

Optimisasi Pengelompokan Judul Proyek Aplikasi Konsentrasi Menggunakan Algoritma K-Means

Ardimansyah, Mudarsep, Baharuddin Rahman

Universitas Dipa Makassar

Jl. Perintis Kemerdekaan Km.9; Telp. 0411- 587194

e-mail: ardiman@undipa.ac.id, mudarsep@undipa.ac.id, baharuddin.rahman@undipa.ac.id

Abstrak

Pengelompokan judul proyek aplikasi merupakan langkah penting dalam pengembangan perangkat lunak untuk mengatur sumber daya dan mempercepat proses pengembangan. Algoritma K-Means telah digunakan secara luas dalam pengelompokan data, termasuk pengelompokan judul proyek. Namun, kinerja K-Means dapat dipengaruhi oleh faktor seperti inisialisasi centroid dan jumlah cluster yang ditentukan secara sembarang. Oleh karena itu, penelitian ini bertujuan untuk mengoptimalkan penggunaan algoritma K-Means dalam pengelompokan judul proyek aplikasi. Metode yang diusulkan mencakup pemilihan jumlah cluster yang optimal, strategi inisialisasi centroid yang lebih baik, dan evaluasi kualitas pengelompokan. Pengujian dilakukan menggunakan dataset judul proyek aplikasi, dan hasilnya menunjukkan peningkatan signifikan dalam efisiensi dan akurasi pengelompokan. Penelitian ini memberikan kontribusi penting dalam meningkatkan pengelolaan proyek perangkat lunak dan mempercepat pengembangan aplikasi.

Kata kunci: Pengelompokan judul proyek aplikasi, Algoritma K-Means, Optimisasi, Inisialisasi centroid, Evaluasi kualitas.

Abstract

Grouping application project titles is an important step in software development to organize resources and expedite the development process. The K-Means algorithm has been widely used in data clustering, including project title grouping. However, the performance of K-Means can be influenced by factors such as centroid initialization and arbitrarily determined cluster numbers. Therefore, this research aims to optimize the use of the K-Means algorithm in grouping application project titles. The proposed methods include selecting the optimal number of clusters, better centroid initialization strategies, and evaluating clustering quality. Testing was conducted using a dataset of application project titles, and the results show a significant improvement in clustering efficiency and accuracy. This research contributes significantly to enhancing software project management and expediting application development.

Keywords: *Application project title grouping, K-Means algorithm, Optimization, Centroid initialization, Quality evaluation.*

1. Pendahuluan

Dalam pengembangan aplikasi perangkat lunak, pengelompokan judul proyek menjadi kategori yang tepat adalah langkah awal yang penting. Pengelompokan yang efisien memungkinkan pengembang untuk mengalokasikan sumber daya dengan lebih baik, mempercepat proses pengembangan, dan meningkatkan hasil akhir. Namun, pengelompokan manual seringkali memakan waktu dan kurang efisien ketika dihadapkan pada jumlah judul proyek yang besar.

Algoritma K-Means adalah salah satu algoritma pengelompokan yang populer dalam analisis data dan pembelajaran mesin. Algoritma ini bekerja dengan mengelompokkan data ke dalam k kelompok berdasarkan karakteristik yang serupa, di mana k merupakan jumlah kelompok yang telah ditentukan sebelumnya. Dalam konteks pengelompokan judul proyek aplikasi, algoritma K-Means dapat digunakan untuk mengelompokkan judul proyek yang memiliki kesamaan topik atau fitur tertentu.

Meskipun algoritma K-Means telah terbukti efektif dalam berbagai aplikasi, namun kinerjanya dapat dipengaruhi oleh beberapa faktor, seperti inisialisasi centroid yang acak atau jumlah cluster yang

ditentukan secara sembarang. Oleh karena itu, optimisasi penggunaan algoritma K-Means menjadi relevan dalam konteks ini.

