

# KLASIFIKASI EFEK KERUSAKAN GEMPA BUMI BERDASARKAN SKALA MODIFIED MERCALLI INTENSITY MENGGUNAKAN ALGORITMA MULTICLASS SUPPORT VECTOR MACHINE

Raka Wahyu Pratama, Yulison Herry Chrisnanto, Gunawan Abdillah

Informatika, Universitas Jenderal Achmad Yani Cimahi

Jl. Terusan Jend. Sudirman, Cibeber, Kec. Cimahi Sel., Kota Cimahi, Jawa Barat 40531

rakawahyup62@gmail.com

## ABSTRAK

Gempa bumi merupakan bencana alam yang sering terjadi di Indonesia karena letak geografisnya yang berada di pertemuan tiga lempeng tektonik besar. Data gempa bumi diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) dalam 2 tahun terakhir pada periode tahun 2022 hingga 2023, dan terdapat 1.689 record data yang tersedia, dengan atribut seperti lokasi, magnitudo, kedalaman, dan efek gempa. Pada data yang diperoleh menunjukkan ketidakseimbangan yang signifikan pada antar kelas dengan itu dilakukan teknik oversampling menggunakan SMOTE untuk menyeimbangkan jumlah sampel pada setiap kelas. Konsep Support Vector Machine (SVM) yang paling dasar, metode ini dapat melakukan klasifikasi biner dengan memisahkan titik data menjadi dua kelas, namun tidak mendukung klasifikasi multikelas secara bawaan. Penelitian ini bertujuan untuk mengimplementasikan pendekatan algoritma Multiclass Support Vector Machine (SVM) dengan membandingkan kernel non-linear dalam mengklasifikasikan dampak gempa bumi berdasarkan skala Modified Mercalli Intensity (MMI). Eksperimen pengujian pada penelitian ini membandingkan Multiclass Support Vector Machine (SVM) One vs Rest dan One vs One dengan kernel non-linear seperti polynomial, rbf, dan sigmoid. Hasil penelitian ini menunjukkan bahwa pendekatan algoritma Multiclass Support Vector Machine (SVM) One vs Rest dengan kernel rbf lebih baik dari pendekatan dan kernel lainnya dengan hasil akurasi 51.13%, presisi 42.60%, recall 52.46%, dan f1-score 42.51%.

**Kata kunci :** gempa bumi, multiclass support vector machine, MMI, klasifikasi.

## 1. PENDAHULUAN

Negara Indonesia merukan salah satu negara yang sangat rentan terjadi bencana alam, salah satu bencana yang terjadi yaitu bencana alam gempa bumi, peristiwa tersebut diakibatkan oleh beberapa faktor yakni keberadaan letak kedudukan geologi wilayah Indonesia yang ada di pertemuan antara 3 lempeng litosferik besar, yaitu diantaranya adalah Lempeng Indo-Australia, Lempeng Eurasia, dan, Lempeng Pasifik, serta terletak di jalur lingkaran api atau ring of fire. Gaya kontak antara lempengan tersebut menekan serta menggeser berbagai patahan yang terpecah seantero wilayah Indonesia, seperti di daratan ataupun di dasar lautan Indonesia [1][2].

Namun, bencana gempa bumi biasanya diakibatkan oleh pergerakan lempengan bumi di sepanjang bidang patahan [3]. Kewaspadaan adalah keadaan yang perlu dibentuk pada golongan masyarakat Indonesia. Pengalaman membuktikan bahwa kehancuran yang disebabkan oleh bencana bisa diantisipasi secara cepat dan tanggap bila semua golongan masyarakat lebih siaga menghadapi sebuah bencana [4].

Prediksi dampak bencana gempa bumi sangat diperlukan bagi beberapa golongan yang paling utama dari proses mitigasi bencana yang hendak direalisasikan bagi pihak terkait khususnya oleh pemerintah dalam melakukan Standard Operating Procedure (SOP) penanganan bencana alam seperti bencana gempa bumi tersebut [5]. Besaran bencana

gempa bumi dapat ditunjukkan dengan ukuran skala Richter (M) atau juga bisa dengan skala Modified Mercalli Intensity (MMI). Skala Richter dapat mengukur Magnitude sebuah gempa berdasarkan amplitudo yang terjadi sehingga lebih objektif. Sedangkan Skala Mercalli dibagi menjadi 12 subskala berdasarkan laporan orang-orang yang mengalami gempa dan evaluasi tingkat kerusakan yang dihasilkan oleh gempa bumi tersebut.

Klasifikasi merupakan salah satu aspek penelitian pada machine learning. Analisis klasifikasi merupakan sebuah metode untuk memperoleh model yang terbaik dari classifier guna memprediksikan suatu kelas dari sebuah sesuatu atau informasi yang tidak memiliki label kelas yang diketahui [6]. Maka dari itu data mining bisa menjadi sebuah solusi untuk memprediksi kejadian yang akan datang contohnya fenomena bencana gempa bumi [7].

