

**PERBANDINGAN KINERJA ALGORITMA SUPPORT
VECTOR MACHINE DAN DECISION TREE
DENGAN PENDEKATAN UNDERSAMPLING
DALAM MEMPREDIKSI PENYAKIT
JANTUNG BERDASARKAN POLA
HIDUP**

LAPORAN NON-REGULER

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



Disusun oleh :

GUSTI AYU PUTU FEBRIYANTI

21.11.4326

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

**PERBANDINGAN KINERJA ALGORITMA SUPPOR VECTOR
MACHINE DAN DECISION TREE DENGAN PENDEKATAN
UNDERSAMPLING DALAM MEMPREDIKSI PENYAKIT
JANTUNG BERDASARKAN POLA HIDUP**

LAPORAN NON-REGULER

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



Disusun oleh :

GUSTI AYU PUTU FEBRIYANTI

21.11.4326

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

JALUR NON-REGULER

PERBANDINGAN KINERJA ALGORITMA SUPPORT VECTOR MACHINE
DAN DECISION TREE DENGAN PENDEKATAN UNDERSAMPLING
DALAM MEMPREDIKSI PENYAKIT JANTUNG
BERDASARKAN POLA HIDUP

yang disusun dan diajukan oleh

GUSTI AYU PUTU PERHYANTI

21.71.4324

telah disetujui oleh Dosen Pembimbing

pada tanggal 26 Januari 2025

Dosen Pembimbing,



Anna Baita, S.Kom., M.Kom
NIK. 190302290

HALAMAN PENGESAHAN

JALUR NON-REGULER

PERBANDINGAN KINERJA ALGORITMA SUPPORT VECTOR MACHINE
DAN DECISION TREE DENGAN PENDEKATAN UNDERSAMPLING
DALAM MEMPREDIKSI PENYAKIT JANTUNG
BERDASARKAN POLA HIDUP

yang disusun dan diajukan oleh

GUSTI AYU PUTU FEBRIYANTI
21.11.4326

Telah dipertahankan di depan Dewan Penguji
pada tanggal 30, Januari 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Muhammad Tofa N., S.Kom., M.Kom
NIK. 190302281

Arifiyanto Hadinegoro, S.Kom., M.T
NIK. 190302289

Anna Baita, S.Kom., M.Kom
NIK. 190302290



Laporan ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 30, Januari 2025

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : GUSTI AYU PUTU FEBRIYANTI

NIM : 21.11.4326

Menyatakan bahwa Laporan dengan judul berikut:

PERBANDINGAN KINERJA ALGORITMA SUPPORT VECTOR MACHINE DAN DECISION TREE DENGAN PENDEKATAN UNDERSAMPLING DALAM MEMPREDIKSI PENYAKIT JANTUNG BERDASARKAN POLA HIDUP

Dosen Pembimbing : Anna Baita, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan kegiatan SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidak-benaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 30 Januari 2025

Yang Menyatakan,



Gusti Ayu Putu Febriyanti

HALAMAN PERSEMBAHAN

~ Grateful hearts create boundless possibilities ~

Dengan penuh rasa syukur yang mendalam, penulis panjatkan kepada Tuhan Yang Maha Esa atas segala berkat, karunia, dan rahmat-Nya yang tak terhingga. Berkat kasih dan pertolongan-Nya yang senantiasa menyertai di setiap langkah, sehingga penulis mampu melalui segala rintangan dan menyelesaikan laporan non-skripsi ini hingga tuntas. Tiada kata yang dapat menggambarkan betapa besar rasa terima kasih ini atas kekuatan, ketabahan, dan petunjuk yang selalu diberikan oleh-Nya dalam setiap proses perjalanan hidup ini.

