



**PENERAPAN SMOTE-NCL UNTUK MENGATASI  
KETIDAKSEIMBANGAN KELAS PADA KLASIFIKASI PENYAKIT  
JANTUNG KORONER MENGGUNAKAN ALGORITMA SUPPORT  
VECTOR MACHINE**

**Skripsi**

**Untuk Memenuhi Persyaratan  
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

**Oleh**

**MARIANA DEWI**

**1711016320011**

**PROGRAM STUDI S-1 ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMBUNG MANGKURAT  
BANJARBARU  
JUNI 2023**



**PENERAPAN SMOTE-NCL UNTUK MENGATASI  
KETIDAKSEIMBANGAN KELAS PADA KLASIFIKASI PENYAKIT  
JANTUNG KORONER MENGGUNAKAN ALGORITMA SUPPORT  
VECTOR MACHINE**

**Skripsi**

**Untuk Memenuhi Persyaratan  
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

**Oleh**

**Mariana Dewi**

**1711016320011**

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

**JUNI 2023**

**SKRIPSI**

**PENERAPAN SMOTE-NCL UNTUK MENGATASI  
KETIDAKSEIMBANGAN KELAS PADA KLASIFIKASI PENYAKIT  
JANTUNG KORONER MENGGUNAKAN ALGORITMA SUPPORT  
VECTOR MACHINE**

Oleh :

**MARIANA DEWI**

**1711016320011**

Telah dipertahankan di depan Dosen Penguji pada tanggal 9 Juni 2023,

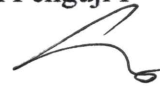
Susunan Dosen Penguji :

Pembimbing I



**Triando Hamonangan Saragih,**  
**S.Kom., M.Kom**  
**NIP. 199308242019031012**

Dosen Penguji I



**Radityo Adi Nugroho, S.T., M.Kom**  
**NIP. 198212042008011006**

Pembimbing II



**Rudy Herteno, S.Kom., M.Kom**  
**NIP. 198809252022031003**

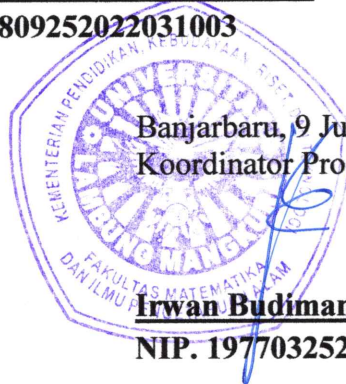
Dosen Penguji II



**Dodon Turianto Nugrahadi,**  
**S.Kom., M.Eng**  
**NIP. 198001122009121002**

Banjarbaru, 9 Juni 2023

Koordinator Program Studi Ilmu Komputer



**Irwan 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 keserjanaan 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, 9 Juni 2023

Yang Menyatakan,



Mariana Dewi

NIM. 1711016320011

## ABSTRAK

**PENERAPAN SMOTE-NCL UNTUK MENGATASI KETIDAKSEIMBANGAN KELAS PADA KLASIFIKASI PENYAKIT JANTUNG KORONER MENGGUNAKAN ALGORITMA *SUPPORT VECTOR MACHINE*** (Oleh: Mariana Dewi; Pembimbing: Triando Hamonangan Saragih, S.Kom., M.Kom. dan Rudy Herteno, S.Kom., M.Kom.; 2023; 60 halaman)

Penyakit jantung koroner (PJK) terjadi akibat penyumbatan atau penyempitan pada pembuluh darah jantung karena adanya endapan lemak dan kolesterol sehingga mengakibatkan suplai darah ke jantung menjadi terganggu. PJK masih merupakan masalah kesehatan yang penting dan berdampak secara sosioekonomi karena biaya obat-obatan yang cukup mahal dan lamanya waktu perawatan serta pengobatannya. Upaya pencegahan melalui deteksi dini dan upaya pengendaliannya sangat penting untuk dilakukan. Salah satu cara untuk mendeteksi penyakit jantung koroner dengan memanfaatkan teknologi komputasi, yaitu melakukan klasifikasi menggunakan algoritma tertentu. Pada penelitian ini dilakukan klasifikasi dengan menggunakan algoritma *Support Vector Machine* (SVM) serta penanganan ketidakseimbangan data menggunakan SMOTE dan SMOTE-NCL. Data yang digunakan dalam penelitian ini adalah data *Coronary Heart Disease* yang memiliki dua buah kelas, yaitu kelas 0 (negatif PJK) dan kelas 1 (positif PJK) dengan permasalahan data yang tidak seimbang. Penelitian ini dilakukan dengan membandingkan kinerja dari klasifikasi SVM tanpa dilakukan penyeimbangan data, klasifikasi SVM dengan penyeimbangan data SMOTE, dan klasifikasi SVM dengan penyeimbangan data SMOTE-NCL. Hasil yang didapatkan dari penelitian ini adalah pada klasifikasi SVM dengan penyeimbangan data SMOTE-NCL menghasilkan kinerja terbaik jika dibandingkan dengan model klasifikasi lain dengan nilai akurasi sebesar 85,10%.