Data mining ialah metode untuk mengubah, mengolah atau mengartikan pengetahuan dari data yang berukuran sangat besar. Banyaknya perkembangan metode machine learning yang dapat membantu klasifikasi salah satunya yaitu algoritma Support Vector Machine (SVM). SVM merupakan salah satu metode machine learning bersumber pada konsep pembelajaran statistik serta bisa memecahkan masalah non-linear yang tinggi, regresi, dan lain-lain di ruang sampel dan bisa diaplikasikan untuk sarana identifikasi sebuah sistem prediksi. Konsep dari Support Vector Machine ini sebenarnya ialah mencari

sebuah hyperline untuk memisahkan antara dua kelas yang berbeda [2] dan ini menjadi kekurangan jika ingin memisahkan lebih dari dua kelas maka diperlukan penyesuaian yang memengaruhi waktu training serta ukuran memori yang diperlukan mengatasi klasifikasi non-binner ini dibutuhkan pendekatan Multiclass Support Vector Machine (SVM) yaitu dengan pendekatan One Versus One dan One Versus Rest [8].

Beberapa penelitian yang menggunakan Multiclass Support Vector Machine (SVM) untuk klasifikasi telah dilakukan, antara lain penelitian mengenai Klasifikasi Data Tweet pada penelitian ini terdapat tiga kelas dan menggunakan pendekatan One Versus All dengan menggunakan lima fitur. Nilai tertinggi didapatkan menggunakan parameter gamma 0.7 ialah 80.59% [8]. Lalu penelitian selanjutnya yaitu mengimplementasi metode Multiclass Support Vector Machine (SVM) untuk sistem rekomendasi obat berdasarkan penyakit, penelitian ini menggunakan 200 data obat, 2 atribut utama, dan 3 kelas. Sesuai dengan pengujian 3-fold cross validation hasil yang diperoleh yaitu 96,6% pada rata-rata 88,33% [9].

Dari penjelasan di atas maka penelitian yang dilakukan yaitu mengetahui kinerja dari metode Multiclass Support Vector Machine (SVM) terhadap klasifikasi efek bencana gempa bumi dan mengetahui nilai akurasi yang didapatkan. Algoritma tersebut akan diterapkan pada dataset gempa bumi dari BMKG yang untuk mengklasifikasikan efek kerusakan dari gempa bumi di Indonesia sebagai metode yang dipakai berdasarkan nilai skala MMI gempa bumi.

**2. TINJAUAN PUSTAKA**

**2.1. Gempa Bumi**

Gempa bumi adalah suatu peristiwa alam yang menimbulkan getaran di permukaan bumi akibat pelepasan energi secara mendadak dari suatu titik gempa. Energi yang dilepaskan tiba-tiba tersebut merambat melalui tanah dalam bentuk gelombang guncangan. Peristiwa di mana gelombang guncangan ini mencapai permukaan bumi disebut sebagai gempa bumi [10]. Gempa bumi dapat diukur menggunakan skala Richter (M) atau skala Modified Mercalli (MMI). Skala Richter mengukur Magnitude gempa berdasarkan amplitudo, sehingga memberikan ukuran yang lebih objektif [4]. Efek dari gempa bumi berdasarkan skala Richter bisa dilihat pada Tabel 1 :

Tabel 1. Skala richter

Skala Richter	Efek Gempa
< 2,0	Gempa kecil, tidak terasa
2,0 – 2,9	Tidak terasa, terekam alat
4,0 – 4,9	Dapat diketahui getarannya, kerusakan tidak signifikan
5,0 – 5,9	Menyebabkan kerusakan pada area kecil
6,0 – 6,9	Dapat merusak area hingga jarak sekitar 160 Km
7,0 – 7,9	Menyebabkan kerusakan pada area luas

Sumber : [https://id.wikipedia.org/wiki/Skala\\_Richter](https://id.wikipedia.org/wiki/Skala_Richter)

Sedangkan Skala Mercalli dibagi menjadi 12 subskala berdasarkan laporan dari individu yang mengalami gempa dan evaluasi tingkat kerusakan yang disebabkan oleh gempa tersebut. Sehingga, skala Mercalli dianggap sangat subjektif dan kurang presisi jika dibandingkan dengan perhitungan magnitudo gempa yang menggunakan metode lain[11]. Didapatkan hubungan antara skala Richter dan skala MMI seperti pada Tabel 2 :

Tabel 2. Skala MMI

Skala Richter	MMI	Efek Gempa
< 2	I–II	Pada umumnya tidak terasa
3	III	Terasa di dalam rumah, tidak ada kerusakan
4	IV–V	Terasa oleh banyak orang, barang-barang bergerak, Tidak adak kerusakan struktural
5	VI–VII	Terjadi beberapa kerusakan struktural, seperti Retak-retak pada dinding.
6	VII–VIII	Kerusakan menengah, seperti hancurnya dinding
7	IX–X	Kerusakan besar, seperti runtuhnya bangunan
8	XI–XII	Rusak total atau hampir hancur total

Sumber : <https://www.bmkg.go.id/gempabumi/skala-mmi.bmkg>

**2.2. Synthetic Minority Oversampling Technique (SMOTE)**

*Synthetic Minority Oversampling Technique* (SMOTE) adalah metode *oversampling* yang digunakan untuk menangani ketidakseimbangan data pada masalah klasifikasi. Metode ini terbukti efektif dalam meningkatkan kinerja prediksi algoritma klasifikasi pada dataset yang memiliki jumlah sampel yang tidak seimbang antar kelas. SMOTE mampu meningkatkan kinerja prediksi algoritma klasifikasi, bahkan sedikit lebih baik dibandingkan dengan metode *oversampling* lain yang dipertimbangkan. SMOTE bekerja dengan menghasilkan contoh kelas minoritas baru secara acak dari tetangga kelas minoritas terdekat dalam sampel, dengan menggunakan jarak *Euclidean* antara titik data dalam ruang fitur untuk menentukan tetangga terdekat[14].

