

BAB III METODE PENELITIAN

3.1. Objek Penelitian

Objek Penelitian ini yaitu jumlah kunjungan wisatawan berdasarkan asal daerah wilayah kabupaten Karawang di objek wisata alam Puncak Sempur Karawang tahun 2021. Data yang didapatkan berupa hasil wawancara yang akan dikonversikan dalam bentuk tabel data, agar memudahkan dalam pengaplikasiannya ke dalam *data mining*. Setelah dikonversikan, data tersebut akan diproses dengan menggunakan metode *clustering data mining* dengan algoritma K-Means dan K-Medoid. Sampai nantinya akan ditemukan daerah mana saja yang dominan mengunjungi objek wisata alam Puncak Sempur.

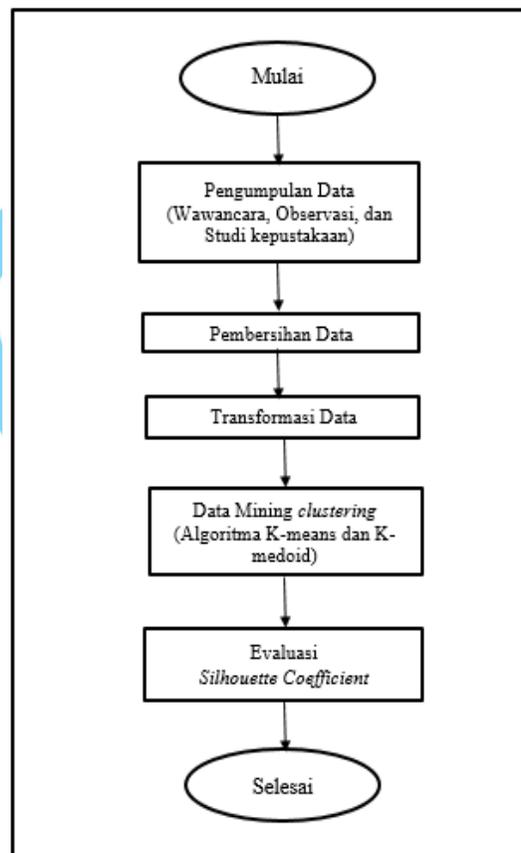
Lokasi penelitian dilaksanakan di objek wisata alam Puncak Sempur Kabupaten Karawang. Puncak Sempur yang merupakan salah satu kaki dari gunung yang berada di kabupaten Karawang yaitu Gunung Sanggabuana, mayoritas pengunjung pasti akan melakukan kemah atau *camping* ceria, karena terdapat lahan yang luas dan wisatawan akan mendapatkan suguhan *sunset* yang indah. Wisatawan pun ada yang berkunjung hanya sekedar mengunjungi kuliner yang ada di puncak sempur, berburu spot foto yang instagramable, dan curug ciomas yang berada di puncak sempur. Penelitian ini dilaksanakan sejak bulan Desember 2021. Berikut adalah tabel rincian penelitian :

| No | Item | 2021 | 2022 | | | | | |
|----|---|------|------|-----|-----|-----|-----|-----|
| | | Des | Jan | Feb | Mar | Apr | Mei | Jun |
| 1 | Studi Literatur | ■ | ■ | ■ | ■ | ■ | ■ | ■ |
| 2 | Pengumpulan Data | ■ | ■ | ■ | ■ | ■ | ■ | ■ |
| 3 | Pembersihan Data | ■ | ■ | ■ | ■ | ■ | ■ | ■ |
| 4 | Transformasi Data | ■ | ■ | ■ | ■ | ■ | ■ | ■ |
| 5 | Proses <i>Data mining</i> <i>Clustering</i> Algoritma k-means dan k-medoid | ■ | ■ | ■ | ■ | ■ | ■ | ■ |
| 6 | Evaluasi | ■ | ■ | ■ | ■ | ■ | ■ | ■ |

Gambar 3 1 Rincian Penelitian

3.2. Prosedur Penelitian

Alur penelitian yang dilakukan yaitu dimulai dari pengumpulan data, analisis data, *clustering data mining* yang terakhir pengujian dan mendapatkan hasil yang di inginkan. Berikut diagram prosedur penelitian :



