



Komparasi Algoritma Menggunakan Random Forest dan Support Vector Machine Pada Dataset Kematian Akibat Gagal Jantung

LAPORAN TUGAS AKHIR

Dareff Danny Oesman	41820010043
Khalil Zufar	41820010074
Wicaksana Three Julius H	41820010079

**UNIVERSITAS
MERCU BUANA**

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2024**

HALAMAN PERNYATAAN KARYA SENDIRI

HALAMAN PENYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Dareff Danny Oesman
NIM : 41820010043
Program Studi : Sistem Informasi
Judul Laporan Magang/Skripsi/Tesis/Disertasi : Komparasi Algoritma Menggunakan Random Forest dan Support Vector Machine Pada Dataset Kematian Akibat Gagal Jantung

Menyatakan bahwa Laporan Magang/Skripsi/Tesis/Disertasi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Magang/Skripsi/Tesis/Disertasi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 25 Juni 2024

UNIVERSITAS
MERCU BUANA



Dareff Danny Oesman

HALAMAN PENGESAHAN

UNIVERSITAS MERCU BUANA

LEMBAR PENGESAHAN

Nama Mahasiswa (1) : Khalil Zufar
NIM (41820010074)
Nama Mahasiswa (2) : Dareff Danny Oesman
NIM (41820010043)
Nama Mahasiswa (3) : Wicaksana Julius Three Haryanto
NIM (41820010079)
Judul Tugas Akhir : Komparasi Algoritma Menggunakan Random Forest dan Support Vector Machine Pada Dataset Kematian Akibat Gagal Jantung

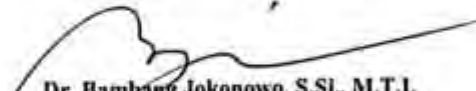
Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 25 - 06 - 2024

Menyetujui

Pembimbing : Yunita Sartika Sari, S.Kom, M.Kom ()
NIDN : 0309068903
Ketua Penguji : Dr. Bambang Jokonowo, S.Si., M.T.I. ()
NIDN : 0320037002
Penguji 1 : Sulis Sandiwarno, S.Kom, M.Kom ()
NIDN : 0302028803
Penguji 2 : Ifan Prihandi, S.Kom, M.Kom ()
NIDN : 0313098901

Mengetahui,


Dr. Bambang Jokonowo, S.Si., M.T.I.
Dekan Fakultas Ilmu Komputer


Dr. Ruci Meivang, M.Kom
Ka.Prodi Sistem Informasi

KATA PENGANTAR

Puji syukur kita panjatkan Tuhan Yang Maha Esa atas rahmatnya kami dapat menyelesaikan penulisan yang berjudul “Komparasi Algoritma Menggunakan Random Forest dan Support Vector Machine Pada Dataset Kematian Akibat Gagal Jantung” dengan baik dan tepat waktu. Laporan ini dibuat untuk memenuhi Tugas Akhir di universitas Mercu Buana. Penulis menyadari bahwa tanpa bantuan dan bimbingan dari pihak yang terlibat kami sebagai penulis tidak dapat menyelesaikan dengan baik. Oleh karena itu, Kami mengucapkan terima kasih kepada:

1. Yunita Sartika Sari, S.Kom, M.Kom, selaku dosen Pembimbing Tugas Akhir yang telah menyediakan waktu, tenaga, dan pikiran untuk merahkan penulisan dalam penyusunan Tugas Akhir ini;
2. Dr. Ruci Meiyanti, M.Kom, selaku Kepala Program Studi Sistem Informasi Universitas Mercu Buana Jakarta;
3. Orang Tua dan Keluarga yang sudah mendukung penulis dari segi finansial maupun material.

Kami sadar penulisan ini masih jauh dari kata sempurna, kami mohon untuk memberikan kritik dan saran yang bersifat membangun demi kesempurnaan penulisan yang akan datang. Semoga penulisan ini dapat berguna bagi pembaca Kami sadar penulisan ini masih jauh dari kata sempurna, kami mohon untuk memberikan kritik dan saran yang bersifat membangun demi kesempurnaan penulisan yang akan datang. Semoga penulisan ini dapat berguna bagi pembaca

Jakarta, 25 Juni 2024

Dareff Danny Oesman

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Dareff Danny Oesman
NIM : 41820010043
Program Studi : Sistem Informasi
Judul Laporan Magang/Skripsi/Tesis/Disertasi : Komparasi Algoritma Menggunakan Random Forest dan Support Vector Machine Pada Dataset Kematian Akibat Gagal Jantung

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

MERCU BUANA

Jakarta, 25 Juni 2024

Yang menyatakan,



(Dareff Danny Oesman)

ABSTRAK

Nama : Dareff Danny Oesman
NIM : 41820010043
Program Studi : Sistem Informasi
Judul Laporan Skripsi : Komparasi Algoritma Menggunakan Random Forest dan Support Vector Machine Pada Dataset Kematian Akibat Gagal Jantung
Pembimbing : Yunita Sartika Sari, S.Kom, M.Kom

Gagal jantung merupakan salah satu penyebab utama kematian dunia, dengan 17,9 juta kematian akibat penyakit kardiovaskular tercatat pada tahun 2019. Penelitian ini bertujuan untuk membandingkan keakuratan prediksi kematian akibat gagal jantung antara dua algoritma machine learning, yaitu Random Forest dan Support Vector Machine (SVM). Penelitian ini diharapkan dapat memberikan kontribusi yang dapat membantu para profesional medis dalam pengambilan keputusan yang lebih baik di dunia medis. Penelitian ini menggunakan dataset yang berasal dari platform Kaggle, yang berisi data tentang pasien dengan risiko gagal jantung, dan data tersebut diolah melalui beberapa tahap seperti pre-processing diantaranya, data cleaning, data scaling dan pengurangan data. Algoritma Random Forest dan SVM diterapkan pada data yang diolah dan kinerjanya dievaluasi menggunakan Confusion Matrix untuk mengukur akurasi, presisi, recall dan F1 Score. Hasil penelitian menunjukkan bahwa algoritma Random Forest memiliki tingkat akurasi yang lebih tinggi daripada SVM.