Penelitian ini bertujuan untuk mengoptimalkan proses pengelompokan judul proyek aplikasi menggunakan algoritma K-Means. Optimisasi ini dapat mencakup pemilihan jumlah cluster yang optimal, strategi inisialisasi centroid yang lebih baik, atau penggunaan metode evaluasi yang sesuai untuk mengukur kualitas pengelompokan. Dengan melakukan optimisasi ini, diharapkan hasil pengelompokan judul proyek aplikasi menjadi lebih efisien dan akurat, yang pada gilirannya akan meningkatkan produktivitas dan kualitas dalam pengembangan aplikasi perangkat lunak.

2. Metode Penelitian

2.1 Pemilihan Dataset

Dataset adalah kumpulan data terstruktur dan terorganisir yang digunakan untuk analisis dan pemodelan. Pemilihan dataset yang tepat sangat penting untuk menghasilkan model yang akurat dan berkinerja baik[1].

2.2 Implementasi Algoritma K-Means

Algoritma K-Means adalah algoritma clustering yang membagi data menjadi k kelompok berdasarkan jarak Euclidean. Algoritma ini tergolong mudah dipahami dan diimplementasikan[2].

2.3 Pemilihan Jumlah Cluster

Jumlah cluster adalah jumlah kelompok yang ingin dibuat dalam proses clustering. Menentukan jumlah cluster yang tepat merupakan hal yang penting untuk menghasilkan pengelompokan yang optimal[3].

2.4 Strategi Inisialisasi Centroid

Centroid adalah titik pusat dari sebuah cluster. Strategi inisialisasi centroid yang tepat dapat membantu algoritma K-Means mencapai konvergensi dengan lebih cepat dan menghasilkan pengelompokan yang lebih baik[4].

2.5 Validasi dan Analisis Hasil

Validasi adalah proses untuk memastikan bahwa hasil clustering valid dan akurat. Analisis hasil adalah proses untuk menginterpretasikan hasil clustering dan mendapatkan insights yang bermanfaat[5].

3. Hasil dan Pembahasan

Dalam penelitian ini akan digunakan data sebanyak 20 judul proyek aplikasi yang akan akan dikelompokkan menggunakan algoritma K-Means.

Berikut adalah contoh lengkap hasil clustering dalam bentuk tabel untuk 20 data judul:

Tabel 1 Data Judul Proyek Aplikasi

No.	Judul	Panjang Judul	Jumlah Kata
1	Sistem Manajemen Toko Online	28	4
2	Aplikasi Pencatatan Keuangan Pribadi	36	4
3	Platform E-learning Interaktif	30	3
4	Aplikasi Pemesanan Makanan Online	33	4
5	Aplikasi Pengelolaan Jadwal Dokter	34	4
6	Sistem Informasi Geografis Pariwisata	37	4
7	Aplikasi Catatan Kesehatan Digital	34	4
8	Aplikasi Olahraga & Kesehatan	29	3
9	Aplikasi Manajemen Proyek Kolaboratif	37	4
10	Platform Streaming Musik	24	3
11	Aplikasi Pembelajaran Bahasa Asing	34	4
12	Sistem Manajemen Inventaris Toko Retail	39	5
13	Aplikasi Pencarian Pekerjaan	28	3
14	Aplikasi Katalog Produk Fashion	31	4
15	Aplikasi Reservasi Hotel Online	31	4

No.	Judul	Panjang Judul	Jumlah Kata
16	Sistem Manajemen File Cloud	27	4
17	Aplikasi Penyewaan Mobil	24	3
18	Aplikasi Keuangan Bisnis	24	3
19	Aplikasi Membuat & Berbagi Resep Masakan	40	5
20	Sistem Manajemen Layanan Pelanggan	34	4

Selanjutnya berdasarkan data yang ada maka akan dilakukan Langkah-langkah berikut:

1. **Pemilihan Jumlah Cluster**
 Pada penelitian ini judul akan dikelompokkan menjadi tiga kelompok, maka akan ditentukan bahwa jumlah cluster yaitu 3
2. **Inisialisasi**
 Pada tahap ini dimulai dengan inisialisasi centroid secara acak. Mari pilih 3 klaster dan inisialisasikan centroid-nya:

Tabel 2 Inisialisasi Centroid

Centroid	Panjang Judul	Jumlah Kata Kunci
Centroid 1	28	4
Centroid 2	24	3
Centroid 3	31	4

