

PENERAPAN ALGORITMA K-MEANS UNTUK PENGELOMPOKAN DAERAH RAWAN BENCANA DI INDONESIA

Amril Mutoi Siregar

Universitas Buana Perjuangan Karawang

amrilmutoi@ubpkarawang.ac.id

ABSTRACT

Indonesia is a country located in the equator, which has beautiful natural. It has a mountainous constellation, beaches and wider oceans than land, so that Indonesia has extraordinary natural beauty assets compared to other countries. Behind the beauty of natural it turns out that it has many potential natural disasters in almost all provinces in Indonesia, in the form of landslides, earthquakes, tsunamis, Mount Meletus and others. The problem is that the government must have accurate data to deal with disasters throughout the province, where disaster data can be in categories or groups of regions into very vulnerable, medium, and low disaster areas. It is often found when a disaster occurs, many found that the distribution of long-term assistance because the stock for disaster-prone areas is not well available. In the study, it will be proposed to group disaster-prone areas throughout the province in Indonesia using the k-means algorithm.

The expected results can group all regions that are very prone to disasters. Thus, the results can be Province West java, central java very vulnerable categories, provinces Aceh, North Sumatera, West Sumatera, east Java and North Sulawesi in the medium category, provinces Bengkulu, Lampung, Riau Island, Babel, DIY, Bali, West Kalimantan, North Kalimantan, Central Sulawesi, West Sulawesi, Maluku, North Maluku, Papua, west Papua including of rare categories. With the results obtained in this study, the government can map disaster-prone areas as well as prepare emergency response assistance quickly. In order to reduce the death toll and it is important to improve the services of disaster victims. With accurate data can provide prompt and appropriate assistance for victims of natural disasters.

Keyword: *K-means Algorithm, Clustering, Data Mining, Natural disasters.*

ABSTRAK

Indonesia adalah negara yang terletak digaris khatulistiwa, yang memiliki keindahan alam yang indah. Memiliki konsturpengunungan, pantai dan lebih luas lautan daripada daratan, sehingga Indonesia memiliki asset keindahan alam yang luar biasa dibandingkan negara lain. Dibalik keindahan alam ternyata memiliki banyak potensi bencana alam hampir seluruh propinsi di Indonesia, berupa bencana tanah longsor, gempa bumi, tsunami, gunung Meletus dan lain lain. Dengan pemasalahan yang hadapai pemerintah harus memiliki data yang akurat dan tepat untuk menangani bencana diseluruh propinsi, dimana data bencana dapat di kategori atau dikelompok wilayah ke dalam sangat rawan, sedang, rendah yang mengalami bencana. Sering ditemukan ketika terjadi bencana, banyak ditemukan penyaluran bantuan yang lama dikarena stok untuk daerah rawan bencana tidak tersedia dengan baik. Dalam penelitian akan mangajukan pengelompokan daerah yang rawan bencana diseluruh propinsi di Indonesia menggunakan algoritma k-means.

Hasil yang diharapkan bisa mengelompok seluruh daerah yang kategori sangat rawan bencana. Jadi hasil yang dapat Propinsi Jawa barat, Jawa tengah kategori sangat rawan, propinsi Aceh, Sumut, Sumbar, Jawa timur, Sulut kategori sedang, propinsi jarang adalah propinsi Bengkulu, Lampung, Kepri, Babel, DIY, Bali, Kalbar, Kalut, Sulteng, Sulbar, Maluku, Maluku Utara. Papua, Papua Barat termasuk kategori jarang. Dengan hasil yang

didapatkan dalam penelitian ini adalah agar pemerintah dapat memetakan wilayah wilayah rawan bencana sekaligus mempersiapkan bantuan tanggap darurat dengan cepat. Agar dapat mengurangi korban meninggal dan penting untuk meningkatkan pelayanan korban bencana. Dengan data yang akurat dan tepat dapat memberikan pertolongan yang cepat dan tepat bagi korban bencana alam.

Kata kunci: Algoritma *K-means*, *Clustering*, *Data Mining*, Bencana Alam,

1. PENDAHULUAN

1.1 Latar Belakang Masalah

Indonesia berada pada dilempengan dunia, dan dikelilingi dengan dua Samudra, hampir setiap tahun mengalami bencana alam, bencana tanah longsor, gempa bumi, tsunami, gunung Meletus dan lain lain. Bencana alam adalah salah satu fenomena yang terjadi setiap saat, dimana dan kapanpun bisa terjadi sehingga menimbulkan risiko atau bahaya terhadap kehidupan manusia, baik kerugian material maupun spiritual dan korban jiwa (Nugroho.dkk, 2009).

