

KLASIFIKASI STATUS BENCANA ALAM GEMPA BUMI DENGAN ALGORITMA K-NEAREST NEIGHBOUR (K-NN)

Rizal Wahyu Purnama¹, Kartiko Ardi Widodo², Bima Romadhon Parada Dian Palevi³

Teknik Elektro S1, Institut Teknologi Nasional Malang, Malang Indonesia

¹rizalwahyupurnama21@gmail.com, ²tiko_@lecturer.itn.ac.id, ³bimarpd@lecturer.itn.ac.id

Abstrak— Gempa bumi merupakan fenomena alam yang terjadi secara tiba-tiba dan berpotensi menimbulkan kerugian bagi manusia, baik secara materil maupun korban jiwa. Untuk meminimalisir dampak tersebut, diperlukan deteksi dini serta klasifikasi bencana gempa untuk meminimalkan risiko yang ditimbulkan. Penelitian ini melakukan klasifikasi tingkat risiko kerusakan akibat gempa bumi menggunakan algoritma K-Nearest Neighbor (K-NN) kedalam tiga kelas, yaitu risiko ringan, sedang, dan parah berbasis parameter kedalaman dan magnitudo. Data yang digunakan diperoleh melalui situs web Kaggle, dengan sumber BMKG dan USGC dalam rentang tahun 2008 hingga 2023. Setelah melalui tahap pra-pemrosesan, termasuk pembersihan dan pelabelan data, kemudian dataset sebanyak 92.887 dibagi menjadi data latih dan data uji dengan rasio 80:20. Tujuan dari penelitian ini adalah untuk mengevaluasi akurasi algoritma K-NN berdasarkan data yang tersedia serta menentukan nilai K yang optimal. Hasil penelitian menunjukkan bahwa algoritma K-NN mampu memprediksi kelas risiko gempa dengan tingkat akurasi tinggi, yaitu sebesar 99.82% untuk K=3, 99.70%, untuk K=5, dan 99.57% untuk K=7 dengan menggunakan parameter magnitudo dan kedalaman.

Kata Kunci: gempa bumi, klasifikasi resiko, algoritma k-nn,

I. PENDAHULUAN

Indonesia merupakan salah satu negara yang terletak di kawasan pertemuan tiga lempeng tektonik dunia dan Cincin Api Pasifik, wilayah yang dikenal memiliki aktivitas tektonik yang sangat tinggi. Hal ini menyebabkan wilayah Indonesia sering kali dilanda berbagai jenis bencana alam, seperti gempa bumi, tsunami, dan erupsi gunung berapi [1]. Gempa bumi, terjadi akibat pergeseran dan benturan antara lempeng kerak bumi yang memicu gelombang seismik dan getaran di permukaan bumi [2]. Pergeseran lempeng ini tidak hanya memicu gempa bumi, tetapi juga dapat memicu gelombang tsunami dan aktivitas vulkanik yang memicu letusan gunung berapi [3]. Gempa bumi merupakan salah satu bencana alam yang sering terjadi di Indonesia dan menyebabkan kerugian yang sangat besar bagi masyarakat baik materi maupun korban jiwa [1]. Seperti kejadian gempa bumi di Provinsi Jawa Barat, Kabupaten Cianjur tahun 2022, gempa bumi yang berkekuatan 5,6 magnitudo dengan kedalaman 10 kilometer, menyebabkan kerusakan fasilitas umum dan bangunan di wilayah Cianjur [4]. Oleh karena itu, diperlukan deteksi dini serta prediksi jenis atau pola dari gelombang seismik yang

terjadi agar penanganan mitigasi bencana dapat dilakukan dengan cepat dan tepat.

Saat ini, kecerdasan buatan (AI) berkembang pesat serta memberikan dampak yang baik diberbagai bidang, termasuk pada bidang kebencanaan seperti mitigasi, pengolahan dan analisa data gempa bumi. Dengan memanfaatkan teknik pengumpulan data dan pembelajaran mesin, AI dapat digunakan untuk mengklasifikasikan berbagai jenis informasi tentang gempa bumi [5]. Seperti jenis gelombang gempa, tingkat risiko gempa, dan status kejadian gempa berdasarkan parameter yang tertentu yang digunakan, antara lain lokasi kejadian (garis bujur dan lintang), kedalaman, kekuatan gempa (magnitudo).

Klasifikasi merupakan salah satu proses dalam pembelajaran mesin terbimbing (*supervised learning*) yang bertujuan memetakan data ke dalam kategori tertentu berdasarkan data yang telah tersedia. Klasifikasi efektif digunakan untuk menyelesaikan permasalahan kompleks [6]. Dalam konteks kebencanaan klasifikasi digunakan untuk menggolongkan suatu dampak risiko gempa bumi, kedalaman gempa bumi, kekuatan gempa bumi, jenis gelombang gempa bumi dan lain sebagainya berdasarkan data parameter yang ada. Salah satu metode yang dapat digunakan untuk mengklasifikasikan bencana gempa bumi adalah algoritma *K-Nearest Neighbor* (K-NN). Algoritma ini bekerja dengan membandingkan karakteristik data baru dengan data latih, kemudian memprediksi kelas berdasarkan jarak mayoritas nilai K dari tetangga terdekat pada data latih [7][8]. *K-Nearest Neighbor* (K-NN) juga memiliki kelebihan yaitu implementasinya mudah dan sederhana, serta dikenal karena keefektifannya dalam berbagai kasus klasifikasi, termasuk dalam konteks seismologi dan geofisika [8]. Dengan menerapkan algoritma *K-Nearest Neighbor* (K-NN), proses klasifikasi status bencana gempa bumi dapat dilakukan secara cepat dan akurat, ketika menggunakan data seismik yang terdeteksi melalui sistem *embedded*. Penggunaan sistem *embedded* pada klasifikasi status gempa bumi bertujuan untuk memperoleh data seismik, seperti tanggal, waktu, lokasi (*longitude/latitude*) kejadian gempa, kedalaman gempa, dan magnitudo gempa, yang digunakan sebagai data latih dan data uji. Sistem ini memiliki keunggulan dalam daya tahan, kemampuan bekerja di lokasi terpencil, serta efisiensi energi.

Keunggulan tersebut memungkinkan pengolahan data secara real-time, sehingga deteksi dan klasifikasi gempa dapat dilakukan dengan cepat dan akurat. Dengan demikian, sistem ini dapat dimanfaatkan sebagai peringatan dini terhadap bencana alam [9].

Pada penelitian terdahulu yang melakukan metode klasifikasi data gempa bumi, seperti yang dilakukan oleh Laia et al., menggunakan algoritma *Naïve Bayes* untuk mengklasifikasikan data gempa di Pulau Sumatera berdasarkan hiposentrum. Penelitian ini menunjukan metode klasifikasi gempa bumi di pulau sumatera cukup baik, tetapi pada penelitian ini terbatas pada satu parameter klasifikasi [10]. Kemudian penggunaan sistem *embedded* dalam penelitian yang dilakukan oleh Ulil Albab et al., yang mengembangkan sistem informasi gempa berbasis Raspberry Pi. Sistem ini bertujuan untuk menyampaikan informasi gempa melalui platform berbasis web. Pada penelitian ini informasi yang diberikan hanya menampilkan klasifikasi gempa “kuat” dan “lemah” berdasarkan nilai pembacaan sensor [11]. Penelitian selanjutnya, yang dilakukan oleh Putri et al., menunjukkan bahwa algoritma *K-Nearest Neighbor* (K-NN) dapat diimplementasikan pada sistem *embedded* untuk aplikasi klasifikasi. Penelitian tersebut menggunakan algoritma *K-Nearest Neighbor* (K-NN) untuk mengklasifikasikan status gizi bayi berdasarkan parameter seperti jenis kelamin, umur, dan berat badan. Serta memperoleh hasil pengujian akurasi yang baik, pada pengujian sensor dan algoritma *K-Nearest Neighbor* (K-NN) [12]. Dari penelitian terdahulu bahwa penerapan metode klasifikasi dalam konteks kebencanaan memiliki keterbatasan dari segi parameter serta kategori klasifikasi yang dihasilkan.

Dalam penelitian ini, penulis membahas tentang penerapan algoritma *K-Nearest Neighbor* (K-NN) untuk menklasifikasikan dan memprediksi gempa bumi berdasarkan parameter yang tersedia seperti magnitudo dan kedalaman, serta mengevaluasi kinerja algoritma *K-Nearest Neighbor* (K-NN) melalui perhitungan akurasi dan peformanya. Sistem ini dirancang untuk memprediksi risiko kerusakan yang disebabkan oleh gempa bumi, seperti risiko ringan, risiko sedang, dan risiko parah, berdasarkan kedalaman serta magnitudo gempa yang diperoleh dari katalog gempa BMKG dan USGS tahun 2008-2023. Dengan demikian, informasi mengenai jenis gempa dan potensi kerusakan yang ditimbulkan dapat disampaikan secara mudah, akurat, dan tepat, sehingga dapat mendukung pengambilan keputusan yang cepat dalam situasi darurat. Harapannya, sistem ini mampu membantu penanganan bencana yang lebih efisien serta meminimalkan kerugian yang ditimbulkan.

