

**ANALISIS KOMPARATIF MODEL PEMBELAJARAN
MESIN RANDOM FOREST DAN SUPPORT VECTOR
MACHINE DALAM MEMPREDIKSI PROBABILITAS
GEMPA DI JAWA BARAT**

SKRIPSI SARJANA KOMPUTER

Oleh :

Annisa Fitriana

227006516064



**PROGRAM STUDI SISTEM INFORMASI FAKULTAS
TEKNOLOGI KOMUNIKASI DAN INFORMATIKA
UNIVERSITAS NASIONAL**

2026

**ANALISIS KOMPARATIF MODEL PEMBELAJARAN
MESIN RANDOM FOREST DAN SUPPORT VECTOR
MACHINE DALAM MEMPREDIKSI PROBABILITAS
GEMPA DI JAWA BARAT**

SKRIPSI SARJANA

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer
dari Fakultas Teknologi Komunikasi dan Informatika

Oleh

Annisa Fitriana

227006516064



**PROGRAM STUDI SISTEM INFORMASI FAKULTAS
TEKNOLOGI KOMUNIKASI DAN INFORMATIKA
UNIVERSITAS NASIONAL**

2026

Halaman Pernyataan Orisinalitas

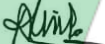
HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Bilamana di kemudian hari ditemukan bahwa karya tulis ini menyalahi peraturan yang ada berkaitan etika dan kaidah penulisan karya ilmiah yang berlaku, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Yang Menyatakan,

Nama : Annisa Fitriana

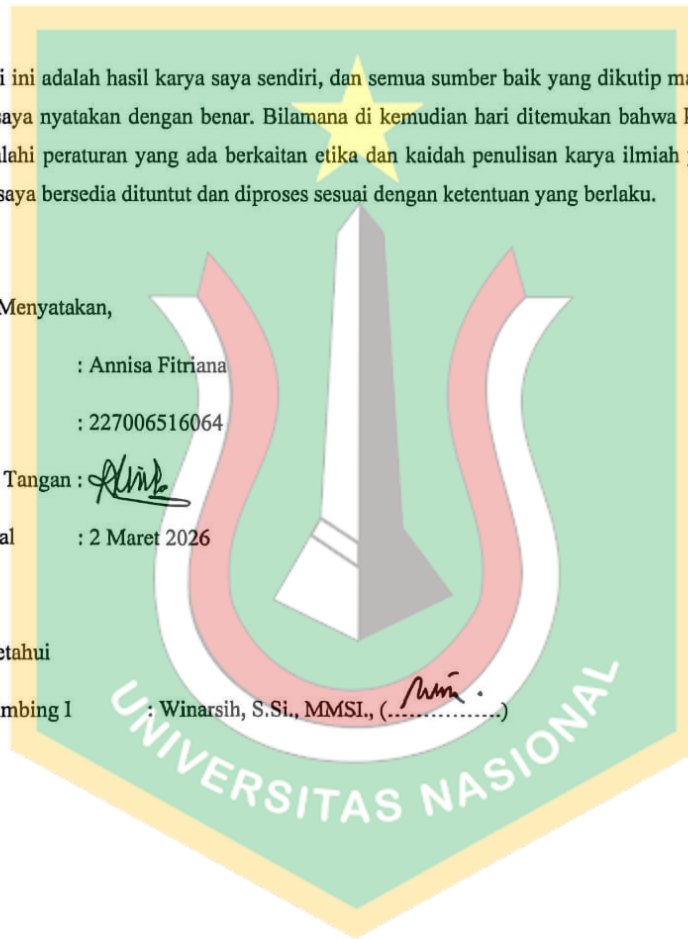
NIM : 227006516064

Tanda Tangan : 

Tanggal : 2 Maret 2026

Mengetahui

Pembimbing I : Winarsih, S.Si., MMSI., ()



