4.8. Melakukan <i>Summary Statistics</i>	30
Gambar 4.9. Mendefinisikan Untuk Melakukan Visualisasi	31
Gambar 4.10. Melakukan Visualisasi Menu Kopi	31
Gambar 4.11. Melihat Urutan Nilai Dari yang Paling Banyak	32
Gambar 4.12. Melakukan Visualisasi Item	32

Gambar 4.13. Melihat Urutan Item Mulai Dari yang Terbanyak	33
Gambar 4.14. Melihat Nilai Mean dan Modus Dari Data.....	33
Gambar 4.15. Memilih Variabel yang Akan Digunakan	34
Gambar 4.16. Merubah Tipe Data Objek Menjadi <i>OneHotEncoder</i>	34
Gambar 4.17. Mengecek <i>Missing Value</i>	35
Gambar 4.18. Mengecek Kembali <i>Missing Value</i>	35
Gambar 4.19. Membagi Data <i>Train</i> dan <i>Test</i>	36
Gambar 4.20. <i>Clustering K=2</i>	37
Gambar 4.21. Membuat Label dari Setiap <i>Cluster</i>	37
Gambar 4.22. Mengecek Data <i>Train</i>	38
Gambar 4.23. Menggunakan fungsi <i>silhoutte_scores</i>	38
Gambar 4.24. Visualisasi <i>Silhoutte Scores</i>	39
Gambar 4.25. Membuat Data Banding Untuk Analisis <i>Cluster</i>	40
Gambar 4.26. Melihat <i>Missing Value</i> Dan Menanganinya.....	40
Gambar 4.27. Melihat Data Banding	41
Gambar 4.28. Menganalisis Setiap <i>Cluster</i>	41

DAFTAR LAMPIRAN

1. Surat Keterangan Riset Perusahaan.
2. Surat Rekomendasi dari Pembimbing.
3. Lembar Plagiarisme.
4. Biodata Mahasiswa.
5. Kartu Bimbingan Dosen Pembimbing 1.
6. Kartu Bimbingan Dosen Pembimbing 2.