Bencana longsor merupakan salah satu jenis bencana yang terjadi di Indonesia. Seperti terjadinya pendangkalan, terganggunya jalur lalu lintas, merusak lahan pertanian, dan pemukiman. Kenyataannya Badan Nasional Penanggulangan Bencana (BNPB) belum bisa memaksimalkan pelayanan korban bencana, misalnya lamanya datangnya bantuan dikarenakan stok peralatan dan makanan yang terbatas di daerah bencana alam. Padahal dengan adanya pemetaan risiko bencana menjadi sangat penting dalam penataan penanggulangan bencana yang terarah, tepat, dan (Nugraha, 2013). Diberitakan BBC Indonesia kejadian bencana bahwa tidak ada alasan Pemerintah daerah dan masyarakat tidak tahu daerah rawan bencana/longsor. Dampaknya adalah setidaknya sebanyak 274 kabupaten dan kota di Indonesia berada di daerah bahaya longsor dan 40,9 juta jiwa terancam bencana bencana longsor, sebut Badan Nasional Penanggulangan Bencana (BNPB). Seperti dilansir liputan6.com, belum usai penanganan dampak tsunami di Selat Sunda, bencana alam kembali menerjang wilayah Indonesia. Pada Senin, 31 Desember 2018 sekitar pukul 17.30 WIB, longsor menerjang Dusun Garehong, Desa Simaresmi di Kecamatan Cisolok, Kabupaten Sukabumi, Jawa Barat. Bencana alam menjelang malam pergantian tahun itu mengakibatkan 15 orang meninggal dunia. Hingga Rabu siang, 2 Januari 2019, 20 orang masih dilaporkan hilang tertimbun longsor.

Badan Nasional Penanggulangan Bencana (BNPB) memprediksi akan banyak bencana yang melanda Indonesia pada 2019. Kepala Pusat Data, Informasi dan Humas BNPB Sutopo Purwo Nugroho mengungkapkan, bencana hidrometeorologi akan mendominasi di tahun ini. Dilansir Kompas.com, budaya Sadar Bencana Masyarakat Indonesia Masih Rendah. Kepala Pusat Data, Informasi dan Hubungan Masyarakat Badan Nasional Penanggulangan Bencana (BNPB) Sutopo Purwo Nugroho mengatakan budaya sadar bencana masyarakat Indonesia masih rendah. Padahal, budaya sadar bencana penting mengingat ancaman bencana bisa datang kapan pun di waktu yang tidak terduga. Berdasarkan penelitian BNPB, pengetahuan kebencanaan masyarakat Indonesia meningkat setelah peristiwa tsunami Aceh 2004, tetapi, sikap sadar bencana belum menjadi perilaku sehari-hari masyarakat. Sutopo mengatakan, hal ini menjadi tantangan tersendiri. "Menjadi sikap, menjadi budaya sehari-hari itu sulit dan tidak bisa langsung dilakukan seketika perlu upaya yang terus menerus dan lintas generasi," kata Sutopo di kantor BNPB, Utan Kayu, Jakarta Timur, Selasa (25/12/2018).

BEKASI, KOMPAS.com memberitakan Tsunami Selat Sunda yang menerjang Kabupaten Pandeglang, Kabupaten Serang, Banten, dan Lampung Selatan pada Sabtu (22/12/2018) banyak menelan korban yang merupakan pendatang atau tamu penginapan. Dari acara *family gathering* PLN Unit Induk Transmisi Jawa Bagian Barat saja tercatat ada 41

peserta yang meninggal dunia. Anggota Satlantas Polres Pandeglang Brigadir Aria mengatakan, banyaknya korban meninggal duni dari pendatang atau tamu penginapan disebabkan kurangnya sosialisasi dari pihak hotel terkait bencana alam. Menurut Aria, seharusnya pihak hotel atau penginapan bisa memberikan sosialisasi terkait apa yang harus dilakukan tamu hotel jika ada bencana alam. seperti pengarahan soal jalur evakuasi dan lainnya.

1.2 Penelitian Terdahulu

Penelitian ini mempunyai keterkaitan beberapa peneliti sebelumnya yaitu penelitian yang berjudul Implementasi Algoritma *Clustering* dengan *Singular Vector Decomposition* untuk Menunjang Keputusan dalam Meningkatkan Produktivitas Tanaman Jagung (Siregar AM, 2018), Implementasi *K-Means Clustering* pada *Rapidminer* untuk Analisis Daerah Rawan Kecelakaan (Rahmat, *et all*, 2017), Implementasi Algoritma *K-Means Clustering* untuk Menentukan Strategi *Marketing President University* (J.O Ong, 2013).