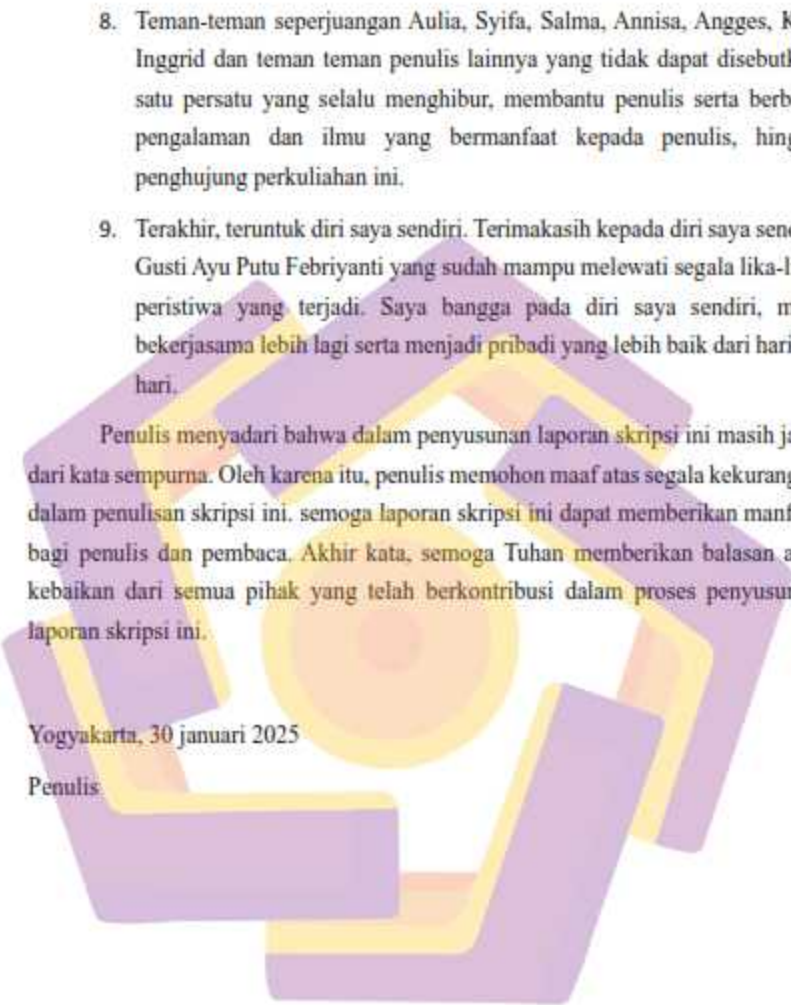
Karya ini penulis persembahkan sebagai tanda bukti cinta dan sayang yang luar biasa kepada kedua orangtua penulis, Ayah dan Mama, terimakasih banyak atas pengorbanan, perjuangan, kesabaran, dukungan, dan kasih sayang yang tulus telah diberikan selama ini dalam mendidik, mendoakan, mendukung dan memberikan semangat di setiap langkah perjalanan yang penulis lalui untuk mencapai mimpinya.

Karya ini juga penulis persembahkan kepada kedua adik penulis Asti dan Aditya yang senantiasa menghibur di saat penulis merasa sedih, dan selalu menjadi penyemangat terbaik. Serta tak lupa penulis ucapkan terimakasih banyak kepada diri sendiri, terimakasih telah bertahan sejauh ini dan tidak pernah berhenti berusaha dan berdoa untuk menyelesaikan penelitian ini.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Tuhan Yang Maha Esa atas Berkat dan Karunia-Nya sehingga penulis mampu menyelesaikan laporan skripsi dengan judul “Perbandingan Kinerja Algoritma Support Vector Machine dan Decision Tree dengan Pendekatan Undersampling dalam Memprediksi Penyakit Jantung Berdasarkan Pola Hidup”. laporan skripsi ini disusun untuk memenuhi syarat memperoleh gelar Sarjana Komputer Universitas Amikom Yogyakarta. selama proses penyusunan tugas akhir ini, penulis menyadari bahwa terdapat beberapa kendala, tetapi dengan adanya dukungan, bantuan, bimbingan, serta saran dari berbagai pihak, sehingga penulis mampu menyelesaikan laporan skripsi ini. Oleh karena itu, penulis menyampaikan terimakasih kepada:

