

**KOMPARASI OPTIMASI ALGORITMA K-NEAREST NEIGHBOR
DAN SUPPORT VECTOR MACHINE DENGAN GRID SEARCH CV
PADA PREDIKSI STROKE**

JALUR SCIENTIST

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

WAHYU APRILLIANDHIKA

21.11.3869

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

**KOMPARASI OPTIMASI ALGORITMA K-NEAREST NEIGHBOR
DAN SUPPORT VECTOR MACHINE DENGAN GRID SEARCH CV
PADA PREDIKSI STROKE**

JALUR SCIENTIST

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



disusun oleh

WAHYU APRILLIANDHIKA

21.11.3869

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

HALAMAN PERSETUJUAN

JALUR SCIENTIST

**KOMPARASI OPTIMASI ALGORITMA K-NEAREST NEIGHBOR DAN
SUPPORT VECTOR MACHINE DENGAN GRID SEARCH CV PADA
PREDIKSI STROKE**

yang disusun dan diajukan oleh

Wahyu Aprilliandhika

21.11.3869

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 16 Oktober 2024

Dosen Pembimbing,



Ferian Fauzi Abdulloh, M.Kom

NIK. 190302276

HALAMAN PENGESAHAN

JALUR SCIENTIST

KOMPARASI OPTIMASI ALGORITMA K-NEAREST NEIGHBOR DAN
SUPPORT VECTOR MACHINE DENGAN GRID SEARCH CV PADA
PREDIKSI STROKE

yang disusun dan diajukan oleh

Wahyu Aprilliandhika

21.11.3869

Telah dipertahankan di depan Dewan Penguji
pada tanggal 16 Oktober 2024

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Theopilus Bayu Sasongko, S.Kom., M.Eng
NIK. 190302375

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

Ferian Fauzi Abdulloh, M.Kom
NIK. 190302276



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 16 Oktober 2024

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Wahyu Aprilliandhika
NIM : 21.11.3869

Menyatakan bahwa Skripsi dengan judul berikut:

KOMPARASI OPTIMASI ALGORITMA K-NEAREST NEIGHBOR DAN SUPPORT VECTOR MACHINE DENGAN GRID SEARCH CV PADA PREDIKSI STROKE

Dosen Pembimbing : Ferian Fauzi Abdulloh, 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 penelitian 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 ketidakbenaran 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, 16 Oktober 2024

Yang Menyatakan,



Wahyu Aprilliandhika

HALAMAN PERSEMBAHAN

Dengan mengucapkan rasa syukur kepada Allah SWT, saya mempersembahkan laporan non-skripsi jalur scientist ini kepada:

1. Kedua orang tua saya yaitu, ayah Suriono dan bunda Sugiaty, yang telah memberikan cinta, doa, dukungan, serta pengorbanan tanpa henti sepanjang perjalanan hidup saya. Terima kasih atas kasih sayang yang tulus dan tanpa batas, yang senantiasa menjadi cahaya dalam setiap langkah saya.
2. Kakek dan nenek saya yaitu, Kakek Sardi Sakat Wongso dan Nenek Wagisah yang selalu memberikan cinta, dan dukungan. Kehadiran kalian selalu menjadi kekuatan dalam hidup saya.
3. Kakek Alm. Ngadimin dan Nenek Alm. Suratik meski kalian telah berpulang, warisan cinta, doa, serta nilai-nilai luhur yang kalian tanamkan tetap mengalir dalam hidup saya. Kehadiran kalian tetap terus saya rasakan, memberi dorongan dan semangat dalam setiap langkah yang saya tempuh.
4. Adik-adik saya, Sakti Adnan Magani dan Zaky Fadhil Alfarizil, yang selalu memberikan semangat dan kebersamaan, serta menjadi bagian dari perjalanan hidup saya.
5. Bapak Ferian Fauzi Abdulloh, M.Kom, dosen pembimbing saya, yang dengan kesabaran, arahan, dan membimbing saya dengan baik.
6. Seluruh keluarga besar, yang selalu memberikan dukungan moral, dorongan, serta semangat tanpa henti, yang memungkinkan saya untuk menyelesaikan laporan ini dengan baik.
7. Teman-teman seperjuangan yang selalu hadir dengan dukungan, tawa, dan kebersamaan sepanjang masa kuliah.
8. Almamater tercinta, Universitas Amikom Yogyakarta, tempat di mana saya menimba ilmu dan mengembangkan diri selama ini.