1.3 Tinjauan Pustaka

Menurut Connolly dan Begg (2010), *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 yang sangat penting. Sedangkan menurut Segall, *et.all.* (2008), *Data mining* biasa juga disebut dengan “data atau *knowledge discovery*” atau menemukan pola tersembunyi pada data. *Data mining* adalah proses dari menganalisa data dari prespektif yang berbeda dan menyimpulkannya ke dalam informasi yang berguna. Adapun menurut Han dan Kamber (2006:5), *Data mining* didefinisikan sebagai proses mengekstrak atau menambang pengetahuan yang dibutuhkan dari sejumlah data besar. Pada prosesnya *data mining* akan mengekstrak informasi yang berharga dengan cara menganalisis adanya pola-pola ataupun hubungan keterkaitan tertentu dari data-data yang berukuran besar.

Data mining berkaitan dengan bidang ilmu seperti *Database System*, *artificial intelligence*, *Statistic*, *Machine Learning*, *Information Retrieval*, dan Komputasi Tingkat Tinggi. Selain itu *data mining* didukung oleh ilmu lain seperti *Neural Network*, *K-means*, *C4.5 Decision tree*, *KNN*, *Naïve Bayes* dan lain-lain. Pengenalan Pola, *Spatial Data Analysis*, *Image Database*, *Signal Processing*. Beberapa survey tentang proses pemodelan dan metodologi menyatakan bahwa, “*Data mining* digunakan sebagai penunjuk, dimana *data mining* menyajikan intisari atas sejarah, deskripsi dan sebagai standar petunjuk mengenai masa depan dari sebuah proses model *data mining* (Mariscal, Marba’n dan Ferna’ndes, 2010). Definisi sederhana dari *data mining* adalah ekstraksi informasi atau pola yang penting atau menarik dari data yang ada di *database* yang besar. Dalam jurnal ilmiah, *data mining* juga dikenal dengan nama *Knowledge Discovery in Databases* (KDD). Kehadiran *data mining* dilatarbelakangi dengan problema data *explosion* yang dialami akhir-akhir ini dimana banyak organisasi telah mengumpulkan data sekian tahun lamanya (data pembelian, data penjualan, data nasabah, data transaksi).

Cluster adalah kumpulan objek data yang mirip satu sama lain dalam kelompok yang sama dan berbeda dengan objek data dikelompok lainnya. Sedangkan, *Clustering* atau analisis cluster adalah proses pengelompokkan satu set benda-benda fisik atau abstrak ke dalam kelas objek yang sama. Tujuannya adalah untuk menghasilkan pengelompokan objek yang mirip satu sama lain dalam kelompok-kelompok. Semakin besar kemiripan objek dalam suatu cluster dan semakin besar perbedaan tiap cluster maka kualitas analisis cluster semakin baik. (muhammad-mulyana.blogspot.co.id)

Clustering juga dapat mengidentifikasi kelompok alami dari data berdasarkan kemiripan atribut. disebut juga *segmentation*. *Unsupervised Method* adalah tidak satupun

atribut yang digunakan untuk memandu proses pembelajaran (tidak ada label) seperti Algoritma: *k-Means*, *k-Medoid*, *Fuzzy C-Means*, *Subtractive Mountain*, Hierarki.

Algoritma *K-means* merupakan salah satu algoritma *clustering*(pengelompokan). *K-means clustering* merupakan metode *clustering non-hirarki* yang mengelompokkan data dalam bentuk satu atau lebih *cluster*. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster*/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan *cluster* yang lain, sehingga data yang berada dalam satu *cluster* memiliki tingkat variasi kecil (Agusta, 2007). Adapun langkah-langkah dari algoritma *K-means* adalah sebagai berikut :(Tan, Steinbech, & Kumar, 2006)

1. Tentukan K data sebagai *centroid*, K adalah jumlah *cluster* yang diinginkan ditentukan oleh peneliti
2. Tiap titik data kemudian dicari *centroid* terdekatnya.
3. Setiap himpunan titik (data) yang menjadi *centroid* disebut *cluster*.
4. Hitung kembali *centroid* dari setiap *cluster*.
5. Ulangi langkah 1-4 sampai *centroid* tidak berubah.

Metode *clustering* menggunakan Algoritma *K-means*, ukuran kedekatan data dihitung menggunakan jarak *Euclidean*. Algoritma *K-means* bertujuan untuk meminimumkan jarak total *Euclidean* diantara setiap titik Xi, dan cluster terdekat yakni cj (Wu & Kumar, 2009). Jarak *Euclidean* ditentukan dengan menggunakan persamaan 1 berikut ini (Agusta, 2007).

$$d_{ij} = \sqrt{\sum_{k=0}^n \{x_{ik} - x_{jk}\}^2} \quad (1)$$

Keterangan:

d_{ij} = Jarak antara data ke-i dan data ke-j

n = Dimensi data

x_{ik} = Koordinat data ke-i pada dimensi k

x_{jk} = Koordinat data ke-j pada dimensi k

EuclideanDistance Formula untuk menghitung jarak antar dengan *EuclideanDistance* untuk dua titik dalam satu, dua, dan tiga dimensi secara berurutan ditunjukkan pada persamaan (2), (3), dan (4) sebagai berikut (jurnal.unmuhjember.ac.id).