1. Prof. Dr. M. Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, M.Kom., Ph.D. selaku Dekan Fakultas Ilmu Komputer.
3. Ibu Windha Mega Pradnya Duhita, M.Kom selaku Ketua Program Studi Informatika.
4. Ibu Anna Baita, M. Kom. selaku dosen pembimbing yang telah memberikan ilmu, bimbingan, nasihat, dukungan serta solusi kepada penulis selama proses penyusunan jurnal hingga laporan ini.
5. Seluruh bapak dan ibu dosen serta staff karyawan Universitas Amikom Yogyakarta yang telah memberikan ilmu dan bimbingan kepada penulis.
6. Kedua orangtua penulis, Ayahanda tercinta Gusti Komang Wirata dan Mama Gusti Ayu Made Ardiasih yang selalu memberikan doa, dukungan semangat, motivasi, hingga memberikan segala fasilitas keperluan yang diperlukan oleh penulis.
7. YARIS team (Aulia, Riza, Irvan dan Syifa) selaku teman-teman tim project selama perkuliahan yang selalu memberikan dukungan, bantuan dan motivasi hingga perkuliahan selesai.

- 
8. Teman-teman seperjuangan Aulia, Syifa, Salma, Annisa, Angges, Kak Ingrid dan teman teman penulis lainnya yang tidak dapat disebutkan satu persatu yang selalu menghibur, membantu penulis serta berbagi pengalaman dan ilmu yang bermanfaat kepada penulis, hingga penghujung perkuliahan ini.
 9. Terakhir, teruntuk diri saya sendiri. Terimakasih kepada diri saya sendiri Gusti Ayu Putu Febriyanti yang sudah mampu melewati segala lika-liku peristiwa yang terjadi. Saya bangga pada diri saya sendiri, mari bekerjasama lebih lagi serta menjadi pribadi yang lebih baik dari hari ke hari.

Penulis menyadari bahwa dalam penyusunan laporan skripsi ini masih jauh dari kata sempurna. Oleh karena itu, penulis memohon maaf atas segala kekurangan dalam penulisan skripsi ini. semoga laporan skripsi ini dapat memberikan manfaat bagi penulis dan pembaca. Akhir kata, semoga Tuhan memberikan balasan atas kebaikan dari semua pihak yang telah berkontribusi dalam proses penyusunan laporan skripsi ini.

