

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Indonesia merupakan negara yang mempunyai struktur alam yang terdiri dari pertemuan lempeng-lempeng tektonik yaitu Eurasia, Indo-Australia dan Pasifik. Hal tersebut menjadikan kawasan Indonesia memiliki kondisi geologi yang sangat kompleks. Kondisi ini mengakibatkan banyak daerah-daerah di Indonesia yang sangat rawan terhadap bencana alam (Indriasari & Sidhi, 2011). Salah satu daerah yang rawan terjadi bencana alam adalah Provinsi Sumatera Selatan. Daerah-daerah rawan bencana di Provinsi Sumatera Selatan hampir semuanya berada di daerah yang tingkat populasi sangat tinggi. Selain bencana alam yang disebabkan oleh faktor geologi, terdapat juga bencana yang dipicu oleh kerusakan alam akibat tindakan manusia seperti banjir dan tanah longsor.

Dengan kondisi geologi seperti itu diperlukan adanya informasi terkait status kebencanaan suatu daerah khususnya di Provinsi Sumatera Selatan. Informasi tersebut dapat berupa pengelompokan daerah rawan bencana sehingga dapat berguna untuk pengambilan keputusan dalam proses penanggulangan bencana. Salah satu cara untuk pengelompokan daerah rawan bencana dapat dilakukan dengan menerapkan metode data mining.

Data mining adalah suatu proses ekstraksi atau penggalian data yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari *database* yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting (Connolly & Begg, 2010). Data mining berfungsi untuk proses pengambilan keputusan dari volume data yang besar yang disimpan dalam *database*, *data warehouse*, atau informasi yang disimpan dalam repositori. Berdasarkan polanya data mining dikelompokkan menjadi deskripsi, estimasi, prediksi, klasifikasi, clustering, dan asosiasi (Larose, 2005). Data Mining adalah inti dari proses *Knowledge Discovery in Database* (KDD), yang melibatkan algoritma dalam mengeksplorasi data, mengembangkan model dan menemukan pola yang sebelumnya tidak diketahui. KDD adalah proses

terorganisir untuk mengidentifikasi pola yang valid, baru, berguna, dan dapat dimengerti dari sebuah dataset yang besar dan kompleks (Maimon, 2010).

Terdapat banyak algoritma yang digunakan untuk melakukan data mining salah satunya adalah algoritma *K-Means*. Algoritma *K-Means* adalah salah satu algoritma yang digunakan untuk mengklasterisasi atau mengelompokkan data ke dalam satu *cluster*/kelompok, dimana data yang mempunyai karakteristik yang sama dimasukkan dalam satu kelompok dan data yang memiliki karakteristik berbeda akan dimasukkan ke dalam kelompok yang lain (Oscar, 2013).

Dari hasil uraian di atas maka peneliti tertarik untuk melakukan pengelompokan daerah rawan bencana di Provinsi Sumatera Selatan dengan menerapkan data mining dan algoritma *K-Means*. Sehingga dengan adanya penelitian ini diharapkan mampu memberikan informasi berupa kelompok daerah rawan bencana berdasarkan tingkat bahayanya sehingga dapat menjadi masukan bagi pemerintah dalam proses penanggulangan bencana di Provinsi Sumatera Selatan dengan lebih efisien.

Berdasarkan latar belakang tersebut, maka diangkat judul “Penerapan Data Mining untuk Pengelompokan Daerah Rawan Bencana di Sumatera Selatan Menggunakan Algoritma K-Means”.

## **1.2 Rumusan Masalah**

Adapun rumusan masalah dalam penelitian berdasarkan uraian latar belakang adalah bagaimana hasil dari pengelompokan daerah rawan bencana di Sumatera Selatan dengan menggunakan algoritma *K-Means*?

## **1.3 Batasan Masalah**

Supaya tidak menyimpang dari tujuan penelitian, maka perlu dibuat batasan masalah di atas, yaitu sebagai berikut:

1. Data yang digunakan adalah data bencana Sumatera Selatan tahun 2014-2019.

2. Data yang digunakan merupakan data dari Badan Penanggulangan Bencana Daerah (BPBD) Provinsi Sumatera Selatan.
3. Data yang diolah dalam penelitian ini menggunakan algoritma *K-Means*.
4. *Tool* yang dipakai dalam penerapan data mining ini adalah *RapidMiner*.
5. Jumlah *cluster* yang digunakan pada penelitian ini yaitu sebanyak 3 *cluster*.

#### **1.4 Tujuan Penelitian**

Adapun tujuan yang ingin dicapai dari penelitian ini adalah untuk mengetahui hasil dari pengelompokan daerah rawan bencana di Sumatera Selatan dengan menggunakan algoritma *K-Means*

#### **1.5 Manfaat Penelitian**

Adapun manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Dapat memberikan informasi berupa kelompok daerah rawan terjadi bencana di Provinsi Sumatera Selatan.
2. Dapat menjadi masukan dan informasi tambahan bagi pemerintah dalam proses penanggulangan bencana di Provinsi Sumatera Selatan.