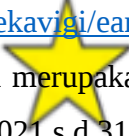


BAB III

METODE PENELITIAN

3.1. Objek Penelitian

Bahan yang akan digunakan merupakan data kejadian gempa bumi di Indonesia yang berisi 92887 data dan bersumber dari Badan Meteorologi, Klimatologi dan Geofisika (BMKG) yang didapatkan dengan cara mengakses website Kaggle.com dengan alamat tautan <https://www.kaggle.com/datasets/kekavigi/earthquakes-in-indonesia>. Data yang digunakan pada penelitian ini merupakan data kejadian gempa bumi di Pulau Jawa dari tanggal 3 Januari 2021 s.d 31 Desember 2022



	tgl	ot	lat	lon	depth	mag	remark	strike1	dip1	rake1	strike2	dip2	rake2
0	2008/11/01	21:02:43.058	-9.18	119.06	10	4.9	Sumba Region - Indonesia	NaN	NaN	NaN	NaN	NaN	NaN
1	2008/11/01	20:58:50.248	-6.55	129.64	10	4.6	Banda Sea	NaN	NaN	NaN	NaN	NaN	NaN
2	2008/11/01	17:43:12.941	-7.01	106.63	121	3.7	Java - Indonesia	NaN	NaN	NaN	NaN	NaN	NaN
3	2008/11/01	16:24:14.755	-3.30	127.85	10	3.2	Seram - Indonesia	NaN	NaN	NaN	NaN	NaN	NaN
4	2008/11/01	16:20:37.327	-6.41	129.54	70	4.3	Banda Sea	NaN	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...
92882	2023/01/26	02:25:09.288	3.24	127.18	10	4.0	Talaud Islands - Indonesia	NaN	NaN	NaN	NaN	NaN	NaN
92883	2023/01/26	02:15:03.893	2.70	127.10	10	3.9	Northern Molucca Sea	NaN	NaN	NaN	NaN	NaN	NaN
92884	2023/01/26	01:57:08.885	-7.83	121.07	10	3.8	Flores Sea	NaN	NaN	NaN	NaN	NaN	NaN
92885	2023/01/26	01:46:21.009	3.00	127.16	10	4.1	Northern Molucca Sea	NaN	NaN	NaN	NaN	NaN	NaN
92886	2023/01/26	00:00:35.181	-8.87	118.95	10	2.4	Sumbawa Region - Indonesia	NaN	NaN	NaN	NaN	NaN	NaN

92887 rows x 13 columns

Gambar 3. 1 Data Keseluruhan

	tanggal	depth	magnitude	wilayah
70214	2021-01-03	19	3	Java - Indonesia
70226	2021-01-04	53	3	South of Java - Indonesia
70227	2021-01-04	10	3	South of Java - Indonesia
70232	2021-01-04	25	3	Java - Indonesia
70241	2021-01-05	36	3	Java - Indonesia
...	...	...	...	...
91650	2022-12-29	11	3	Java - Indonesia
91667	2022-12-30	19	2	Java - Indonesia
91693	2022-12-30	10	3	Java - Indonesia
91699	2022-12-31	26	1	Java - Indonesia
91703	2022-12-31	10	4	South of Java - Indonesia

[2072 rows x 4 columns]

Gambar 3. 2 Data Pulau Jawa 2021-2022

3.2. Lokasi dan Waktu Penelitian

Lokasi penelitian dilakukan di Kampus Universitas Buana Perjuangan Karawang, Jalan Ronggo Waluyo Sirnabaya, Puseurjaya, Kecamatan Telukjambe Timur, Kabupaten Karawang, Jawa Barat. Waktu penelitian