II. KAJIAN PUSTAKA

A. Penelitian Terdahulu

Beberapa penelitian yang dijadikan sebagai rujukan dalam penelitian ini adalah sebagai berikut:

Penelitian terdahulu yang dilakukan oleh Yohanes Setiawan pada tahun 2023 yang berjudul “Data Mining berbasis Nearest Neighbor dan Seleksi Fitur untuk Deteksi Kanker Payudara”. Penelitian ini bertujuan untuk mendeteksi kanker payudara berbasis nearest neighbor dengan seleksi fitur. Metode seleksi fitur yang digunakan adalah filter based, yakni Correlation based, Information Gain, dan ReliefF. Hasil

uji coba menunjukkan bahwa MLM-KHNN pada $K=3$ dengan metode seleksi fitur Information Gain memiliki nilai recall terbaik sebesar 94% disertai metrik-metrik lainnya yang berada diatas 70% [13].

Kemudian, penelitian serupa yang dilakukan oleh Ruth Maharani Aquilia Hutagaol, Vilanata Tesalonika Lana, Zena Azzahra tahun 2024 berjudul “Penerapan Machine Learning dalam Prediksi Klasifikasi Big Data Kedalaman Gempa Bumi di Indonesia Tahun 2015-2024”. Penelitian ini bertujuan untuk mengevaluasi prediksi dengan mengklasifikasikan gempa bumi berdasarkan kedalamannya menggunakan variabel-variabel yang digunakan. Penelitian ini menggunakan machine learning, menggunakan 3 metode klasifikasi seperti *Naïve Bayes*, KNN, dan Random Forest. Hasil penelitian berdasarkan nilai accuracy, precision, recall, dan F1 score, didapatkanlah metode yang terbaik dalam mengklasifikasikan gempa bumi adalah metode Random Forest [14].

Berdasarkan penelitian Terdahulu, dapat disimpulkan bahwa algoritma *K-Nearest Neighbors* (K-NN) memiliki beberapa keunggulan, antara lain kemudahan dalam implementasi, efektivitas dalam mengenali pola baru, dan kebutuhan komputasi yang relatif rendah serta. Metode ini sangat bermanfaat dalam berbagai aplikasi pengklasifikasian, khususnya ketika bekerja dengan dataset yang kecil hingga sedang. Namun, pada dataset berukuran besar, *K-Nearest Neighbor* (K-NN) kurang efektif karena membutuhkan waktu komputasi yang lebih lama untuk menghitung jarak dari setiap titik data.

B. Gempa Bumi

Gempa bumi terjadi akibat adanya getaran yang berasal dari dalam bumi yang muncul secara tiba-tiba, disebabkan oleh pergeseran atau tumbukan antar lempeng di kerak bumi. Getaran tersebut kemudian merambat ke segala arah hingga mencapai permukaan bumi dan dapat menimbulkan dampak berupa kerusakan bangunan, gangguan aktivitas masyarakat, hingga korban jiwa. Pergeseran lempeng ini juga menyebabkan terbentuknya sesar atau patahan di dalam lapisan bumi, yang menjadi salah satu sumber utama aktivitas seismik di Indonesia [1][3].

BMKG menjelaskan bahwa gempa bumi terjadi karena akumulasi energi dari pergerakan lempeng-lempeng tektonik yang saling berinteraksi. Ketika tegangan yang terakumulasi melebihi kekuatan batuan, terjadi patahan yang melepaskan energi dalam bentuk gelombang seismik yang merambat ke segala arah dan menyebabkan getaran di permukaan bumi. Indonesia merupakan daerah rawan gempa bumi karena dilalui oleh pertemuan tiga lempeng tektonik utama yaitu, Indo-Australia, Eurasia, dan Pasifik [15]. Terdapat beberapa jenis gempa bumi, yaitu: (a) gempa tektonik yang terjadi akibat pergeseran lempeng tektonik, (b) gempa vulkanik yang terjadi akibat aktivitas vulkanik gunung berapi, dan (c) gempa runtuhan yang disebabkan oleh runtuhnya gua bawah tanah, biasanya akibat kegiatan pertambangan.

C. Parameter Gempa Bumi

Menurut Balai Besar Meteorologi, Klimatologi, dan Geofisika (BMKG), terdapat beberapa parameter yang dapat digunakan untuk mendeteksi gempa bumi, antara lain[15]:

1. Waktu kejadian (*origin time*): waktu terjadinya gempa bumi secara tepat, mencakup tanggal, hari, jam, menit, detik.
2. Lokasi pusat kejadian gempa bumi (episenter): titik koordinat di permukaan bumi yang berada tepat di atas pusat gempa. Biasanya dinyatakan dalam lintang (*latitude*) dan bujur (*longitude*).
3. Kedalaman pusat gempa bumi (*depth*): lokasi awal gempa bumi di dalam bumi.
4. Kekuatan guncangan gempa bumi (magnitudo): Merupakan ukuran energi yang dilepaskan oleh gempa bumi.

1) Kedalaman

Kedalaman pusat gempa atau yang sering disebut hiposentrum, merupakan titik awal gempa bumi di dalam permukaan bumi. Parameter kedalaman berperan penting dalam menentukan besarnya dampak yang ditimbulkan. Terdapat klasifikasi dampak gempa berdasarkan jarak kedalaman hiposentrum dari permukaan bumi, seperti yang tertera pada Tabel 1[16].

Tabel 1. Klasifikasi Kedalaman

No	Rentang kedalaman dari permukaan bumi (Km)	Kategori Kedalaman Gempa	Dampak
1	≤ 70 Km	Gempa Dangkal	Menyebabkan kerusakan parah
2	$>70-300$ Km	Gempa Menengah	Menyebabkan Kerusakan sedang hingga ringan
3	>300 Km	Gempa Dalam	Menyebabkan kerusakan ringan atau bisa hanya menyebabkan getaran

2) Kekuatan Guncangan

Energi yang dilepaskan akibat patahan batuan keras di dalam lapisan bumi diukur dalam satuan magnitudo, yang merepresentasikan besarnya kekuatan gempa. Magnitudo merupakan besaran logaritmik yang menggambarkan kekuatan suatu gempa bumi. Magnitudo secara umum diklasifikasikan ke dalam dua jenis skala, yaitu Skala Richter (SR) dan Magnitudo Momen (Mw). Skala Richter (SR) digunakan mengukur amplitudo gelombang seismik secara lokal, sedangkan Magnitudo Momen (Mw) dihitung berdasarkan momen seismik yang meliputi luas patahan besar slip dan kekuatan batuan. Magnitudo Momen (Mw) dianggap representatif untuk menghitung total energi yang dilepaskan gempa berskala besar [17].

a) Skala Richter

Skala Richter diciptakan oleh Charles F. Richter Bersama rekannya pada tahun 1935, yang mengukur kekuatan gempa berdasarkan besar gelombang amplitudo. Skala ini bersifat logaritmik, yaitu setiap kenaikan 1 skala berarti gempa 10 kali lebih kuat. Terbatas pada gempa menengah[17].

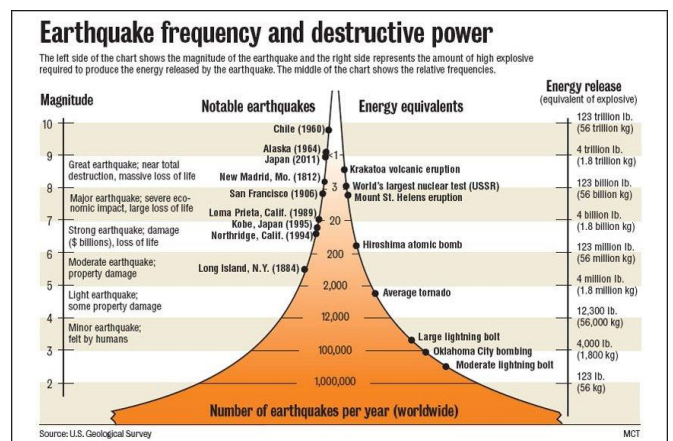
Tabel 2. Klasifikasi Magnitudo

No	Magnitudo (SR)	Intensitas	Perkiraan Dampak/Kerusakan
1	< 3.0	Sangat kecil	Hanya terdeteksi oleh alat seismograf.
2	$3.0 - 3.9$	Kecil	Jarang menyebabkan kerusakan, tapi bisa dirasakan
3	$4.0 - 4.9$	Ringan	jarang menimbulkan kerusakan serius.
4	$5.0 - 5.9$	Sedang	Bisa menyebabkan kerusakan ringan pada bangunan yang tidak tahan gempa.
5	$6.0 - 6.9$	Kuat	Dapat menyebabkan kerusakan sedang hingga berat di daerah padat penduduk.
6	$7.0 - 7.9$	Sangat kuat	Menimbulkan kerusakan parah di area yang luas, berpotensi memicu tsunami.
7	≥ 8.0	Dahsyat	Dapat menghancurkan infrastruktur didekat pusat gempa dalam radius ratusan kilometer, dan berpotensi tsunami.