3. **Clustering**
 Menghitung jarak antara setiap titik data dengan semua centroid, dan menentukan klaster yang memiliki centroid terdekat untuk setiap titik data. Misalkan kita menggunakan Euclidean distance untuk menghitung jarak. Berikut adalah hasil perhitungan:

Tabel 3 Clustering

No.	Judul	Panjang Judul	Jumlah Kata Kunci	Jarak ke Centroid 1	Jarak ke Centroid 2	Jarak ke Centroid 3	Terdekat	Klaster
1	Sistem Manajemen Toko Online	28	4	0	4.12	3	0	1
2	Aplikasi Pencatatan Keuangan Pribadi	36	4	8	12.04	5	5	3
3	Platform E-learning Interaktif	30	3	2.24	6	1.41	1.41	3
4	Aplikasi Pemesanan Makanan Online	33	4	5	9.06	2	2	3
5	Aplikasi Pengelolaan Jadwal Dokter	34	4	6	10.05	3	3	3
6	Sistem Informasi Geografis Pariwisata	37	4	9	13.04	6	6	3
7	Aplikasi Catatan Kesehatan Digital	34	4	6	10.05	3	3	3
8	Aplikasi Olahraga & Kesehatan	29	3	1.41	5	2.24	1.41	1
9	Aplikasi Manajemen Proyek	37	4	9	13.04	6	6	3

No.	Judul	Panjang Judul	Jumlah Kata Kunci	Jarak ke Centroid 1	Jarak ke Centroid 2	Jarak ke Centroid 3	Terdekat	Klaster
	Kolaboratif							
10	Platform Streaming Musik	24	3	4.12	0	7.07	0	2
11	Aplikasi Pembelajaran Bahasa Asing	34	4	6	10.05	3	3	3
12	Sistem Manajemen Inventaris Toko Retail	39	5	11.05	15.13	8.06	8.06	3
13	Aplikasi Pencarian Pekerjaan	28	3	1	4	3.16	1	1
14	Aplikasi Katalog Produk Fashion	31	4	3	7.07	0	0	3
15	Aplikasi Reservasi Hotel Online	31	4	3	7.07	0	0	3
16	Sistem Manajemen File Cloud	27	4	1	3.16	4	1	1
17	Aplikasi Penyewaan Mobil	24	3	4.12	0	7.07	0	2
18	Aplikasi Keuangan Bisnis	24	3	4.12	0	7.07	0	2
19	Aplikasi Membuat & Berbagi Resep Masakan	40	5	12.04	16.12	9.06	9.06	3
20	Sistem Manajemen Layanan Pelanggan	34	4	6	10.05	3	3	3

4. Perbarui Centroid

Menghitung rata-rata dari semua titik data dalam setiap kluster untuk menentukan posisi centroid yang baru.

Tabel 4 Perbarui Centroid

Centroid	Panjang Judul	Jumlah Kata Kunci
Centroid 1	28	3.5
Centroid 2	24	3
Centroid 3	34.61	4.07

5. Iterasi

Ulangi langkah 2-4 hingga tidak ada perubahan yang signifikan dalam posisi centroid atau titik data tidak berpindah kelompok.

Tabel 5 Hasil Clustering

No.	Judul	Kluster
1	Sistem Manajemen Toko Online	1
2	Aplikasi Pencatatan Keuangan Pribadi	3
3	Platform E-learning Interaktif	1
4	Aplikasi Pemesanan Makanan Online	3

No.	Judul	Kluster
5	Aplikasi Pengelolaan Jadwal Dokter	3
6	Sistem Informasi Geografis Pariwisata	3
7	Aplikasi Catatan Kesehatan Digital	3
8	Aplikasi Olahraga & Kesehatan	1
9	Aplikasi Manajemen Proyek Kolaboratif	3
10	Platform Streaming Musik	2
11	Aplikasi Pembelajaran Bahasa Asing	3
12	Sistem Manajemen Inventaris Toko Retail	3
13	Aplikasi Pencarian Pekerjaan	1
14	Aplikasi Katalog Produk Fashion	1
15	Aplikasi Reservasi Hotel Online	1
16	Sistem Manajemen File Cloud	1
17	Aplikasi Penyewaan Mobil	2
18	Aplikasi Keuangan Bisnis	2
19	Aplikasi Membuat & Berbagi Resep Masakan	3
20	Sistem Manajemen Layanan Pelanggan	3

Melalui iterasi langkah-langkah di atas, algoritma K-Means akan menghasilkan pengelompokan yang optimal berdasarkan titik data yang diberikan.

