Jurnal Dinamika Informatika
Volume 9, No 1, Februari 2020
ISSN 1978-1660 : 51-60
ISSN *online* 2549-8517

Analisis Gempa Bumi Pada Pulau Jawa Menggunakan Clustering Algoritma K-Means

Analysis Earthquake in Java Island Using Clustering K-Means Algorithm

Ferin Reviantika[1], Chita Nauly Harahap[2], Yufis Azhar[3]
[1,2,3]Jurusan Teknik Informatika, Universitas Muhammadiyah Malang
e-mail: **[1]ferinrevi@webmail.umm.ac.id**,
[2]chitanaulyharahap@webmail.umm.ac.id, [3]yufis@umm.ac.id

Intisari

Pulau Jawa adalah pulau dengan penduduk terpadat di Indonesia dengan jumlah populasi lebih dari 145 juta. Hal itu menyebabkan mitigasi bencana alam sangat penting dilakukan untuk mengurangi kerugian. Gempa bumi menjadi salah satu bencana alam yang sering terjadi pada Pulau Jawa. Data gempa bumi pulau jawa diambil dari twitter Badan Meteorologi dan Geofisika (BMKG) yaitu @infobmkg dengan info #Gempa, data bemula dari Desember 2018 hingga Maret 2020 Atribut yang digunakan ada empat yaitu Tanggal, Lintang Selatan, Bujur Timur dan Magnitudo. Gempa bumi dikelompokan berdasarkan daerah terjadinya menggunakan algoritma K-Means dan algoritma DBSCAN sebagai pembandingnya. Hasil clustering ditampilkan dengan peta dan algoritma K-Means lebih unggul dengan nilai indeks sillhoute 0.54 sedangkan algoritma DBSCAN memiliki nilai indeks sillhoute 0.17.

Kata kunci— Gempa Bumi, Penggalian Data, K-Means, DBSCAN, Indeks Sillhoute

Abstract

Java Island is the most populated island in Indonesia with a population over 145 million. It causes mitigation of natural disaster is very important to decrease the damage. Earthquake become one of the natural disasater often happens in Java Island. The data is taken from Indonesian Agency Meteorology, Climatology and Geophysics (BMKG) twitter @infoBMKG with used #Gempa, data start from December 2018 to Maret 2020. The variabels are used is four namely Date, Latitude, Longitude, and Magnitude. Earthquake clustering based on location of earthquake used K-Means algorithm with using DBSCAN algorithm as comparison. The results are visualized using maps and algorithm K-Means perfom more better with 0.54 point of index sillhoute while the algorithm DBSCAN perfom with 0.17 point of index sillhoute.

***Keywords*—** Earthquake, Data Mining, K-Means algorithm, DBSCAN algorithm, Index Sillhoute

PENDAHULUAN

Bencana alam merupakan suatu peristiwa alam yang akan berdampak besar bagi kelangsungan hidup manusia [1]. Salah satu bencana alam yang rawan terjadi di Indonesia yaitu gempa bumi, gempa bumi merupakan peristiwa adanya getaran atau guncangan yang terjadi pada permukaan bumi [2]. Gempa bumi dapat disebabkan oleh pergerakan kerak bumi atau lempeng bumi, gunung berapi, dan tabrakan meteor. Gempa bumi juga berasal dari getaran lapisan batuan patah yang memiliki energi. Menurut beberapa sumber, gempa bumi yang bermula dari pergerakan kerak bumi menjadi salah satu penyebab gempa paling sering terjadi [3]. Besarnya goncangan bumi beragam mulai dari yang sangat kecil sehingga sulit dirasakan, sampai ke goncangan yang sangat dahsyat. Salah satu efek yang ditimbulkan dari gempa bumi yaitu mampu meruntuhkan bangunan yang kokoh karena adanya perecpatan tanah [4]. Pengukuran kekuatan gempa bumi dilakukan dengan menggunakan alat seismograf dan kekuatan gempa bumi disebut magnitudo [5].

Pulau Jawa sering kali disebut pulau terpadat di Indonesia, penduduk pulau Jawa yang mencapai angka 141 juta. Disamping itu pulau Jawa juga merupakan pusat ekonomi Indonesia. Hal ini menyebabkan jika terjadi bencana alam seperti gempa bumi menimbulkan kerugian yang cukup besar, kerugian yang terjadi tidak hanya materil tetapi juga kerugian immateril. Penelitian ini bertujuan untuk masyarakat agar mengetahui daerah-daerah yang rawan terhadap pusat gempa bumi. Sehingga masyarakat dapat mengetahui informasi secara detail dan menyeluruh.