$$\sqrt{(x - y)^2} = |x - y| \quad (2)$$

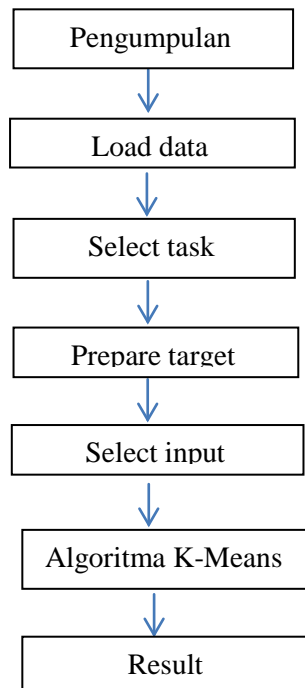
$$d(p, q) = \sqrt{(p1 - q1)^2 + (p2 - q2)^2} \quad (3)$$

$$d(p, q) = \sqrt{(p1 - q1)^2 + (p2 - q2)^2 + (p3 - q3)^2} \quad (4)$$

2. METODOLOGI

2.1 Metode Penelitian

Penelitian ini diusulkan dalam pengelompokan dengan kategori sangat, sedang, jarang terjadi untuk dengan algoritma *K-Means*. Penelitian ini menggunakan *tools* dalam percobaan dengan Rapidminer 9.1. tahapan yang dilakukan dalam penelitian ini sebagai berikut:



Gambar 1 Metode Penelitian yang digunakan

2.2 Pengumpulan Dataset

Dataset penelitian ini menggunakan yang tersedia <https://bnpb.go.id>. Dataset terdiri dari beberapa *attribute* yaitu nama propinsi, jumlah meninggal, luka luka, mengungsi, rumah rusak berat, sedang, ringan dan terendam. Data yang diambil badan nasional penganggulangan bencana pada tahun 2013 sampai 2018, dan 34 propinsi yang ada Indonesia.

Tabel 1 Contoh Dataset yang digunakan

Propinsi	Meninggal	Luka-luka	Menderita	Rusak Berat	Rusak Sedang	Rusak Ringan	Terendam
Aceh	190	3896	1538296	11311	7124	17566	185719
sumatera utara	137	118	621502	1507	1163	4024	95765
sumatera barat	102	110	106724	740	962	4038	38933
riau	22	13	720317	402	511	5616	50995
jambi	40	16	234458	346	258	2356	54513
Sumatera Selatan	42	52	453148	836	260	2261	83311
bengkulu	52	11	9464	343	260	711	2571
lampung	29	52	38256	414	360	3082	22712
KEPULAUAN BANGKA BELITUNG	26	7	40405	56	210	346	3625
kepulauan riau	13	18	2052	171	356	784	118
dki jakarta	82	17	895608	51	93	1204	51418
jawa barat	401	568	3299637	9345	12662	29193	345883
jawa tengah	488	567	1034449	5639	4396	29016	245559
di yogyakarta	38	44	150133	141	74	730	499
Jawa Timur	260	1748	1606199	19493	9734	28955	237705
banten	25	18	560099	878	503	2608	59876
ball	39	43	127608	229	436	415	710
Nusa Tenggara Barat	29	1478	1115279	3773	6966	14846	42168
nusa tenggara timur	32	27	68714	1487	633	3311	14847
kalimantan barat	39	4	148325	119	285	597	22158
kalimantan tengah	7	5	389081	61	1	3411	46470
Kalimantan Selatan	9	21	266107	1033	435	1550	61029
kalimantan timur	29	236	368544	308	53	597	127525
kalimantan utara	4	10	6659	32	15	33	7537
sulawesi utara	106	29	121166	845	447	2962	21620
sulawesi tengah	26	78	338308	911	202	840	11467
Sulawesi Selatan	40	101	259298	800	1529	3145	119349
sulawesi tenggara	17	26	79253	2083	1267	11019	15923
gorontalo	20	2	842985	61	65	495	28828
sulawesi barat	15	170	15010	720	1442	2721	4436
maluku	27	60	25870	840	354	634	3293
maluku utara	4	80	39301	491	505	1017	6405
papua barat	11	62	7875	1667	341	1912	2668
nanua	71	473	19117	439	14	168	9131

Sumber : <https://bnpb.go.id>

2.3 Load data

Untuk membuat sebuah model langkah pertama adalah memilih kumpulan data/dataset yang digunakan. Tahap ini load dataset dari repositori, setelah ini lanjut ke langkah berikutnya.

2.4 Select task

Setelah load dataset dengan sempurna, selanjutnya menentukan masalah apa yang akan diselesaikan dalam penelitian ini. Dimana masalah yang akan diselesaikan adalah pengelompokan (*clustering*). Tugas ini ingin mengelompokkan data ke dalam ke klaster. Tujuannya bukan memprediksi nilai nilai dalam satu kolom tetapi untuk menemukan set titik data yang berdekatan.