**2.3. Multiclass Support Vector Machine**

SVM adalah salah satu metode seleksi yang membandingkan parameter standar di antara serangkaian nilai diskrit yang disebut sebagai set kandidat, dan memilih satu yang memberikan hasil akurasi klasifikasi terbaik[10]. Mekanisme serta konsep dari Support Vector Machine ini sebenarnya ialah mencari sebuah hyperline untuk memisahkan antara dua kelas yang berbeda. Hyperline terbaik ini berada di tengah-tengah dua himpunan objek dari dua kelas yang berbeda[2].

Dengan berfokus pada suatu tugas, kekuatan dan sifat komputasi yang mudah dilakukan, SVM telah

berhasil dan dapat dianggap sebagai state-of-the-art classifier saat ini. Pada tipe SVM yang paling sederhana, SVM ini mendukung klasifikasi biner dan memisahkan titik data menjadi dua kelas dan tidak mendukung klasifikasi multikelas secara asli. Untuk klasifikasi lebih dari dua kelas, prinsip yang sama dipakai setelah memecah masalah multiklasifikasi menjadi beberapa masalah klasifikasi biner.

Idenya adalah untuk memetakan titik data ke ruang berdimensi tinggi untuk mendapatkan pemisahan linier timbal balik antara setiap dua kelas. Ini disebut pendekatan One vs One, yang memecah masalah multikelas menjadi beberapa masalah klasifikasi biner. Pengklasifikasi biner untuk setiap pasangan kelas. Pendekatan lain yang dapat digunakan adalah One vs Rest. Dalam pendekatan tersebut, pengelompokan diatur ke pengklasifikasi biner untuk setiap kelas[15].

### 2.4. Confusion Matrix

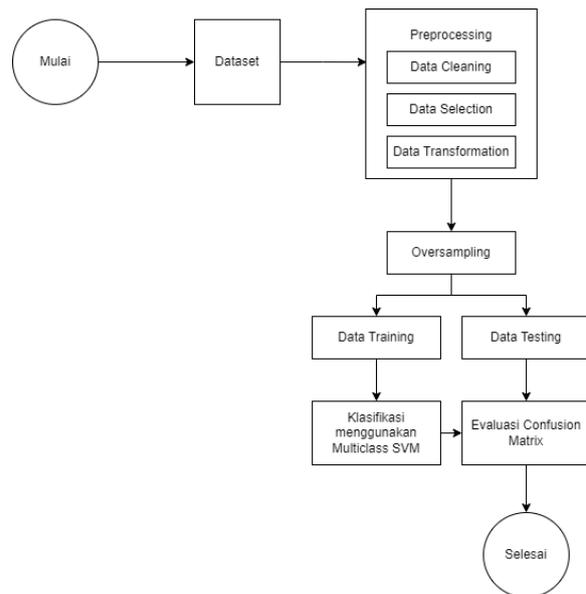
*Confusion matrix* merupakan tabel yang menunjukkan jumlah prediksi yang akurat dan tidak akurat dari setiap kelas. Dalam konteks prediksi dampak gempa bumi menggunakan SVM, confusion matrix digunakan untuk mengevaluasi kinerja model. Beberapa elemen yang dapat dievaluasi dalam confusion matrix pada kasus ini melibatkan:

- True Positive (TP): jumlah prediksi yang tepat dengan kelas positif.
- False Positive (FP): jumlah prediksi yang tidak tepat dengan kelas positif.
- True Negative (TN): jumlah prediksi yang tepat dengan kelas negatif.
- False Negative (FN): jumlah prediksi yang tidak tepat dengan kelas negatif.

Dari confusion matrix ini, dapat dihitung beberapa metrik evaluasi seperti akurasi, presisi, recall, dan F1 score. Sebagai contoh, akurasi adalah rasio jumlah prediksi yang tepat dibandingkan dengan jumlah keseluruhan prediksi. Presisi adalah rasio jumlah prediksi yang tepat dengan kelas positif dibandingkan dengan total prediksi yang tepat maupun tidak tepat dengan kelas positif. Recall mengukur rasio jumlah prediksi yang tepat dengan kelas positif dibandingkan dengan jumlah keseluruhan prediksi yang tepat atau tidak tepat dengan kelas positif. F1 score adalah nilai rata-rata harmonis dari presisi dan recall[15].

### 3. METODE PENELITIAN

Berikut gambaran alur penelitian yang dilakukan seperti pada Gambar 1.