Gempa bumi yang sering terjadi pada pulau Jawa termasuk dalam kategori gempa dangkal (kedalaman 0-35-70 km) berdasakarkan episenternya dan hanya memiliki 3-5 skala richter. Gempa yang dirasakan dipermukaan secara matematis merupakan gempa yang memiliki kekuatan lebih dari 3,5 skala richter. Dan gempa yang lebih dari 7 skala richter biasanya berpotensi terjadinya tsunami. Walaupun dengan skala richter yang kecil namun gempa dangkal ini dapat merusakkan bangunan yang berada pada permukaannya. Melihat dari banyaknya gempa bumi yang terjadi membuat upaya mitigasi sangatlah diperlukan [2]. Oleh karena itu, untuk mendeteksi persebaran gempa bumi di pulau Jawa menggunakan data mining dengan teknik k-means clustering.

Data Mining (Penggalian Data) adalah metode pengolahan data berskala besar yang memiliki tujuan memperoleh informasi dari kumpulan data dan mengubah informasi tersebut menjadi struktur informasi baru yang dapat dengan mudah dipahami [6]. Salah satu metode dalam data mining membahas tentang clustering. Clustering atau klasterisasi adalah proses metode pengelompokan data yang mempartisi data ke dalam beberapa cluster atau kelompok sehingga data dalam satu cluster memiliki tingkat kemiripan yang maksimum dan data antar cluster memiliki kemiripan yang minimum [7]. Metode clustering memliki beberapa algoritma yang dapat dimanfaatkan untuk implementasi diantaranya adalah algoritma K-means [2].

Penelitian sebelumnya sudah pernah dilakukan pada tahun 2017 oleh Pepi Novianti, Dyah Setyorini, Ulfasari Raflesia berjudul K-Means Cluster Analysis in

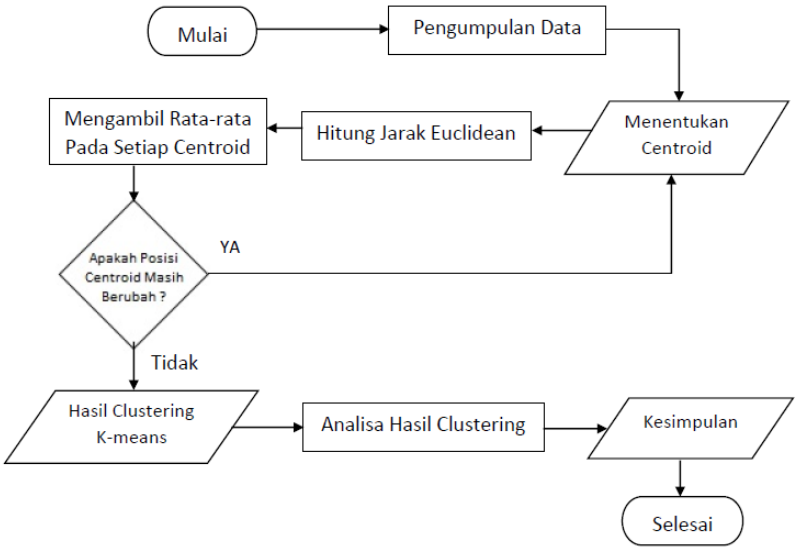
Jurnal Dinamika Informatika
Volume 9, No 1, Februari 2020
ISSN 1978-1660 : 51-60
ISSN *online* 2549-8517

Earthquake Epicenter Clustering penelitian yang berfokus pada daerah Bengkulu [8]. Penelitian lainnya pada tahun yang sama oleh Kamat dan Kamath berjudul Earthquake Cluster Analysis: K-means Approach yang berfokus pada gempa bumi yang terjadi di daratan Eropa [9]. Mereka menjelaskan dengan sangat baik tentang penggunaan algoritma K-means untuk klasterisasi gempa bumi.

Oleh karena itu, penelitian ini bertujuan agar dapat melihat kawasan yang sering terjadi bencana gempa bumi dengan begitu proses mitigasi bencana gempa bumi dapat diterapkan pada kawasan tersebut. Karena Pulau Jawa merupakan pulau dengan penduduk terbanyak di Indonesia, tentu saja membuat dampak yang ditimbulkan dari bencana gempa bumi lebih besar. Dengan mengelompokan data persebaran gempa bumi di pulau Jawa menggunakan k-means clustering. Sehingga dapat membantu masyarakat mengetahui titik-titik pusat gempa secara detail dan menyeluruh. Selain itu, dataset juga dikelompokkan menjadi tiga kelompok yang berbeda berdasarkan tahun terjadinya yaitu 2018, 2019, dan 2020.