**Kata Kunci:** jantung koroner, klasifikasi, SVM, SMOTE, SMOTE-NCL

## **ABSTRACT**

**IMPLEMENTATION OF SMOTE-NCL TO OVERCOME CLASS IMBALANCE IN CORONARY HEART DISEASE CLASSIFICATION USING SUPPORT VECTOR MACHINE ALGORITHM** (By: Mariana Dewi; Advisor: Triando Hamonangan Saragih, S.Kom., M.Kom. and Rudy Herteno, S.Kom., M.Kom.; 2023; 60 pages)

Coronary heart disease (CHD) occurs as a result of blockage or narrowing of the heart's blood vessels due to the buildup of fat and cholesterol, which disrupts blood supply to the heart. CHD remains a significant health problem with socio-economic implications due to the high cost of medication, prolonged treatment duration, and management. Prevention efforts through early detection and control are crucial. One way to detect coronary heart disease is by utilizing computational technology, specifically through classification using certain algorithms. In this study, classification was performed using the Support Vector Machine (SVM) algorithm, along with addressing data imbalance using SMOTE and SMOTE-NCL techniques. The data used in this research consisted of Coronary Heart Disease data with two classes: class 0 (negative for CHD) and class 1 (positive for CHD), with an imbalanced data issue. This study compared the performance of SVM classification without data balancing, SVM classification with SMOTE data balancing, and SVM classification with SMOTE-NCL data balancing. The results obtained from this research indicate that SVM classification with SMOTE-NCL data balancing yielded the best performance compared to other classification models, achieving an accuracy of 85.10%.

**Keywords:** coronary heart disease, classification, SVM, SMOTE, SMOTE-NCL.

## PRAKATA

Puji syukur penulis panjatkan kepada Allah SWT karena atas berkat, rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul **PENERAPAN SMOTE-NCL UNTUK MENGATASI KETIDAKSEIMBANGAN KELAS PADA KLASIFIKASI PENYAKIT JANTUNG KORONER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE** 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 Triando Hamonangan Saragih, S.Kom., M.Kom. selaku dosen pembimbing utama dan Bapak Rudy Herteno, S.Kom., M.Kom. selaku dosen pembimbing pendamping yang telah memberikan waktu dan bimbingannya kepada penulis dalam pengerjaan skripsi.
4. Bapak Muliadi, S.Kom., M.Kom. selaku Dosen Pembimbing Akademik.
5. Bapak Irwan Budiman, S.T., M.Kom. selaku Koordinator Program Studi Ilmu Komputer FMIPA ULM, atas bantuan dan ijin 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.
8. Teman-teman yang telah banyak membantu dalam memberikan saran, motivasi dan yang selalu ada ketika aku berbagi keluh kesah selama proses penyelesaian 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, 9 Juni 2023



Mariana Dewi



## DAFTAR ISI

HALAMAN JUDUL.....	i
LEMBAR PENGESAHAN .....	ii
PERNYATAAN.....	iii
ABSTRAK .....	iv
<i>ABSTRACT</i> .....	v
PRAKATA .....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR .....	xii
DAFTAR LAMPIRAN.....	xiii
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian .....	3
1.4 Manfaat Penelitian .....	3
1.5 Batasan Masalah .....	3
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Kajian Terdahulu .....	5
2.2 Landasan Teori .....	9
2.2.1 Jantung Koroner .....	9
2.2.2 Data Mining.....	10
2.2.3 Preprocessing .....	11
2.2.4 Data Tidak Seimbang .....	12
2.2.5 Synthetic Minority Oversampling Technique.....	14
2.2.6 Neighborhood Cleaning Rule .....	15
2.2.7 Klasifikasi.....	16
2.2.8 Support Vector Machine.....	17
2.2.9 Evaluasi .....	20
<b>BAB III METODE PENELITIAN</b>	
3.1 Alat Penelitian .....	22
3.2 Bahan Penelitian .....	22

3.3	Variabel Penelitian.....	22
3.4	Prosedur Penelitian .....	22

**BAB IV HASIL DAN PEMBAHASAN**

4.1	Hasil.....	26
4.1.1	Pengumpulan Data.....	26
4.1.2	<i>Preprocessing</i> .....	29
4.1.3	Pembagian Data.....	30
4.1.4	Penyeimbangan Data .....	30
4.1.5	Klasifikasi.....	41
4.1.6	Evaluasi .....	50
4.2	Pembahasan .....	56