Gambar 3 2 Flowchart Tahapan Penelitian

3.2.1. Pengumpulan Data

Tahap pertama dari penelitian ini yaitu mengumpulkan data dari sumber data atau objek yang ditentukan. Seperti yang sudah disebutkan sebelumnya data dalam penelitian ini yaitu jumlah kunjungan wisatawan ke objek wisata alam puncak sempur. Tahapan ini terdiri dari 3 metode yaitu :

1. **Wawancara**, tahapan wawancara ini dilakukan kepada Kepala pengelola objek wisata puncak sempur, guna mengetahui jumlah pengunjung asal

Kabupaten Karawang pada tahun 2021. Berdasarkan hasil wawancara terhadap kepala pengelola wisata alam Puncak sempur, data kunjungan wisatawan ada 1.797 pengunjung. Pengunjung terdiri dari yang berkemah dan tidak berkemah. Hasil wawancara lainnya yaitu, pernyataan dari Bapak wawan sebagai pengelola wisata alam puncak sempur, bahwa puncak sempur berada di ketinggian 712 mdpl dengan luas tanah 300 Ha dan wisatawan sudah dapat menikmati keindahan kabupaten Karawang. “Puncak sempur yang berada di desa Cintelaksana, Kecamatan Tegalwaru, Kabupaten Karawang ini sudah resmi menjadi objek wisata alam sejak tahun 2015, yang diresmikan oleh dr. Cellica Nurrachdiana selaku Bupati kabupaten Karawang” ujar Bapak Yofa kepala pengelola objek wisata puncak sempur.

2. **Observasi**, mengunjungi langsung objek wisata alam puncak sempur agar dapat mengetahui kondisi asli yang bisa menjadi faktor dalam pengelompokan penelitian ini. Menurut hasil dari observasi yang telah dilakukan peneliti Fasilitas yang ada di puncak sempur sudah cukup memadai mulai dari warung kopi, musholla, toilet, sampai dengan tempat penyewaan alat *camping*, sehingga untuk wisatawan yang ingin secara mendadak berkemah atau tidak mempunyai alat kemah, akan terbantu oleh tempat penyewaan alat kemah yang disediakan oleh pengelola puncak sempur.
3. **Studi kepustakaan**, wisata alam yang bisa dikatakan *viral* di Kabupaten Karawang ini memberi ide kepada peneliti untuk mengetahui sejauh manakah puncak sempur ini diketahui oleh masyarakat di kabupaten Karawang. *Clustering* asal daerah wisatawan yang di dapat dari buku kunjungan pada tahun 2021 dan akan dijadikan Data set oleh peneliti, untuk mengetahui seberapa jauhnya objek wisata ini di kenal oleh wisatawan, dengan membagi menjadi 3 *cluster* yaitu C1 untuk Ramai pengunjung, C2 untuk sedang pengunjung, dan C3 untuk sepi pengunjung. Membaca jurnal atau penelitian yang telah ada sebelumnya dengan *browsing* di internet menjadi salah satu metode pengumpulan data dalam penelitian ini karena dapat menambah data pendukung lainnya.

3.2.2. Pembersihan Data

Banyak data yang tidak diperlukan saat dilakukannya pengumpulan data, dengan cara memisahkannya dengan manual menggunakan Microsoft Excel 2013 antara data yang dibutuhkan atau tidak untuk tahap selanjutnya. Data dibersihkan dalam bentuk data sampah, kurang nya pengejaan sampai dengan melengkapi data yang kurang lengkap.

Data yang diperlukan dalam penelitian ini hanyalah jumlah kunjungan wisata dan asal daerah Kabupaten Karawang, sehingga data lainnya dapat di hapus atau di bersihkan agar data menjadi efektif saat proses perhitungan. Berikut beberapa *sample* data sebelum dan sesudah dibersihkan dan keseluruhan data terdapat pada lampiran.

