

PERNYATAAN

Saya menyatakan bahwa skripsi ini adalah hasil karya saya sendiri dan bukan plagiat dari orang lain. Apabila di kemudian hari terbukti melakukan plagiat dalam skripsi ini, maka saya bersedia menerima sanksi sesuai ketentuan peraturan perundang-undangan.

Bojonegoro, 22 Agustus 2023

Yang membuat pernyataan



Yuliana Fuji Rahayu

NIM :2520190031


HALAMAN PERSETUJUAN

Nama : Yuliana Fuji Rahayu
NIM : 2520190031
Program Studi/Fakultas : Statistika/Sains dan Teknologi
Judul : Implementasi *Data Mining* Melalui Pengelompokan dan Asosiasi Untuk Peringatan Dini Bencana di Kabupaten Bojonegoro

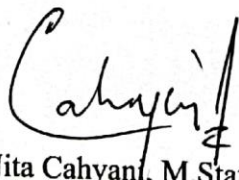
Menyatakan bahwa Mahasiswa tersebut telah disetujui dan memenuhi syarat untuk diajukan dalam ujian Skripsi.

Bojonegoro, 18 Juli 2023

Pembimbing I


Denny Murdiansyah, M.Si.
NIDN.0726058702

Pembimbing II


Nita Cahyan, M.Stat
NIDN. 0704038906

HALAMAN PENGESAHAN

Nama : Yuliana Fuji Rahayu
NIM : 2520190031
Judul skripsi : Implementasi *Data Mining* Melalui Pengelompokan dan Asosiasi Untuk Peringatan Dini Bencana di Kabupaten Bojonegoro

Telah dipertahankan di hadapan penguji pada tanggal 21 Agustus 2023

Dewan Penguji

Tim Pembimbing

Penguji I

Pembimbing I



Nur Mahmudah, M.Stat

NIDN. 0715039201



Denny Nurdiansyah, M.Si.

NIDN.0726058702

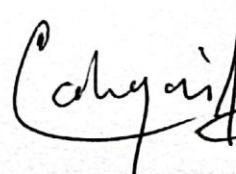
Penguji II

Pembimbing II



Dr. H. M. Ridlwan Hambali, Lc., M.A.

NIDN. 2117056803



Nita Cahyani, M.Stat.

NIDN. 0704038906

Mengetahui,

Mengetahui,

Dekan Fakultas Sains dan Teknologi

Ketua Program Studi



FAKULTAS
FST
UNUGIRI
Sunu Wahyudhi, M.Pd.
NIDN.0709058902



S T
FST
UNUGIRI
Nita Cahyani, M.Stat.
NIDN. 0704038906

MOTTO

Sesungguhnya sesudah kesulitan itu ada kemudahan, maka apabila kamu telah selesai (dari sesuatu urusan), kerjakanlah dengan sungguh-sungguh (urusan) yang lain, dan hanya kepada Tuhanmulah hendaknya kamu berharap.

(Al-Insyirah: 6-8)

PERSEMBAHAN

Ayah dan Ibuku tercinta

A. Fauzi (Ayah) dan Naeni (Ibu)

Terima kasih untuk setiap do'a yang selalu dipanjatkan agar aku senantiasa diberikan kemudahan dalam menjalani hidup. Terima kasih untuk setiap usaha yang dilakukan demi memberikan yang terbaik untukku. Terutama untuk ibuku yang tercinta terimakasih telah memberiku semangat dan kepercayaan untuk segala hal yang sedang aku usahakan.

UNUGIRI

KATA PENGANTAR

Puji syukur dipanjatkan kepada Allah SWT karena atas segala nikmatnya dan karunia-Nya penyusunan skripsi yang berjudul “Implementasi *Data Mining* Untuk Pengelompokan dan Asosiasi Bencana di Kabupaten Bojonegoro”. Dapat dirampungkan dengan baik. terselesainya skripsi tersebut tidak lepas dari berbagai bantuan serta dukungan dari berbagai pihak. Oleh karena itu, ucapan terima kasih disampaikan kepada :

1. Bapak K. H. Jauharul Ma'arif, M.Pd.I selaku Rektor Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
2. Bapak Sunu Wahyudhi, M.Pd selaku Dekan Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
3. Ibu Nita Cahyani, M.Stat selaku Ketua Program Studi Statistika Universitas Nahdlatul Ulama Sunan Giri Bojonegoro
4. Bapak Denny Nurdiansyah, M.Si selaku Dosen Pembimbing I, dan Ibu Nita Cahyani, M.Stat selaku Dosen Pembimbing II, yang telah memberikan ilmu yang bermanfaat dalam penyempurnaan penulisan pada skripsi ini.
5. Bapak Denny Nurdiansyah, M.Si selaku Dosen Wali yang telah memberikan bimbingan serta nasehat selama masa perkuliahan.
6. Terutama kepada kedua orang tua, keluarga, dan teman-teman yang telah memberikan bantuan sejak memulai perkuliahan hingga skripsi ini terselesaikan.