KATA PENGANTAR

Segala puji syukur kehadirat Allah SWT atas segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan laporan non-skripsi jalur scientist ini dengan judul "Komparasi Optimasi Algoritma K-Nearest Neighbor dan Support Vector Machine dengan Grid Search CV Pada Prediksi Stroke". Laporan non-skripsi jalur scientist ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Sarjana (S1) Informatika Fakultas Ilmu Komputer di Universitas Amikom Yogyakarta.

Selama proses penyusunan laporan ini, penulis mendapatkan banyak bantuan, arahan, doa, dan bimbingan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Bapak Ferian Fauzi Abdulloh, M.Kom, selaku dosen pembimbing, yang dengan sabar memberikan arahan, bimbingan, dan waktu dalam penyelesaian laporan non-skripsi jalur scientist ini.
2. Ayah Suriono dan Bunda Sugianti, yang telah memberikan cinta, doa, dukungan, serta pengorbanan tanpa henti sepanjang hidup penulis.
3. Adik Sakti Adnan Magani dan Zaky Fadhil Alfarizil, yang selalu memberikan semangat dan keceriaan dalam setiap kesempatan kepada penulis.
4. Teman-teman seperjuangan di Universitas Amikom Yogyakarta, yang telah menjadi pendukung, dan partner dalam menghadapi setiap tantangan akademik.
5. Seluruh dosen Universitas Amikom Yogyakarta, yang telah berbagi ilmu, pengalaman, dan nasihat selama masa perkuliahan.
6. Seluruh keluarga besar yang telah memberi dukungan penuh dan doa, sehingga penulis dapat menyelesaikan laporan ini.

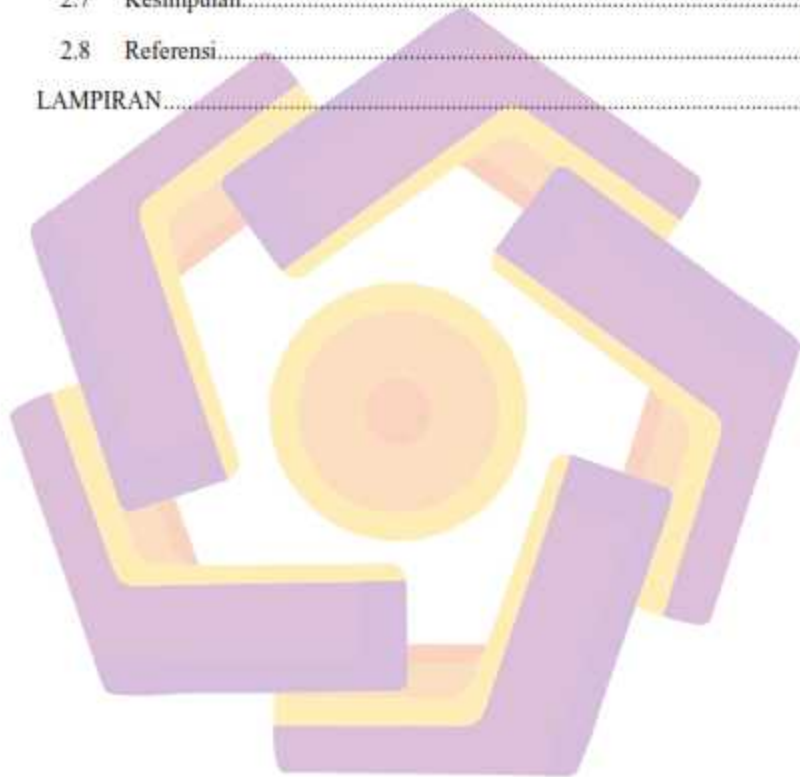
Yogyakarta, 10 Oktober 2024

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN.....	xi
INTISARI	xii
ABSTRACT	xiii
BAB I IDENTITAS JURNAL.....	1
BAB II ISI KARYA ILMIAH.....	2
2.1 Abstrack	2
2.2 Intisari.....	3
2.3 Pendahuluan	3
2.4 Metode Penelitian.....	6
2.4.1 Preprocessing	7
2.4.2 Model	9
2.4.3 Evaluasi.....	10
2.5 Hasil dan Pembahasan.....	12
2.5.1 Preprocessing Data.....	12