Yogyakarta, 30 januari 2025

Penulis

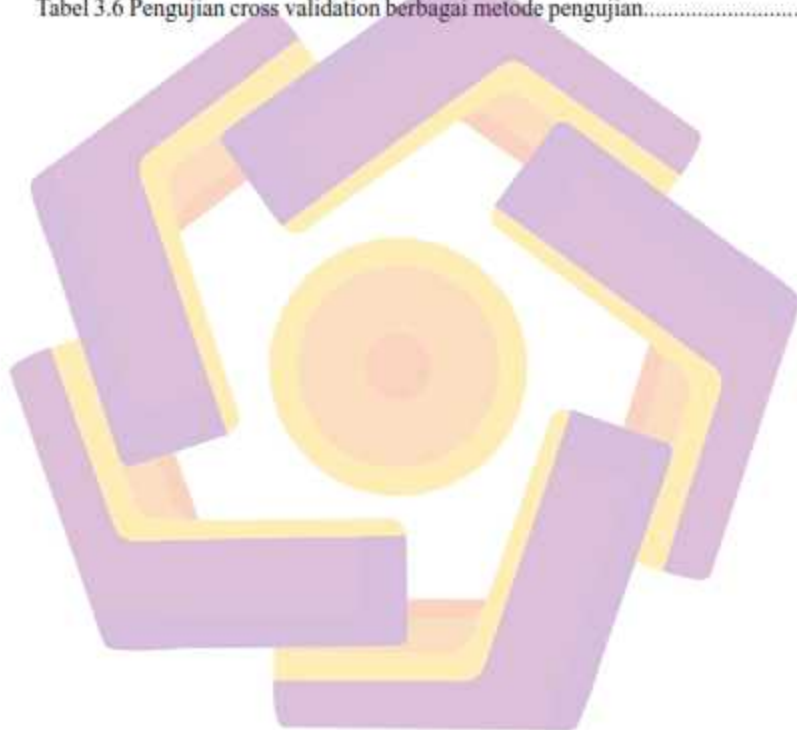
DAFTAR ISI

HALAMAN JUDUL.....	II
HALAMAN PERSETUJUAN.....	III
HALAMAN PENGESAHAN.....	IV
HALAMAN PERNYATAAN KEASLIAN KARYA.....	V
HALAMAN PERSEMBAHAN.....	VI
KATA PENGANTAR.....	VII
DAFTAR ISI.....	IX
DAFTAR TABEL.....	XI
DAFTAR GAMBAR.....	XII
DAFTAR LAMPIRAN.....	XIII
DAFTAR LAMBANG DAN SINGKATAN.....	XIV
DAFTAR ISTILAH.....	XV
INTISARI.....	XVI
ABSTRACT.....	XVII
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah.....	3
1.4 Tujuan.....	3
BAB II TINJAUAN PUSTAKA.....	4
2.1. Studi Literatur.....	4
2.2. Landasan Teori.....	7
2.2.1. Preprocessing Data.....	7
2.2.2 Undersampling.....	7
2.2.3 Support Vector Machine.....	8
2.2.4 Decision Tree.....	8

2.2.5 Hyperparameter Tuning.....	8
2.2.6 Cross Validation.....	10
BAB III METODE PENELITIAN.....	11
3.1. Metode.....	11
3.2 Hasil dan Pembahasan.....	12
3.2.1 Data Collecting.....	12
3.2.2 Preprocessing Data.....	13
3.2.3 Undersampling.....	14
3.2.4 Modeling.....	15
3.2.5 Hyperparameter Tuning dan K-fold.....	18
BAB IV KESIMPULAN.....	24
4.1. Kesimpulan.....	24
4.2. Saran.....	25
REFERENSI.....	26
CURICULUM VITAE.....	30
LAMPIRAN DAN BUKTI PENDUKUNG.....	31

