



**IMPLEMENTASI SVM-SMOTE DALAM MENGATASI KELAS TIDAK  
SEIMBANG UNTUK PREDIKSI BAHAYA ASTEROID PADA BUMI  
MENGUNAKAN ALGORITMA KLASIFIKASI *RANDOM FOREST***

**Skripsi**

**Untuk Memenuhi Persyaratan**

**Dalam Menyelesaikan Strata-1 Ilmu Komputer**

**Oleh**

**AHMAD NOVAL**

**NIM 1711016210001**

**PROGRAM STUDI S-1 ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMBUNG MANGKURAT  
BANJARBARU**

**JUNI 2023**



**IMPLEMENTASI SVM-SMOTE DALAM MENGATASI KELAS TIDAK  
SEIMBANG UNTUK PREDIKSI BAHAYA ASTEROID PADA BUMI  
MENGUNAKAN ALGORITMA KLASIFIKASI *RANDOM FOREST***

**Skripsi**

**Untuk Memenuhi Persyaratan**

**Dalam Menyelesaikan Strata-1 Ilmu Komputer**

**Oleh**

**Ahmad Noval**

**NIM 1711016210001**

**PROGRAM STUDI S-1 ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMBUNG MANGKURAT  
BANJARBARU**

**JUNI 2023**

**SKRIPSI**

**IMPLEMENTASI SVM-SMOTE DALAM MENGATASI KELAS TIDAK SEIMBANG UNTUK PREDIKSI BAHAYA ASTEROID PADA BUMI MENGGUNAKAN ALGORITMA KLASIFIKASI *RANDOM FOREST***

Oleh :

**AHMAD NOVAL  
1711016210001**

Telah dipertahankan di depan Dosen Penguji pada tanggal 22 Juni 2023,  
Susunan Dosen Penguji :

Pembimbing I



Triando Hamonangan Saragih,  
S.Kom., M.Kom.  
NIP. 19930824201901012

Dosen Penguji I



Muliadi, S.Kom., M.Cs.  
NIP. 197804222010121002

Pembimbing II



Fatma Indriani, S.T, M.I.T.  
NIP. 198404202008122004

Dosen Penguji II



Rahmat Ramadhani, S.Kom., M.Sc.  
NIP. 19920330201901110001



Banjarbaru, 22 Juni 2023  
Kelas Program Studi Ilmu Komputer

Jawan Budiman, S.T., M.Kom  
NIP. 197703252008121001

## **PERNYATAAN**

Dengan ini Saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan Saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Banjarbaru, Juni 2023

Yang Menyatakan,



**Ahmad Noval**  
**NIM. 1711016210001**

## ABSTRAK

### IMPLEMENTASI SVM-SMOTE DALAM MENGATASI KELAS TIDAK SEIMBANG UNTUK PREDIKSI BAHAYA ASTEROID PADA BUMI MENGGUNAKAN ALGORITMA KLASIFIKASI *RANDOM FOREST*

(Oleh : Ahmad Noval; Pembimbing: Triando Hamonangan Saragih S.Kom., M.Kom. dan Fatma Indriani, S.T., M.I.T., Ph.D.; 2023; 68 halaman)

Asteroid adalah bebatuan yang mengelilingi matahari dan biasanya ditemukan di sabuk asteroid antara Mars dan Jupiter. Mereka datang dalam berbagai ukuran, dari beberapa meter hingga ratusan kilometer. Asteroid dapat menimbulkan ancaman karena ketika bertabrakan dengan Bumi, mereka dapat memberikan efek yang signifikan. Tabrakan asteroid dapat menimbulkan efek langsung seperti gempa bumi, semburan udara, dan efek termal yang dapat merusak area yang luas. Selain itu, asteroid juga dapat memicu bahaya sekunder seperti tsunami dan perubahan iklim jangka panjang. Untuk menghadapi ancaman asteroid, dilakukan beberapa upaya untuk mendeteksi bahaya dari ancaman asteroid. Banyaknya jumlah asteroid yang berada disekitar Bumi, membuat pendeteksian bahaya secara manual menjadi kurang efektif. Pendeteksian tersebut dapat dilakukan secara otomatis dengan menggunakan algoritma komputasi, yaitu dengan cara melakukan klasifikasi pada bahaya dari asteroid menggunakan fitur-fiturnya. Pada penelitian ini dilakukan klasifikasi dengan menggunakan algoritma *Random Forest* dan *Decision Tree* serta *oversampling* dengan menggunakan SVM-SMOTE. Data yang digunakan pada penelitian ini adalah data *Nearest Earth Objects* (asteroid) yang memiliki kelas *major* sebanyak 81.996 data dan memiliki kelas *minor* sebanyak 8840 data. Penelitian ini dilakukan dengan membandingkan klasifikasi *Random Forest* serta *Decision Tree* dengan dan tanpa *oversampling*. Hasilnya didapatkan bahwa klasifikasi *Random Forest* menggunakan parameter terbaik dan SVM-SMOTE menghasilkan performa terbaik dibandingkan dengan model klasifikasi lain, dengan nilai akurasi sebesar 89,69%, presisi sebesar 47,68%, *recall* sebesar 61,65%, *f1-score* sebesar 53,77%, dan AUC sebesar 0,7718.