Pada Tabel 2, menjelaskan tentang kerusakan akibat kekuatan guncangan dari energi yang dilepaskan oleh gempa bumi. Gempa dengan magnitudo kurang dari 5 diklasifikasikan sebagai gempa ringan karena dampak umumnya terbatas pada getaran yang tidak merusak. Gempa yang dengan magnitudo antara 5 hingga 6 tergolong sedang, yang berpotensi menyebabkan kerusakan pada bangunan tua atau tidak tahan gempa. Sedangkan gempa dengan magnitude lebih dari 6 hingga 8 termasuk kategori berat, karena dapat menyebabkan kerusakan parah, tsunami, hingga korban jiwa terutama pada wilayah padat penduduk dan kedalaman dangkal [16].

b) Skala Magnitudo Momen

Skala Magnitudo Momen (Mw) digunakan untuk kekuatan gempa yang bermagnitudo besar pada tahun 1976. Magnitudo momen juga bersifat logaritmik, yang berarti setiap kenaikan satu skala magnitudo terjadi peningkatan kekuatan gempa 32 kali lipat. Sebagai contoh, gempa bermagnitudo 7.0 melepaskan kekuatan 32 kali lebih besar dari gempa yang bermagnitudo 6.0, dan 1000 kali lebih besar dari magnitudo 5.0. Besarnya kekuatan gempa biasanya digambarkan dengan ekuivalen ledakan *Trinitrotoluene* (TNT). Seperti gambar 1 [17].



Gambar 1. Korelasi Magnitudo

D. Klasifikasi

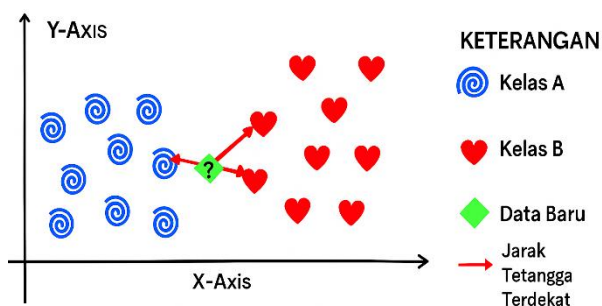
Klasifikasi adalah proses dalam pembelajaran mesin (*machine learning*) yang digunakan untuk memetakan data ke dalam kelas atau kategori tertentu berdasarkan pola, atau ciri-ciri dari data yang telah tersedia sebelumnya. Proses ini termasuk ke dalam metode pembelajaran terawasi (*supervised learning*), karena memerlukan data berlabel atau kategori yang telah diketahui sebelumnya sebagai dasar pelatihan model. Dengan adanya data berlabel, sistem dapat mempelajari karakteristik dari tiap kelasnya, seperti distribusi nilai, fitur dominan, ataupun perbedaan kelas [5].

Dalam praktiknya, klasifikasi bekerja dengan membangun model dari data latih yang berisikan input (parameter atau fitur) serta output (kelas atau label) yang sudah diketahui sebelumnya, lalu model akan memprediksi kelas dari data baru yang belum diketahui label atau kelasnya. Semakin banyak label atau kelas yang berhasil dipelajari, semakin baik kemampuan model dalam mengklasifikasikan atau menggolongkan data baru. Metode klasifikasi banyak digunakan untuk mengolongkan hal-hal rumit seperti mendiagnosis penyakit berdasarkan rekam medis, mendeteksi spam e-mail, dan lain sebagainya [13].

E. Algoritma K-Nearest Neighbor

K-Nearest Neighbor (K-NN) adalah salah satu algoritma dalam pembelajaran mesin (*machine learning*) yang termasuk ke dalam metode pembelajaran terawasi (*supervised learning*). Algoritma ini digunakan untuk menyelesaikan permasalahan klasifikasi maupun regresi, dengan prinsip dasar bahwa objek data baru akan diklasifikasikan berdasarkan jarak, dengan sejumlah K data tetangga terdekat didalam dataset pelatihan. Keputusan klasifikasi pada K-NN bergantung pada mayoritas kelas dari tetangga terdekat dari nilai K yang sudah ditentukan [6].

Kelebihan utama K-Nearest Neighbor (K-NN) adalah kesederhanaan implementasinya dan tidak memerlukan proses pelatihan (*training*) yang kompleks karena bersifat *lazy learner*, yaitu memprediksi berdasarkan jarak atau vektor. Selain itu K-NN memiliki kekurangan seperti lambat memprediksi dataset besar serta bergantung pada pemilihan nilai K yang tepat [7][8].



Gambar 2. Algoritman K-NN

1) Parameter Nilai K

Nilai K dalam algoritma K-Nearest Neighbor (K-NN) merupakan parameter yang menunjukkan jumlah tetangga

terdekat yang digunakan untuk dasar pertimbangan dalam memprediksi kelas dari suatu data baru. Pemilihan nilai K yang tepat sangat penting karena secara langsung memengaruhi tingkat akurasi hasil klasifikasi. Jika nilai K terlalu kecil, model cenderung mengalami *overfitting*, kondisi dimana model terlalu bergantung pada jarak atau tetangga terdekat sehingga sensitif terhadap noise. Sebaliknya, jika nilai K terlalu besar, kondisi dimana model berisiko mengalami *underfitting*, dimana model akan terlalu banyak mempertimbangkan banyak tetangga, termasuk yang letaknya yang lebih jauh dari data uji, sehingga mengurangi keakuratan model dalam mengenali label atau kelas [7].

Dalam praktiknya, tidak ada metode pasti yang dapat menjamin penentuan nilai K terbaik untuk semua jenis data, karena pemilihan nilai K sangat bergantung pada karakteristik dari dataset, seperti jumlah data, distribusi kelas, dan lain sebagainya. Beberapa pendekatan umum yang digunakan untuk menentukan nilai K antara lain seperti, metode *cross-validation*, yaitu teknik evaluasi performa model dengan membagi data pelatihan ke dalam beberapa subset dan menguji model secara bergantian untuk memperoleh estimasi akurasi yang tepat. Selain itu, pemilihan nilai K yang ganjil sering direkomendasikan untuk menghindari ambiguitas atau hasil klasifikasi yang imbang dalam proses klasifikasi [18].

2) Perhitungan Jarak K-NN

Dalam proses klasifikasi dengan menggunakan algoritma K-Nearest Neighbor (K-NN), tahapan terpenting yang dilakukan yaitu menghitung jarak antara data uji dengan seluruh data latih. Pengukuran jarak berfungsi untuk menentukan seberapa dekat suatu data uji dengan data yang sudah ada pada dataset pelatihan. Terdapat berbagai cara dalam menghitung jarak pada algoritma K-Nearest Neighbor (K-NN), salah satunya dengan menggunakan rumus *Euclidean Distance* dimana rumus tersebut umum digunakan untuk mengukur jarak pada algoritma K-NN.

Untuk dapat menghitung jarak pada algoritma K-Nearest Neighbor (K-NN) dengan rumus *Euclidean Distance* dijabarkan sebagai berikut [6]:

$$d(p_i, q_i) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (1)$$

dimana:

$d(p_i, q_i)$: jarak data baru

p_i : data latih

q_i : data uji

i : variabel data

n : dimensi data

F. Python

Python merupakan bahasa pemrograman yang sering digunakan dalam machine learning, data science, maupun data mining karena memiliki sejumlah keunggulan, antara lain efisiensi, bersifat interpreted, interaktif, berorientasi objek, mudah dibaca (*readability*), dan didukung oleh komunitas yang besar. Python dapat digunakan pada berbagai sistem operasi, seperti Windows, Linux, dan mac OS. Selain itu bahasa python lebih sederhana dibandingkan bahasa pemrograman lainnya [19].