2.5 Prepare target/Select input

Tahap ini menentukan target atau memilih masukan. Kuncinya adalah pencairan pola, tanpa melihat variasi dalam data dan beberapa pola dapat melihat data yang tidak berguna. Dalam persiapan target dan memilih input adalah beberapa yang dipertimbangkan yaitu kolom yang terlalu dekat dengan kolom target, nilai dalam kolom harus berbeda, nilai Kolom semuaidentik, nilai kolom tidak ada yang kosong.

2.6 Algoritma K-Means

Dalam penelitian ini menggunakan algoritma *K-Meansclustering hirarki* untuk pengelompokan propinsi propinsi kedalam 5 kelompok/ *cluster*.

2.7 Hasil yang Tampilkan

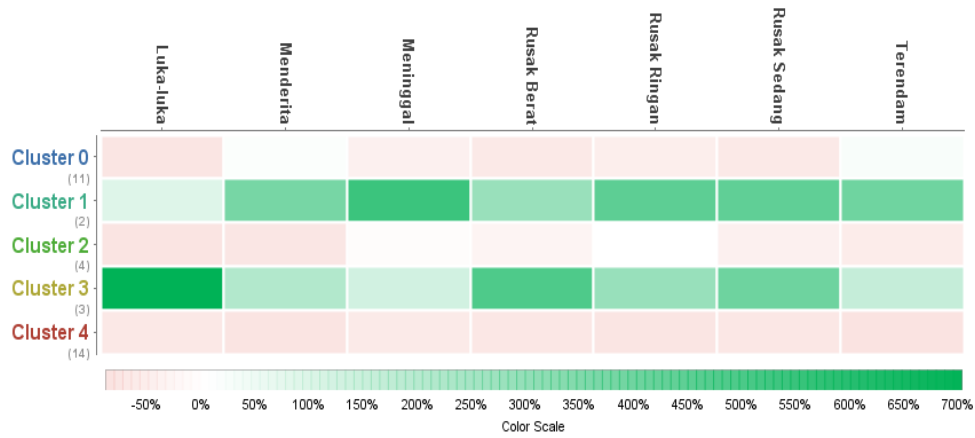
Langkah terakhir dari Model ini adalah output yang ditampilkan secara otomatis. *Output*nya tergantung pada data dan pilihan yang dibuat. Ini akan menunjukkan proses yang melakukan semua *preprocessing* data yang diperlukan dalam optimisasi model. Dapat digunakan model atau sebagai titik awal untuk optimasi lebih lanjut. Hasil *clustering* dapat ditampilkan adalah sebagai berikut:

1. *Heat Map* adalah mengidentifikasi atribut yang paling penting untuk setiap *cluster*.
2. *Cluster Tree* adalah menampilkan pohon keputusan yang menggambarkan perbedaan utama antar *cluster*.
3. *Centroid Chart* adalah menunjukkan nilai-nilai untuk *cluster centroid* dalam grafik paralel
4. *Centroid Table* adalah menunjukkan nilai-nilai untuk *centroid cluster* dalam sebuah tabel.
5. *Scatter Plot* adalah pilihan *cluster*, menampilkan plot sebaran dalam hal dua atribut paling penting.
6. *Data Clustered* adalah menampilkan tabel dengan semua data, termasuk label *cluster* untuk setiap titik data
7. Korelasi adalah matriks yang menunjukkan korelasi antara atribut.