**BAB V PENUTUP**

5.1	Kesimpulan.....	60
5.2	Saran .....	60

**DATAR PUSTAKA**

**LAMPIRAN**

## DAFTAR TABEL

<b>Tabel</b>	<b>Halaman</b>
Tabel 1. Keaslian Penelitian.....	6
Tabel 2. Perancangan Penelitian .....	9
Tabel 3. Derajat Ketidakseimbangan .....	13
Tabel 4. <i>Confusion Matrix</i> Kelas <i>Biner</i> .....	20
Tabel 5. Kategori Klasifikasi berdasarkan Nilai AUC.....	21
Tabel 6. Deskripsi Dataset .....	26
Tabel 7. Contoh Dataset.....	28
Tabel 8. Data sebelum Pelabelan .....	29
Tabel 9. Data setelah Pelabelan .....	29
Tabel 10. Pembagian Data Training dan Data Testing .....	30
Tabel 11. Data sebelum Oversampling SMOTE.....	30
Tabel 12. Contoh Perhitungan Data Replika Data 1 dengan Data 6.....	33
Tabel 13. Contoh Perhitungan Data Replika Data 1 dengan Data 10.....	33
Tabel 14. Contoh Perhitungan Data Replika Data 1 dengan Data 2.....	34
Tabel 15. Data Hasil Oversampling SMOTE .....	34
Tabel 16. Perbandingan Data Sebelum dan Sesudah Penerapan SMOTE.....	35
Tabel 17. Data sebelum Oversampling-Undersampling SMOTE-NCL .....	36
Tabel 18. Data setelah Oversampling-Undersampling SMOTE-NCL .....	39
Tabel 19. Perbandingan Data Sebelum dan Sesudah Penerapan SMOTE-NCL ..	40
Tabel 20. Data untuk Perhitungan Klasifikasi SVM .....	41
Tabel 21. Matriks K Hasil Perhitungan Kernel <i>Polynomial</i> .....	44
Tabel 22. Hasil Prediksi .....	47
Tabel 23. Nilai Parameter Uji pada Klasifikasi SVM.....	47
Tabel 24. Performa Klasifikasi SVM.....	47
Tabel 25. Nilai Parameter Uji pada Klasifikasi SVM dengan SMOTE.....	48
Tabel 26. Performa Klasifikasi SVM dengan SMOTE.....	48
Tabel 27. Nilai Parameter Uji pada Klasifikasi SVM dengan SMOTE-NCL .....	49
Tabel 28. Performa Klasifikasi SVM dengan SMOTE-NCL .....	49

Tabel 29. Confusion Matrix Klasifikasi SVM .....	50
Tabel 30. Confusion Matrix Klasifikasi SVM dengan SMOTE .....	52
Tabel 31. Confusion Matrix Klasifikasi SVM dengan SMOTE-NCL.....	53
Tabel 32. Perbandingan Kinerja Klasifikasi SVM dan SMOTE-SVM .....	55
Tabel 33. Perbandingan Kinerja Klasifikasi SVM dan SMOTE-NCL-SVM .....	55
Tabel 34. Perbandingan Kinerja Klasifikasi SMOTE-SVM dan SMOTE-NCL-SVM .....	55
Tabel 35. Perbandingan Hasil Klasifikasi .....	57
Tabel 36. Perbandingan Nilai AUC .....	59

## DAFTAR GAMBAR

<b>Gambar</b>	<b>Halaman</b>
Gambar 1. Ovesampling SMOTE .....	15
Gambar 2. Kombinasi SMOTE-NCL .....	16
Gambar 3. Alur Penelitian.....	23
Gambar 4. Grafik Perbandingan Persentase Data .....	27
Gambar 5. Grafik Perbandingan Data Hasil SMOTE .....	35
Gambar 6. Grafik Perbandingan Data Hasil SMOTE-NCL.....	40
Gambar 7. Grafik Performa Klasifikasi SVM.....	48
Gambar 8. Grafik Performa Klasifikasi SVM dengan SMOTE .....	49
Gambar 9. Grafik Performa Klasifikasi SVM dengan SMOTE-NCL .....	50
Gambar 10. Perbandingan Hasil Klasifikasi .....	57
Gambar 11. Perbandingan Nilai AUC .....	59

## **DAFTAR LAMPIRAN**

Lampiran 1. Dataset Coronary Heart Disease

Lampiran 2. Dataset Setelah Label Encoding

Lampiran 3. Kode Program