METODE PENELITIAN

Penelitian ini dilakukan menggunakan tahapan algoritma K-Means dengan membandingkan hasil clustering pada hasil akhirnya. Flowchart algoritma K-means seperti pada gambar 1.



Gambar 1 Flowchart Algoritma K-Means

Pengumpulan Data

Dataset diperoleh melalui proses pengumpulan data yang dilakukan secara manual dengan cara mengambil data dari twitter Badan Meteorologi dan Geofisikia (BMKG) yaitu @infoBMKG https://twitter.com/infobmkg lalu mengambil data info #Gempa yang hanya terjadi pada pulau Jawa. Data gempa yang berhasil terekam mulai dari tanggal 09/12/2018 sampai dengan tanggal 05/03/2020 yaitu, sebanyak 100 data gempa bumi. Pada dataset terdapat enam atribut yaitu, Tanggal,

Daerah, Magnitudo, Posisi pusat gempa bumi (Lintang Selatan (LS) dan Bujur Timur (BT), serta Kedalaman (km) seperti terlihat pada Tabel 1.

Tabel 1 Data Gempa Bumi Pada Pulau Jawa

No	Tanggal	Daerah	Magnitudo	Lintang Selatan (LS)	Bujur Timur (BT)	Kedalaman (km)
1	05/03/2020	Sukabumi - Jawa Barat	3.3	-7.07	106.46	8
2	21/02/2020	Tasikmalaya - Jawa Barat	4.9	-8.17	107.5	11
3	16/02/2020	Karawang - Jawa Barat	3.5	-6.1	107.56	270
4	14/02/2020	Kediri - Jawa Timur	3.9	-7.71	111.98	7
5	13/02/2020	Jember - Jawa Timur	5.0	-11.39	113.27	10
6	10/02/2020	Pandeglang - Banten	4.8	-6.81	105.32	5
7	08/02/2020	Garut - Jawa Barat	3.3	-7.69	107.71	1
8	06/02/2020	Bangkalan - Madura	6.3	-6.43	113.05	636
9	05/02/2020	Sukabumi - Jawa Barat	4.2	-7.33	106.63	109
10	05/02/2020	Bandung - Jawa Barat	2.9	-7.13	107.52	12

Prepocessing Data (Cleaning Data)

Data yang diperoleh merupakan data bersih yang tidak mengandung missing value. Sehingga hanya dilakukan selection pada kolom yang dipakai untuk penelitian. Dataset berisi 100 data gempa bumi di pulau Jawa, namun dalam penelitian ini hanya data gempa bumi yang dikategorikan sedang atau lebih kuat yaitu yang memiliki rentang magnitudo ≥ 3 sehingga data gempa bumi yang berhasil diproses sebanyak 92 data gempa bumi dari tahun 2018 sampai 2020.

Proses Clustering

Merupakan tahap pengolahan dataset yang dimodelkan dengan algoritma K-Means melalui perhitungan sehingga menghasilkan pengelompokan data [6]. Adapun tahapan pada algoritma K-Means adalah:

1) Tentukan jumlah cluster.
2) Tentukan titik centroid (pusat kluster) secara acak. Untuk mendapatkan titik centroid dilakukan pendekatan heuristik dengan Metode Siku untuk mengidentifikasi nilai optimal untuk K.

Jurnal Dinamika Informatika
Volume 9, No 1, Februari 2020
ISSN 1978-1660 : 51-60
ISSN *online* 2549-8517

3) Hitung jarak Euclidean antara setiap titik data dan centroid cluster yang dipilih.

Jarak Euclidian digunakan untuk menentukan jarak antara titik pusat data dan pusat massa. Dimana p_1 adalah atribut pertama dari data gempa bumi berposisi $> 110.5^0$ dan p_2 atribut kedua dari data gempa bumi berposisi $> 110.5^0$ sedangkan p_n atribut $ke-n$ dari data gempa bumi berposisi $> 110.5^0$. Demikian pula pada q_1 adalah atribut pertama dari data gempa bumi berposisi $< 110.5^0$ dan q_2 atribut kedua dari data gempa bumi berposisi $< 110.5^0$ sedangkan q_n atribut $ke-n$ dari data gempa bumi berposisi $< 110.5^0$. Dalam hal ini, atribut pertama adalah Lintang Selatan (LS) dan atribut kedua adalah Bujur Timur (BT) dengan jarak ini dapat mengurutkan titik data menjadi cluster. Perhitungan dalam mencari Jarak Euclidian dapat dihitung menggunakan persamaan (1) dan (2) [6], yaitu:

a. Persamaan untuk dua dimensi

$$d(p,q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2} \quad (1)$$

b. Persamaan untuk n-dimensi

$$d(p,q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^n} \quad (2)$$

4) Tentukan centroid baru untuk setiap cluster dengan mengambil rata-rata dari semua poin yang ditugaskan ke cluster