Setelah melakukan proses pengelompokan judul proyek aplikasi menggunakan algoritma K-Means, kami mendapatkan hasil berupa tiga kelompok utama sebagai berikut:

1. Kelompok 1 dengan judul proyek :
 - “Sistem Manajemen Toko Online”,
 - “Platform E-learning Interaktif “,
 - “Aplikasi Olahraga & Kesehatan”,
 - “Aplikasi Pencarian Pekerjaan”,
 - “Aplikasi Katalog Produk Fashion”,
 - “Aplikasi Reservasi Hotel Online”,
 - “Sistem Manajemen File Cloud”.
2. Kelompok 2 dengan judul proyek :
 - “Platform Streaming Musik”,
 - “Aplikasi Penyewaan Mobil“,
 - “Aplikasi Keuangan Bisnis”.
3. Kelompok 3 dengan judul proyek :
 - “Aplikasi Pencatatan Keuangan Pribadi”,
 - “Aplikasi Pemesanan Makanan Online“,
 - “Aplikasi Pengelolaan Jadwal Dokter”,
 - “Sistem Informasi Geografis Pariwisata”,
 - “Aplikasi Catatan Kesehatan Digital”,
 - “Aplikasi Manajemen Proyek Kolaboratif”,
 - “Aplikasi Pembelajaran Bahasa Asing”,
 - “Sistem Manajemen Inventaris Toko Retail”,
 - “Aplikasi Membuat & Berbagi Resep Masakan”,
 - “Sistem Manajemen Layanan Pelanggan”.

Dengan demikian, proses pengelompokan menggunakan algoritma K-Means memberikan cara yang efektif untuk mengorganisir judul proyek aplikasi ke dalam kategori yang sesuai dengan topiknya, yang dapat meningkatkan efisiensi dalam pengelolaan proyek perangkat lunak.

4. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan, penulis menyimpulkan bahwa hasil akhir penelitian adalah penelitian ini menunjukkan potensi penggunaan algoritma K-Means dalam pengelompokan judul proyek aplikasi. Dengan mengoptimalkan parameter-parameter yang relevan, algoritma ini dapat menjadi alat yang efektif dalam meningkatkan efisiensi dan akurasi pengelompokan judul proyek aplikasi, serta memberikan kontribusi yang signifikan dalam industri pengembangan aplikasi.

Daftar Pustaka

- [1] Rahayu, P. W., Sudipa, I. G. I., Suryani, S., Surachman, A., Ridwan, A., Darmawiguna, I. G. M., & Maysanjaya, I. M. D. (2024). Buku Ajar Data Mining. PT. Sonpedia Publishing Indonesia.
- [2] Geni, B. Y., Kurnia, O., Hayati, N., Thoriq, M., & Manurung, K. H. (2024). Analisa Algoritma K-Means Untuk Menentukan Strategi Marketing. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 8(1), 211-220..
- [3] Ishak, R., & Bengnga, A. (2024). Clustering Prestasi Akademik Lulusan Menggunakan Metode K-Means. *Jambura Journal of Electrical and Electronics Engineering*, 6(1), 76-81.
- [4] Rizki, U., Jakak, P. M., Prayogi, M. B., & Rahman, M. (2024). Penerapan Metode K-Means Clustering Dalam Pengembangan Strategi Promosi Berbasis Data Penerimaan Mahasiswa Baru (Studi Kasus: Universitas Nurul Huda). *Jurnal Nasional Ilmu Komputer*, 5(1), 25-43.
- [5] Maulana, M. I., Rahayudi, B., & Setiawan, N. Y. (2024). Analisis Pengelompokan Ulasan Pengguna menggunakan K-Means Clustering untuk Evaluasi Aplikasi My SAPK BKN. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 8(14).