Halaman Pengesahan

HALAMAN PENGESAHAN

TUGAS AKHIR

ANALISIS KOMPARATIF MODEL PEMBELAJARAN MESIN RANDOM
FOREST DAN SUPPORT VECTOR MACHINE DALAM MEMPREDIKSI
PROBABILITAS GEMPA DI JAWA BARAT



Annisa Fitriana
227006516064

Dosen Pembimbing 1

(Winarsih, S.Si., MMSI)

Pernyataan Keaslian Tugas Akhir

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

**ANALISIS KOMPARATIF MODEL PEMBELAJARAN MESIN RANDOM FOREST
DAN SUPPORT VECTOR MACHINE DALAM MEMREDIKSI PROBABILITAS
GEMPA DI JAWA BARAT**

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, 2 Maret 2026



Annisa Fitriana

227006516064

Lembar Persetujuan Review Akhir

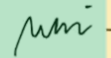
LEMBAR PERSETUJUAN REVIEW AKHIR

Tugas Akhir dengan judul :

ANALISIS KOMPARATIF MODEL PEMBELAJARAN MESIN RANDOM FOREST DAN SUPPORT VECTOR MACHINE DALAM MEMREDIKSI PROBABILITAS GEMPA DI JAWA BARAT

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Review Akhir Semester Ganjil 2025-2026 pada tanggal 24 Februari Tahun 2026

Dosen Pembimbing I



Winarsih, S.Si., MMSI

NIDN. 0310057602

Ketua Program Studi



Dr. Ir. Andriandhussili, S.Kom., MMSI.

NIDN. 0303097902

Lembar Persetujuan Judul

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama	: Annisa Fitriana	
NPM	: 227006516064	
Fakultas/Akademi	: Fakultas Teknologi Komunikasi dan Informatika	
Program Studi	: Sistem Informasi	
Tanggal Sidang	: 24 Februari 2026	
JUDUL DALAM BAHASA INDONESIA :		
Analisis Komparatif Model Pembelajaran Mesin Random Forest dan Support Vector Machine dalam Memprediksi Probabilitas Gempa di Jawa Barat		
JUDUL DALAM BAHASA INGGRIS :		
Comparative Analysis of Random Forest and Support Vector Machine Learning Models in Predicting Earthquake Probability in West Java		
TANDA TANGAN DAN TANGGAL		
Pembimbing 1	Ka. Prodi	Mahasiswa
TGL: 2 Maret 2026	TGL: 2 Maret 2026	TGL: 2 Maret 2026
 Winarsoh		 (Annisa fitriana)

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat, karunia, dan kelancaran yang diberikan-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Komparatif Model Pembelajaran Mesin Random Forest Dan Support Vector Machine dalam Memprediksi Probabilitas Gempa di Jawa Barat” dengan baik. Skripsi ini disusun sebagai salah satu syarat kelulusan untuk memperoleh gelar Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika, Universitas Nasional.

Penelitian dan penyusunan skripsi ini tentu tidak terlepas dari bimbingan, dukungan, dan bantuan dari berbagai pihak. Oleh karena itu, dengan rasa hormat penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada Ibu Winarsih, S.Si., MMSI., selaku Dosen Pembimbing Tugas Akhir, yang telah meluangkan banyak waktu, tenaga, dan pikiran untuk memberikan bimbingan, arahan, memaklumi segala kekurangan penulis, serta tak henti-hentinya memberikan motivasi selama proses penelitian hingga selesainya penyusunan skripsi ini.

Selain itu, penulis juga ingin mengucapkan terima kasih yang setulus-tulusnya kepada:

1. Ibu Dr. Ir. Andrianingsih, S.Kom., MMSI., selaku Ketua Program Studi Sistem Informasi yang telah memberikan arahan, dukungan akademik, dan memfasilitasi kelancaran perkuliahan.
2. Badan Amil Zakat Nasional (BAZNAS) yang telah memberikan dukungan nyata melalui program beasiswa, sehingga proses pendidikan dan penyusunan tugas akhir ini dapat berjalan dengan lancar.
3. Ayah dan Ibu selaku orang tua tercinta yang tak pernah lelah memberikan doa, kasih sayang, serta dukungan moral maupun material yang tak terhingga setiap saat.