Ulangi proses dua dan tiga hingga posisi centroid tidak lagi berubah dan penetapan cluster tetap sama. Sedangkan untuk enggunaan atribut tanggal dengan cara mengubah format menjadi tahun kemudian, dibagi menjadi tiga format tahun yaitu 2018, 2019, dan 2020 dan atribut daerah akan dihilangkan karena menimbulkan adanya redudansi data.

HASIL DAN PEMBAHASAN

Data

Dataset yang digunakan adalah data gempa bumi di Pulau Jawa pada tahun 2018 hinga tahun 2020 sebanyak 100 data. Pada data tersebut terdapat enam atribut, keterangan setiap atributnya dapat dilihat pada Tabel 1.

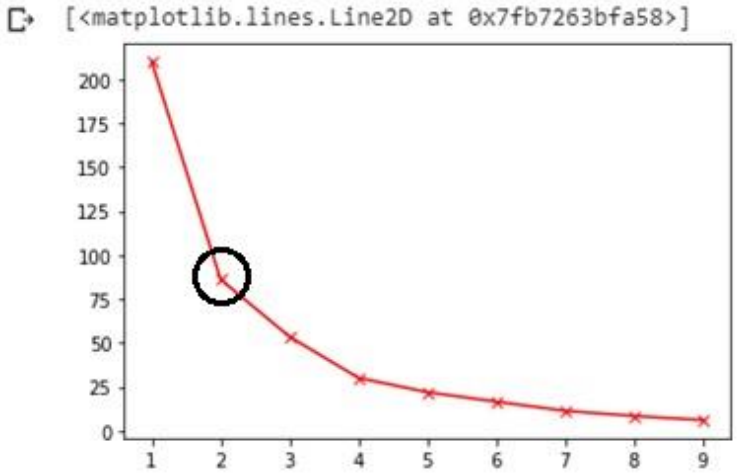
Tabel 1 Keterangan Atribut Data Gempa Bumi

Atribut	Keterangan
Tanggal	Tanggal terjadinya gempa bumi
Daerah	Daerah terjadinya gempa bumi yang meliputi Kabupaten dan Provinsi lokasi gempa bumi
Magnitudo	Besar nya kekuatan gempa bumi yang terjadi dalam satuan skala richter
Lintang Selatan	Posisi lintang dimana lokasi gempa bumi terjadi
Bujur Timur	Posisi bujur dimana lokasi gempa bumi terjadi
Kedalaman	Kedalaman terjadinya gempa bumi dalam satuan kilometer (KM)

Dari enam atribut yang tersedia, hanya empat atribut yang digunakan pada proses clustering, yaitu atribut tanggal, magnitudo, lintang selatan, dan bujur timur. Setelah itu, penggunaan data gempa bumi yang digunakan hanya gempa bumi yang berkekuatan ≥ 3 skala richter.

Hasil Cluster

Dalam dataset jumlah cluster yang dibutuhkan tidak pasti. Sehingga dalam penelitian menggunakan pendekatan heuristic dengan elbow method untuk mengidentifikasi nilai optimal K (jumlah cluster). Grafik pada gambar 2 menunjukan engsel yang jelas seperti elbow dan titik tersebut dianggap nilai K yang didapat secara optimal yaitu K = 2. Selanjutnya diperoleh hasil centroid menggunakan metode K-Means pada Tabel 2.