2.5.2	Hasil Pelatihan dan Pengujian Model	15
2.5.3	Hasil optimasi model dengan GridSearchCV	17
2.5.4	Hasil perbandingan model KNN dan SVM	19
2.6	Diskusi.....	23
2.7	Kesimpulan.....	24
2.8	Referensi.....	25
LAMPIRAN	30



DAFTAR TABEL

Tabel 1.1 Identitas Publikasi	1
Tabel 2.1 Dataset Stroke	8
Tabel 2.2 Confussion Matrix	11
Tabel 2.3 Jumlah Data Stroke	12
Tabel 2.4 Data Bernilai Unknown	12
Tabel 2.5 Kolom yang Bertipe Data Binary dan Kategorikal	13
Tabel 2.6 Data X_{train}	16
Tabel 2.7 Data y_{train}	16
Tabel 2.8 Data X_{test}	16
Tabel 2.9 Data y_{test}	17
Tabel 2.10 Hasil knn dan svm sebelum optimasi Grid Search	17
Tabel 2.11 Hasil Optimasi Menggunakan Grid Search CV	18
Tabel 2.12 Perbandingan Hasil Optimasi	22

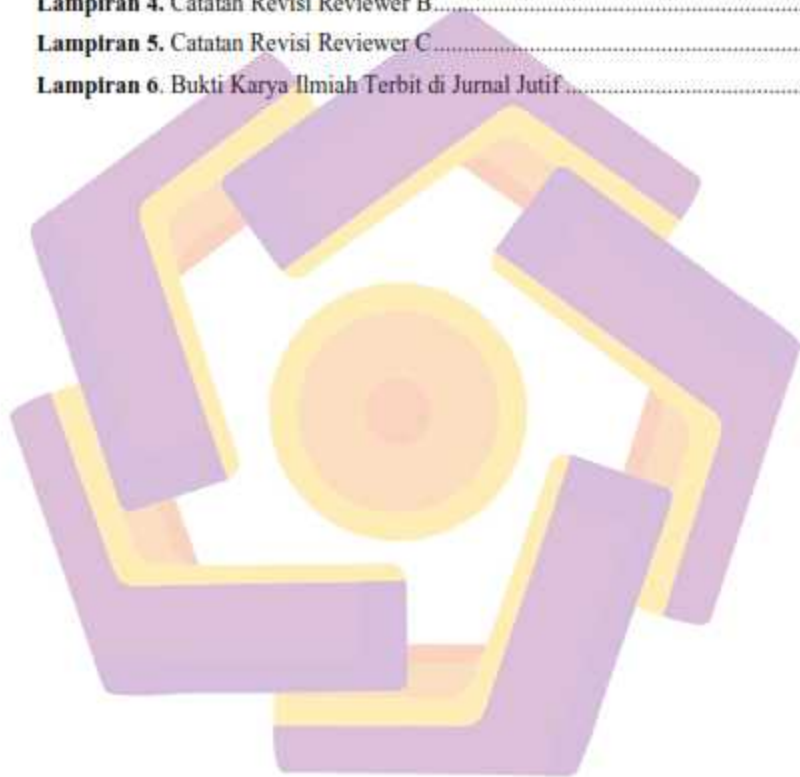
DAFTAR GAMBAR

Gambar 2.1 Diagram Penelitian	7
Gambar 2.2 Sebelum SMOTE.....	14
Gambar 2.3 Setelah SMOTE.....	15
Gambar 2.4 Hasil Prediksi Model KNN Sebelum Optimasi.....	19
Gambar 2.5 Hasil Prediksi Model KNN Setelah Optimasi.....	20
Gambar 2.6 Hasil Prediksi Model SVM Sebelum Optimasi.....	20
Gambar 2.7 Hasil Prediksi Model SVM Setelah Optimasi.....	21
Gambar 2.8 Perbandingan Hasil Sebelum dan Sesudah Optimasi	21