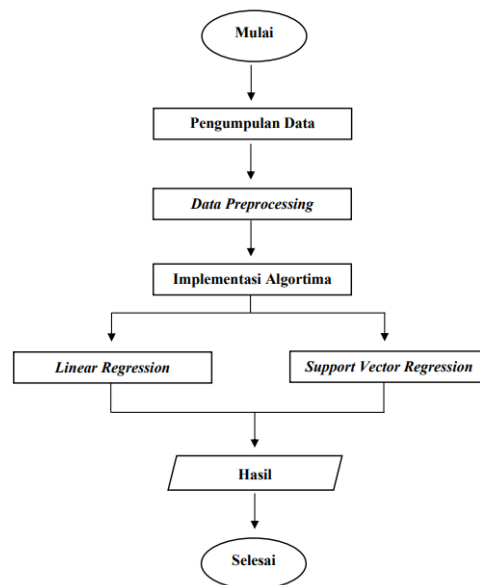
dilakukan selama 11 bulan dari Oktober 2022 s.d. maret 2023. Rincian waktu penelitian dijabarkan dalam tabel 3.1

Tabel 3. 1 Waktu Penelitian

No	Keterangan	Okt 2022				Nov 2023				Des 2023				Jan 2023				Feb 2023				Mar 2023			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Pengumpulan data																								
2	<i>Pre processing</i>																								
3	Perhitungan Manual																								
4	Penerapan Algoritma																								
5	Evaluasi																								

3.3. Prosedur Penelitian

Prosedur penelitian dilaksanakan secara bertahap, dimulai dari tahap pengumpulan data, pra-pemrosesan data, pembuatan model algoritma, dan evaluasi. Tahapan prosedur penelitian didefinisikan dalam bentuk diagram alir pada Gambar 3.3.



Gambar 3. 3 Prosedur Penelitian

3.3.1. Pengumpulan Data

Pengumpulan data dilakukan dengan mengakses website Kaggle pada laman tautan <https://www.kaggle.com/datasets/kekavigi/earthquakes-in-indonesia>. Data tersebut merupakan data gempa bumi di Indonesia dan data yang akan digunakan merupakan data kejadian Pulau Jawa tahun 2021-2022.

3.3.2. Data Pre-Processing

Preprocessing merupakan tahap untuk memperbaiki kualitas data sebelum dilakukan proses *mining*. Pada tahap ini dilakukan pembersihan data pada sebuah data yang akan digunakan agar data menjadi berkualitas, seperti membuang data yang tidak diperlukan dan memperbaiki kesalahan pada data.

1. Data Selection

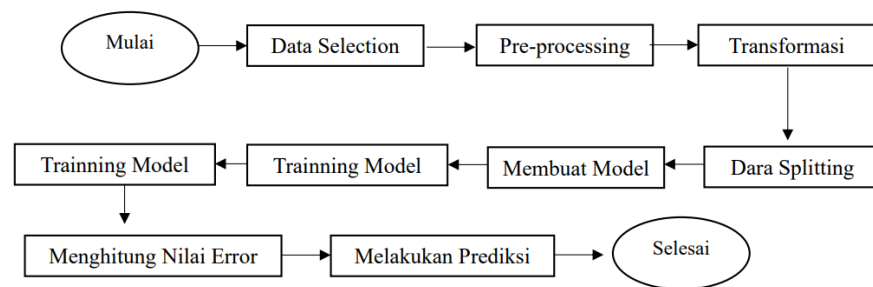
Data Selection yaitu proses untuk menyederhanakan data yaitu dengan memilih atribut-atribut yang akan mendukung proses pembentukan model prediksi.

2. Transformasi Data

Proses ini yaitu mengubah atau merombak data sehingga data sesuai, yang berguna untuk menyajikan data dengan tujuan yang ingin dicapai dan prosesnya merubah data yang dipilih agar sesuai dengan proses data mining.

3.4.Implementasi Algoritma

Pada tahap ini dilakukan pembuatan model untuk melakukan sebuah prediksi. Metode yang digunakan yaitu *Linear Regression* dan *Support Vector Regression*. *Linear Regression* adalah analisis statistika yang memodelkan hubungan beberapa variabel menurut bentuk hubungan persamaan *Linear Regression* (Nafi'iyah & Aulia, 2022) . sedangkan *Support Vector Regression* merupakan pengembangan dari metode SVM yang digunakan pada kasus regresi (Arfan & ETP, 2020).