Gambar 1. Alur penelitian

### 3.1. Pengumpulan Data

Data gempa bumi yang digunakan dalam penelitian ini berasal dari sumber terbuka, yakni Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) yang didapatkan pada website Repositori BMKG dalam 2 tahun terakhir pada periode tahun 2022 hingga 2023, dan terdapat 1.689 record data yang tersedia. Data yang didapatkan berbentuk time series yang mempunyai variabel yang dengan 11 atribut yang terdiri dari No, Event ID, Origin Time, Latitude, Longitude, Magnitude, Pusat Gempa, Depth, MMI, Kota, dan Provinsi. Atribut MMI akan dijadikan sebuah label, hanya saja untuk mereduksi kelas yang terlalu banyak akan dilakukan pengkategorian atribut MMI menjadi “Tidak ada kerusakan”, “Kerusakan Ringan”, “Kerusakan Sedang” dan “Kerusakan Besar”. Data ini diperlukan untuk nantinya diproses menggunakan Algoritma Support Vector Machine (SVM).

### 3.2. Preprocessing Data

Pada tahap preprocessing ini merupakan tahap untuk mempersiapkan data yang masih mentah sebelum melanjutkan ke tahap selanjutnya. Data harus diolah sebelum digunakan dalam proses training. Ini termasuk tahap data cleaning, data selection, data transformation. Penjelasan dari setiap tahap preprocessing yaitu :

- a. Data Cleaning: tahapan ini dimaksudkan untuk meningkatkan kualitas data agar sesuai dengan persyaratan analisis yang akan dijalankan. Proses ini meliputi tahapan-tahapan seperti pemeriksaan missing values dan koreksi data yang tidak memenuhi format yang telah ditentukan.
- b. Data Selection: proses ini adalah pemilihan subset dari dataset yang lebih besar untuk digunakan dalam analisis atau pemodelan. Tujuan dari data selection adalah untuk fokus pada data yang paling relevan atau representatif,

mengurangi kompleksitas, dan meningkatkan efisiensi analisis. Pemilihan data dapat dilakukan berdasarkan berbagai kriteria, seperti fitur-fitur tertentu, rentang waktu tertentu, atau kelas tertentu dalam konteks pengklasifikasian.

- c. Data Transformation: Pada proses ini data digabungkan untuk dilakukan inisialisasi sebuah data yang pada atributnya akan diselaraskan dengan teknik normalisasi data supaya data tidak bias, juga penambahan sebuah data juga dibutuhkan untuk memperlancar proses data mining.

### 3.3. Oversampling

Apabila terdapat perbedaan yang signifikan dalam jumlah sampel antara satu kelas dengan kelas lainnya, oversampling menjadi pendekatan yang diterapkan dalam analisis data untuk mengatasi ketidakseimbangan kelas tersebut. Dengan melakukan oversampling, tujuannya adalah menciptakan keseimbangan yang lebih baik antara jumlah sampel pada kelas minoritas dan mayoritas, dimana lebih banyak sampel dari kelas minoritas akan disediakan.

### 3.4. Klasifikasi Menggunakan Metode Multiclass Support Vector Machine

Alasan digunakannya algoritma SVM (Support Vector Machine) karena mempunyai performa yang cukup baik dan waktu prosesnya yang cepat. Konsep dari Support Vector Machine ini sebenarnya ialah mencari sebuah hyperline untuk memisahkan antara dua kelas yang berbeda ini menjadi kekurangan jika ingin melakukan pemisahan kelas dengan lebih dari dua kategori, perlu dilakukan penyesuaian yang memengaruhi waktu pelatihan dan kapasitas memori yang diperlukan mengatasi masalah klasifikasi non-binner ini dibutuhkan pendekatan Multiclass Support Vector Machine (SVM) yaitu dengan pendekatan One Versus One dan One Versus Rest.

### 3.5. Uji Validasi

Untuk mengevaluasi performa dari algoritma Multiclass SVM dalam melakukan klasifikasi maka akan dilakukan pengujian menggunakan confusion matrix. Confusion matrix digunakan sebagai instrumen untuk mengevaluasi efektivitas algoritma dalam melakukan klasifikasi dan memberikan prediksi yang sesuai dengan kenyataan. Dalam evaluasi ini menggunakan confusion matrix untuk melakukan pengukuran terhadap indikator yakni accuracy, recall, precision dan F1-score.

## 4. HASIL DAN PEMBAHASAN

### 4.1. Perolehan Data

Perolahan data Gempa Bumi didapatkan dari website Repositori BMKG yang bisa diakses di <https://repogempa.bmkg.go.id>. Data yang didapatkan pada penelitian ini memiliki keterbatasan dalam memperoleh data. Data gempa hanya bisa diambil dalam waktu perbulan saja. Dalam waktu 1 bulan data

yang didapatkan hanya sedikit sehingga dilakukan pengumpulan data dari tiap bulan dalam kurun waktu tahun 2022-2023 dan menggabungkan menjadi satu sehingga data yang di peroleh terdapat 1689 record.