Dalam penerapan machine learning, Python memiliki berbagai pustaka yang mendukung pengembangan kegiatan analisis data, antara lain; NumPy digunakan untuk komputasi numerik dan operasi array, Pandas digunakan untuk manipulasi dan analisis data berbentuk Tabel (dataframe), Matplotlib dan Seaborn digunakan untuk visualisasi data dalam bentuk grafik yang informatif, Scikit-Learn mendukung implementasi berbagai algoritma machine learning, baik untuk klasifikasi, regresi, maupun clustering [20].



Gambar 3. Python

G. Google Colaboratory

Google Colab atau Google Colaboratory adalah platform berbasis cloud dari Google yang digunakan untuk menjalankan kode Python dalam format Jupyter Notebook. Platform ini banyak dimanfaatkan untuk pengembangan aplikasi seperti machine learning, data science, dan analisis data. Google Colab mendukung berbagai pustaka Python populer, terintegrasi dengan GDrive, dan memungkinkan akses GPU serta TPU secara gratis. Selain itu, pengguna dapat berkolaborasi secara real-time dan menambahkan elemen interaktif seperti Gambar dan teks. Dengan kemudahan akses dan fitur lengkap, Google Colab bisa menjadi salah satu pilihan untuk eksperimen dan pengembangan berbasis Python [21].



Gambar 4. Google Colaboratory

III. METODOLOGI PENELITIAN

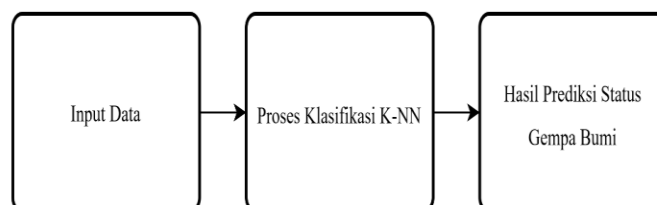
A. Spesifikasi Sistem

Spesifikasi sistem untuk pengklasifikasian status gempa bumi memerlukan beberapa hal seperti:

1. Dataset yang digunakan berupa data sekunder yang diperoleh dari web kaggle yang bersumber dari BMKG tahun 2008 sampai 2023 yang berisikan data sebanyak 92.887 data.
2. Menggunakan variasi pengujian nilai K (K=3, K=5, dan K=7) untuk menentukan akurasi model yang optimal pada algoritma K-NN.

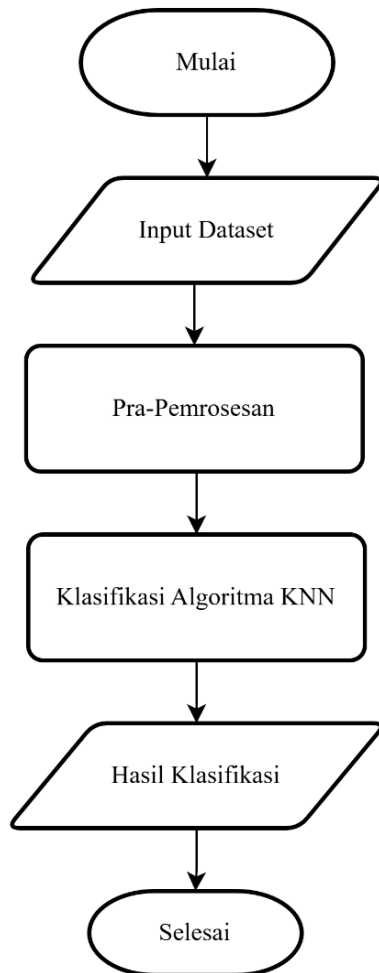
3. *Euclidean Distance* digunakan untuk menghitung jarak serta menentukan kelas pada algoritma K-NN.
4. Menggunakan software google colaboratory untuk mengolah data.
5. Menggunakan bahasa pemrograman python.
6. Pelabelan atau pemberian kelas kedalam tiga golongan, yaitu risiko ringan, risiko sedang, dan risiko parah.
7. Visualisasi hasil menggunakan laptop atau pc.

B. Blok Diagram



Gambar 5. Blok Diagram

Pada Gambar 5, menunjukan alur blok diagram sistem yaitu, data baru yang diniputkan sesuai parameter yang dibutuhkan akan diolah ke dengan algoritma *K-Nearest Neighbors* (K-NN) untuk diklasifikasikan berdasarkan mayoritas jarak terdekat dari data pelatihan sebelumnya, setelah itu hasil pengolahan data akan ditampilkan pada layar monitor seperti risiko kerusakan ringan, sedang, dan parah yang diakibatkan gempa bumi.



Gambar 6. Flowchart Klasifikasi

Pada Gambar 6, menunjukkan diagram alur proses klasifikasi yaitu, input dataset, pra-pemrosesan, proses klasifikasi dengan algoritma *K-Nearest Neighbor* (K-NN), hasil klasifikasi yang berupa prediksi dan tingkat akurasi. Yang kemudian akan dijelaskan lebih lanjut sebagai berikut:

1) Input Dataset

Pada tahapan ini, penulis akan menginputkan dataset mentah atau raw yang akan digunakan sebagai dasar pelatihan dan pengujian dari algoritma *K-Nearest Neighbor* (K-NN). Dataset yang digunakan merupakan data sekunder, didapatkan dari situs web Kaggle berdasarkan sumber BMKG dan USGS mulai tahun 2008–2023 sebanyak 92.887 [22]. Dataset tersebut umumnya berisikan informasi tentang waktu dan tanggal kejadian, lokasi pusat gempa, kedalaman, magnitudo, dan lain sebagainya.

Tabel 3. Data Raw

Pada Tabel 3, menunjukan isi dari dataset raw yang

no	tgl	ot	lat	lon	dep th	m ag	remark
1	11/1/2008	02:43.1	-9.18	119.06	10	4.9	Sumba Region - Indonesia
2	11/1/2008	58:50.2	-6.55	129.64	10	4.6	Banda Sea
3	11/1/2008	43:12.9	-7.01	106.63	121	3.7	Java - Indonesia
4	11/1/2008	24:14.8	-3.3	127.85	10	3.2	Seram - Indonesia
5	11/1/2008	20:37.3	-6.41	129.54	70	4.3	Banda Sea
...	...	...	...	...	...	...	...
92884	1/26/2023	15:03.9	2.7	127.1	10	3.9	Northern Molucca Sea
92885	1/26/2023	57:08.9	-7.83	121.07	10	3.8	Flores Sea
92886	1/26/2023	46:21.0	3	127.16	10	4.1	Northern Molucca Sea
92887	1/26/2023	00:35.2	-8.87	118.95	10	2.4	Sumbawa Region - Indonesia

berisikan jumlah data yang diperoleh yaitu sebanyak 92.887 data, yang berisikan parameter tanggal, origin time, latitude, longitude, depth, magnitudo, remark dan lain sebagainya.

2) Pra-Pemrosesan Data

Pada tahapan ini, bertujuan untuk memastikan bahwasanya data yang digunakan dalam tahap pelatihan dan pengujian dalam kondisi yang sesuai dengan tujuan klasifikasi. Tahapan pra-pemrosesan meliputi:

1. Pengumpulan data, yaitu mengumpulkan data-data dari berbagai sumber, seperti BMKG, USGS, Kaggle, dan sumber lainnya.
2. Pembersihan data, yaitu menghapus atau menghilangkan baris serta kolom yang tidak diperlukan.
3. Normalisasi data, menyelaraskan fitur parameter supaya berada dalam rentang yang sama untuk dianalisa.
4. Pelabelan data, yaitu memberikan label secara manual berdasarkan kategori yang relevan, seperti risiko kerusakan gempa, kedalaman gempa, dan lain sebagainya.
5. Split data, membagi data latih dan data uji dengan rasio 80:20, dimana 80% sebagai data latih dan 20% sebagai data uji untuk menguji performa model.

Tabel 4. Data Bersih

no	depth	mag	Klasifikasi risiko
1	10	4.9	Sedang
2	10	4.6	Sedang
3	121	3.7	Ringan
4	10	3.2	Ringan

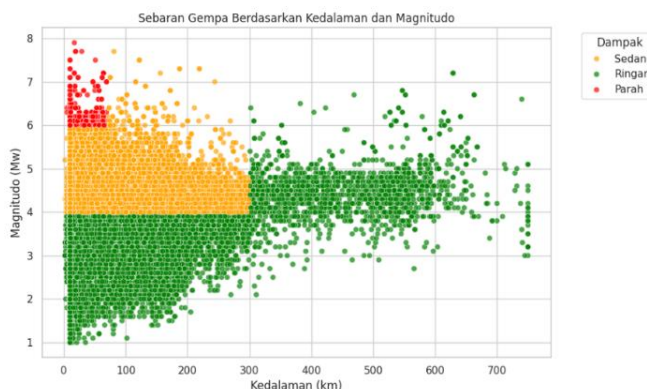
5	70	4.3	Sedang
...	...	...	...
92884	10	3.9	Ringan
92885	10	3.8	Ringan
92886	10	4.1	Sedang
92887	10	2.4	Ringan

Pada Tabel 4, merupakan hasil dari pra-pemrosesan yang menghilangkan kolom yang tidak diperlukan serta menambahkan kolom yang dibutuhkan seperti klasifikasi risiko. Sehingga menjadi data bersih yang kemudian akan diproses dengan algoritma *K-Nearest Neighbor*.