4. Seluruh teman seperjuangan di Program Studi Sistem Informasi yang selalu memberikan semangat, menjadi tempat berdiskusi saya selama penelitian ini.
5. Diri penulis sendiri, yang telah memilih untuk terus berjuang pantang menyerah, menjaga komitmen, bersabar melewati berbagai revisi, dan mau menyelesaikan apa yang telah dimulai hingga mencapai titik akhir ini.

Akhir kata, semoga Tuhan Yang Maha Esa membalas segala kebaikan dan bantuan yang telah diberikan oleh semua pihak dengan balasan yang jauh lebih baik. Penulis menyadari bahwa skripsi ini masih memiliki banyak kekurangan, sehingga kritik dan saran yang bersifat membangun sangat diharapkan. Semoga skripsi ini dapat memberikan kontribusi nyata dan manfaat bagi pengembangan ilmu pengetahuan, khususnya dalam bidang sistem informasi geospasial dan mitigasi bencana.

Jakarta, 20 Februari 2026

Annisa Fitriana



ABSTRAK

Provinsi Jawa Barat memiliki kerentanan seismik yang sangat tinggi akibat aktivitas tektonik dan keberadaan sesar aktif darat. Namun, upaya prakiraan kejadian kegempaan seringkali terkendala oleh pola data geofisika yang sangat kompleks, tidak linier, serta didominasi oleh kelas mayoritas di mana kejadian berskala besar jauh lebih jarang terjadi dibandingkan kondisi aman. Pendekatan statistik konvensional terbukti kurang peka dalam menangkap anomali tersebut, sehingga dibutuhkan intervensi kecerdasan buatan yang adaptif sekaligus mampu memetakan dampaknya terhadap masyarakat. Penelitian ini bertujuan untuk mengevaluasi kinerja dua algoritma klasifikasi tingkat lanjut dalam memprakirakan kemunculan lindu berskala di atas 4.5 bermagnitudo, serta mengintegrasikan model paling optimal ke dalam platform pemetaan spasial berbasis web guna menilai tingkat paparan ancaman terhadap warga. Metode penelitian kuantitatif diterapkan dengan memanfaatkan rekam jejak historis dari Badan Meteorologi, Klimatologi, dan Geofisika selama kurun waktu sepuluh tahun (2015-2025) yang diagregasikan ke dalam grid beresolusi tinggi. Evaluasi difokuskan pada penanganan ketimpangan data historis menggunakan teknik penyeimbang kelas secara terprogram, di mana pengujian dilakukan secara runtun waktu guna mencegah kebocoran informasi. Hasil evaluasi metrik membuktikan bahwa pendekatan pohon keputusan jamak tampil jauh lebih superior dibandingkan metode *hyperplane* tunggal. Arsitektur pertama berhasil membukukan nilai metrik *Area Under the Curve* sebesar 0,8171, ketepatan keseluruhan mencapai 0,7703, dan tingkat sensitivitas tangkapan data di angka 0,5364. Di sisi lain, arsitektur pembandingan hanya mampu mencatatkan sensitivitas yang sangat lemah yakni 0,3526 meskipun akurasi berada di level 0,7568. Analisis lanjutan menyingkap fakta bahwa variabel kedalaman rata-rata menjadi prediktor paling menentukan dengan skor kepentingan sebesar 0,34. Kesimpulan dari studi ini menegaskan kelayakan pendekatan *ensemble* dalam membedah data kebencanaan yang timpang. Penggabungan komputasi cerdas dengan sensus demografi sukses merumuskan skor kerawanan secara visual, menghasilkan instrumen taktis yang sangat relevan untuk mendukung peringatan dini serta perumusan kebijakan evakuasi oleh otoritas penanggulangan bencana daerah.