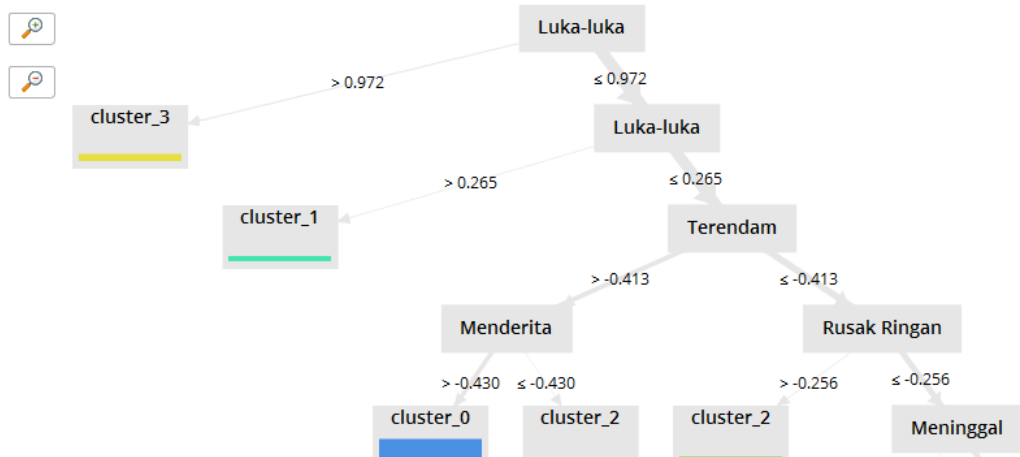
3. ANALISIS DAN PERANCANGAN SISTEM

Hasil penelitian ini didapat beberapa jenis informasi untuk pengelompokan. RapidMiner merupakan software/perangkat lunak untuk pengolahan data. Dengan menggunakan prinsip dan algoritma data mining, RapidMiner dapat mengekstrak pola-pola dari data set yang besar dengan mengkombinasikan metode statistika, kecerdasan buatan dan database. Dan memudahkan penggunaanya dalam melakukan perhitungan data yang sangat banyak dengan menggunakan banyak operator. Hasil dapat ditampilkan secara visual dengan grafik, dari hasil penelitian ini dapat disimpulkan bahwa analisis data bencana, dengan menggunakan metode pengelompokan data bencana menjadi 5 kelompok/*cluster*, terdiri dari

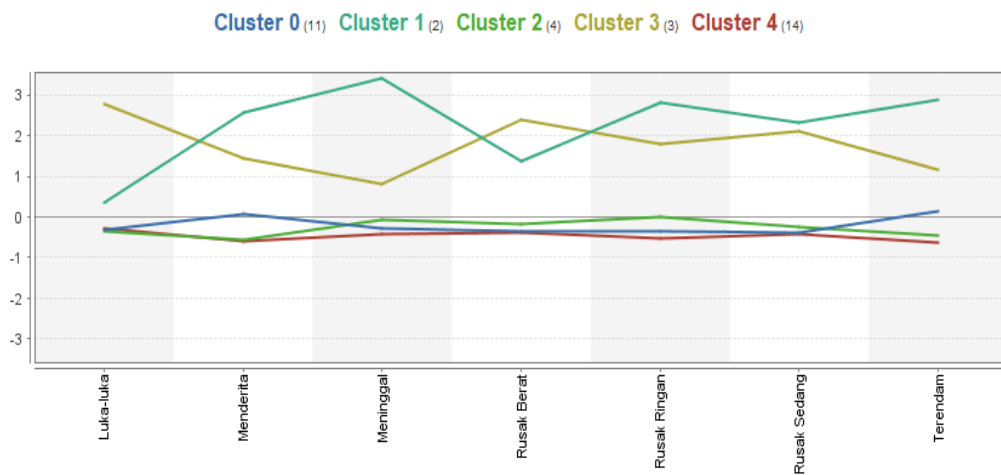
34 propinsi dengan data kejadian tahun 2013-2018 dengan menggunakan 8 atribut. Berikut hasil identifikasi Atribut yang paling penting untuk setiap cluster.



Gambar 2 . K-Means Heat Map



Gambar 3K-Means Cluster Tree



Gambar 4K-Means Centroid Chart

Cluster	Luka-luka	Menderita	Meninggal	Rusak Berat	Rusak Ringan	Rusak Sedang	Terendam
Cluster 0	-0.325	0.079	-0.288	-0.355	-0.340	-0.380	0.138
Cluster 1	0.362	2.564	3.397	1.378	2.809	2.306	2.885
Cluster 2	-0.334	-0.545	-0.077	-0.175	-0.003	-0.252	-0.449
Cluster 3	2.782	1.444	0.794	2.388	1.786	2.111	1.168
Cluster 4	-0.297	-0.582	-0.407	-0.380	-0.515	-0.412	-0.643