Data yang didapatkan berbentuk time series yang mempunyai variabel yang dengan 11 atribut yang terdiri dari No, Event ID, Origin Time, Latitude, Longitude, Magnitude, Pusat Gempa, Depth, MMI, Kota, dan Provinsi. Atribut MMI akan dijadikan sebuah label, hanya saja untuk mereduksi kelas yang terlalu banyak akan dilakukan pengkategorian atribut MMI menjadi "Tidak ada kerusakan", "Kerusakan Ringan", "Kerusakan Sedang" dan "Kerusakan Besar". Data ini diperlukan untuk nantinya diproses menggunakan Algoritma Support Vector Machine (SVM).

### 4.2. Tahap Preprocessing

Pada tahap preprocessing ini merupakan tahap untuk mempersiapkan data yang masih mentah sebelum melanjutkan ke tahap selanjutnya. Data harus diolah sebelum digunakan dalam proses training. Ini termasuk tahap data cleaning, data selection data transformation.

#### 4.2.1. Data Cleaning

Proses ini adalah penghapusan data yang tidak valid, pemeriksaan missing values, duplikat atau tidak relevan. Dari data yang didapatkan terdapat data yang hilang pada atribut MMI sebanyak 11 record.

#### 4.2.2. Data Selection

Dalam proses ini, seleksi data yang dilakukan ialah menghapus atribut-atribut yang tidak digunakan dalam proses klasifikasi seperti No, Event ID, Origin Time, Kota dan Provinsi. Atribut tersebut merupakan identifikasi unik untuk setiap kejadian gempa bumi, namun tidak memunculkan informasi langsung tentang karakteristik gempa yang berkaitan dengan intensitas atau dampaknya. Sehingga atribut yang digunakan pada klasifikasi ini yaitu Latitude, Longitude, Magnitude, Depth, MMI. Karena Karena atribut-atribut ini penting untuk mengetahui lokasi episenter gempa, mengetahui seberapa besar gempa, dan dapat memengaruhi dampak gempa di permukaan. Total terdapat 5 atribut dan 1689 record data yang terdapat pada dataset.

#### 4.2.3. Data Transformation

Dalam data ini ada atribut yang perlu dilakukan perubahan tipe data dan salah satu atribut juga perlu dilakukan proses pengkategorian agar label kelas tidak terlalu banyak. Atribut tersebut ialah atribut MMI. Atribut Modified Mercalli Intensity (MMI) digunakan sebagai parameter untuk mengukur tingkat intensitas gempa bumi. MMI menghasilkan nilai kontinu yang merefleksikan sejauh mana dampak gempa pada suatu wilayah. Dalam rangka meningkatkan interpretasi dan keterbacaan data, dilakukan transformasi data MMI ke dalam kategori kerusakan gempa bumi. Proses

transformasi dimulai dengan menetapkan batas ambang untuk membagi rentang nilai MMI ke dalam empat kategori kerusakan yang berbeda, seperti pada Tabel 3.

Tabel 3. Transformasi data

MMI	Efek Kerusakan
1-2	Tidak Ada Kerusakan
3-4	Kerusakan Ringan
5-6	Kerusakan Sedang
7-10	Kerusakan Berat

Namun atribut "Efek\_Kerusakan" pada dataset awalnya merupakan atribut kualitatif yang menyatakan tingkat efek kerusakan akibat gempa bumi, dengan nilai-nilai seperti "Tidak Ada Kerusakan," "Kerusakan Ringan," "Kerusakan Sedang," dan "Kerusakan Berat." Agar dapat digunakan sebagai input dalam model Multiclass Support Vector Machine (SVM) yang memerlukan representasi numerik, dilakukan transformasi menggunakan metode Label Encoding. Proses ini melibatkan pemberian label unik berupa bilangan bulat pada setiap nilai unik dalam atribut "Efek\_Kerusakan." Proses Label Encoding mengubah nilai-nilai pada atribut "Efek\_Kerusakan" seperti pada Tabel 4.

Tabel 4. Label encoding

Efek Kerusakan	Value
Tidak Ada Kerusakan	0
Kerusakan Ringan	1
Kerusakan Sedang	2
Kerusakan Berat	3

Dengan transformasi ini, atribut "Efek\_Kerusakan" kini telah direpresentasikan dalam bentuk numerik, memungkinkan penggunaan efektif dalam analisis dan pemodelan selanjutnya.

### 4.3. Data Imbalanced

Dalam data mining diperlukan data yang balance antar kelasnya, dalam tahapan ini akan diperiksa pada atribut kelas yang akan diklasifikasikan yaitu pada Kelas Efek Kerusakan yang memiliki empat kelas yaitu Tidak Ada Kerusakan, Kerusakan Ringan, Kerusakan Sedang, dan Kerusakan Berat. Dapat dilihat pada Tabel 5.

Tabel 5. Jumlah kelas pada data gempa bumi

Efek Kerusakan	Jumlah Kelas
Tidak Ada Kerusakan	588
Kerusakan Ringan	1077
Kerusakan Sedang	24

Pada jumlah tiap kelas pada Tabel 5, terdapat kelas Tidak Ada Kerusakan dengan jumlah data 588, kelas Kerusakan Ringan dengan jumlah data 1077 dan kelas Kerusakan Sedang dengan jumlah data 24. Jumlah pada tiap kelas ini mengindikasikan bahwa dataset yang digunakan Imbalance atau tidak

seimbang. Namun pada keterbatasan record data yang tercatat selama periode 2 tahun terakhir tidak ada rekaman kejadian gempa bumi dengan kelas Kerusakan Berat.