Gambar 3. 4 Alur Metode

3.4.1. Data Selection

Data Selection merupakan proses penyeleksian/pemilihan atribut data yang relevan untuk digunakan pada penelitian ini, atribut yang dianggap tidak berguna akan dihapus dan tidak akan digunakan. Dalam penelitian ini dataset terdapat 2061 data yang mana data tersebut memiliki 7 atribut, data atribut tersebut terdiri dari tgl, depth, mag dan remark.

	tgl	depth	mag	remark
0	2008/11/01	10	4.9	Sumba Region - Indonesia
1	2008/11/01	10	4.6	Banda Sea
2	2008/11/01	121	3.7	Java - Indonesia
3	2008/11/01	10	3.2	Seram - Indonesia
4	2008/11/01	70	4.3	Banda Sea
...	...	...	...	...
92882	2023/01/26	10	4.0	Talaud Islands - Indonesia
92883	2023/01/26	10	3.9	Northern Molucca Sea
92884	2023/01/26	10	3.8	Flores Sea
92885	2023/01/26	10	4.1	Northern Molucca Sea
92886	2023/01/26	10	2.4	Sumbawa Region - Indonesia

92887 rows × 4 columns

Gambar 3. 5 *Data Selection*

3.4.2. Pre-processing

Preprocessing merupakan tahap untuk memperbaiki kualitas data sebelum dilakukan proses mining. Pada tahap ini dilakukan pembersihan data pada sebuah data yang akan digunakan agar data menjadi berkualitas, seperti membuang data yang tidak diperlukan dan memperbaiki kesalahan pada data. Pada umumnya, data yang diperoleh dari sumbernya memiliki isian-isian yang tidak sempurna seperti data yang hilang, missing value atau

hanya sekedar salah ketik. Data-data yang tidak relevan itu juga lebih baik dibuang karena keberadaanya bisa mengurangi akurasi dari hasil data mining nantinya. Pembersihan data juga akan mempengaruhi performansi dari sistem data mining karena data yang akan ditangani akan berkurang jumlah dan kompleksitasnya. Untuk mengetahui ada atau tidaknya data yang salah, maka dilakukan visualisasi data

3.4.3. Transformation

Pada tahap transformation dilakukan perubahan tipe data yang digunakan. Perubahan tipe data dilakukan agar data dapat diproses saat proses mining. Proses transformasi merupakan proses yang sangat tergantung pada jenis atau pola informasi yang akan dicari dalam data. Pada tahap ini transformasi dilakukan pada atribut prediktor (x), sehingga data siap digunakan dalam proses selanjutnya

3.4.4. Data Splitting

Pada tahap Data Splitting dilakukan pemisahan data yang digunakan untuk memisahkan dari independent variabel dan dependent variabel

3.4.5. Pembuatan Model

Pada tahap ini dilakukan pembuatan model untuk melakukan sebuah prediksi. Metode yang digunakan yaitu *Linear Regression* dan *Support Vector Regression*.

3.4.6. Training Model

Pada tahap ini dilakukan proses training terhadap model yang sudah dibuat terhadap data train yang telah ditentukan saat data splitting

3.4.7. Testing Model

Setelah model melakukan training, selanjutnya model digunakan untuk melakukan sebuah prediksi. Data hasil prediksi nantinya akan dibandingkan dengan data testing yang telah ditentukan saat proses data splitting

3.4.8. Menghitung Nilai Error

Hasil prediksi pada tahap sebelumnya, akan dibandingkan dengan data testing untuk mencari nilai error dari setiap prediksi yang telah dilakukan.

Perhitungan didasarkan pada nilai penyimpangannya meliputi, mean square error (MSE), root mean square error (RMSE), dan mean absolute percentage error (MAPE). Dari hasil ini, maka nantinya dapat kita ketahui metode yang lebih baik dalam melakukan sebuah prediksi.

3.4.9. Melakukan Prediksi

Setelah menemukan metode yang lebih baik diantara linear regresi atau polynomial regresi, tahap selanjutnya yaitu melakukan prediksi terhadap kekuatan gempa di Pulau Jawa. Prediksi dapat dilakukan dengan menentukan atribut prediktor (x) untuk dimasukkan kedalam model yang telah dibuat pada pembuatan model.

