

## **IMPLEMENTATION OF DATA MINING FOR CLUSTERING DISASTER-PRONE AREAS IN SOUTH SUMATRA USING K-MEANS ALGORITHM**

### **ABSTRACT**

*Indonesia is a country that has a natural structure consisting of the meeting of tectonic plates. This has resulted in many regions in Indonesia being very vulnerable to natural disasters, one of which is the Province of South Sumatra. This study aims to group disaster-prone areas in South Sumatra by utilizing the data mining process using the Clustering technique. The algorithm used for cluster formation is the K-Means algorithm. K-Means is one of the non-hierarchical clustering data methods that can group disaster data into several clusters based on the similarity of the data, so that disaster data with the same characteristics are grouped in one cluster and those with different characteristics are grouped in another cluster. The results of this study are groups of disaster-prone areas based on three categories, namely low disaster-prone areas, moderate disaster-prone areas and high disaster-prone areas. This research is expected to be additional information for the government in the process of disaster management in South Sumatra.*

**Keywords:** *Natural Disaster, Data Mining, K-Means, Clustering, Disaster Prone*

# **PENERAPAN DATA MINING UNTUK PENGELOMPOKAN DAERAH RAWAN BENCANA DI SUMATERA SELATAN MENGGUNAKAN ALGORITMA K-MEANS**

## **ABSTRAK**

Indonesia merupakan negara yang memiliki struktur alam yang terdiri dari pertemuan lempeng-lempeng tektonik. Hal tersebut mengakibatkan banyak daerah di Indonesia yang sangat rawan terhadap bencana alam, salah satunya Provinsi Sumatera Selatan. Penelitian ini bertujuan untuk melakukan pengelompokan daerah rawan bencana di Sumatera Selatan dengan memanfaatkan proses data mining dengan menggunakan teknik *Clustering*. Adapun algoritma yang digunakan untuk pembentukan *cluster* adalah algoritma *K-Means*. *K-Means* merupakan salah satu metode data *non-hierarchical clustering* yang dapat mengelompokkan data bencana ke dalam beberapa *cluster* berdasarkan kemiripan dari data tersebut, sehingga data bencana yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster* dan yang memiliki karakteristik yang berbeda dikelompokkan dalam *cluster* yang lain. Hasil dari penelitian ini berupa kelompok daerah rawan bencana berdasarkan tiga kategori yaitu daerah rawan bencana rendah, daerah rawan bencana sedang dan daerah rawan bencana tinggi. Penelitian ini diharapkan dapat menjadi informasi tambahan bagi pemerintah dalam proses penanggulangan bencana yang ada di Sumatera Selatan.

**Kata Kunci:** Bencana Alam, Data Mining, *K-Means*, Pengelompokan, Rawan Bencana