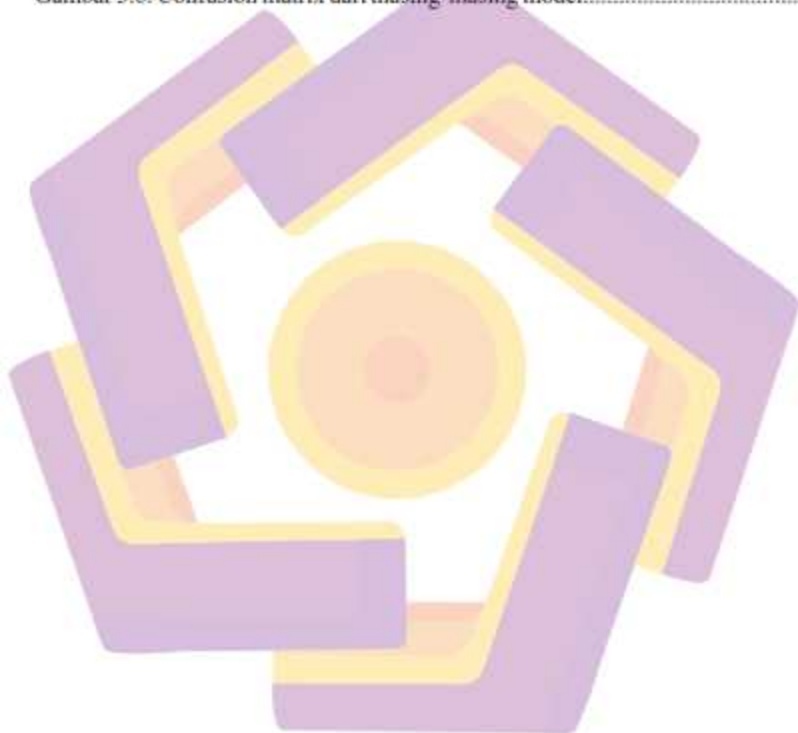
DAFTAR TABEL

Tabel 3.1 Fitur pada dataset yang digunakan.....	13
Tabel 3.2 Jumlah data undersampling.....	15
Tabel 3.3 Perbandingan nilai akurasi sebelum dan setelah di terapkan undersampling pada SVM dan DT.....	16
Tabel 3.4 Parameter yang digunakan pada hyperparameter tuning.....	18
Tabel 3.5 Best parameter tuning dengan grid search.....	19
Tabel 3.6 Pengujian cross validation berbagai metode pengujian.....	21



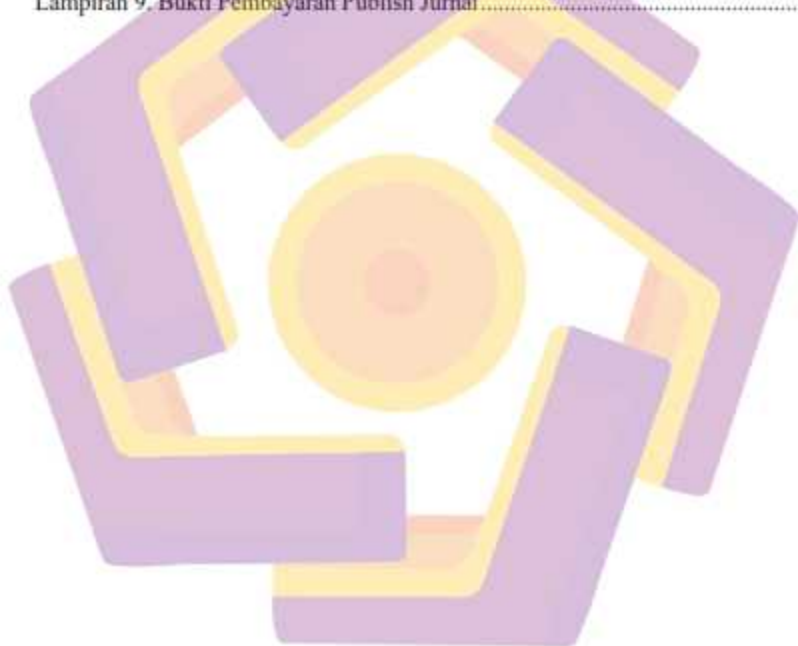
DAFTAR GAMBAR

Gambar 3.1 Tahapan penelitian.....	11
Gambar 3.2 Bar chart distribusi class pasien.....	12
Gambar 3.3 Jumlah data undersampling.....	15
Gambar 3.4 Grafik persentase precision, recall dan f1-score SVM.....	17
Gambar 3.5 Grafik persentase precision, recall, dan F1-score DT.....	18
Gambar 3.6. Confusion matrix dari masing-masing model.....	22