Tabel 5. Jumlah Data Kelas

No	Klasifikasi risiko	Jumlah
1	Ringan	63889
2	Sedang	28739
3	Parah	259
Total		92887

Tabel 5, merupakan jumlah kelas risiko dari data yang sudah dilabeli dengan jumlah risiko tertinggi pada kelas ringan sebanyak 63.889, kemudian kelas sedang sebanyak 28.739, dan kelas parah sebanyak 256, dengan total dataset sebanyak 92.887.



Gambar 7. Sebaran Kelas Gempa

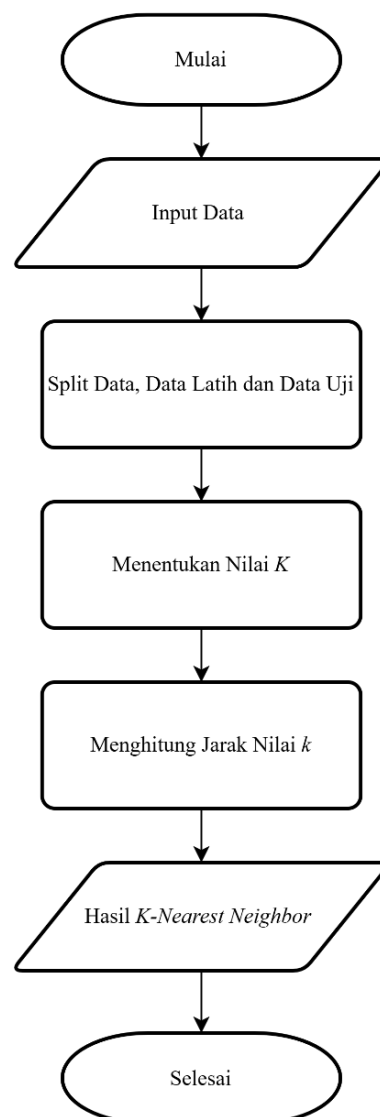
Pada Gambar 7, menunjukkan jumlah distribusi sebaran risiko gempa bumi, dengan warna hijau menunjukkan risiko ringan, kuning menunjukkan risiko sedang, merah menunjukkan risiko parah. Distribusi tersebut berasal dari parameter magnitudo dan kedalaman, dengan pelabelan dilakukan secara manual berdasarkan sumber yang relevan. Informasi ini merupakan perkiraan dampak risiko yang terjadi pada titik episenter.

3) Algoritma *K-Nearest Neighbor*

Gambar 8, merupakan diagram alur tahapan proses klasifikasi menggunakan algoritma *K-Nearest Neighbor* (K-NN) sebagai berikut;

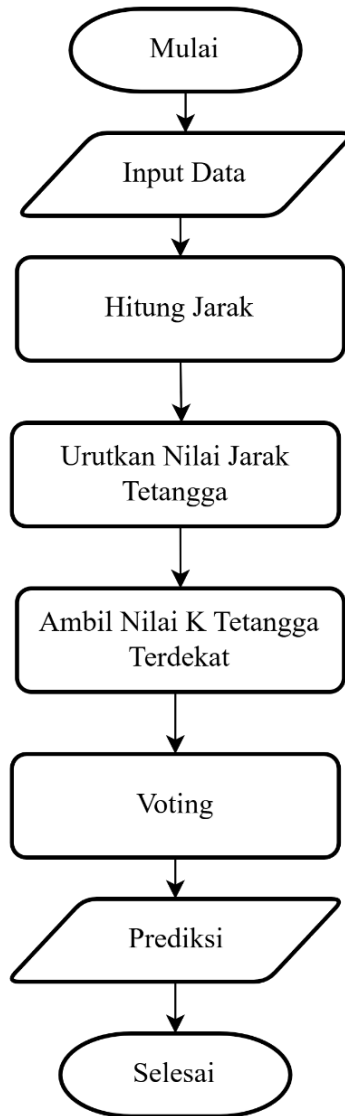
1. Mulai
2. Input data, merupakan tahapan memasukkan dataset yang telah melalui tahap pra-pemrosesan, sehingga menjadi data yang bersih dan siap digunakan dalam proses klasifikasi.

3. Split data, ialah membagi dataset tersebut menjadi dua bagian, yaitu data latih (training data) dan data uji (testing data). Pembagian ini umumnya dilakukan dengan perbandingan 80:20, tergantung pada kebutuhan dan ukuran dataset.
4. Setelah itu menentukan nilai K seperti ($K=3,5,7$), yang merupakan jumlah tetangga terdekat yang digunakan sebagai bahan pertimbangan dalam proses klasifikasi.
5. Perhitungan nilai K, proses ini melibatkan perhitungan jarak antara data uji (data baru) dan seluruh data poin (data latih) dengan menggunakan *euclidean distance*. Kemudian akan dipilih mayoritas jarak terdekat dari data uji sebanyak nilai K yang diinputkan.
6. Hasil klasifikasi ditampilkan dalam bentuk prediksi kelas dari data uji, dan tingkat akurasi dari algoritma *K-Nearest Neighbor*.
7. Selesai.



Gambar 8. Flowchart Algoritma K-NN

a) Menghitung Jarak Nilai K



Gambar 9. Flowchart Menghitung Jarak

Dari gambar 9, merupakan proses perhitungan jarak pada klasifikasi dengan algoritma *K-Nearest Neighbor* (K-NN) untuk prediksi kelas.

1. Mulai
2. Input data, seperti data latih, data uji, serta nilai K yang sudah ditentukan.
3. Menghitung jarak antara data baru atau data uji ke seluruh data latih. Biasanya menggunakan rumus perhitungan jarak *Euclidean Distance* atau yang lainnya.
4. Mengurutkan semua hasil dari jarak yang dihitung dari yang terdekat hingga terjauh, untuk menentukan tetangga atau kelas terdekat.
5. Memilih tetangga terdekat sebanyak nilai K yang diinputkan, misalnya K=3, yaitu memilih 3 jarak kandidat tetangga terdekat dari data uji atau data baru.

6. Voting bertujuan untuk memilih kelas mayoritas dengan jarak terdekat dari data uji atau data baru sebanyak nilai K yang ditentukan, misal K=3 memilih kelas mayoritas dari 3 jarak tetangga terdekat.
7. Prediksi merupakan hasil voting mayoritas tetangga terdekat.
8. Selesai.

4) Hasil Klasifikasi

Pada tahap ini, hasil klasifikasi yang diperoleh merupakan prediksi terhadap data uji berdasarkan algoritma *K-Nearest Neighbor* (K-NN). Umumnya, hasil tersebut disertai dengan perhitungan performa model berupa *confusion matrix* atau matrik evaluasi, seperti akurasi, presisi, dan *recall*, yang digunakan untuk mengevaluasi kinerja algoritma secara menyeluruh. Berikut beberapa rumusny:

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

dimana:

TP (True Positive): Data positif yang diprediksi benar.

TN (True Negative): Data negatif yang diprediksi benar.

FP (False Positive): Data negatif yang diprediksi salah sebagai positif.

FN (False Negative): Data positif yang diprediksi salah sebagai negatif.

Jika terdapat klasifikasi multi-kelas atau lebih dari dua kelas, maka akurasi dihitung berdasarkan jumlah prediksi yang tepat, yaitu ketika label prediksi sesuai dengan data aktual. Nilai tersebut kemudian dibandingkan dengan jumlah keseluruhan data uji. Adapun rumusnya adalah sebagai berikut:

$$\text{Akurasi} = \frac{\sum TP}{n} \quad (5)$$

dimana:

$\sum TP$: Jumlah data diprediksi benar (True Positive)

n : Jumlah keseluruhan data uji

D. Skema Percobaan

Guna memperoleh hasil penelitian yang komprehensif, maka dilakukanlah beberapa skema percobaan untuk menggambarkan penggunaan klasifikasi dengan algoritma *K-Nearest Neighbor* (K-NN) dalam mengklasifikasikan risiko dari gempa bumi. Seperti berikut ini:

1. Skema percobaan pertama, klasifikasi dengan algoritma *K-Nearest Neighbor* dengan menginputkan dataset yang telah dibersihkan dan siap digunakan. Kemudian dataset dibagi menjadi dua (*split data*) dengan rasio 80:20, yaitu 80% data latih dan 20% data uji. Percobaan ini menggunakan tiga variasi nilai K, yaitu K=3, K=5, dan K=7, untuk menguji tingkat akurasi dari masing-masing model nilai K yang diuji.
2. Skema percobaan kedua, mengimplementasikan algoritma *K-Nearest Neighbor* dengan menggunakan

berbagai model nilai K untuk menklasifikasi kelas data baru berdasarkan data latih yang dibuat sebelumnya.

