

BAB I

PENDAHULUAN

1.1. Latar Belakang

Perguruan tinggi merupakan jenjang pendidikan setelah pendidikan menengah atas atau SMA sederajat, perguruan tinggi mencakup program pendidikan diplomat, sarjana, magister, dan doktor yg diselenggarakan oleh perguruan tinggi. Dengan adanya perguruan tinggi maka semakin meningkatlah kualitas tenaga kerja. Dan salah satu cara untuk meningkatkan kualitas tenaga kerja yaitu melihat kinerja perguruan tinggi tersebut (Amri et al., 2019). Direktur Jenderal Pendidikan Tinggi menjelaskan bahwa pentingnya klasterisasi perguruan tinggi untuk menyediakan landasan bagi pengembangan kebijakan pembangunan, pembinaan perguruan tinggi serta untuk mendorong perguruan tinggi dalam meningkatkan kualitas pelaksanaan tridharma perguruan tinggi secara berkelanjutan. Selain itu, klasterisasi perguruan tinggi berfungsi untuk menyediakan informasi kepada masyarakat umum tentang kualitas kinerja perguruan tinggi di Indonesia. Dalam klasterisasi yang dilakukan pemerintah hanya menggunakan metode survey dan pemberian nilai (skor) klasterisasi berdasarkan capaian perguruan tinggi terhadap setiap indikator pada masing-masing aspek (Neni Herlina, 2020).

Pemeringkatan atau perangkingan perguruan tinggi oleh Webometrics memiliki tujuan sistem untuk menyediakan informasi yang handal, akurat, multidimensi, dan bermanfaat kepada masyarakat tentang kinerja perguruan tinggi di seluruh Indonesia (Hermawan, 2018). Yang kemudian data tersebut akan penulis kembangkan lagi dengan cara klasterisasi hasil dari perangkingan perguruan tinggi Webometrics kedalam 3 klaster yaitu teratas, menengah, dan terendah. Contohnya perguruan a,b, dan c masuk kedalam klaster peringkat teratas, perguruan tinggi d,e, dan f masuk kedalam klaster peringkat menengah, dan perguruan tinggi g,h, dan i kedalam klaster peringkat terendah.

Tujuan penelitian ini yakni untuk mengolah data tersebut dan dibutuhkan suatu teknik agar hasil pengolahan atau informasi yang dihasilkan tepat guna. Salah satu teknik yang bisa dimanfaatkan yaitu data mining. Data mining merupakan proses menemukan pola serta pengetahuan yang menarik dari suatu data dalam jumlah yang besar. Data mining mendeskripsikan sebuah pengumpulan teknik-teknik dengan tujuan untuk menemukan pola-pola yang tidak diketahui pada data yang sudah dikumpulkan (Nabila et al., 2021). Tahapan proses dalam data mining terdapat *Clustering* atau analisis pengelompokan. Proses ini ialah membagi data pada suatu himpunan ke pada beberapa grup yang mempunyai kesamaan. Kesamaan ini wajib dikelompokkan dalam suatu bagian yang lebih besar dari kesamaan data menggunakan data pada kelompok lain. *Clustering* adalah salah satu alat bantu pada data mining. *Clustering* bertujuan untuk mengelompokkan objek ke dalam *cluster*. *Cluster* adalah sekelompok atau kumpulan objek data yang serupa. Data dikelompokkan ke dalam *cluster* yang sama, dan setiap *cluster* dibedakan oleh ketidakmiripannya dengan objek dalam *cluster* yang berbeda. Potensi *Clustering* bisa dipergunakan untuk mengetahui struktur pada data. Algoritma untuk melakukan *cluster* dalam penelitian ini merupakan K-Means. K-Means adalah algoritma yang digunakan untuk mengelompokkan secara partisi. Setiap kelompok membagi data ke dalam kelompok yang berbeda. Dalam K-Means menerapkan metode data *Clustering non hierarchical*. Metode ini membagi data menjadi satu atau lebih *cluster*/kelompok. Dalam K-Means menerapkan algoritma dalam memisahkan data ke dalam kelompok yang berbeda-beda dengan melakukan partisi. jarak antara data ke *clusternya* harus diminimalisir.

Algoritma K-Means melakukan *Clustering* yang bergantung pada data yang didapat dan konklusi yang ingin dicapai pada akhir proses. Sehingga pada penggunaan algoritma K-Means ada aturan yaitu berapa jumlah *cluster* yang dibutuhkan, hal ini karena K-Means memiliki kemampuan untuk mengelompokkan data pada jumlah yang

cukup besar dengan saat komputasi yang cukup cepat dan efisien (Dewi et al., 2019). Pemanfaatan algoritma ini diujikan pada penggalian ilmu pengetahuan dari database perangkingan perguruan tinggi pada situs web Webometrics serta dikelompokkan kedalam beberapa kluster. Dengan Algoritma *K-Means Clustering* bisa memberikan informasi mengenai kluster perguruan tinggi yang ada di Indonesia.

Berdasarkan latar belakang di atas mendorong peneliti untuk mengambil penelitian tersebut agar menjadi masukan berupa informasi kepada pemerintah serta masyarakat dan supaya penelitian ini dapat di kembangkan lagi oleh peneliti lainnya.

1.2. Rumusan Masalah

Adapun rumusan masalah dari penelitian ini yaitu:

1. Bagaimana menganalisis data perangkingan Webometrics agar dapat mengetahui pengelompokkan perguruan tinggi di Indonesia?
2. Bagaimana penerapan data mining untuk *Clustering* data ranking Webometrics menggunakan algoritma *K-Means*?

1.3. Tujuan Penelitian

Adapun tujuan penelitian ini yaitu:

1. Mengimplementasikan Algoritma *K-Means Clustering* untuk menganalisis data perangkingan Webometrics guna mengetahui pengelompokkan perguruan tinggi di Indonesia.
2. Memberikan informasi secara detail/terperinci serta perhitungan secara umum menggunakan algoritma *K-Means* mengenai pengelompokkan data perguruan tinggi di Indonesia.

1.4. Batasan Penelitian

Adapun batasan dari penelitiannya yaitu:

1. Data yang digunakan yaitu data perguruan tinggi di Indonesia berdasarkan perankingan Webometrics periode Juli 2022 sejumlah 2624 perguruan tinggi.
2. Atribut yang digunakan yaitu *impact rank*, *openness rank*, dan *excellent rank*.
3. Dikelompokkan menjadi 3 *cluster* yaitu C0 sebagai peringkat rendah, C1 sebagai peringkat atas, dan C2 sebagai peringkat menengah.
4. Analisis data menggunakan bahasa pemrograman *Python*.

1.5. Manfaat Penelitian

1. Dapat mengetahui *cluster* perguruan tinggi dari peringkat teratas hingga terendah.
2. Dapat memberikan informasi perhitungan secara umum dan secara detail mengenai pengelompokkan perguruan tinggi kepada masyarakat/khalayak ramai.
3. Dapat bermanfaat sebagai bahan *referensi* bagi penelitian selanjutnya atau bagi pihak-pihak yang akan melakukan penelitian klasterisasi perguruan tinggi dan juga Algoritma *K-Means Clustering* serta sebagai literatur untuk menambah ilmu pengetahuan.