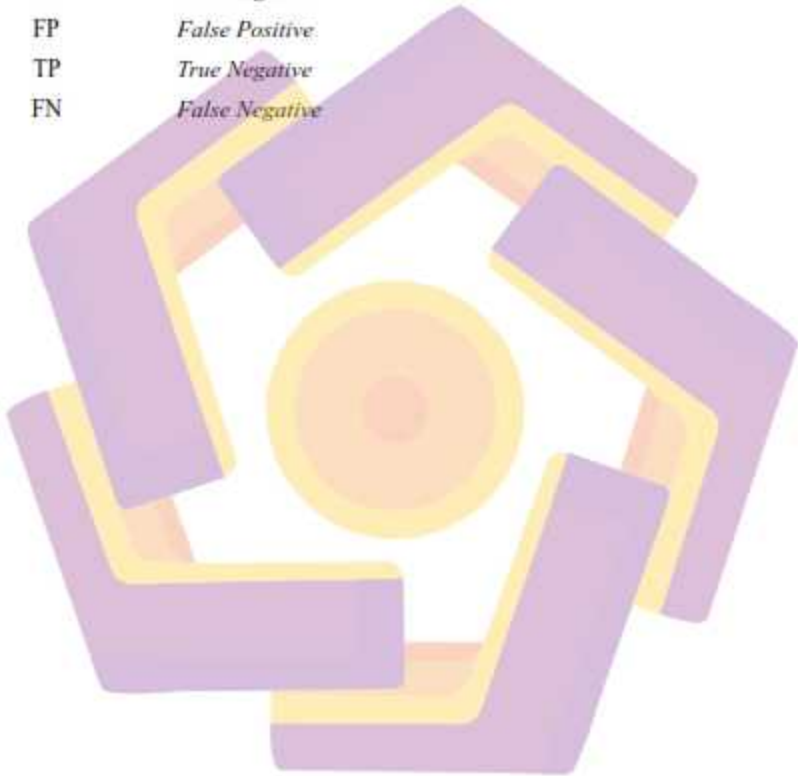
DAFTAR LAMPIRAN

Lampiran 1. LoA dari Journal of Applied Informatics and Computing	31
Lampiran 2. Review dari reviewer A.....	32
Lampiran 3. Review dari reviewer B	32
Lampiran 4. Balasan penulis serta revisi yang telah dilakukan sesuai review dari kedua reviewer	33
Lampiran 5. Bukti artikel telah di acc	34
Lampiran 6. Bukti artikel telah masuk ke terbitan jurnal yang akan datang.....	34
Lampiran 7. Sertifikat akreditasi JAIC	35
Lampiran 8. Index JAIC.....	35
Lampiran 9. Bukti Pembayaran Publish Jurnal.....	36



DAFTAR LAMBANG DAN SINGKATAN

SVM	<i>Support Vector Machine</i>
DT	<i>Decision Tree</i>
CV	<i>Cross Validation</i>
TN	<i>True Negative</i>
FP	<i>False Positive</i>
TP	<i>True Negative</i>
FN	<i>False Negative</i>



DAFTAR ISTILAH

K-fold	Teknik untuk mengevaluasi performa model pembelajaran mesin
Hyperparameter tuning	Proses mencari kombinasi nilai yang optimal untuk parameter-parameter tertentu yang tidak diajarkan oleh model selama proses training
Akurasi	Nilai hasil untuk menentukan keakuratan model.
Confusion Matrix	Tabel yang menggambarkan kinerja sample klasifikasi yang digunakan untuk mengukur kinerja pada kumpulan data.
Precision	Tingkat probabilitas dengan hasil positif dan ditentukan dengan benar
Recall	Tingkat probabilitas dengan hasil yang positif dan ditentukan dengan benar.
F1-Score	Nilai hasil untuk mengetahui nilai rata-rata dari perbandingan keluaran precision maupun keluaran recall.
True Negative (TN)	Prediksi hasil yang bernilai negatif, data yang sebenarnya negatif
False Positive (FP)	Prediksi hasil yang bernilai positif, data yang sebenarnya negatif.
False Negative (FN)	Prediksi hasil yang bernilai negatif, data yang sebenarnya positif
True Positive (TP)	Prediksi hasil yang bernilai positif, data sebenarnya yang positif
Grid Search	Proses pengoptimalan nilai akurasi dengan menemukan parameter terbaik