Kata Kunci: Gagal jantung, Machine Learning, Prediksi Kematian

ABSTRACT

Name : Dareff Danny Oesman
NIM : 41820010043
Study Program : Sistem Informasi
Title Thesis : Comparison of Algorithms Using Random Forest and Support
Vector Machine on the Death Due to Heart Failure Dataset
Counsellor : Yunita Sartika Sari, S.Kom, M.Kom

Heart failure is one of the main causes of death in the world, with 17.9 million deaths due to cardiovascular disease recorded in 2019. This study aims to compare the accuracy of predicting deaths from heart failure between two machine learning algorithms, namely Random Forest and Support Vector Machine (SVM). It is hoped that this research can provide contributions that can help medical professionals make better decisions in the medical world. This research uses a dataset originating from the Kaggle platform, which contains data about patients at risk of heart failure, and this data is processed through several stages such as pre-processing including, data cleaning, data scaling and data reduction. The Random Forest and SVM algorithms are applied to the processed data and their performance is evaluated using the Confusion Matrix to measure accuracy, precision, recall and F1 Score. The research results show that the Random Forest algorithm has a higher level of accuracy than SVM.

Keyword: Heart Failure, Machine Learning, Death Prediction

DAFTAR ISI

JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN	iii
KATA PENGANTAR	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	v
ABSTRAK	v
ABSTRACT	vii
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	14
1.1 Latar Belakang	14
1.2 Rumusan Masalah	15
1.3 Tujuan	15
1.4 Batasan Masalah	15
1.5 Manfaat Penelitian	15
1.6 Sistematika Penulisan	15
BAB II TINJAUAN PUSTAKA	17
2.1 Teori/Konsep Terkait	17
2.1.1 Kardiovaskular	17
2.1.2 Data Mining	17
2.1.3 Pre-Processing	18
2.1.4 Random Forest	18
2.1.5 <i>Support Vector Machine</i>	19
2.1.6 <i>Confusion Matrix</i>	20
2.1.7 <i>Google Colaboratory</i>	22
2.2 Penelitian Terdahulu	23
2.2.1 Literature Review	23
2.3 Analisis Literature Review	42
BAB III METODE PENELITIAN	43
3.1 Deskripsi Sumber Data	43
3.2 Teknik Pengumpulan Data	43
3.3 Diagram Alir Penelitian	44
3.3.1 Literatur Review	44
3.3.2 Menentukan Judul	44
3.3.3 Identifikasi Masalah	45
3.3.4 Pengumpulan Data	45
3.3.5 Preprocessing Data/ Data Cleaning	45

3.3.6 Data Scaling	45
3.3.7 Data Reduction	46
3.3.8 Support Vector Machine	46
3.3.9 Random Forest	46
3.3.10 Confusion Matrix	46
BAB IV HASIL DAN PEMBAHASAN	47
4.1 Hasil Pengujian	47
4.1.1 Data Understanding	47
4.1.2 Data Preprocessing	49
4.1.3 Data Processing	52
4.1.4 Encoding	53
4.1.5 Tools yang digunakan	64
BAB V PENUTUP	65
5.1. Kesimpulan	65
5.2. Saran	65
DAFTAR PUSTAKA	66
LAMPIRAN	70
Lampiran Kartu Asistensi Tugas Akhir	70
Curriculum Vitae	71
Lampiran BNSP	72
Lampiran Form Revisi	73
Lampiran Formulir Verifikasi Data	75



DAFTAR TABEL

Tabel 1 Confusion Matrix 20



DAFTAR GAMBAR

Gambar 1. Random Tree	19
Gambar 2. Support Vector Machine.....	20
Gambar 3. Diagram Alir Penelitian.....	44
Gambar 4. Import Package.....	47
Gambar 4.1. Read Data	47
Gambar 4.2. Informasi Data	48
Gambar 4.3. Statistik Deskriptif.....	48
Gambar 4.4. Null Detection.....	49
Gambar 4.4.1. Memeriksa Outliers.....	50
Gambar 4.4.2. Mendefinisikan Outliers	50
Gambar 4.4.3. Baris Outliers	51
Gambar 4.4.4. Removing Outliers.....	51
Gambar 4.5. Heatmap Correlation plot Of the heart failure prediction.....	52
Gambar 4.6. Dividing Into Numerical and Categorical.....	53
Gambar 4.7. Encoding Catagorical Features.....	53
Gambar 4.8. Coding Visualisasi categorical.....	54
Gambar.4.9. Visualisasi Categorical	55
Gambar 4.9.2. Coding Numerical Features.....	55
Gambar 4.9.3. Visualisasi Numerical.....	56
Gambar 4.9.4. Visualiasi HeartDisease.....	57
Gambar 4.9.5. Categorical features vs target variable.....	58
Gambar 4.9.6. Selection Categorical Features	59
Gambar 4.9.7. Selection Numerical Features.....	60
Gambar 5.0. Coding Training And Prediciton	61

UNIVERSITAS
MERCU BUANA

DAFTAR LAMPIRAN

LAMPIRAN.....	70
Lampiran Kartu Asistensi Tugas Akhir.....	70
Curriculum Vitae.....	71
Lampiran BNSP.....	72
Lampiran Form Revisi.....	73
Lampiran Formulir Verifikasi Data.....	75



UNIVERSITAS
MERCU BUANA