Kata kunci: Asteroid, *Hazardous*, *Random Forest*, *Decision Tree*, *Oversampling*, SVM-SMOTE, Klasifikasi.

## **ABSTRACT**

**IMPLEMENTATION OF SVM-SMOTE IN HANDLING IMBALANCED CLASS FOR PREDICTING HAZARDOUS ASTEROID ON EARTH USING RANDOM FOREST CLASSIFICATION ALGORITHM** (By: Ahmad Noval; Advisor: Triando Hamonangan Saragih, S.Kom., M.Kom. and Fatma Indriani, S.T., M.I.T., Ph.D.; 2023; 68 Pages)

*Asteroids are rocks that's orbit the sun, usually they located in asteroid belt between Mars and Jupiter. They come in various size, from a few metres to a hundred kilometres. They can pose a threat when they collide with Earth. Such collision can cause immediate impact likes earthquakes, airbursts, and thermal effects that can cause extensive damage. Besides that, asteroid can also trigger secondary threat such as tsunami and long-term climate changes. To overcome the threat of asteroids, several efforts are made to detect them. Due to large number of asteroids near Earth, makes manual detection become less effective. An automated detection can overcome that problem using computational algorithms, to classify and identity the hazardous of asteroids based on their features. In this study, classification is performed using Random Forest and Decision Tree algorithm, along with oversampling techniques using SVM-SMOTE. The dataset that is used in this study is Nearest Earth Objects, which consisted of 81,996 data in majority class and 8,840 data in minority class. This study aimed to compare the performance of Random Forest and Decision Tree with and without using oversampling. The results showed that Random Forest with optimized parameters and implementing SVM-SMOTE give the best performance compared to other classification models, with accuracy 89,69%, precision 47,68%, recall 61,65%, f1-score 53,77%, and AUC 0,7718.*

*Keywords: Asteroid, Hazardous, Random Forest, Decision Tree, Oversampling, SVM-SMOTE, Classification.*

## PRAKATA

Puji syukur penulis panjatkan kepada Allah SWT. karena atas berkat, rahmat, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul **IMPLEMENTASI SVM-SMOTE DALAM MENGATASI KELAS TIDAK SEIMBANG UNTUK PREDIKSI BAHAYA ASTEROID PADA BUMI MENGGUNAKAN ALGORITMA KLASIFIKASI *RANDOM FOREST*** untuk memenuhi syarat dalam menyelesaikan Pendidikan program S-1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat. Tak lupa penulis panjatkan sholawat dan salam ke hadirat Rasulullah Muhammad SAW. beserta para sahabat, keluarga, dan pengikut beliau hingga *yaumul qiama*.

Pada lembar ini penulis ingin menyampaikan ucapan terimakasih kepada pihak-pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan skripsi ini, adapun yang dimaksud adalah sebagai berikut:

1. Allah SWT. karena atas limpahan karunia-Nya dan bantuan-Nya lah penulis dapat menyelesaikan skripsi.
2. Keluarga terutama kedua orang tua dan saudara Saya yang selalu memberikan do'a, semangat, bantuan, dan dukungan dalam proses penyelesaian skripsi.
3. Bapak Dodon Turianto Nugrahadi, S.Kom., M.Eng. selaku Dosen Pembimbing Akademik.
4. Bapak Triando Hamonangan Saragih, S.Kom., M.Kom. selaku Dosen Pembimbing Utama dan Ibu Fatma Indriani, S.T, M.I.T., Ph.D. selaku Dosen Pembimbing Pendamping yang telah memberikan waktu dan bimbingannya kepada penulis dalam pengerjaan skripsi.
5. Bapak Irwan Budiman, S.T., M.Kom. selaku Koordinator Program Studi Ilmu Komputer FMIPA ULM, atas bantuan dan izin beliau skripsi ini dapat diselesaikan.
6. Seluruh Dosen dan Staf Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama berkuliah di Ilmu Komputer yang sangat bermanfaat.