Penyusunan skripsi ini masih jauh dari kata sempurna, sehingga besar harapan terkait adanya kritik serta saran yang membangun demi perbaikan penulisan skripsi di masa mendatang. Skripsi ini diharapkan dapat memberikan manfaat kepada berbagai pihak.

Bojonegoro, 20 Juni 2023

Penulis,

Yuliana Fuji Rahayu

ABSTRACT

Rahayu, Yuliana Fuji. 2023. *Implementation of Data Mining for Disaster Grouping and Association in Bojonegoro Regency*. Thesis, Department of Statistics, Faculty of Science and Technology, Nahdlatul Ulama Sunan Giri University. Main Advisor Denny Nurdiansyah, M.Si. and Advisor Nita Cahyani, M.Stat. Disasters are defined as a series of events or occurrences that can result in human suffering, property loss, environmental damage, facilities and infrastructure and affect people's lives and livelihoods. Based on data obtained from the Fire Department in Bojonegoro Regency and BPBD of Bojonegoro Regency, throughout 2019 in Bojonegoro Regency there were 392 disasters with a total of 695 affected villages, 20,801 affected victims, and a total material loss of Rp 15.44 billion. The impact of disasters can be reduced or can be minimized by studying disaster events and patterns formed from disaster data, namely by doing data mining such as mapping, clustering, and association rules. This research was conducted with the aim of seeing the relationship between one disaster and another disaster to measure how likely a disaster is to occur. Disaster events can be difficult to anticipate when there is a lack of knowledge and understanding of disasters such as in Bojonegoro Regency in the hope that disasters can be more easily calculated in the future. The methods applied in this research are mapping, clustering with the K-Means algorithm and Association Rule with the Apriori algorithm. Secondary data obtained from the Regional Disaster Management Agency of Bojonegoro Regency in the range of January 5, 2019 to January 18, 2023 with 8 types of disasters namely, House Fires, Extreme Weather, Landslides, Forest and Land Fires, Overflow Floods, Flash Floods, Drought, and Other Events. The results of this study showed that the K-Means model divided the data into 5 groups of 28 sub-districts in Bojonegoro Regency. Cluster 0 has 13 sub-district members, cluster 1 has only 1 sub-district member, cluster 2 has 4 sub-district members, cluster 3 has 6 sub-district members, cluster 4 has 4 sub-district members. The association analysis results obtained with a minimum support of 10% and a minimum confidence of 50% resulted in four association rules. The conclusion of this research shows that sub-districts that require more handling are Ngasem and Bojoengoro sub-districts, from 4 association rules formed, the fourth rule has the highest confidence value of 78.79% if there is a drought disaster, there will be a forest and land fire disaster.