Kata Kunci: Algoritma Klasifikasi, Kerentanan Penduduk, Ketimpangan Kelas, Mitigasi Bencana, Sistem Informasi Geografis

ABSTRACT

West Java exhibits high seismic vulnerability due to active tectonic subductions and numerous inland faults. Predicting such seismic events is a highly complex computational challenge, primarily because geophysical anomalies are non-linear and historical earthquake data suffers from extreme class skewness. In these records, significant tremors occur far less frequently than safe periods. Consequently, traditional mathematical models often become heavily biased toward the majority class, failing to detect actual impending threats. This research aims to rigorously evaluate two advanced classification architectures in forecasting the occurrence of earthquakes with a magnitude above 4.5. Furthermore, it seeks to integrate the most optimal model into a web-based spatial mapping platform to dynamically assess civilian risk exposure. The study employs a quantitative methodology, processing a decade of historical catalog records (2015–2025) from the national meteorological agency into high-resolution spatial grids. To overcome the severe historical data disproportion, the algorithm training phase applies a programmatic class weight balancing technique. The models are then validated using a strict time-series aware splitting protocol to prevent data leakage and simulate real-world conditions accurately. The quantitative results demonstrate that the ensemble-based architecture significantly outperformed the margin-based classifier in navigating the skewed dataset. The multiple decision tree approach achieved an Area Under the Curve metric of 0.8171, an overall accuracy of 0.7703, and a critical sensitivity rate of 0.5364. Conversely, the comparative architecture recorded an accuracy of 0.7568 but suffered a heavily degraded sensitivity rate of merely 0.3526. Feature importance analysis further revealed that the average depth of prior tremors was the most deterministic predictor, registering an importance score of 0.34. In conclusion, ensemble algorithms are proven to be highly effective for analyzing unbalanced geological disaster data. By combining these computational insights with actual demographic census figures, the developed system successfully visualizes specific red zones where high seismic threats

intersect with dense populations. This integration provides regional disaster management authorities with a robust, tactical instrument for strategic spatial planning and precise early warning mitigation.

Keywords: *Class Imbalance, Classification Algorithm, Disaster Mitigation, Geographic Information System, Population Vulnerability*



DAFTAR ISI

Halaman Pernyataan Orisinalitas	iii
Halaman Pengesahan	iv
Pernyataan Keaslian Tugas Akhir.....	v
Lembar Persetujuan Review Akhir.....	vi
Lembar Persetujuan Judul	vii
Abstrak	x
<i>Abstract</i>	xi
Daftar Isi.....	xiii
Daftar Gambar.....	xvii
Daftar Tabel	xviii
Bab I Pendahuluan	19
1.1 Latar Belakang	19
1.2 Rumusan Masalah	22
1.3 Tujuan Penelitian	23
1.4 Manfaat Penelitian	23
1.5 Batasan Masalah.....	24
1.6 Mata Kuliah yang Mendasari Penelitian	26
Bab II Tinjauan Pustaka	28
2.1 Landasan Teori.....	28
2.1.1 Machine Learning	28
2.1.1.1 Random Forest	28

