

SKRIPSI
IMPLEMENTASI DATA MINING UNTUK PENGELOMPOKAN BUKU
MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING



ANGELINA NOPE
NIM: 205410055
PROGRAM STUDI INFORMATIKA
PROGRAM SARJANA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA
2023

SKRIPSI
IMPLEMENTASI DATA MINING UNTUK PENGELOMPOKAN BUKU
MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING

Diajukan sebagai salah satu syarat untuk menyelesaikan studi



Program Sarjana
Program Studi Informatika
Fakultas Teknologi Informasi
Universitas Teknologi Digital Indonesia
Yogyakarta

Disusun Oleh

ANGELINA NOPE

NIM : 205410055

PROGRAM STUDI INFORMATIKA
PROGRAM SARJANA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA

2023

HALAMAN PERSETUJUAN

UJIAN SKRIPSI

Judul : Implementasi Data Mining Untuk Pengelompokan Buku
Menggunakan Algoritma K-Means Clustering

Nama : Angelina Nope

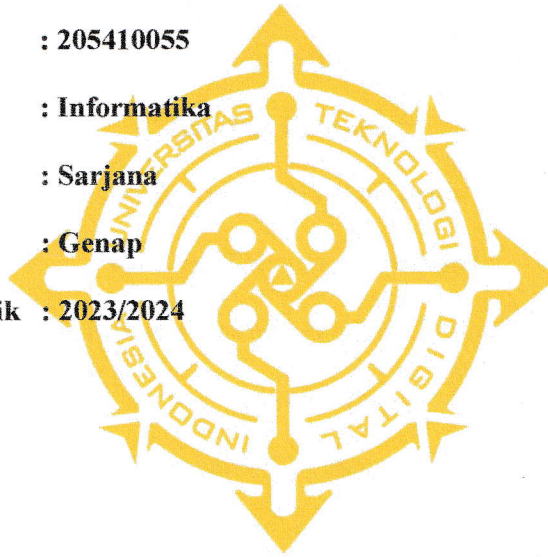
NIM : 205410055

Program Studi : Informatika

Program : Sarjana

Semester : Genap

Tahun Akademik : 2023/2024



Telah diperiksa dan disetujui untuk diujikan di hadapan Dewan Penguji Skripsi

Yogyakarta, 22 Februari 2024

Dosen Pembimbing,

Maria Mediatrix Sebatubun S.Kom., M.Eng

NIDN: 0514089101

HALAMAN PENGESAHAN

SKRIPSI

**IMPLEMENTASI DATA MINING UNTUK PENGELOMPOKAN BUKU
MENGUNAKAN ALGORITMA K-MEANS CLUSTERING**

Telah dipertahankan di depan Dewan Penguji Skripsi dan dinyatakan
diterima untuk memenuhi sebagian persyaratan guna memperoleh Gelar

Sarjana Komputer

Program Studi Informatika

Fakultas Teknologi Informasi

Universitas Teknologi Digital Indonesia

Yogyakarta

Yogyakarta, 22 Februari 2024

Dewan Penguji

NIDN

Tanda Tangan

Femi Dwi Astuti, S.Kom., M.Cs.

0516088701

Maria Mediatrix Sebatubun S.Kom., M.Eng.

0514089101

Mengetahui

Ketua Program Studi Informatika


Dini Fakta Sari, S.T., M.T.

NPP : 121172



27 FEB 2024

PERNYATAAN KEASLIAN SKRIPSI

Dengan ini penulis menyatakan bahwa naskah skripsi ini belum pernah diajukan untuk memperoleh gelar Sarjana Komputer di suatu Perguruan Tinggi, dan sepanjang pengetahuan penulis tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara sah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 22 Februari 2024



Angelina Nope

NIM: 205410055

HALAMAN PERSEMBAHAN

Puji Syukur penulis panjatkan kepada Tuhan Yesus Kristus, yang telah memberikan anugerah kesehatan dan kesanggupan, sehingga penulis masih diberikan kesempatan untuk menyelesaikan skripsi ini sebagai salah satu syarat untuk mendapatkan gelar sarjana. Akhirnya terselesaikan juga tugas akhir ini dan untuk itu penulis ingin mempersembahkannya untuk orang-orang yang penulis cintai dan sayangi, yaitu:

1. Tuhan Yesus Kristus yang teramat sangat baik memberikan hikmat, kesanggupan, dan kesehatan sehingga penulis dapat menyelesaikan skripsi ini.
2. Keluarga tercinta Bapa, Mama, Ka Charles, Om Pina, Any, Ma ally, Ma Ella, Lady, Boy dan Petra yang telah memberikan doa, dukungan dan menjadi tempat berkeluh kesah tetapi juga menjadi sumber semangat penulis dalam menyelesaikan skripsi ini.
3. Teman-teman Informatika 1, ML Skuy, Minimal Lulus, NTT Bersaudara, Hawila, Imel, Jati, Nova, Dooni yang telah membantu dan menjadi teman selama perkuliahan.
4. Keluarga maroon, keluarga besar HIMAFORKA yang telah menjadi bagian perubahan-perubahan baik selama ini. Pelajaran, semangat, motivasi, persahabatan dan pengalaman yang belum tentu bisa penulis dapatkan di tempat lain.
5. Keluarga besar PMK Exodus UTDI yang telah menjadi bagian penulis mengembangkan diri dan mendekatkan diri kepada Tuhan. Persaudaraan,

semangat dan motivasi yang penulis dapatkan akan selalu dikenang sepanjang masa.