Keywords: Disaster, Mapping, K-Means, Association Rule

UNUGIRI

ABSTRAK

Rahayu, Yuliana Fuji. 2023. *Implementasi Data Mining Untuk Pengelompokan dan Asosiasi Bencana Di Kabupaten Bojonegoro*. Skripsi. Program Studi Statistika Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri. Pembimbing Utama Denny Nurdiansyah, M.Si. dan Pembimbing Pendamping Nita Cahyani, M.Stat. Bencana didefinisikan sebagai rangkaian peristiwa atau kejadian yang dapat mengakibatkan penderitaan manusia, kerugian harta benda, kerusakan lingkungan, sarana dan prasarana serta mempengaruhi kehidupan dan penghidupan masyarakat. Berdasarkan data yang diperoleh dari Dinas Pemadam Kebakaran di Kabupaten Bojonegoro dan BPBD Kabupaten Bojonegoro, sepanjang tahun 2019 di Kabupaten Bojonegoro terjadi bencana sebanyak 392 kejadian dengan jumlah desa terdampak sebanyak 695 desa, korban terdampak 20.801 kepala keluarga (KK), dan total kerugian material mencapai Rp 15.44 miliar. Dampak bencana dapat dikurangi atau dapat diminimalkan dengan mempelajari kejadian bencana dan pola yang terbentuk dari data bencana, yaitu dengan melakukan penggalian data (*data mining*) seperti pemetaan, pengelompokan, dan *association rule*. Penelitian ini dilakukan dengan tujuan untuk melihat hubungan antara bencana satu dengan bencana lain mengukur seberapa mungkin bencana terjadi. Kejadian bencana bisa sulit diantisipasi ketika kurangnya pengetahuan dan pemahaman tentang bencana seperti di Kabupaten Bojonegoro dengan harapan bencana dapat lebih mudah diperhitungkan kedepannya. Metode yang diterapkan dalam penelitian ini adalah pemetaan, *clustering* dengan algoritma *K-Means* dan *Association Rule* dengan algoritma Apriori. Data sekunder didapatkan dari Badan Penanggulangan Bencana Daerah Kabupaten Bojonegoro dalam rentang 5 Januari 2019 sampai 18 Januari 2023 dengan 8 jenis bencana yaitu, Kebakaran Rumah, Cuaca Ekstrem, Tanah Longsor, Kebakaran Hutan dan Lahan, Banjir Luapan, Banjir Bandang, Kekeringan, dan Kejadian Lain-lain. Hasil kajian ini diketahui model *K-Means* membagi data menjadi 5 kelompok dari 28 Kecamatan di Kabupaten Bojonegoro. Cluster 0 mempunyai anggota 13 kecamatan, cluster 1 hanya mempunyai 1 anggota kecamatan, cluster 2 mempunyai anggota 4 kecamatan, cluster 3 mempunyai anggota 6 kecamatan, cluster 4 mempunyai anggota 4 kecamatan. Hasil analisis asosiasi yang didapatkan dengan minimum *support* 10% dan minimum *confidence* 50% menghasilkan empat aturan asosiasi. Kesimpulan dari penelitian ini menunjukkan kecamatan yang memerlukan penanganan lebih yaitu Kecamatan Ngasem dan Kecamatan Bojoengoro, dari 4 aturan asosiasi yang terbentuk aturan ke-empat mempunyai nilai *confidence* tertinggi yaitu 78.79% apabila terjadi bencana kekeringan maka akan terjadi bencana kebakaran hutan dan lahan.

Kata Kunci: Bencana, Pemetaan, *K-Means*, *Association Rule*

DAFTAR ISI

	Halaman
HALAMAN SAMPUL	i
HALAMAN SAMPUL DALAM	ii
PERNYATAAN	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PENGESAHAN	v
MOTTO	vi
PERSEMBAHAN	vi
KATA PENGANTAR	vii
ABSTRACT	viii
ABSTRAK	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR BAGAN	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	5
1.4 Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI	7
2.1 Dasar Teori.....	7
2.2 Kerangka Teori	11
2.2.1 Bencana	11
2.2.2 Distribusi Frekuensi.....	12
2.2.3 Pemetaan.....	12
2.2.4 Data Mining	13
2.2.5 Algoritma K-means	14
2.2.6 Metode Apriori	16
BAB III METODOLOGI PENELITIAN	17

3.1 Sumber Data.....	17
3.2 Variabel Penelitian.....	17
3.3 Langkah Analisis	19
3.3.1 Pemetaan.....	20
3.3.2 Clustering	20
3.3.3 Association Rule.....	21
3.4 Diagram Alir	22
BAB IV HASIL DAN PEMBAHASAN	24
4.1 Distribusi Frekuensi	24
4.2 Pemetaan Bencana	26
4.2.1 Bencana Kebakaran Rumah.....	26
4.2.2 Bencana Cuaca Ekstrem	28
4.2.3 Bencana Banjir Luapan	29
4.2.4 Bencana Tanah Longsor	30
4.2.5 Bencana Banjir Bandang	32
4.2.6 Bencana Kebakaran Hutan dan Lahan.....	33
4.2.7 Bencana Kekeringan.....	34
4.2.8 Bencana Kejadian Lain-lain	36
4.3 Implementasi Clustering dengan RapidMiner	37
4.3.1. Permodelan Data Pada RapidMiner	37
4.3.2. Penentuan Jumlah Cluster	38
4.3.3. Melakukan Pengelompokan Dengan K-Means Dan Bantuan Software RapidMiner	39
4.3.4. Peta Hasil Clustering	47
4.4 Melakukan Proses Association Rule Algoritma Apriori.....	49
4.4.1. Pembentukan 1 Itemset.....	49
4.4.2. Pembentukan 2 itemset.....	50
4.4.3. Pembentukan Aturan Asosiasi.....	52
4.4.4. Pengujian Dengan RapidMiner	55
BAB V KESIMPULAN	58
DAFTAR PUSTAKA	60
LAMPIRAN.....	63