7. Teman-teman keluarga Ilmu Komputer Angkatan 2017 yang memberikan bantuan dan dukungan yang senantiasa menjadi alat pacu penulis untuk selalu berusaha dan semangat untuk menyelesaikan skripsi.

Akhir kata penulis menyadari sepenuhnya bahwa penulisan skripsi ini masih jauh dari kesempurnaan, maka penulis mengharapkan bantuan berupa kritik dan saran yang membangun dari semua pihak sangat diharapkan demi penyempurnaan dan mutu penulisan skripsi ini. Semoga tulisan ini dapat bermanfaat bagi semua pihak untuk dapat menambah pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT.

Banjarbaru, Juni 2023

Penulis



## DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
PERNYATAAN.....	iii
ABSTRAK .....	iv
<i>ABSTRACT</i> .....	v
PRAKATA.....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR .....	xii
DAFTAR LAMPIRAN.....	xiii
BAB I.....	1
PENDAHULUAN .....	1
1.1    Latar Belakang.....	1
1.2    Rumusan Masalah .....	3
1.3    Batasan Masalah.....	4
1.4    Tujuan Penelitian.....	4
1.5    Manfaat Penelitian.....	4
BAB II.....	5
TINJAUAN PUSTAKA.....	5
2.1    Kajian Terdahulu .....	5
2.2    Keaslian Penelitian .....	7
2.3    Landasan Teori .....	11
BAB III .....	18
METODE PENELITIAN.....	18
3.1    Alat Penelitian .....	18
3.2    Bahan Penelitian.....	18
3.3    Variabel Penelitian .....	18
3.4    Prosedur Penelitian.....	18
BAB IV .....	22

HASIL DAN PEMBAHASAN.....	22
4.1 Hasil.....	22
4.1.1 Pengumpulan Data .....	22
4.1.2 <i>Preprocessing</i> .....	24
4.1.3 Pembagian Data .....	25
4.1.4 Sampling Data.....	26
4.1.5 Klasifikasi Data.....	28
4.1.6 <i>Random Forest</i> .....	28
4.1.7 <i>Random Forest</i> dengan SVM-SMOTE.....	40
4.1.8 Uji Parameter <i>Random Forest</i> dengan SVM-SMOTE.....	41
4.1.9 <i>Decision Tree</i> .....	43
4.1.10 <i>Decision Tree</i> dengan SVM-SMOTE.....	51
4.1.11 Uji Parameter <i>Decision Tree</i> dengan SVM-SMOTE.....	52
4.2 Evaluasi .....	54
4.3 Pembahasan .....	62
BAB V.....	68
PENUTUP.....	68
5.1 Kesimpulan.....	68
5.2 Saran .....	68
DAFTAR PUSTAKA .....	69
LAMPIRAN.....	72

## DAFTAR TABEL

Tabel 1. Keaslian Penelitian.....	8
Tabel 2. Penelitian yang Diusulkan .....	10
Tabel 3. Confusion Matrix .....	16
Tabel 4. Keterangan Rumus Confusion Matrix .....	16
Tabel 5. Nilai Kualitas Classifier .....	17
Tabel 6. Contoh Dataset Asli .....	22
Tabel 7. Deskripsi Fitur Dataset.....	23
Tabel 8. Data Setelah Dilakukan Preprocessing .....	25
Tabel 9. Hasil Pembagian Data .....	25
Tabel 10. Jumlah Data Sebelum Oversampling .....	26
Tabel 11. Jumlah Data Setelah Oversampling .....	27
Tabel 12. Data Latih Setelah Oversampling .....	27
Tabel 13. Data contoh untuk klasifikasi Random Forest .....	29
Tabel 14. Nilai entropy dan gain pada pohon pertama .....	30
Tabel 15. Data atribut est_diameter_min yang bernilai $\geq 0,1$ .....	31
Tabel 16. Nilai entropy dan gain pada node 1.2 .....	31
Tabel 17. Data atribut relative_velocity yang bernilai $\leq 60000$ .....	32
Tabel 18. Nilai entropy dan gain pada node 2.2 .....	32
Tabel 19. Data kedua untuk model Random Forest.....	34
Tabel 20. Data ketiga untuk model Random Forest.....	35
Tabel 21. Hasil majority voting .....	37
Tabel 22. Confusion Matrix Random Forest.....	38
Tabel 23. Nilai Parameter .....	39
Tabel 24. Performa Klasifikasi Random Forest.....	39
Tabel 25. Performa Klasifikasi Random Forest + SVM-SMOTE .....	40
Tabel 26. Pengujian parameter Random Forest .....	42
Tabel 27. Nilai Parameter Random Forest Terbaik .....	42
Tabel 28. Performa Klasifikasi Random Forest (Tuning) + SVM-SMOTE.....	42
Tabel 29. Data untuk klasifikasi Decision Tree .....	44
Tabel 30. Hasil perhitungan fitur est_diameter_min.....	44
Tabel 31. Hasil perhitungan entropy dan gain untuk root node .....	45
Tabel 32. Data absolute_magnitude $\leq 21,14$ .....	45
Tabel 33. Hasil perhitungan fitur est_diameter_min.....	46
Tabel 34. Hasil perhitungan entropy dan gain .....	46
Tabel 35. Data miss_distance $\leq 60017850,92$ .....	47
Tabel 36. Hasil perhitungan entropy dan gain .....	47
Tabel 37. Hasil Prediksi Pohon Keputusan.....	48
Tabel 38. Confusion Matrix Decision Tree.....	48