Gambar 5K-Means Centroid Table

Tabel 2 Hasil Clustering Penelitian

Propinsi	Meninggal	Luka luka	Mengungsi	Rusak Berat	Sedang	Ringan	Terendam	Hasi Cluster
Sumut	-0.240257	0.24621	0.58740678	-0.120296	-0.157741	-0.140422	0.442414578	cluster_0
Riau	-0.380912	0.394405	-0.4632613	-0.39699006	0.030545	-0.356985	-0.104540065	cluster_0
Jambi	-0.376893	-0.334246	-0.2988089	-0.41101256	-0.355015	-0.441019	-0.061560694	cluster_0
Sumsel	-0.328669	-0.006273	-0.2805364	-0.28831565	-0.36625	-0.440355	0.290264184	cluster_0
DKI	-0.375554	0.657292	0.08491333	-0.48488111	-0.491262	-0.495824	-0.099372278	cluster_0
Banten	-0.374214	0.154123	-0.4358526	-0.27779877	-0.325211	-0.359642	0.003959023	cluster_0
Kalteng	-0.391629	-0.102355	-0.600305	-0.48237709	-0.23024	-0.526382	-0.159821945	cluster_0
Kalsel	-0.370196	-0.286781	-0.5820325	-0.23898648	-0.45034	-0.382228	0.018045213	cluster_0
Kaltim	-0.082187	-0.133155	-0.3993076	-0.42052783	-0.563052	-0.50911	0.830426181	cluster_0
Sulsel	-0.26303	-0.296993	-0.2988089	-0.29733011	-0.2617	-0.018854	0.73054007	cluster_0
Gorontalo	-0.395647	0.578372	-0.4815338	-0.48237709	-0.575115	-0.505124	-0.375354083	cluster_0
Jabar	0.3625521	4.262654	2.99937533	1.842353781	2.818993	3.678989	3.498103686	cluster_1
Jateng	0.3612125	0.865514	3.7942286	0.914364487	2.79806	0.9334241	2.27244638	cluster_1
Sumbar	-0.250973	-0.525811	0.26763822	-0.31235423	-0.156085	-0.207184	-0.251901398	cluster_2
NTT	-0.362158	-0.582815	-0.3718989	-0.12530404	-0.242067	-0.316462	-0.54615982	cluster_2
Sulut	-0.359479	-0.504152	0.3041832	-0.28606203	-0.283343	-0.378242	-0.463414147	cluster_2
Sulteng	-0.363498	-0.567009	-0.5089426	0.023935472	0.669557	-0.105878	-0.533014339	cluster_2
Aceh	4.8206586	1.621142	1.07162774	2.334643839	1.44387	1.8395336	1.541381698	cluster_3
Jatim	1.9432509	1.722977	1.71116485	4.38343186	2.790845	2.7064492	2.176494141	cluster_3
NTB	1.5815656	0.986736	-0.3993076	0.447114616	1.122176	1.7870536	-0.212379436	cluster_3
Bengkulu	-0.383591	-0.671673	-0.189174	-0.41176376	-0.549569	-0.440355	-0.696135589	cluster_4
Lampung	-0.328669	-0.628493	-0.3993076	-0.39398523	-0.269151	-0.40714	-0.450073193	cluster_4
Babel	-0.38895	-0.62527	-0.4267164	-0.4836291	-0.592737	-0.456962	-0.683258882	cluster_4
Kepri	-0.374214	-0.682789	-0.5454875	-0.45483288	-0.540935	-0.408468	-0.726103866	cluster_4
DIY	-0.339385	-0.460709	-0.3170814	-0.46234494	-0.547322	-0.502135	-0.721449193	cluster_4
Bali	-0.340725	-0.49449	-0.3079452	-0.44030958	-0.584577	-0.381896	-0.718871408	cluster_4
Kalbar	-0.392968	-0.463421	-0.3079452	-0.46785378	-0.563052	-0.432051	-0.456841406	cluster_4
Kalut	-0.384931	-0.67588	-0.6277137	-0.48963874	-0.629756	-0.521732	-0.635466017	cluster_4
Sulteng	-0.29384	-0.1785	-0.4267164	-0.26953551	-0.534312	-0.45962	-0.587453246	cluster_4
Sulbar	-0.170599	-0.663356	-0.5272151	-0.31736226	-0.311846	-0.047752	-0.673350903	cluster_4
Maluku	-0.317952	-0.647069	-0.4175801	-0.28731404	-0.558676	-0.409133	-0.687314923	cluster_4
Maluku utara	-0.291161	-0.626926	-0.6277137	-0.37470429	-0.513378	-0.358978	-0.64929565	cluster_4
Papua Barat	-0.315273	-0.674056	-0.56376	-0.0802317	-0.407527	-0.413451	-0.694950541	cluster_4
Papua	0.1683137	-0.657204	-0.0155854	-0.38772519	-0.613789	-0.522064	-0.615992135	cluster_4

Attributes	Luka-luka	Menderita	Meninggal	Rusak Berat	Rusak Ringan	Rusak Sedang	Terendam
Luka-luka	1	0.525	0.407	0.746	0.591	0.669	0.502
Menderita	0.525	1	0.730	0.711	0.804	0.888	0.869
Meninggal	0.407	0.730	1	0.670	0.867	0.732	0.863
Rusak Berat	0.746	0.711	0.670	1	0.856	0.868	0.755
Rusak Ringan	0.591	0.804	0.867	0.856	1	0.909	0.860
Rusak Sedang	0.669	0.888	0.732	0.868	0.909	1	0.825
Terendam	0.502	0.869	0.863	0.755	0.860	0.825	1

Gambar 6K-Means Korelasi

Setelah melakukan penelitian, hasil ditampilkan dalam bentuk tabel, berikut hasil *clustering* dengan 34 propinsi di Indonesia dengan menggunakan 5 *cluster* adalah sebagai berikut:

Tabel 3 Hasil jumlah per Cluster

Nama Cluster	Jumlah Cluster
Cluster 0	11
Cluster 1	2
Cluster 2	4
Cluster 3	3
Cluster 4	14