6. Keluarga Besar PD Kasih yang selalu mendukung dalam doa dan selalu memberikan semangat kepada penulis untuk tidak menyerah dalam mengerjakan skripsi ini, terima kasih banyak semoga Tuhan memberkati selalu.
7. Sahabat-sahabat penulis Ofri, Enjel, Intan, Yusni, Heny, Itho, Boy Djingi yang selalu mengirimkan semangat dan doa, terima kasih kalian anak-anak baik
8. Teman-teman pelayan GBI Blessing yang selalu tertawa bersama dan saling mendoakan serta mengingatkan untuk semangat mengerjakan skripsi hingga bisa selesai dengan baik
9. Kepada diri saya sendiri yang sudah mau berjuang untuk menyelesaikan skripsi ini, terima kasih karena tidak menyerah, terima kasih sudah kuat, terima kasih karena sudah mencintai diri ini dengan sangat baik

MOTTO

“Segala perkara dapat ku tanggung didalam

DIA yang memberi kekuatan kepadaku”

- Filipi 4:13 -

*“If somebody hurt you, it’s okay to cry a river,
Just remember to build a bridge and get over it.”*

- Taylor Swift -

KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa atas Anugerah-Nya, baik itu berupa sehat fisik maupun akal pikiran, sehingga penulis dapat menyelesaikan Skripsi ini meskipun jauh dari kata sempurna.

Selesainya Skripsi ini tidak terlepas dari bantuan, bimbingan serta dukungan dari berbagai pihak, oleh karena itu melalui kesempatan ini dengan segala kean hati penulis mengucapkan banyak terimakasih kepada:

1. Ibu Sri Redjeki, S.Si., M.Kom., Ph.D. selaku Rektor Universitas Teknologi Digital Indonesia.
2. Dr. Bambang Purnomosidi DP, S. E. Akt., S. Kom., MMSI. Selaku Dekan Fakultas Teknologi Informasi Universitas Teknologi Digital Indonesia.
3. Ibu Dini Fakta Sari, S.T., M.T. dan Ibu Femi Dwi Astuti, S.Kom., M.Cs. , Ketua dan Sekretaris Program Studi Informatika Universitas Teknologi Digital Indonesia.
4. Ibu Maria Mediatrix Sebatubun S.Kom., M.Eng selaku dosen pembimbing yang telah memberikan bimbingan selama pengerjaan skripsi.
5. Ibu Femi Dwi Astuti, S.Kom., M.Cs selaku dosen penguji pada sidang skripsi.
6. Keluarga besar HIMAFORKA, UKM PMK EXODUS, UKM KESENIAN yang telah menjadi rumah dan memberikan banyak pelajaran dan pengalaman.
7. Teman-teman mahasiswa Universitas Teknologi Digital Indonesia yang tidak bisa penulis sebutkan satu persatu yang telah terlibat dalam membantu terselesaikannya skripsi ini.

Menyadari akan kekurangan dan kesalahan penulis dalam skripsi ini karena keterbatasan ilmu penulis, maka penulis mengharapkan masukan berupa saran yang membangun dari semua pihak.

Semoga skripsi ini dapat memberikan manfaat, baik bagi penulis sendiri, maupun bagi masyarakat luas.

Yogyakarta, 22 Februari 2024



Angelina Nope

NIM: 205410055

DAFTAR ISI

SKRIPSI	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
MOTTO	vii
KATA PENGANTAR	viii
DAFTAR GAMBAR	xii
INTISARI	xiv
ABSTRACT	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Ruang Lingkup	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI	7
2.1 Tinjauan Pustaka	7
2.2 Dasar Teori	11
2.2.1 <i>Data Mining</i>	11
BAB III METODE PENELITIAN	25
3.1 Bahan/Data	25
3.2 Peralatan	25
3.2.1 Kebutuhan Perangkat Keras	25
3.2.2 Kebutuhan Perangkat Lunak	25
3.2.3 Kebutuhan <i>Input</i>	26
3.2.4 Kebutuhan Proses	26
3.2.5 Kebutuhan <i>Output</i>	29
3.3 Analisis dan Rancangan Sistem	29
3.3.1 <i>Use Case Diagram</i>	30
3.3.2 Perancangan Antarmuka	31

BAB IV IMPLEMENTASI DAN PEMBAHASAN	34
4.1 Implementasi	34
4.1.1 <i>Cleaning data</i>	34
4.1.2 <i>Elbow Method</i>	34
4.1.3 Penerapan Algoritma <i>K-Means Clustering</i>	36
4.1.4 Implementasi SVM	43
4.2 Uji Coba Sistem	44
4.1.1 Antarmuka Aplikasi Web	44
4.1.2 Hasil Uji Coba	46
BAB V PENUTUP	49
5.1 Kesimpulan	49
5.2 Saran	50
DAFTAR PUSTAKA	51
LAMPIRAN	52