### 4.4. SMOTE Oversampling

SMOTE (Synthetic Minority Over-sampling Technique) adalah teknik oversampling yang digunakan untuk menangani ketidakseimbangan kelas dalam dataset. Pada umumnya, tujuan SMOTE adalah membuat jumlah sampel pada setiap kelas menjadi seimbang atau mendekati seimbang, terutama pada kasus klasifikasi biner atau multikelas. Pemilihan jumlah data setelah SMOTE umumnya didasarkan pada kelas mayoritas (kelas dengan jumlah data terbanyak) untuk mencapai keseimbangan.

Dalam penelitian ini, kelas Tidak Ada Kerusakan memiliki jumlah data awal sebanyak 588, sementara kelas Kerusakan Ringan memiliki 1077, dan kelas Kerusakan Sedang memiliki 24. Setelah penerapan SMOTE dipilih jumlah data 1077 untuk setiap kelas. Pemilihan jumlah ini kemungkinan besar didasarkan pada kelas dengan jumlah data terbanyak setelah penerapan SMOTE, yang dalam hal ini adalah kelas Kerusakan Ringan. Alasan pemilihan jumlah 1077 untuk setiap kelas setelah SMOTE adalah untuk mencapai keseimbangan antara ketiga kelas. Dengan membuat jumlah sampel pada setiap kelas menjadi sama, dapat menghindari dominasi kelas mayoritas yang dapat mengakibatkan model menjadi tidak seimbang dan kurang mampu mengenali pola pada kelas minoritas. Seperti pada Tabel 6.

Tabel 6. Hasil oversampling SMOTE

Kondisi	Tidak ada kerusakan	Kerusakan Ringan	Kerusakan Sedang
Sebelum	588	1077	24
Sesudah	1077	1077	1077

### 4.5. Proses Klasifikasi Menggunakan Multiclass Support Vector Machine

Pada tahap klasifikasi data ini, data masuk ke dalam tahap klasifikasi dengan menerapkan pendekatan algoritma Multiclass SVM One vs One dan One vs Rest menggunakan Kernel Non-Linear seperti polynomial, RBF, dan sigmoid. Untuk pembagian data latih dan data uji d, yaitu 70:30. Data latih akan dijadikan sebagai pelatihan untuk algoritma SVM dan untuk data uji akan dijadikan sebagai pengujiannya. Selanjutnya adalah proses implementasi dengan membangun model pendekatan Multiclass SVM menggunakan beberapa kernel Non-Linear untuk klasifikasi.

Tabel 7 Data gempa bumi

Latitude	Longitude	Magnitude	Depth	Efek Kerusakan
1.54	127.83	5.1	10	Kerusakan ringan
-0.96	119.89	2.3	10	Tidak ada kerusakan

Latitude	Longitude	Magnitude	Depth	Efek Kerusakan
-0.71	98.5	6.9	25	Kerusakan sedang
-7.01	105.26	6.7	10	Kerusakan sedang
-3.45	128.84	4.9	10	Kerusakan ringan

#### 4.6. Hasil Eksperimen

Pada eksperimen pertama, model klasifikasi dampak gempa bumi menggunakan pendekatan One-Vs-One (OVO) dengan kernel polynomial diuji. Model ini berhasil memprediksi dengan benar 392 dari total 970 data uji, memberikan tingkat akurasi sebesar 40.41%. Metrik evaluasi lainnya juga diperhatikan, di mana presisi mencapai 37.79%, recall sebesar 41.67%, dan f1-score mencapai 36.76%. Hasil ini memberikan wawasan awal tentang kinerja model dengan menggunakan kernel polynomial dan pendekatan OVO.

Eksperimen kedua menggunakan pendekatan OVO dengan kernel rbf. Model ini menunjukkan peningkatan dalam prediksi dengan benar sebanyak 432 dari total 970 data uji, memberikan akurasi sebesar 44.54%. Presisi, recall, dan f1-score masing-masing mencapai 43.61%, 45.15%, dan 43.92%. Hasil ini menandakan bahwa kernel rbf memberikan kontribusi positif terhadap kinerja model.

Eksperimen ketiga melibatkan pendekatan OVO dengan kernel sigmoid. Model ini berhasil memprediksi 348 data dengan benar dari total 970 data uji, namun, akurasi yang diperoleh sebesar 35.88% menunjukkan tantangan dalam penggunaan kernel sigmoid. Presisi, recall, dan f1-score berturut-turut mencapai 46.42%, 36.05%, dan 32.16%. Evaluasi ini dapat menjadi pertimbangan untuk mengoptimalkan parameter model.