Gambar 3 4 Data sesudah dibersihkan

| DATA KUNJUNGAN WISATA PUNCAK SEMPUR KABUPATEN KARAWANG 2021 | | | | | | | |
|---|------------------|-----------------|----------------|---------------|--------|------------|-------------------|
| NO | TANGGAL | NAMA KETUA | JUMLAH ANGGOTA | ALAMAT | TUJUAN | | Keterangan |
| | | | | | CAMP | TIDAK CAMP | |
| 1 | 01-Januari-2021 | Pungki | 120 orang | LOJI | ✓ | | Archiless Goes |
| 2 | 02-Januari-2021 | ucay | 3 orang | CACABAN | ✓ | | |
| 3 | 02-Januari-2021 | adzuar kumar | 2 orang | CIDORO | | ✓ | |
| 4 | 04-Januari-2021 | ahmad julpa | 5 orang | BADAMI | ✓ | | |
| 5 | 04-Januari-2021 | ilmi Nurhidayah | 5 orang | BOGOR | ✓ | | |
| 6 | 04-Januari-2021 | zikri | 2 orang | SUKASARI | | ✓ | |
| 7 | 10-Januari-2021 | anggi susena | 2 orang | LOJI | | ✓ | |
| 8 | 11-Januari-2021 | ade | 3 orang | LOJI | ✓ | | |
| DATA KUNJUNGAN WISATA PUNCAK SEMPUR KABUPATEN KARAWANG 2021 | | | | | | | |
| NO | TANGGAL | NAMA KETUA | JUMLAH ANGGOTA | ALAMAT | CAMP | TIDAK CAMP | KECAMATAN |
| 10 | 18-Januari-2021 | rosid | 1 orang | BOGOR | ✓ | | |
| 11 | 19-Januari-2021 | Edi | 6 orang | TELUK JAMBE | ✓ | | |
| 12 | 01-Februari-2021 | Edi | 6 orang | TELUK JAMBE | ✓ | | |
| 1 | 13-Februari-2021 | fahruzi | 3 orang | CIKARANG | ✓ | | Tegalwaru |
| 14 | 01-Maret-2021 | Wiwit | 2 orang | Loji | | | Tegalwaru |
| 2 | 15-Maret-2021 | adzuar kumar | 2 orang | Cidoro | ✓ | | Tegalwaru |
| 3 | | Ahmad Julpa | 5 orang | Gambus | | | Teluk Jambe Barat |
| 4 | | Zikri | 2 orang | Sukasari | | | Purwasari |
| 5 | | Anggi Susena | 2 orang | Loji | | | Tegalwaru |
| 6 | | Ade | 3 orang | Loji | | | Tegalwaru |
| 7 | | Linda Sovia | 2 orang | Sukasari | | | Purwasari |
| 8 | | Edi | 6 orang | Teluk Jambe | | | Teluk Jambe Barat |
| 9 | | Wiwit | 2 orang | Loji | | | Tegalwaru |
| 10 | | Nandar | 4 orang | Klari | | | Klari |
| 11 | | Nikmah | 5 orang | Badami | | | Teluk Jambe Timur |
| 12 | | Caca | 1 orang | Tempuran | | | Tempuran |
| 13 | | Syamsiyah | 6 orang | Teluk Jambe | | | Teluk Jambe Barat |
| 14 | | Sandi | 5 orang | Parakan Badak | | | Tegalwaru |
| 15 | | Eziz | 4 orang | Tegal Waru | | | Tegalwaru |

3.2.3. Transformasi Data

Mengubah atau merombak data sehingga data sesuai. Fitur ditransformasi berguna untuk menyajikan data tergantung tujuan yang ingin dicapai dan prosesnya ubah data yang dipilih agar sesuai dengan data dengan proses *data mining*.

3.2.4. *Data mining Clustering* menggunakan algoritma k-means dan k-medoid

Penelitian ini diselesaikan dengan dua model algoritma *clustering* yaitu algoritma k-means dan algoritma k-medoid, di mana hasilnya yang akan dibandingkan dan akan mengetahui tingkat akurasi yang pas pada penelitian ini. Jadi semua data yang sudah melalui tahapan-tahapan sebelumnya akan dihitung atau dianalisis dengan 2 algoritma tersebut. Software yang digunakan dalam proses *data mining* ini adalah Rapidminer dan Jupyter Lab yang sudah terinstall pada laptop dengan spesifikasi Processor Intel ® Celeron® N4000, CPU @ 1.10Ghz (2 CPUs), 1.1Ghz, dan RAM 4.00 GB.

