

## BAB I

### PENDAHULUAN

#### 1.1 Latar Belakang

Pendidikan yaitu sebuah proses pembelajaran bagi setiap individu untuk mencapai pengetahuan dan pemahaman yang lebih tinggi mengenai suatu objek dan spesifikasi tertentu (KBBI). Kementerian Pendidikan dan Kebudayaan (Kemendikbud) mewajibkan pendidikan di Indonesia selama 12 tahun dari SD, SMP, dan SMA. Mutu pendidikan SMA harus selalu di tingkatkan sehingga mampu menciptakan lulusan yang berkualitas karena pada tahapan inilah peserta didik sudah mulai di bentuk kemampuan individunya sesuai tujuan .

Pada tahun 2020, tercatat bahwa SMA di Kabupaten Kediri baik negeri maupun swasta berjumlah SMA 47 sekolah (Cabang Dinas Pendidikan Wilayah Kediri). Setiap sekolah memiliki karakteristik yang berbeda. Data yang cukup banyak serta adanya perbedaan karakteristik tiap sekolah menjadikan sekolah SMA di Kabupaten Kediri perlu untuk di kelompokkan. Pengelompokan atau sering disebut clustering merupakan sebuah metode di dalam data mining yang bertujuan untuk mengelompokkan data dengan karakteristik yang sama kedalam satu kelompok dan data dengan karakteristik berbeda ke dalam kelompok lainnya.

Penelitian pengelompokan data campuran sudah beberap kali di lakukan, Ichsan (2018) melakukan penelitian dengan judul pengelompokan Kabupaten/Kota di Jawa Timur berdasarkan pembangunan kualitas sumberdaya manusia dan pembangunnn ekonomi dengan menggunakan metode *ensemble* ROCK.

Prakoso (2017) melakukan pengelompokan SMA di Sidoarjo menggunakan *similarity weight and filter method* (SWFM). Alvionita (2017) melakukan penelitian metode *ensemble* ROCK dan SWFM untuk mengelompokan data campuran numerik dan kategorik pada kasus aksesori jeruk dan hasil penelitian menunjukan bahwa metode *ensemble* ROCK memberikan kinerja pengelompokan lebih baik dari pada metode *ensemble* SWFM.

Secara umum metode dalam analisis kluster dibagi dua yakni Hierarchical Clustering (metode hirarki) dan Non-Hierarchical Clustering (metode tak hirarki). Pada penelitian ini akan menggunakan metode Non-Hierarchical Clustering pada pengelompokan data numerik. Metode ini merupakan metode yang jumlah clusternya ditentukan terlebih dahulu, sehingga objek-objek akan dikelompokkan pada  $k$  kelompok yang telah ditentukan oleh peneliti. Metode Non-Hierarchical Clustering yang akan digunakan adalah algoritma K-Mean dan K-Medoid.

Pengelompokan data kategorik akan menggunakan metode *RObust Clustering using linKs* (ROCK). Metode ROCK menggunakan konsep link sebagai ukuran kemiripan untuk membentuk cluster-nya. Metode ROCK dapat menangani outlier dengan cukup efektif. Tahapan selanjutnya metode ROCK akan digunakan kembali pada final cluster untuk penggabungan cluster-cluster yang telah dihasilkan melalui kedua algoritma tersebut dengan menganggapnya sebagai data baru bertipe kategorik.

## 1.2 Rumusan Masalah

Rumusan masalah yang akan dibahas dalam penelitian ini adalah :

- a. Bagaimana karakteristik data faktor-faktor sarana prasarana dan

SDM pada SMA di Kabupaten Kediri ?

- b. Bagaimana pengelompokan menggunakan *ensembel Robust clustering using links (ROCK)* pada faktor-faktor sarana prasarana dan SDM pada SMA di Kabupaten Kediri ?
- c. Bagaimana validitas hasil pengelompokan menggunakan *ensembel Robust clustering using links (ROCK)* pada faktor-faktor sarana prasarana dan SDM pada SMA di Kabupaten Kediri ?

### 1.3 Tujuan Penelitian

- a. Mendeskripsikan karakteristik data faktor-faktor sarana prasarana dan SDM pada SMA di Kabupaten Kediri.
- b. Mengelompokkan faktor-faktor sarana prasarana dan SDM pada SMA di Kabupaten Kediri menggunakan *ensembel Robust clustering using links (ROCK)*.
- c. Menghitung validitas hasil pengelompokan menggunakan *ensembel Robust clustering using links (ROCK)* pada faktor-faktor sarana prasarana dan SDM pada SMA di Kabupaten Kediri.

### 1.4 Manfaat Penelitian

#### 1. Manfaat Teoritis

Hasil dari penelitian ini dapat menjadi rujukan dalam mengembangkan penelitian lebih lanjut. Selain itu, penelitian ini dapat menjadi sebuah nilai tambah khasanah pengetahuan ilmiah di Indonesia.

## 2. Manfaat Praktis

### a. Bagi penulis

Mengetahui dan dapat mengaplikasikan metode pengelompokan data khususnya cluster ensemble ROCK.

### b. Bagi pemerintah

Dengan hasil pengelompokan ini diharapkan dapat membantu pemerintah dalam mempersiapkan kebijakan- kebijakan yang akan di ambil khususnya pada pendidikan SMA dengan melihat karakteristik kelompok SMA tersebut.

### c. Bagi masyarakat

Dapat mengetahui tentang pengelompokan cluster *ensembl* ROCK dan dari menelitian dapat menjadikan pengetahuan, serta dapat digunakan untuk referensi bagi penelitian selanjutnya.

## 1.5 Batasan masalah

Batasan masalah dalam penelitian ini adalah :

- a. Membahas uji pengelompokan dengan metode *cluster ensembl* ROCK.
- b. Penelitian ini dilakukan untuk mengelompokan SMA di Kabupaten Kediri dengan indikator sarana prasarana dan SDA menggunakan 8 variabel.
- c. Data yang digunakan adalah data indikator sarana prasarana dan SDA pada SMA di Kabupaten Kediri tahun 2020.

- d. Penelitian ini menguji seluruh SMA di Kabupaten Kediri baik negeri maupun swasta kecuali SMAN 5 Taruna Brawijaya dan SMAS Muhammadiyah dikarenakan data sekolah tersebut tidak tersedia.



