

BAB I PENDAHULUAN

1.1 Latar Belakang

Teras Dakwah merupakan salah satu yayasan yang bergerak di bidang dakwah dan sosial kemasyarakatan di wilayah kota Yogyakarta. Di bidang dakwah, Teras Dakwah memiliki program mengaji dan mengkaji ilmu agama. Saat ini terdapat 3 program mengaji di Teras Dakwah, yaitu Ngaji Teras, Sekolah Hadits Bersanad (SHB), dan Ngaji Amida (*Akhwat* dan *Ummi* Muda). Sedangkan untuk bidang sosial kemasyarakatan, Teras Dakwah memiliki program sembako bahagia, *ambulance* gratis, layanan kesehatan untuk lansia, tanggap bencana dan lain sebagainya.

Bidang dakwah dan sosial kemasyarakatan memiliki hubungan erat. Teras Dakwah mengajak jamaah yang menghadiri program mengaji untuk memiliki kepekaan terhadap permasalahan sosial dengan berbagai program di bidang sosial kemasyarakatan dengan diadakannya sistem donatur tetap. Sistem donatur tetap bersifat sukarela dan ditawarkan kepada jamaah yang menghadiri program mengaji ketika mengisi formulir data diri singkat di bagian meja depan.

Target jamaah yang mengisi formulir data diri berasal dari Provinsi Yogyakarta. Namun, ada juga jamaah yang berasal dari luar Provinsi Yogyakarta, meskipun berjumlah sedikit dan kebanyakan berasal dari wilayah seperti Klaten, Solo dan Magelang. Data dari jamaah yang telah mengisi formulir kemudian dimasukkan secara manual kedalam *Microsoft Excel* dan disimpan sebagai database yang nantinya akan digunakan untuk menghubungi jamaah yang bersedia menjadi donatur tetap sebagai bentuk konfirmasi.

Selama ini, data donatur yang terkumpul hanya digunakan untuk mengkonfirmasi jamaah yang bersedia untuk menjadi donatur tetap saja. Belum ada

pengembangan lebih lanjut mengenai penanganan data yang dikumpulkan. Oleh sebab itu, peneliti hendak mencoba untuk menerapkan klusterisasi jamaah Ngaji Teras yang bersedia untuk menjadi donatur rutin.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah penulis paparkan di atas, maka didapatkan rumusan masalah adalah sebagai berikut:

Apakah K-Means dapat diterapkan untuk clustering jamaah Ngaji Teras di Teras Dakwah?

1.3 Batasan Masalah

Berdasarkan latar belakang yang sudah penulis paparkan di atas, maka didapatkan beberapa batasan masalah yang akan digunakan dalam penelitian ini adalah sebagai berikut:

1. Penelitian dilakukan di Teras Dakwah, Nitikan, Yogyakarta.
2. Data didapatkan melalui database Teras Dakwah yang telah dikonversi kedalam format (.csv)
3. Data yang digunakan hanya empat kali pertemuan Ngaji Teras.
4. Data yang diproses hanya untuk memetakan usia peserta Ngaji Teras yang bersedia menjadi donatur tetap.
5. Data yang diambil hanyalah data peserta yang berasal dari wilayah Provinsi Yogyakarta.

1.4 Tujuan Penelitian

Berdasarkan latar belakang yang sudah penulis paparkan diatas, maka di dapatkan tujuan dari penelitian ini adalah sebagai berikut:

1. Memberikan pengetahuan kepada pihak Teras Dakwah sebagai pertimbangan untuk pengembangan dikemudian hari.

2. Menambah wawasan pengetahuan peneliti mengenai penerapan dari algoritma K-Means untuk clustering.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini akan dibagi menjadi dua poin, berikut adalah pemaparannya :

1. Bagi Peneliti
 - 1.1 Menerapkan ilmu akademis yang telah di dapatkan selama masa kuliah.
 - 2.1 Menambah wawasan pengetahuan mengenai sub bidang Machine Learning, klustering menggunakan algoritma K-Means.
 - 3.1 Menambah wawasan pengetahuan mengenai cara pengoptimalan algoritma K-Means untuk clustering dengan metode evaluasi silhouette score.
2. Bagi Teras Dakwah
 - 1.1 Mengetahui pemetaan usia jamaah Ngaji Teras yang bersedia untuk menjadi donatur tetap.
 - 2.1 Mendapatkan pengetahuan maupun wawasan baru untuk pengembangan maupun evaluasi bagi Teras Dakwah kedepannya, khususnya di bidang dakwah.
 - 3.1 Mendapatkan pengetahuan maupun wawasan dari hasil cluster terkait penentuan strategi yang akan digunakan dalam melakukan pendekatan dan perawatan jamaah yang bersedia untuk menjadi donatur tetap maupun yang belum bersedia.