2.1.1.2	Support Vector Machine (SVM).....	30
2.1.1.3	Ketidakseimbangan Data (<i>Imbalanced Data</i>).....	31
2.1.2	Peramalan Gempa Probabilistik.....	31
2.1.2.1	Probabilistic Seismic Hazard Analysis (PSHA)	31
2.1.2.2	Model Kejadian Gempa (Hukum Gutenberg-Richter).....	32
2.1.2.3	Hazard Curves dan Return Period.....	32
2.1.3	Sistem Informasi Geografis (SIG) dan WebGIS.....	32
2.1.3.1	Definisi dan Komponen SIG	32
2.1.3.2	WebGIS.....	32
2.1.4	Metrik Evaluasi Model.....	33
2.2	Penelitian Terdahulu.....	34
2.3	Kerangka Pemikiran.....	37
Bab III	Metodologi Penelitian.....	39
3.1	Jenis Penelitian.....	39
3.2	Lokasi Penelitian dan Waktu Penelitian	41
3.2.1	Lokasi Penelitian.....	41
3.2.2	Waktu Penelitian.....	42
3.3	Data dan Variabel Penelitian.....	43
3.3.1	Data Gempa Bumi BMKG.....	43
3.3.2	Data Kependudukan BPS.....	44
3.3.3	Pembentukan Grid Spasial Penelitian	45
3.3.4	Variabel Dependen (Target).....	45
3.3.5	Variabel Independen (Predictors)	46

3.3.6	Indeks Risiko (Risk Exposure)	46
3.4	Populasi dan Sample	47
3.4.1	Populasi Penelitian	47
3.4.2	Unit Analisis	47
3.4.3	Sampel Penelitian	47
3.4.4	Pembagian Data Latih, Validasi, dan Uji	48
3.5	Perangkat Penelitian dan Arsitektur Sistem	49
3.5.1	Perangkat Keras (Hardware)	49
3.5.2	Perangkat Lunak (Software)	50
3.5.3	Integrasi Perangkat dalam Alur Sistem	51
3.6	Teknik Pengambilan Data	51
3.6.1	Teknik Pengambilan Data Gempa Bumi dari BMKG	51
3.6.2	Teknik Pengambilan Data Kependudukan dari BPS	52
3.6.3	Integrasi Data BMKG dan BPS ke dalam Grid	52
3.7	Tahapan Pengolahan Data dan Pemodelan	53
3.7.1	Feature Engineering	53
3.7.2	Pemodelan Probabilitas dan Penanganan <i>Imbalanced Data</i>	55
3.8	Evaluasi Model	55
3.8.1	Metrik Evaluasi	55
3.8.2	Perbandingan Kinerja Model	56
3.9	Pengujian Sistem WebGIS	57
Bab IV	Hasil dan Pembahasan	58
4.1	Deskripsi dan Pemrosesan Dataset	58

4.2	Pembentukan Dataset Pelatihan dan Penanganan Ketimpangan Kelas	58
4.3	Evaluasi Kinerja Model Machine Learning	59
4.4	Verifikasi Prediksi Menggunakan Perhitungan Manual (Simulasi Model)	63
4.4.1	Hasil Komputasi Prediksi.....	64
4.4.2	Validasi Silang dengan Data Aktual	65
4.5	Implikasi Peta Risiko pada Sistem WebGIS	65
BAB V	Penutup	67
5.1	Kesimpulan	67
5.2	Saran.....	68
Daftar Pustaka		70
Daftar Lampiran		72
1.	Cek AI Detector	72
2.	Cek Turnitin	73



DAFTAR GAMBAR

Gambar 2. 1 Diagram Kerangka Pemikiran.....	38
Gambar 3. 1 Diagram Flowchart Preprocessing BMKG	52
Gambar 3. 2 Diagram Integrasi BMKG–BPS ke Grid.....	53
Gambar 3. 3 Diagram Feature Engineering Diagram	54
Gambar 4. 1 Kurva ROC - Perbandingan Performa Random Forest vs SVM.....	61
Gambar 4. 2 Tingkat Kepentingan Fitur Model Random Forest	62



DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	36
Tabel 3. 1 Ringkasan Fitur	54
Tabel 3. 2 Matrik Evaluasi	56
Tabel 4. 1 Hasil Komparasi Metrik Evaluasi Model (Validation Set)	60
Tabel 4. 2 Parameter Input Sampel Grid (Tahun 2021)	64
Table 4. 3 Hasil Perhitungan Prediksi Manual Algoritma	64
Tabel 4. 4 Hasil Validasi Kejadian Aktual BMKG	65