INTISARI

Penyakit jantung merupakan salah satu penyebab utama kematian di dunia dengan faktor risiko seperti aterosklerosis, tekanan darah tinggi, dan merokok. Diagnosis dini penting untuk mengurangi angka kematian dan meningkatkan kualitas hidup pasien. Studi ini mengevaluasi kinerja algoritma Support Vector Machine (SVM) dan Decision Tree (DT) dalam memprediksi risiko penyakit jantung menggunakan teknik undersampling untuk menangani ketidakseimbangan data. Validasi silang k-fold dengan $K=10$ dan hyperparameter tuning diterapkan untuk memperoleh model yang optimal. Hasil menunjukkan bahwa SVM tanpa undersampling memiliki akurasi 92%, sedangkan dengan undersampling menurun menjadi 76%. DT tanpa undersampling memiliki akurasi 91%, dan dengan undersampling menurun menjadi 75%. Teknik undersampling berhasil meningkatkan keseimbangan dalam mengenali kelas minoritas tetapi menurunkan akurasi keseluruhan. Dalam hal ini SVM terbukti lebih andal 1% dalam memprediksi penyakit jantung baik pada dataset tidak seimbang maupun pada dataset seimbang. Namun, penelitian ini memiliki beberapa keterbatasan, termasuk penurunan akurasi akibat undersampling dan performa DT yang lebih rendah karena sensitivitas terhadap overfitting pada data kecil (setelah undersampling). Penelitian ini juga belum mengeksplorasi seleksi fitur yang dapat meningkatkan performa model serta hanya menggunakan satu dataset, sehingga generalisasi hasil ke dataset lain belum teruji. Studi lanjutan disarankan untuk menggunakan teknik balancing alternatif seperti SMOTE, model yang lebih kompleks, dan evaluasi pada dataset berbeda untuk meningkatkan validitas serta aplikasi model. Model pada penelitian ini diharapkan dapat digunakan sebagai sistem prediksi dini penyakit jantung, namun karena dataset yang digunakan dalam penelitian ini adalah dataset sekunder, sehingga hasil prediksi perlu dikombinasikan dengan pemeriksaan lebih lanjut oleh tenaga medis untuk hasil yang lebih akurat.

Kata kunci: DT, Jantung, K-fold, Pencegahan, SVM.

ABSTRACT

Heart disease is one of the leading causes of death in the world with risk factors such as atherosclerosis, high blood pressure, and smoking. Early diagnosis is important to reduce mortality and improve patients' quality of life. This study evaluates the performance of Support Vector Machine (SVM) and Decision Tree (DT) algorithms in predicting heart disease risk using undersampling techniques to handle data imbalance. K-fold cross validation with $K=10$ and hyperparameter tuning were applied to obtain the optimal model. Results show that SVM without undersampling has an accuracy of 92%, while with undersampling it decreases to 76%. DT without undersampling has 91% accuracy, and with undersampling it decreases to 75%. The undersampling technique successfully improved the balance in recognizing minority classes but decreased the overall accuracy. In this case, SVM proved to be 1% more reliable in predicting heart disease in both unbalanced and balanced datasets. However, this study has some limitations, including decreased accuracy due to undersampling and lower DT performance due to sensitivity to overfitting on small data (after undersampling). This study also did not explore feature selection that could improve model performance and only used one dataset, so the generalizability of the results to other datasets has not been tested. Further studies are recommended to use alternative balancing techniques such as SMOTE, more complex models, and evaluation on different datasets to improve the validity and application of the model. The model in this study is expected to be used as an early prediction system for heart disease, but because the dataset used in this study is a secondary dataset, the prediction results need to be combined with further examination by medical personnel for more accurate results.

Keyword: Cardiac, DT, K-fold, Prevention, SVM.