IV. SIMULASI DAN ANALISA

A. Teknik Pengujian

Teknik pengujian pada penelitian ini dibuat untuk mengevaluasi kinerja dari algoritma K-NN dalam klasifikasi gempa bumi, serta menggambarkan penggunaan nilai K yang berbeda.

1. Teknik pengujian pertama, dengan menggunakan dataset bersih pada Tabel 4, kemudian data dibagi menjadi dua, yaitu data pelatihan sebesar 80% dan data pengujian sebesar 20%. Model digunakan dengan nilai K=3, K=5 dan K=7. Proses klasifikasi yang digunakan terdiri dari tiga kelas, yaitu ringan, sedang dan parah. Karena klasifikasi menggunakan tiga kelas, maka akurasi dihitung menggunakan rumus *confusion matrix* multi-kelas pada persamaan (5).
2. Teknik pengujian kedua, mengimplementasikan menggunakan data BMKG secara langsung sebagai data uji untuk proses klasifikasi dengan algoritma K-NN. Menggunakan dataset yang telah terlabeli yang ditampilkan Tabel 4, sebagai data latih. Dengan menggunakan model nilai K.

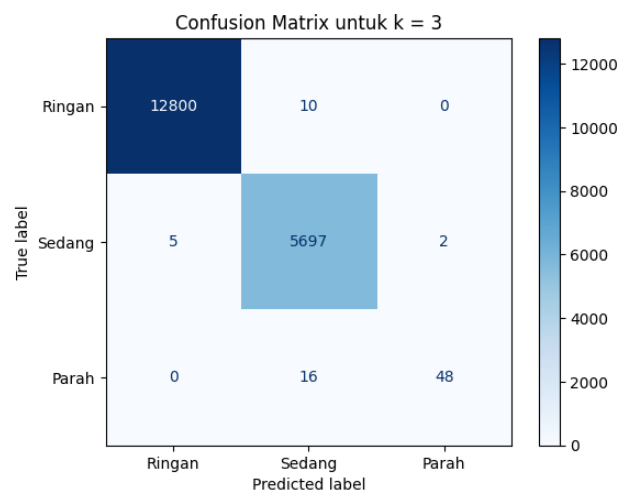
B. Pengujian Model K pada K-NN

a) Pengujian K=3

Dari tampilan Tabel 6, nilai prediksi dan aktual dengan menggunakan K= 3, menampilkan rata-rata data prediksi sesuai dengan data aktual. Dengan total data yang diprediksi sebanyak 18.578 data uji atau 20% dari dataset, kemudian dibuktikan dengan perhitungan menggunakan matrik confusion untuk mencari nilai akurasi seperti pada persamaan (5).

Tabel 6. Pengujian K=3

no	mag	depth	Aktual	Prediksi
1	5	30	Sedang	Sedang
2	2.9	10	Ringan	Ringan
3	2.9	10	Ringan	Ringan
4	4.4	10	Sedang	Sedang
5	2.8	78	Ringan	Ringan
6	3.1	10	Ringan	Ringan
7	3.3	10	Ringan	Ringan
...	...	...	...	...
18575	4.6	74	Sedang	Sedang
18576	4	15	Sedang	Sedang
18577	3.3	18	Ringan	Ringan
18578	3.8	197	Ringan	Ringan



Gambar 10. Visualisasi matrik confusion K=3

Gambar 10, menunjukan visualisasi matrik confusion dengan pengujian K=3, dengan nilai :

$$\sum TP = 18.545$$

$$n = 18.578$$

$$Akurasi = \frac{18545}{18578}$$

$$Akurasi = 99,82\%$$

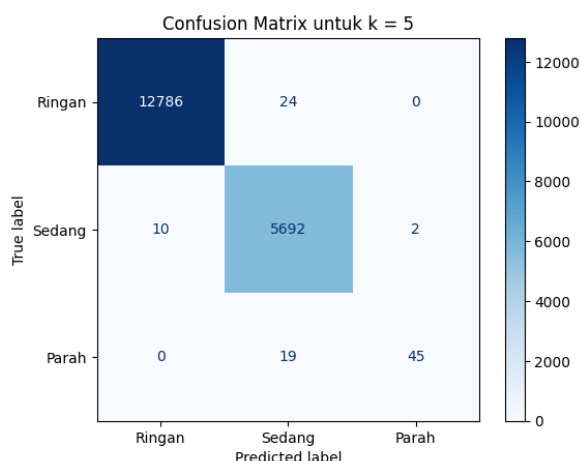
Dari perhitungan klasifikasi K-NN dengan K=3 menghasilkan nilai akurasi sebesar 99.82%, akurasi sebesar 99.67%.

2) Pengujian k=5

Dari tampilan Tabel 7, nilai prediksi dan aktual dengan menggunakan K= 5, menampilkan rata-rata data prediksi sesuai dengan data aktual. Dengan total data yang diprediksi sebanyak 18.578 data uji atau 20% dari dataset, kemudian dibuktikan dengan perhitungan menggunakan matrik confusion untuk mencari nilai akurasi.

Tabel 7. Pengujian K=5

no	mag	depth	Aktual	Prediksi
1	5	30	Sedang	Sedang
2	2.9	10	Ringan	Ringan
3	2.9	10	Ringan	Ringan
4	4.4	10	Sedang	Sedang
5	2.8	78	Ringan	Ringan
6	3.1	10	Ringan	Ringan
7	3.3	10	Ringan	Ringan
...	...	...	...	...
18575	4.6	74	Sedang	Sedang
18576	4	15	Sedang	Sedang
18577	3.3	18	Ringan	Ringan
18578	3.8	197	Ringan	Ringan



Gambar 11. Visualisasi matrik confusion K=5

Gambar 11, menunjukan visualisasi matrik confusion dengan pengujian K=5, dengan nilai :

$$\sum TP = 18.523$$

$$n = 18.578$$

$$Akurasi = \frac{18523}{18578}$$

$$Akurasi = 99,70\%$$

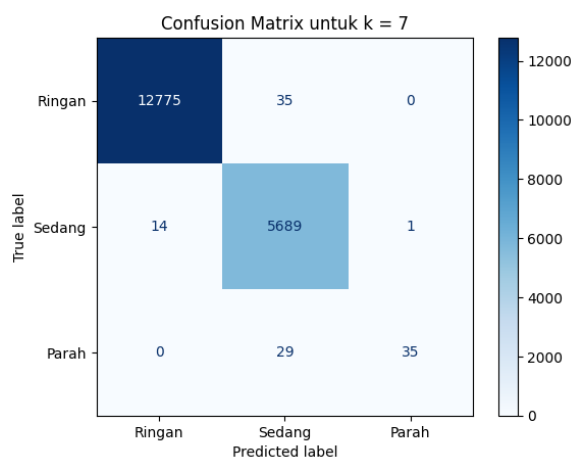
Dari perhitungan klasifikasi K-NN dengan K=5, menghasilkan nilai akurasi sebesar 99.70%.

3) Pengujian k=7

Dari tampilan Tabel 8, nilai prediksi dan aktual dengan menggunakan K= 7, menampilkan rata-rata data prediksi sesuai dengan data aktual. Dengan total data yang diprediksi sebanyak 18.578 data uji atau 20% dari dataset, kemudian dibuktikan dengan perhitungan menggunakan matrik confusion untuk mencari nilai akurasi.