DAFTAR TABEL

Tabel	Halaman
2.1 Penelitian Terdahulu	7
3.1 Definisi Variabel Penelitian Untuk Metode <i>K-Means</i> dan Pemetaan.....	17
3.2 Struktur Data Untuk <i>K-Means</i>	18
3.3 Definisi Variabel Penelitian Untuk Metode Apriori	18
3.4 Struktur Data Penelitian Untuk Apriori	18
4.1 Hasil Distribusi Frekuensi.....	24
4.2 Hasil <i>Average Within Cluster</i>	39
4.3 <i>Output Cluster Model K-Means</i>	39
4.4 Hasil Deskriptif Tiap Kelas	41
4.5 Hasil Kelompok <i>Clustering</i> dan Pemetaan	498
4.6 Nilai <i>Support 1</i> Itemset	49
4.7 Nilai <i>Support 2</i> Itemset	50
4.8 Hasil Perhitungan <i>Confidence</i>	52
4.9 Hasil Aturan Asosiasi.....	55

UNUGIRI

DAFTAR BAGAN

Bagan	Halaman
3.1 Diagram Alir Penelitian.....	22



DAFTAR GAMBAR

Gambar	Halaman
2.1 <i>Clustering</i> dari 3 kelompok.....	15
4.1 Peta Pesebaran Kejadian Bencana Kebakaran Rumah.....	27
4.2 Peta Pesebaran Kejadian Bencana Cuaca Ekstrem	28
4.3 Peta Pesebaran Kejadian Bencana Banjir Luapan	29
4.4 Peta Pesebaran Kejadian Bencana Tanah Longsor	31
4.5 Peta Pesebaran Kejadian Bencana Banjir Bandang	32
4.6 Peta Pesebaran Kejadian Bencana Kebakaran Hutan dan Lahan.....	33
4.7 Peta Pesebaran Kejadian Bencana Kekeringan.....	35
4.8 Peta Pesebaran Kejadian Bencana Kejadian Lain-lain	36
4.9 Hasil Average Within Cluster	38
4.10 Output Diagram Scatter Model Clustering K-Means	39
4.11 Output Performace Vektor Model K-Means	41
4.12 Peta Hasil <i>Clustering</i>	41
4.13 Susunan Operasi Algoritma Apriori.....	56
4.14 Parameter Yang Digunakan Dalam Algoritma Apriori	56
4.15 Hasil Proses Algoritma Apriori.....	57

UNUGIRI

DAFTAR LAMPIRAN

Lampiran	Halaman
1. <i>Output</i> tabel <i>statistic</i> distribusi frekuensi.....	64
2. <i>Output</i> data exampeset Metode K-Means	65
3. Permodelan data pada RapidMiner untuk Metode <i>K-Means</i>	66
4. <i>Output</i> Performance Vector Software Rapidminer untuk Metode K-Means dengan 2 cluster.....	66
5. <i>Output</i> Performance Vector Software Rapidminer untuk Metode K-Means dengan 3 cluster.....	67
6. <i>Output</i> Performance Vector Software Rapidminer untuk Metode K-Means dengan 4 cluster.....	67
7. <i>Output</i> Performance Vector Software Rapidminer untuk Metode K-Means dengan 5 cluster.....	68
8. <i>Output</i> Performance Vector Software Rapidminer untuk Metode K-Means dengan 6 cluster.....	68
9. <i>Output</i> Performance Vector Software Rapidminer untuk Metode K-Means dengan 7 cluster.....	69
10. <i>Output</i> Performance Vector Software Rapidminer untuk Metode K-Means dengan 8 cluster.....	69
10. <i>Output</i> Performance Vector Software Rapidminer untuk Metode K-Means dengan 9 cluster.....	70
11. <i>Output</i> Performance Vector Software Rapidminer untuk Metode K-Means dengan 10 cluster.....	70
12. Hasil Nilai Support 3 Itemset pada <i>Association Rule</i>	71
13. Sebagian dataset yang digunakan untuk distribusi frekuensi.....	72
14. Sebagian dataset yang digunakan untuk pemetaan distribusi	72
15. Sebagian dataset yang digunakan untuk Algoritma K-Means	72
16. Sebagian dataset yang digunakan untuk <i>Association Rule</i>	73