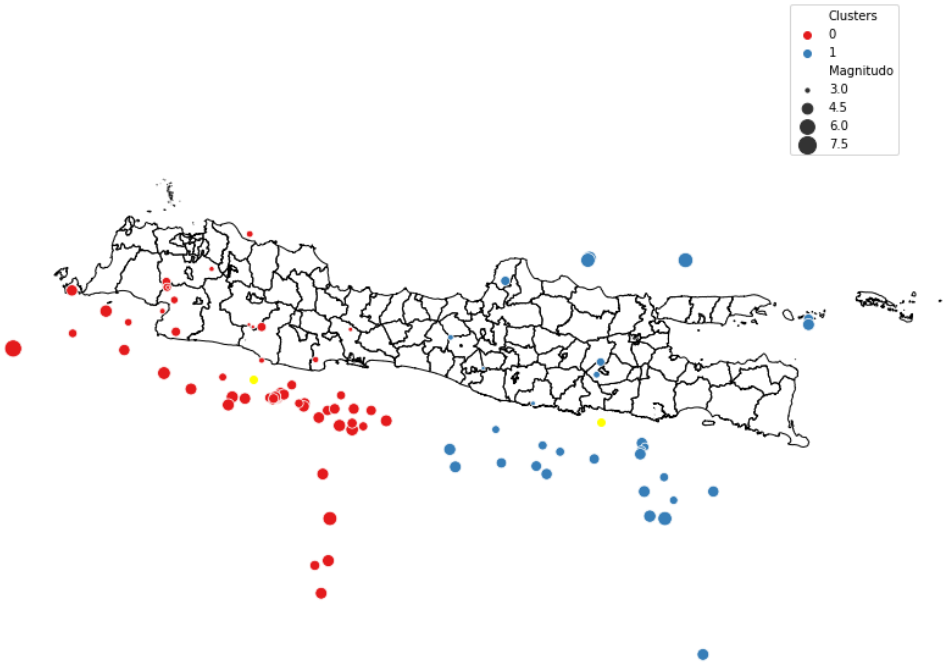
Gambar 2 Grafik *Elbow Method*

Tabel 2 Hasil Centroid Metode K-Means

Cluster	Lintang Selatan	Bujur Timur	Magnitudo
0	-8.471935	111.977742	4.532258
1	-7.927705	107.601475	4.340984

Titik merah merupakan cluster 0 sedangkan titik biru merupakan cluster 1. Lalu besar kecil nya lingkaran dilihat dari besar kecilnya magnitudo gempa tersebut. semakin besar magnitudo gempa semakin besar pula titik plot yang dihasilkan. Titik bewarna kuning tidak menjelaskan tentang terjadinya gempa dikawasan tersebut tetapi menjadi titik pusat centroid dari dua cluster, yang dihasilkan dari perhitungan jarak euclidean antara centroid dan data gempa bumi pulau jawa. Dalam peta pulau Jawa batas antar kabupaten dibedakan dengan warna yang berbeda-beda. Pulau jawa termasuk kedalam tiga pulau terbesar di indonesia. Indonesia adalah negara maritim yang sebgaian besar wilayahnya terdiri dari lautan. Sehingga visualisasi peta dapat dilihat pada gambar 3 epicentrum gempa bumi dipulau Jawa sebagian besar berada pada wilayah laut.

Jurnal Dinamika Informatika
Volume 9, No 1, Februari 2020
ISSN 1978-1660 : 51-60
ISSN *online* 2549-8517



Gambar 3 Peta Persebaran Gempa Bumi

Selanjutnya gempa bumi dibagi menjadi tiga cluster berdasakan tahun 2018, 2019, dan 2020. Visualisasi persebaran gempa ditujukan pada gambar 4. Dalam visualisasi peta dapat diamati bahwa episentrum gempa sudah menyebar pada tahun 2019 ditandai dengan titik berwarna kuning. Gempa menyebar secara menyeluruh di Pulau Jawa, rata-rata titik gempa bumi berada di wilayah laut. Hanya ada beberapa titik gempa bumi didaratan pulau Jawa karena gempa bumi cenderung terjadi di laut tetapi tidak berpotensi tsunami.

Gambar 4 Persebaran Gempa Tahun 2018, 2019, dan 2020

Pada penelitian ini menggunakan metode DBSCAN sebagai pembanding dengan metode yang dilakukan yaitu metode K-Means. Langkah preprocessing yang dilakukan dalam metode DBSCAN sama halnya dengan metode K-Means. Tetapi dalam metode DBSCAN dihasilkan tiga cluster dalam hal pengelompokan data. Nilai centroid dari tiga cluster ditulis pada Tabel 3.