Tabel 8. Pengujian K=7

no	mag	depth	Aktual	Prediksi
1	5	30	Sedang	Sedang
2	2.9	10	Ringan	Ringan
3	2.9	10	Ringan	Ringan
4	4.4	10	Sedang	Sedang
5	2.8	78	Ringan	Ringan
6	3.1	10	Ringan	Ringan
7	3.3	10	Ringan	Ringan
...	...	...	...	...
18575	4.6	74	Sedang	Sedang
18576	4	15	Sedang	Sedang
18577	3.3	18	Ringan	Ringan
18578	3.8	197	Ringan	Ringan



Gambar 12. Visualisasi matrik konfusi K=7

Gambar 12, menunjukan visualisasi matrik confusion dengan pengujian K=7, dengan nilai :

$$\sum TP = 18.499$$

$$n = 18.578$$

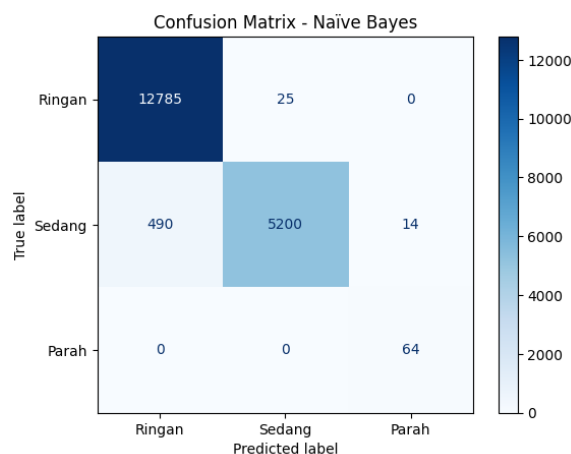
$$Akurasi = \frac{18499}{18578}$$

$$Akurasi = 99,57\%$$

Dari perhitungan klasifikasi K-NN dengan K=7, menghasilkan nilai akurasi sebesar 99.57%.

C. Perbandingan Algoritma

Melakukan perbandingan dengan algoritma lain, seperti klasifikasi *Naïve Bayes* menggunakan dataset yang sama dengan rasio perbandingan 80:20 menghasilkan akurasi sebesar 97.15%. Diperoleh dari perhitungan matrik konfusi.



Gambar 13. Visualisasi matrik konfusi Naïve Bayes

Gambar 13, merupakan visualisasi matrik konfusi *Naïve Bayes* dengan:

$$\sum TP = 18.049$$

$$n = 18.578$$

$$Akurasi = \frac{18.049}{18578}$$

$$Akurasi = 97,15\%$$

D. Analisa Pengujian

Berdasarkan hasil pengujian model K untuk memperoleh tingkat akurasi tertinggi, dengan membagi data (split data) menjadi data uji dan data latih dari katalog gempa 2008 hingga 2023 dengan rasio 20:80. Dari pengujian tersebut diperoleh hasil akurasi pada masing-masing model, pada nilai K=3 diperoleh tingkat akurasi 99.82%, pada nilai K=5 diperoleh akurasi sebesar 99.70%, dan pada nilai K=7 diperoleh tingkat akurasi sebesar 99.57%. Kemudian, perbandingan algoritma *Naïve Bayes* menggunakan metode pengujian serta dataset yang sama dengan algoritma K-NN memperoleh akurasi 99.15%.

Dari hasil tersebut diperoleh bahwa penggunaan algoritma K-NN dalam mengklasifikasikan status gempa bumi lebih baik dibandingkan dengan algoritma *Naïve Bayes*, dengan akurasi tertinggi sebesar 99.82% dengan K=3.

Berdasarkan penggunaan dataset dari katalog gempa 2008 hingga 2023, kelas yang paling banyak diprediksi dengan K-NN ialah kelas ringan. Hal tersebut dapat menjadikan ketimpangan dataset, yang mengakibatkan model cenderung memprediksi kelas mayoritas. Selain itu, kinerja algoritma K-NN tergantung pada nilai K yang optimal.

E. Implementasi K-NN

Percobaan implementasi K-NN dengan menggunakan data gempa terkini BMKG menggunakan model K=3 untuk memprediksi kelas risiko data baru.

```
=== Gempa Terkini BMKG (autogempa.xml) ===
+-----+-----+
| Lokasi          | Pusat gempa berada di laut 13 km Timur Parigi |
+-----+-----+
| Lintang         | 0.85 LS                                         |
+-----+-----+
| Bujur          | 120.29 BT                                       |
+-----+-----+
| Kedalaman       | 5.0 km                                          |
+-----+-----+
| Magnitudo       | 4.7 Mw                                          |
+-----+-----+
| Tanggal        | 04 Sep 2025                                    |
+-----+-----+
| Waktu          | 08:09:58 WIB                                   |
+-----+-----+
| Prediksi Status KNN | Sedang                                         |
+-----+-----+
```

Gambar 14. Tampilan Data Real dengan K=3

Pada gambar 14, menampilkan data gempa bumi yang diperoleh dari BMKG secara realtime dengan parameter kedalaman gempa 5.0 km serta magnitudo 4.7 Mw, di wilayah laut Timur Parigi. Data kemudian dianalisa dengan algoritma *K-Nearest Neighbor* (K-NN) dengan mengambil data kedalaman dan magnitudo. Dari hasil analisa dengan algoritma K-NN menunjukkan gempa yang terjadi dengan kedalaman 5.0 km dan magnitudo 4.7 Mw diprediksi berpotensi mengalami risiko kerusakan sedang pada wilayah episenter.

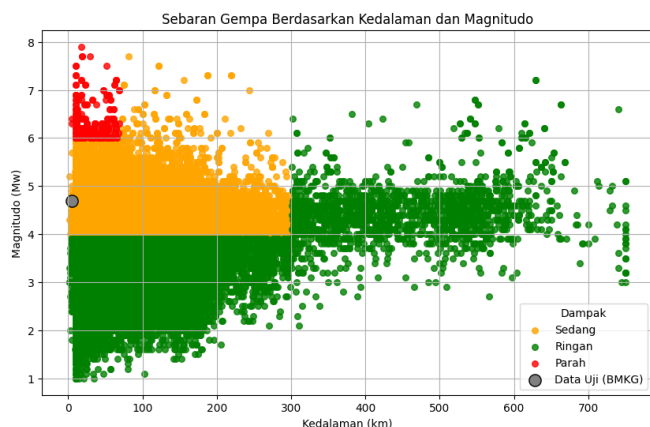
```
=== Hitung Tetangga Terdekat ===
+-----+-----+
| Jarak | Kelas |
+-----+-----+
| 0     | Sedang |
+-----+-----+
| 0     | Sedang |
+-----+-----+
| 0.1   | Sedang |
+-----+-----+
| 0.1   | Sedang |
+-----+-----+
| 0.1   | Sedang |
+-----+-----+

=== Hitung Tetangga Terjauh ===
+-----+-----+
| Jarak | Kelas |
+-----+-----+
| 745   | Ringan |
+-----+-----+
| 745   | Ringan |
+-----+-----+
| 745   | Ringan |
+-----+-----+

=== Jumlah Tetangga (K=3) ===
Sedang: 3 terdekat
Ringan: 0 terdekat
Parah: 0 terdekat
Prediksi Risiko: Sedang
```

Gambar 15. Perhitungan Jarak dengan K=3

Pada gambar 15, memperlihatkan perhitungan jarak antara data baru dengan seluruh data latih menggunakan rumus *Euclidean Distance*. Perhitungan ini dilakukan untuk menentukan jarak antara data baru dan seluruh data latih dengan K=3, kemudian sistem akan mengurutkan dari jarak terdekat atau terpendek hingga terjauh. Selanjutnya, sistem akan memilih tiga tetangga terdekat dari data baru atau sesuai dengan nilai K yang diinputkan, kemudian akan dilakukan proses voting untuk memperoleh mayoritas kelas dari data baru. Dari gambar diatas, 3 tetangga terdekat dan mayoritas kelas dari data baru menunjukkan kelas risiko sedang. Dengan demikian, sistem memprediksi gempa yang terjadi kedalam kelas risiko sedang.



Gambar 16. Kordinat Data Baru dengan Data Latih

Pada gambar 16, menunjukan Kordinat data baru dari BMKG dengan parameter yang digunakan kedalaman 5.0 km dan magnitudo 4.7 Mw, kemudian diklasifikasikan dengan algoritma K-NN ke dalam kelas risiko sedang sesuai dengan mayoritas tetangga terdekat.

V. KESIMPULAN

Berdasarkan hasil dan analisa percobaan dapat disimpulkan sebagai berikut:

Algoritma *K-Nearest Neighbor* (K-NN) efektif dalam mengklasifikasikan status resiko gempa bumi dengan parameter magnitudo dan kedalaman, dengan akurasi tertinggi sebesar 99.82% dengan model $K=3$. Hasil tersebut menunjukan penerapan algoritma K-NN dapat membantu dalam penanggulangan bencana gempa yang cepat dan akurat. Akan tetapi, terdapat ketidak seimbangan distribusi data yang dapat mempengaruhi kinerja pada algoritma K-NN.