Tabel 39. Nilai Parameter .....	49
Tabel 40. Performa Klasifikasi Decision Tree .....	49
Tabel 41. Performa Klasifikasi Decision Tree + SVM-SMOTE .....	51
Tabel 42. Pengujian parameter Decision Tree .....	52
Tabel 43. Nilai Parameter Terbaik .....	53
Tabel 44. Performa Klasifikasi Hyperparameter Decision Tree .....	53
Tabel 45. Confusion Matrix Random Forest.....	54
Tabel 46. Confusion Matrix Random Forest dengan SVM-SMOTE .....	56
Tabel 47. Confusion Matrix Random Forest (Tuning) dengan SVM-SMOTE ....	57
Tabel 48. Confusion Matrix Decision Tree.....	58
Tabel 49. Confusion Matrix Decision Tree dengan SVM-SMOTE .....	59
Tabel 50. Confusion Matrix Decision Tree (Tuning) dan SVM-SMOTE .....	60
Tabel 51. Perbandingan Performa Model Random Forest.....	62
Tabel 52. Perbandingan Performa Model Decision Tree.....	64

## DAFTAR GAMBAR

Gambar 1. Visualisasi Metode Random Forest.....	13
Gambar 2. Struktur Decision Tree .....	14
Gambar 3. Ilustrasi SVM-SMOTE .....	15
Gambar 4. Prosedur Penelitian.....	19
Gambar 5. Perbandingan Antar Kelas.....	23
Gambar 6. Heatmap Feature Correlation .....	24
Gambar 7. Pohon Keputusan I .....	30
Gambar 8. Pohon Keputusan Node 1.2.....	32
Gambar 9. Pohon Keputusan Pertama .....	33
Gambar 10. Pohon Keputusan Kedua .....	34
Gambar 11. Pohon Keputusan Ketiga.....	36
Gambar 12. Performa Model Random Forest .....	39
Gambar 13. ROC Curve Random Forest .....	39
Gambar 14. Performa Random Forest + SVM-SMOTE.....	41
Gambar 15. ROC Curve Random Forest + SVM-SMOTE .....	41
Gambar 16. Performa Random Forest (Tuning) + SVM-SMOTE .....	43
Gambar 17. ROC Curve Random Forest (Tuning) + SVM-SMOTE .....	43
Gambar 18. Root Node .....	45
Gambar 19. Bentuk node 2.....	46
Gambar 20. Bentuk Pohon Keputusan .....	47
Gambar 21. Performa Model Decision Tree .....	50
Gambar 22. ROC Curve Decision Tree .....	50
Gambar 23. Performa Decision Tree + SVM-SMOTE.....	51
Gambar 24. ROC Curve Decision Tree + SVM-SMOTE .....	52
Gambar 25. Performa Decision Tree (Tuning) + SVM-SMOTE .....	53
Gambar 26. ROC Curve Decision Tree (Tuning) + SVM-SMOTE .....	54
Gambar 27. Perbandingan Akurasi Model Random Forest .....	63
Gambar 28. Perbandingan Presisi Model Random Forest .....	63
Gambar 29. Perbandingan Recall Model Random Forest.....	63
Gambar 30. Perbandingan F1-score Model Random Forest.....	64
Gambar 31. Perbandingan AUC Model Random Forest .....	64
Gambar 32. Perbandingan Akurasi Model Decision Tree .....	65
Gambar 33. Perbandingan Presisi Model Decision Tree .....	65
Gambar 34. Perbandingan Recall Model Decision Tree.....	66
Gambar 35. Perbandingan F1-score Model Decision Tree .....	66
Gambar 36. Perbandingan AUC Model Decision Tree .....	66
Gambar 37. Perbandingan Model Random Forest & Decision Tree .....	67

## **DAFTAR LAMPIRAN**

Lampiran 1. Dataset *Nearest Earth Objects* (Asteroid)