DAFTAR GAMBAR

Gambar 2. 1 Tahapan Knowledge Discovery Database (Ependi & Putra,2019)	11
Gambar 2. 2 Support Vector Machine (Parapat, et al.,2018)	22
Gambar 2. 3 Pemetaan Input Space Berdimensi Dua dengan Pemetaan ke Dimensi Tinggi (Rahutomo et al.,2018)	24
Gambar 3. 1 Alur Penelitian	26
Gambar 3. 2 Algoritma K-Means Clustering	28
Gambar 3. 3 Use Case Diagram	30
Gambar 3. 4 Tampilan Halaman Beranda	31
Gambar 3. 5 Tampilan Halaman Hasil	32
Gambar 3. 6 Tampilan Halaman Clustering	33
Gambar 3. 7 Tampilan Halaman Sebaran	33
Gambar 4. 1 Kode Cleaning Data	34
Gambar 4. 2 Kode Menentukan Nilai K dengan Ellbow Method	35
Gambar 4. 3 Hasil elbow method	35
Gambar 4. 4 Kode memanggil variabel yang akan digunakan	36
Gambar 4. 5 Kode Mengubah Data menjadi Array	36
Gambar 4. 6 Kode Scalling Data	37
Gambar 4. 7 Kode Untuk Menampilkan Nilai Centroid	37
Gambar 4. 8 Kode Untuk Menambahkan Variabel Cluster	37
Gambar 4. 9 Kode Untuk Menampilkan Cluster	39
Gambar 4. 10 Kode Untuk Membuat Grafik Visualisasi Cluster	41
Gambar 4. 11 Grafik <i>scatterplot</i> hasil <i>clustering</i>	41
Gambar 4. 12 Sebaran Pie Chart	42
Gambar 4. 13 Kode Untuk Mengukur Akurasi SVM	43
Gambar 4. 14 Hasil Akurasi SVM	43
Gambar 4. 15 Tampilan Interface Halaman Beranda	44
Gambar 4. 16 Tampilan Halaman Hasil	45
Gambar 4. 17 Tampilan Interface Halaman clustering	45
Gambar 4. 18 Tampilan Interface Halaman sebaran	46
Gambar 4. 19 Input judul buku dengan hasil cluster 0	47
Gambar 4. 20 Input judul buku dengan hasil cluster 1	47
Gambar 4. 21 Input judul buku dengan hasil cluster 2	48

DAFTAR TABEL

Tabel 2 1 Perbandingan Penelitian	9
Tabel 2 2 Contoh Perhitungan K-Means	16
Tabel 2 3 Hasil Perhitungan K-Means Iterasi 1	18
Tabel 2 4 Hasil Perhitungan K-Means Iterasi 2	20

INTISARI

Data mining merupakan salah satu cabang penting dalam ilmu komputer yang bertujuan untuk mengekstraksi pola atau pengetahuan yang berguna dari sekumpulan data. Dalam dunia penerbitan dan perpustakaan, analisis data dapat memberikan wawasan yang berharga tentang perilaku pembaca, preferensi buku, dan kepopuleran. Salah satu teknik yang umum digunakan dalam analisis data adalah *K-Means Clustering*, di mana data dikelompokkan ke dalam beberapa *cluster* berdasarkan kemiripan. Dalam studi ini, Implementasi algoritma *K-Means Clustering* menggunakan bahasa pemrograman *Python* untuk melakukan analisis data. Penelitian ini menggunakan *dataset* untuk mendemonstrasikan langkah-langkah dalam menerapkan algoritma *K-Means* dan menganalisis hasil.

Penelitian ini menghasilkan pengelompokan data buku dengan tiga *cluster* yaitu *cluster 0*, *cluster 1* dan *cluster 2*. Implementasi *Data Mining* menggunakan algoritma *K-Means clustering* untuk pengelompokan data buku menghasilkan hasil akhir dari pengelompokan data sebanyak 2123 data buku, dengan 1567 data *cluster 0*, 487 data *cluster 1* dan 69 data *cluster 2*.

Kata Kunci: *Data mining, K-Means Clustering, Python*

ABSTRACT

Data mining is an important branch of computer science that aims to extract useful patterns or knowledge from a set of data. In the world of publishing and libraries, data analysis can provide valuable insights into reader behavior, book preferences, and popularity. One technique commonly used in data analysis is K-Means Clustering, where data is grouped into several clusters based on similarity. In this study, the implementation of the K-Means Clustering algorithm uses the Python programming language to carry out data analysis. This research uses a dataset to demonstrate the steps in applying the K-Means algorithm and analyzing the results.

This research produces grouping of book data with three clusters, namely cluster 0, cluster 1 and cluster 2. Implementation of Data Mining using the K-Means clustering algorithm for grouping book data produces the final result of grouping data of 2123 book data with, 1567 data clusters 0, 487 data cluster 1 and 69 cluster 2 data.

Keywords: *Data mining, K-Means Clustering, Python*