Hasil Cluster dengan K=5 adalah bertujuan agar mendapatkan nilai kelompok yang baik akurasinya, adapun hasil pengelompokan yang paling Rawan adalah sebagai berikut:

1. Meninggal adalah termasuk cluster 1 yaitu propinsi Jawa Barat, Jawa Tengah
2. Luka luka adalah termasuk cluster 3 yaitu propinsi Aceh, Jawa Timur, NTB
3. Menderita/mengungsi adalah termasuk cluster 1 yaitu propinsi Jawa Barat, Jawa Tengah
4. Rusak berat adalah termasuk cluster 3 yaitu propinsi propinsi Aceh, Jawa Timur, NTB
5. Rusak ringan adalah termasuk cluster 1 yaitu propinsi Jawa Barat, Jawa Tengah
6. Rusak sedang adalah termasuk cluster 1 yaitu propinsi Jawa Barat, Jawa Tengah
7. Terendam adalah termasuk cluster 1 yaitu propinsi Jawa Barat, Jawa Tengah,

Berikut propinsi kategori aman dibandingkan propinsi lainnya, yaitu semua cluster 4. Berikut propinsinya Bengkulu, Lampung, Kepri, Babel, DIY, Bali, Kalbar, Kalut, Sulteng, Sulbar, Maluku, Maluku Utara, Papua, Papua Barat.

4. PENUTUP

4.1 Kesimpulan

Setelah dilakukan penelitian ini, dengan dataset jumlah propinsi di Indonesia dengan 8 atribut, dengan data tahun 2013-2018 yang menggambarkan tingkat kerugian baik materiil dan spiritual yang disebabkan bencana dengan mengelompokkan sebanyak 5 cluster. Untuk menganalisa data peneliti menggunakan Rapidminer 9.1, untuk mempercepat mendapatkan hasil, ada beberapa kesimpulan yang dapat diambil adalah:

1. Berdasarkan hasil *clustering K-means* menunjukkan luka luka dan rusak berat masuk kedalam cluster 1 yaitu Jawa Barat dan Jawa Tengah.
2. Berdasarkan hasil *clustering K-means* menunjukkan korban meninggal, mengungsi, rusak sedang, rusak ringan, terendam masuk kedalam cluster 1 yaitu propinsi Aceh, Jawa Timur, NTB.

3. Daerah yang paling aman masuk kedalam *cluster* 4 yaitu propinsi Bengkulu, Lampung, Kepri, Babel, DIY, Bali, Kalbar, Kalut, Sulteng, Sulbar, Maluku, Maluku Utara. Papua, Papua Barat.

4.2 Saran

Dengan hasil penelitian ini dapat dijadikan pemerintah sebagai informasi yang akurat, relevan dan tepat, untuk memetakan rawan bencana dindonesia, agar korban dapat ditangani dengan maksimal. Dan diharapkan peneliti berikut dapat memberikan banyak referensi ilmiah untuk melanjutkan analisis lokasi rawan bencana di Indonesia.

DAFTAR PUSTAKA

- Agusta., 2007., Implementasi Algoritma *Clustering* Dengan *Singular Vector Decomposition* Untuk Menunjang Keputusan Dalam Meningkatkan Produktivitas Tanaman Jagung., Denpasar., bali.
- Connoly, T., & Begg, c., 2005., A practical approach to design, implementation and management (4th ed.), Harlow Addison Wesley.
- Nugraha et al., 2013. Penyusunan dan Penyajian Peta online risiko bencana banjir rob di kota semarang., Yogyakarta., UGM.
- Nugroho et al., 2009., Pemetaan Daerah Rawan Longsor dengan Penginderaan Jauh dan Sistem Informasi Georafis., Surabaya., ITS.
- Siregar AM, 2018., Implementasi algoritma *clustering* dengan *singular vector decomposition* untuk menunjang keputusan dalam meningkatkan produktivitas tanaman jagung., Bogor.,IPB
- Tan, P.N., Steinbech, M.Kumar, V. 2006., Introduction to data mining. Boston., Pearsong Education., Ltd
- Wu, X.,Kumar, V., 2009., The top ten algorithms in data mining. Boca Raton.,CRC Press
<https://www.bbc.com/indonesia/indonesia-46738003>
- <https://www.liputan6.com/news/read/3861499/2500-bencana-alam-ancam-indonesia-tahun-2019>
- <https://muhammad-mulyana.blogspot.co.id>
- <https://nasional.kompas.com/read/2018/12/26/12145181/budaya-sadar-bencana-masyarakat-indonesia-masih-rendah>
- <https://megapolitan.kompas.com/read/2018/12/25/13241341/banyak-korban-pendatang-pemilik-hotel-diimbau-sosialisasikan-soal-bencana>
- <https://jurnal.unmuhjember.ac.id>