1.6 Metode Penelitian

1.6.1 Metode Pengumpulan Data

- a) Studi Literatur

Studi literatur merupakan proses dikumpulkannya teori-teori dari paper, jurnal, hasil penelitian serupa dan buku mengenai algoritma *K-means Clustering*.

b) *Data retrieval*

Dataset diambil dari database Teras Dakwah Ngaji Teras dengan jamaah yang pernah mengisi formulir dan telah dimasukkan kedalam format (.csv)

c) *Interview*

Pengambilan data berupa keterangan dan penjelasan mengenai kehadiran peserta dan pentingnya ketersediaan jamaah Ngaji Teras sebagai donatur tetap. Interview ini hanya dilakukan kepada pengurus Teras Dakwah.

1.6.2 Metode Analisis

1. *Data retrieval*

Pengumpulan dataset dilakukan dengan mengakses database Ngaji Asyik yang berisi data jamaah Ngaji Asyik yang telah di komputerisasi kedalam format (.csv)

2. *Preprocessing*

Tahapan untuk mempersiapkan dataset sebelum di proses ke dalam algoritma yang akan digunakan. Fungsi dari preprocessing data, salah satunya yaitu untuk memastikan agar dataset dalam kondisi bersih sehingga ketika memasuki tahap clustering tidak muncul kesalahan akibat dataset mentah. Pembersihan dataset dilakukan mulai dari menseleksi kolom yang tidak relevan dengan penelitian, menghapus data yang duplikat, memastikan tidak ada data yang kosong dan yang terakhir memastikan bahwa dataset yang akan digunakan sudah dalam

format numerik. Karena pemrosesan data hanya bisa dilakukan apabila nilai dari seluruh variable bernilai numerik.

3. Analisis

- a) Menentukan jumlah *cluster* yang akan digunakan.
- b) Menentukan titik centroid secara acak sebanyak *cluster*.
- c) Memberikan label pada tiap *cluster* yang berdekatan dengan titik centroid.
- d) Menentukan titik centroid baru berdasarkan *cluster* yang telah diberikan label.
- e) Memberikan label ulang kepada tiap *cluster* berdasarkan titik centroid baru apabila terdapat perubahan atau pergeseran jarak data dengan pusat centroid.
- f) Menggunakan metode *silhouette score* untuk melakukan evaluasi terhadap jumlah *cluster* yang telah ditentukan.

4. Implementasi

Implementasi dilakukan dengan menggunakan bahasa python dengan *library* yang dibutuhkan di *Google colab*.

1.6.3 Evaluasi

Evaluasi dari jumlah *cluster* dilakukan menggunakan metode *Silhouette Score* untuk mengetahui apakah jumlah *cluster* yang telah ditentukan di awal telah optimal atau belum.

1.7 Sistematika Penulisan

Sistematika penulisan dari penelitian ini dijelaskan dengan detil dalam beberapa bab. Penjelasannya adalah sebagai berikut:

A. BAB I PENDAHULUAN

Bab yang berisi latar belakang, masalah, rumusan masalah, batasan masalah, maksud dan tujuan penelitian, manfaat penelitian, metodologi penelitian dan sistematika penulisan.

B. BAB II LANDASAN TEORI

Bab yang berisi tinjauan pustaka dan teori-teori pendukung dalam penyusunan implementasi algoritma *K-Means* untuk *clustering* jamaah Ngaji Teras di Teras Dakwah.

C. BAB III METODE PENELITIAN

Bab yang berisi Langkah-langkah yang dilakukan dalam proses implementasi algoritma *K-means* kepada dataset yang telah di dapatkan.

D. BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi tentang hasil dari *preprocessing* data, proses implementasi program, evaluasi dari implementasi yang telah diterapkan.

E. BAB V PENUTUP

Bab ini membahas tentang kesimpulan setelah dilakukannya penelitian serta saran dari peneliti guna memaksimalkan peneliti selanjutnya yang akan mengambil topik serupa.