A. Proses *Clustering* Algoritma K-means

Input data yang berupa jumlah kunjungan wisatawan di objek wisata puncak sempur. Proses kedua buat 3 *clustering* terlebih dahulu dengan simbol C1(Ramai pengunjung), C2(Sedang pengunjung), dan C3(Sepi pengunjung). Proses ketiga menentukan pusat kelompok, titik pusat / *centroid* yang di tentukan secara acak. Proses keempat menghitung jarak terdekat dari masing-masing data ke setiap *centroid*/titik pusat, jika sudah terbentuk data baru dari iterasi 1 maka hitung kembali *centroid* untuk menentukan kembali *centroid* baru. Lalu buat iterasi 2 dengan menghitung kembali jarak terdekat masing-masing data ke setiap *centroid* atau titik pusat. jika hasil *cluster* sama atau tidak berubah maka proses dihentikan atau selesai, namun jika berubah hasil *cluster* maka lakukan kembali proses mulai dari menghitung jarak terdekat data ke setiap *centroid*. Output nya akan terlihat berupa 3 kelompok yang terkategori Ramai pengunjung, Sepi pengunjung, dan Sedang pengunjung.

B. Proses *clustering* Algoritma K-medoid

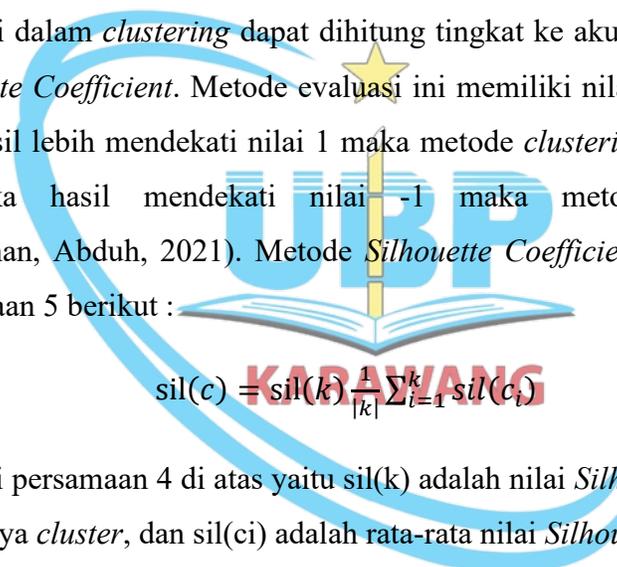
Input data yang berupa jumlah kunjungan wisatawan di objek wisata alam puncak sempur. Kedua pembuatan pusat medoid, penelitian ini menggunakan 3 *cluster* yaitu C1, C2 dan C3. Ketiga pilih masing-masing *centroid* sebagai kandidat medoid baru. Keempat

hitung jarak antar objek dengan masing-masing *centroid*. Kelima hitung total simpangan (S) dengan menghitung nilai total *distance* baru – total *distance* lama. Jika $S < 0$, maka tukar objek dengan data *cluster* untuk membentuk sekumpulan k objek baru sebagai medoid. Lakukan point ke 4 dan ke 5 hingga hasil *cluster* tidak berubah, maka akan didapatkan output berupa 3 kelompok yaitu Ramai pengunjung, Sepi pengunjung, dan Sedang pengunjung.

3.2.5. Evaluasi Pola

Pada tahap ini data yang sudah terlihat hasilnya akan dievaluasi jika hasil sudah seperti yang diharapkan maka akan lanjut ke tahap selanjutnya, namun jika belum maka tahapan kembali lagi ke transformasi data lalu dilakukannya kembali proses *data mining*.

Evaluasi dalam *clustering* dapat dihitung tingkat ke akurasiannya menggunakan *Silhouette Coefficient*. Metode evaluasi ini memiliki nilai dari -1 hingga 1 yang dimana jika hasil lebih mendekati nilai 1 maka metode *clustering* lebih bagus juga dan sebaliknya jika hasil mendekati nilai -1 maka metode *clustering* sangat buruk (Paembunan, Abduh, 2021). Metode *Silhouette Coefficient* ini dapat di terapkan dengan persamaan 5 berikut :


$$sil(c) = \frac{sil(k)}{|k|} \sum_{i=1}^k sil(c_i) \quad (5)$$

Keterangan dari persamaan 4 di atas yaitu $sil(k)$ adalah nilai *Silhouette* semua *cluster*, $|k|$ adalah banyaknya *cluster*, dan $sil(c_i)$ adalah rata-rata nilai *Silhouette*.