DAFTAR LAMPIRAN

Lampiran 1. LOA Jurnal Jutif.....	30
Lampiran 2. Lembar Review.....	31
Lampiran 3. Catatan Reviewer A.....	31
Lampiran 4. Catatan Revisi Reviewer B.....	32
Lampiran 5. Catatan Revisi Reviewer C.....	33
Lampiran 6. Bukti Karya Ilmiah Terbit di Jurnal Jutif.....	33



INTISARI

Stroke menempati peringkat kedua sebagai penyebab kematian global terbanyak, dengan kecacatan sebagai penyebab utama yang menyertai kondisi tersebut. Penyebab kematian pada penderita stroke disebabkan tidak ada sistem yang memprediksi stroke secara optimal, identifikasi apakah seorang pasien mengalami stroke atau tidak menjadi fokus penelitian ini. Maka, tujuan dari penelitian ini membandingkan kinerja prediksi stroke menggunakan dua model klasifikasi yang digunakan, yaitu K-Nearest Neighbors (KNN) dan Support Vector Machine (SVM), dengan dan tanpa menggunakan teknik optimasi GridSearchCV. Dalam eksperimen ini, dataset diproses dan dibagi menjadi data pelatihan dan pengujian dengan menggunakan teknik oversampling SMOTE. Pengujian awal dilakukan tanpa GridSearchCV. Hasil pengujian awal menunjukkan bahwa model KNN memiliki performa lebih baik daripada SVM, dengan akurasi 91% dan 83%. Setelah mengoptimalkan parameter menggunakan GridSearchCV, kedua model mengalami peningkatan signifikan dalam performa. Model KNN meningkatkan akurasi menjadi 95% dengan presisi 91% dan recall 98%, sementara model SVM meningkatkan akurasi menjadi 94% dengan presisi 90% dan recall 99%. Hasil penelitian ini menunjukkan bahwa penggunaan GridSearchCV dalam mengoptimalkan parameter model KNN dan SVM dapat meningkatkan kinerja prediksi stroke secara signifikan. Terdapat perbedaan dalam aspek presisi dan recall antara KNN dan SVM. Model KNN cenderung memiliki recall yang lebih tinggi, sementara model SVM memiliki presisi yang lebih tinggi, dan untuk akurasi algoritma KNN lebih unggul dibandingkan dengan SVM dalam prediksi Stroke.

Kata Kunci: GridSearchCV; Komparasi; K-Nearest Neighbor; Optimasi; Support Vector Machine

ABSTRACT

Stroke ranks second as the leading cause of death globally, with disability being the primary accompanying factor. The cause of death in stroke patients is due to the lack of an optimal stroke prediction system; therefore, identifying whether a patient is experiencing a stroke or not becomes the focus of this research. Thus, the objective of this study is to compare the performance of stroke prediction using two classification models, namely K-Nearest Neighbors (KNN) and Support Vector Machine (SVM), with and without using the GridSearchCV optimization technique. In this experiment, the dataset is processed and divided into training and testing data using the SMOTE oversampling technique. Initial testing is conducted without GridSearchCV. The results of the initial testing show that the KNN model performs better than SVM, with accuracies of 91% and 83%, respectively. After optimizing parameters using GridSearchCV, both models experience a significant performance improvement. The KNN model increases accuracy to 95% with precision of 91% and recall of 98%, while the SVM model increases accuracy to 94% with precision of 90% and recall of 99%. These results indicate that using GridSearchCV to optimize parameters of KNN and SVM models can significantly enhance stroke prediction performance. There are differences in precision and recall between KNN and SVM. The KNN model tends to have higher recall, while the SVM model has higher precision, and for accuracy, the KNN algorithm outperforms SVM in stroke prediction.

Keywords: GridSearchCV; Comparison; K-Nearest Neighbor; Optimization; Support Vector Machine