Tabel 3 Hasil Centroid Metode DBSCAN

Cluster	Lintang Selatan	Bujur Timur	Magnitudo
-1	-8.140769	111.488462	4.746154
0	-7.801552	107.556897	4.315517
1	-8.947619	111.778571	4.442857

Dari pengujian cluster dengan menggunakan nilai centroid pada Tabel 2 dan Tabel 3 didapatkan hasil perbandingan cluster menggunakan metode K-Means dan cluster menggunakan metode DBSCAN. Pada Tabel 4 dijelaskan jumlah data pada setiap clusternya. Meskipun jumlah cluster yang digunakan dalam pengujian berbeda, tetapi perbandingan jumlah data pada masing-masing cluster relatif sama.

Tabel 4 Perbandingan Hasil Cluster K-Means dan Cluster DBSCAN

Metode	Cluster	Jumlah Data
K-Means	0	61
	1	31
DBSCAN	-1	54
	0	30
	1	8

Hasil dari perbandingan perhitungan nilai indeks *Silhoutte* antara metode clustering K- Means dan metode clustering DBSCAN dapat dilihat dalam Tabel 5. Metode K-means memiliki hasil yang lebih baik dari metode DBSCAN.

Tabel 5 Perbandingan metode K-Means dan metode DBSCAN

Metode	Indeks Silhoutte
K-Means	0.54514685
DBSCAN	0.17175428

Jurnal Dinamika Informatika
Volume 9, No 1, Februari 2020
ISSN 1978-1660 : 51-60
ISSN *online* 2549-8517

KESIMPULAN DAN SARAN

Kesimpulan

Kesimpulan dari penelitian yang telah dilakukan ini, diantaranya sebagai berikut.

1) Pembagian cluster pada metode algoritma DBSCAN lebih banyak yaitu tiga cluster sedangkan pada algoritma K-Means dua cluster
2) Berdasarkan perbandingan indeks *Sillhoute* antara metode algoritma K-Means dan algoritma DBSCAN menunjukkan bahwa menggunakan algoritma K-Means lebih unggul dalam clustering data gempa bumi Pulau Jawa.
3) Nilai indeks Silhoute metode algoritma K-Means lebih unggul 0.37339257 dari pada algoritma DBSCAN.

Saran

Saran yang dapat diberikan dari penelitian yang telah dilakukan ini, diantaranya sebagai berikut.

1) Perlu adanya penggunaan metode algoritma pembanding lain. Selain, K-Means dan DBSCAN.
2) Perluasaan data gempa bumi yang tidak hanya mencakup daerah Pulau Jawa dan clustering mitigasi gempa bumi pada pulau lain di Indonesia.

DAFTAR PUSTAKA

[1] A. P. Windarto, S. Tinggi, I. Komputer, T. Bangsa, and A. Wanto, “PENERAPAN ALGORITMA CLUSTERING DALAM MENGELOMPOKKAN BANYAKNYA DESA / KELURAHAN MENURUT UPAYA ANTISIPASI / MITIGASI BENCANA ALAM MENURUT PROVINSI DENGAN K-MEANS,” no. October, 2018.

[2] F. R. Senduk and F. Nhita, “Clustering of Earthquake Prone Areas in Indonesia Using K-Medoids Algorithm,” vol. 4, no. 2016, pp. 65–76, 2019.

[3] A. I. T. A. Sugito, J. Pe, D. Geografi, F. Pe, D. Ilmu, and P. E. Getahua, “Tsunami,” pp. 1–49, 2008.

[4] T. Dengan, M. Mc, R. Mayasyafa, A. Susilo, D. Ph, and M. Ab, “MAKSIMUM GEMPABUMI PROVINSI JAWA,” pp. 1–4, 2014.

[5] M. A. Novianta and E. Setyaningsih, “Early Detection System Design Based on Fluctuation Earthquake Using Magnetic Muhammad Andang Novianta , Emy Setyaningsih,” vol. 16, no. 1, pp. 35–44, 2015.

[6] I. D. Id and T. R. Angraini, “Penentuan Daerah Rawan Titik Api di Provinsi Riau Menggunakan Clustering Algoritma K-Means (Determination of Fire Point Prone Areas in Riau Province Using Clustering K-Means Algorithm),” vol. VI, no. November, pp. 137–147, 2018.

[7] M. Ramageri, “DATA MINING TECHNIQUES AND APPLICATIONS,” vol. 1, no. 4, pp. 301–305.

[8] P. Novianti, D. Setyorini, and U. Rafflesia, "K-Means cluster analysis in earthquake epicenter clustering," vol. 3, no. 2, pp. 81–89, 2017.
[9] R. S. Kamath, "Earthquake Cluster Analysis : K-Means Approach Earthquake Cluster Analysis : K-Means Approach," no. January, 2017.