Eksperimen keempat melibatkan pendekatan One-Vs-Rest (OVR) dengan kernel polynomial. Model ini berhasil memprediksi 479 data dengan benar dari total 970 data uji, memberikan akurasi sebesar 49.38%. Presisi, recall, dan f1-score berturut-turut mencapai 48.17%, 50.44%, dan 44.13%. Hasil ini menunjukkan potensi yang baik dalam penggunaan kernel polynomial dengan pendekatan OVR.

Pada eksperimen kelima, pendekatan OVR dengan kernel rbf diaplikasikan. Model ini berhasil memprediksi 496 data dengan benar dari total 970 data uji, memberikan akurasi sebesar 51.13%. Presisi, recall, dan f1-score berturut-turut mencapai 42.60%, 52.46%, dan 42.51%. Hasil ini menunjukkan bahwa kernel rbf memberikan kontribusi positif terhadap kinerja model pada pendekatan OVR.

Eksperimen keenam menggunakan pendekatan OVR dengan kernel sigmoid. Model ini berhasil memprediksi 358 data dengan benar dari total 970 data uji, namun, akurasi yang diperoleh sebesar 36.91% menunjukkan bahwa kernel sigmoid mungkin kurang cocok untuk tugas klasifikasi ini. Presisi, recall, dan f1-score berturut-turut mencapai 37.30%, 37.32%, dan

36.60%. Evaluasi ini dapat memberikan wawasan tambahan untuk meningkatkan performa model.

#### 4.7. Evaluasi Hasil Eksperimen

Beberapa eksperimen telah dilakukan untuk mengevaluasi kinerja model Support Vector Machine (SVM) dalam mengklasifikasikan gempa bumi berdasarkan Modified Mercalli Intensity (MMI). Tabel 4.1 adalah ringkasan eksperimen yang telah dilakukan beserta hasil evaluasinya.

Tabel 8. Hasil eksperimen

Eks	Pendekatan Kernel	Akurasi %	Presisi %	Recall %	F1-Score %
1	OVO Poly	40.41	37.79	41.47	36.76
2	OVO RBF	44.54	43.61	45.15	43.92
3	OVO Sig	35.88	46.42	36.05	32.16
4	OVR Poly	49.38	48.17	50.44	44.13
5	OVR RBF	51.13	42.60	52.46	42.51
6	OVR Sig	36.91	37.30	37.32	36.60

Tabel di atas memberikan ringkasan evaluasi hasil eksperimen untuk setiap pendekatan dengan kernel yang berbeda. Akurasi model merupakan metrik utama yang digunakan untuk mengukur sejauh mana model mampu memprediksi kelas dengan benar. Pendekatan One-Vs-One (OVO) dan One-Vs-Rest (OVR) memberikan hasil yang komparabel, namun eksperimen nomor 5 dengan pendekatan OVR dan kernel rbf memberikan akurasi tertinggi sebesar 51.13%. Dalam eksperimen dengan kernel rbf, terlihat peningkatan kinerja dibandingkan dengan kernel polynomial dan sigmoid. Hal ini menunjukkan bahwa kernel rbf lebih baik dibandingkan kernel polynomial dan sigmoid untuk tugas klasifikasi ini.

### 5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian dan pengujian terhadap sistem maka dapat disimpulkan bahwa dataset awal menunjukkan ketidakseimbangan yang signifikan pada antar kelas, dengan kelas Kerusakan Ringan mendominasi. Penerapan Synthetic Minority Oversampling Technique (SMOTE) dipilih untuk menyeimbangkan jumlah sampel pada setiap kelas. Setelah penerapan SMOTE, jumlah sampel pada setiap kelas diatur menjadi 1077, dengan pertimbangan untuk mencapai keseimbangan antara kelas. Langkah ini penting untuk menghindari dominasi kelas mayoritas yang dapat menghasilkan model yang tidak seimbang dan kurang mampu mengenali pola pada kelas minoritas.

Berbagai eksperimen dilakukan menggunakan model Support Vector Machine (SVM) dengan berbagai pendekatan dan kernel non-linear, yaitu kernel polynomial, rbf, dan sigmoid. Hasil eksperimen menunjukkan bahwa pendekatan One-Vs-Rest (OVR) dengan kernel rbf memberikan kinerja terbaik, mencapai akurasi sebesar 51.13%. Evaluasi lebih lanjut melibatkan metrik lainnya seperti presisi 42.60%, recall 52.46%, dan f1-score 42.51%, memberikan

wawasan mendalam tentang kemampuan model dalam mengklasifikasikan setiap kelas. Meskipun hasil ini menunjukkan performa yang lebih baik dibanding pendekatan dan kernel non-linear lainnya, perlu dicatat bahwa kinerja model secara keseluruhan belum mencapai tingkat keunggulan yang diharapkan.

Perlu dilakukan pengembangan model lebih lanjut untuk mengoptimalkan parameter dan struktur model SVM agar dapat meningkatkan performa klasifikasi. Pendekatan dengan kernel non-linear, khususnya kernel rbf, memberikan peningkatan signifikan dibandingkan dengan kernel polynomial dan sigmoid. Namun, kinerja model secara umum tidak dapat dianggap sebagai sangat baik, dan masih terdapat ruang untuk peningkatan lebih lanjut.