Saran:

Saran untuk pengembangan penelitian berikutnya sebagai berikut:

1. Penelitian berikutnya, bisa membandingkan kinerja dengan algoritma lainya seperti *Naïve Bayes*, *Support Vektor Machine* (SVM), dan lain sebagainya. Diharapkan memberikan gambaran yang luas tentang kinerja menggunakan dataset yang sama.
2. Diharapkan pada penelitian selanjutnya, untuk memastikan keseimbangan distribusi kelas pada dataset latih. Guna meminimalisir kecenderungan proses klasifikasi terhadap kelas mayoritas, yang dapat menyebabkan penurunan akurasi model.
3. Penelitian selanjutnya, diharapkan terintegrasi dengan sistem *embedded* untuk memperoleh hasil prediksi gempa secara realtime.

VI. DAFTAR PUSTAKA

- [1] H. Hadi, S. Agustina, And A. Subhani, "Penguatan Kesiapsiagaan Stakeholder Dalam Pengurangan Risiko Bencana Alam Gempabumi," *Geodika: Jurnal Kajian Ilmu Dan Pendidikan Geografi*, Vol. 3, No. 1, P. 30, 2019, Doi: 10.29408/Geodika.V3i1.1476.
- [2] C. Tumangkeng, D. R. Wenas, And S. I. Umboh, "Model Kecepatan 1d Gelombang P Dan Relokasi Hiposenter Wilayah Sulawesi Utara Dan Sekitarnya Menggunakan Metode Couple Velocity-Hypocenter," *Charm Sains: Jurnal Pendidikan Fisika*, Vol. 1, No. 2, 2020.
- [3] Dian Agustina, E. Sunandi, And S. Nugroho, "Pendampingan Mitigasi Bencana Gempa Bumi Dan Tsunami Berbasis Pengetahuan Lokal Pada Masyarakat Rentan Bencana Di Kabupaten Mukomuko Bengkulu," *Engagement: Jurnal Pengabdian Kepada Masyarakat*, Vol. 4, No. 1, Pp. 87–99, 2020, Doi: 10.29062/Engagement.V4i1.102.
- [4] F. Setyawan, A. Handayani, B. Maria, D. Swandry, G. Somae, And A. Wibowo, "Analisis Spasial Sebaran Sekolah Dasar Terdampak Bencana Gempa Di Kecamatan Cugenang Kabupaten Cianjur," *Jurnal Pendidikan Geografi Undiksha*, Vol. 11, No. 1, Pp. 87–93, Apr. 2023, Doi: 10.23887/Jjpg.V11i1.56929.
- [5] F. C. Anggian, N. Hidayat, And M. T. Furqon, "Implementasi Metode Modified K-Nearest Neighbor Untuk Klasifikasi Status Gunung Berapi," *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, Vol. 3, No. 12, Pp. 11027–11033, 2019.
- [6] A. M. Argina, "Penerapan Metode Klasifikasi K-Nearest Neighbor Pada Dataset Penderita Penyakit Diabetes," *Indonesian Journal Of Data And Science*, Vol. 1, No. 2, Pp. 29–33, 2020, Doi: 10.33096/Ijodas.V1i2.11.
- [7] F. Tempola, M. Muhammad, And A. Khairan, "Perbandingan Klasifikasi Antara Knn Dan Naive Bayes Pada Penentuan Status Gunung Berapi Dengan K-Fold Cross Validation," *Jurnal Teknologi Informasi Dan Ilmu Komputer*, Vol. 5, No. 5, Pp. 577–584, 2018, Doi: 10.25126/Jtiik.201855983.
- [8] M. F. Oemarki *Et Al.*, "Perbandingan Akurasi Metode Support Vector Machine Dan K-Nearest Neighbour Dalam Prediksi Curah Hujan," No. April, Pp. 160–167, 2024.
- [9] K. K. Siregar, M. F. Tarigan, And M. T. Muhammad Rusdi, S. T., "Penerapan Komunikasi Lora Untuk Sistem Peringatan Dini Gempa Dengan Sensor Accelerometer Berbasis Nodemcu Esp8266," *Konferensi Nasional Sosial Dan Engineering*, Pp. 709–718, 2022.
- [10] F. Laia, T. Duha, M. Laia, A. K. Huda, And A. Jasuma, "Klasifikasi Data Gempa Bumi Di Pulau Sumatera Menggunakan Algoritma Naïve Bayes," *Jurnal Informatika*, Vol. 2, No. 1, Pp. 23–27, 2023, Doi: 10.57094/Ji.V2i1.840.
- [11] Ulil Albab, Rony Darpono, And Fahreza Moch Revikansyah, "Rancang Bangun Sistem Informasi Gempa Menggunakan Raspberry Pi Berbasis Web,"

- Jurnal Ilmiah Sains Teknologi Dan Informasi*, Vol. 1, No. 1, Pp. 11–20, 2023, Doi: 10.59024/Jiti.V1i1.118.
- [12] N. E. A. Putri, D. Syauqy, And M. H. Hanafi, “Sistem Klasifikasi Status Gizi Bayi Dengan Metode K-Nearest Neighbor Berbasis Sistem Embedded,” *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, Vol. 1, No. 9, Pp. 933–939, 2017.
- [13] Y. Setiawan, “Data Mining Berbasis Nearest Neighbor Dan Seleksi Fitur Untuk Deteksi Kanker Payudara,” *Jurnal Informatika: Jurnal Pengembangan It*, Vol. 8, No. 2, Pp. 89–96, 2023, Doi: 10.30591/Jpit.V8i2.4994.
- [14] R. Maharani, A. Hutagaol, V. T. Lana, Z. Azzahra, And R. Kurniawan, “Penerapan Machine Learning Dalam Prediksi Klasifikasi Big Data Kedalaman Gempa Bumi Di Indonesia Tahun 2015-2024,” Vol. 2024, No. Senada, Pp. 42–51, 2024.
- [15] Bmkg Wilayah Iii Denpasar, “Tentang Gempa,” <https://Bbmg3.Bmkg.Go.Id/Tentang-Gempa>. Accessed: May 11, 2025. [Online]. Available: <https://Bbmg3.Bmkg.Go.Id/Tentang-Gempa>
- [16] A. Bambang Siswanto And M. Salim, *Rekayasa Gempa*. 2018. Accessed: Jul. 17, 2025. [Online]. Available: https://Www.Researchgate.Net/Publication/324604027_Rekayasa_Gempa
- [17] Rhino Steel Building System, “How We Measure Earthquakes.” Accessed: Jul. 15, 2025. [Online]. Available: <https://Www.Rhinobldg.Com/Blog/How-We-Measure-Earthquakes/>
- [18] S. W. Binabar And Ivandari, “Optimasi Parameter K Pada Algoritma Knn Untuk Deteksi Penyakit Kanker Payudara,” *Ic-Tech*, Vol. Xii, No. 2, 2017.
- [19] E. Retnoningsih And R. Pramudita, “Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python,” *Bina Insani Ict Journal*, Vol. 7, No. 2, 2020, Doi: 10.51211/Biict.V7i2.1422.
- [20] Runimeirati, Muis Abdul, And Muhammad Figur, “Pelatihan Text Mining Menggunakan Bahasa Pemrograman Python,” *Abdimas Langkanae Jurnal Pengabdian Kepada Masyarakat*, Vol. Vol.3, No.1, Pp. 36–46, Jan. 2023, Doi: <https://doi.org/10.53769/Jpm.V3i1.83>.
- [21] Febby Wilyani, Qonaah Nuryan Arif, And Fitri Aslimar, “Pengenalan Dasar Pemrograman Python Dengan Google Colaboratory,” *Jurnal Pelayanan Dan Pengabdian Masyarakat Indonesia*, Vol. 3, No. 1, 2024, Doi: 10.55606/Jppmi.V3i1.1087.
- [22] Bmkg And Usgs, “Earthquakes In Indonesia,” 2025, *Kaggle*. Doi: 10.34740/Kaggle/Dsv/11641367.

VII. BIODATA PENULIS



Rizal Wahyu Purnama, lahir Pasuruan, 13 Juli 2003. Penulis menyelesaikan Pendidikan Sekolah Menengah Kejuruan di SMK Negeri 1 Singosari tahun 2021. Setelah itu penulis melanjutkan Pendidikan tinggi di perguruan swasta Institut Teknologi Nasional Malang, diprogram studi Teknik Elektro S1 dengan peminatan Teknik Elektronika Kendali dan Instrumentasi. Penulis aktif sebagai asisten Laboratorium Jaringan Komputer. Akhir kata dari penulis mengucapkan terimakasih sebanyak-banyaknya atas selesainya penelitian ini dengan judul “Klasifikasi Bencana Alam Gempa Bumi Dengan Algoritma K-Nearest Neighbor (K-NN)”.