#### DAFTAR PUSTAKA

- [1] T. Tupan, N. R. Widuri, and R. Rachmawati, "Analisis Bibliometrik Publikasi Ilmiah Tentang Prediksi Gempa Bumi Berbasis Data Scopus Periode 2015-2020," *Libr. J. Perpust.*, vol. 8, no. 1, pp. 31–48, 2020, [Online]. Available: <https://journal.iainkudus.ac.id/index.php/Libraria/article/view/7183>
- [2] O. Somantri, S. Purwaningrum, and R. Riyanto, "Model Support Vektor Machine (Svm) Berdasarkan Parameter Windows Untuk Prediksi Kekuatan Gempa Bumi," *JTT (Jurnal Teknol. Ter.)*, vol. 8, no. 1, pp. 17–24, 2022, [Online]. Available: <https://jurnal.polindra.ac.id/index.php/jtt/article/view/352>
- [3] H. A. Nugroho and H. S. Wicaksana, "Acceleration classification of earthquake signals and linear signals of human activity using smartphone accelerometer based on support vector machine algorithm," *J. Teknol. dan Sist. Komput.*, vol. 7, no. 4, pp. 166–171, 2019, doi: 10.14710/jtsiskom.7.4.2019.166-171.
- [4] L. Irawan, L. H. Hasibuan, and F. Fauzi, "Analisa Prediksi Efek Kerusakan Gempa Dari Magnitudo (Skala Richter) Dengan Metode Algoritma Id3 Menggunakan Aplikasi Data Mining Orange," *J. Teknol. Inf. J. Keilmuan dan Apl. Bid. Tek. Inform.*, vol. 14, no. 2, pp. 189–201, 2020, doi: 10.47111/jti.v14i2.1079.
- [5] A. A. Nurkhaliza and A. W. Wijayanto, "Perbandingan Algoritma Klasifikasi Support Vector Machine dan Random Forest pada Prediksi Status Indeks Mitigasi dan Kesiapsiagaan Bencana (IMKB) Satuan Kerja BPS di Indonesia Tahun 2020," *Maret*, vol. 7, no. 1, pp. 54–59, 2022, [Online]. Available: <http://openjournal.unpam.ac.id/index.php/informatika54>
- [6] R. S. Tantika and A. Kudus, "Penggunaan Metode Support Vector Machine Klasifikasi Multiclass pada Data Pasien Penyakit Tiroid," *Bandung Conf. Ser. Stat.*, vol. 2, no. 2, pp. 159–166, 2022, doi: 10.29313/bcss.v2i2.3590.
- [7] D. P. Utomo and B. Purba, "Penerapan Datamining pada Data Gempa Bumi Terhadap Potensi Tsunami di Indonesia," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, p. 846, 2019, doi: 10.30645/senaris.v1i0.91.
- [8] D. N. Fitriana and Y. Sibaroni, "Klasifikasi Data Tweet dengan Menggunakan Metode Klasifikasi Multi-Class Support Vector Machine (SVM) (Studi Kasus : PT.KAI)," *e-Proceeding Eng.*, vol. 7, no. 2, pp. 8493–8505, 2020, [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/12746>
- [9] A. P. Putra, I. Mulyana, S. Maryana, and F. Susanti, "Implementasi Multiclass Support Vector Machine Pada Sistem Rekomendasi Obat Berdasarkan Gejala Penyakit," *Semin. Nas. Sains Teknol. dan Inov. Indones. (SENASTINDO AAU)*, vol. 1, no. 1, pp. 211–222, 2019.
- [10] A. Noor, "Perbandingan Algoritma Support Vector Machine Biasa dan Support Vector Machine berbasis Particle Swarm Optimization untuk Prediksi Gempa Bumi," *J. Hum. Teknol.*, vol. 4, no. 1, pp. 31–37, 2018, doi: 10.34128/jht.v4i1.37.
- [11] BMKG, "Skala MMI." Accessed: Feb. 12, 2023. [Online]. Available: <https://www.bmkg.go.id/gempabumi/skala-mmi.bmkg>
- [12] I. H. Witten, M. A. Hall, and E. Frank, "Data Mining: Practical Machine Learning Third Edition," (*Morgan Kaufmann Ser. Data Manag. Syst. Morgan Kaufmann*, vol. 104, no. June, p. 113, 2005.
- [13] W. Agwil, "Penerapan Jaringan Saraf Tiruan Pada Data Gempa Bumi Di Provinsi Bengkulu," *J. Stat. Univ. Muhammadiyah Semarang*, vol. 8, no. 2, p. 152, 2020, doi: 10.26714/jsunimus.8.2.2020.152-158.
- [14] S. DEMİR and E. K. ŞAHİN, "Evaluation of Oversampling Methods (OVER, SMOTE, and ROSE) in Classifying Soil Liquefaction Dataset based on SVM, RF, and Naïve Bayes," *Eur. J. Sci. Technol.*, no. 34, pp. 142–147, 2022, doi: 10.31590/ejosat.1077867.
- [15] R. Yacouby and D. Axman, "Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models," pp. 79–91, 2020, doi: 10.18653/v1/2020.eval4nlp-1.9.