



Analisis Kinerja Recursive Feature Elimination pada Support Vector Machine untuk Klasifikasi Penyakit Stroke pada Data Tidak Seimbang

Faridatul Jannah, Siska Kurnia Gusti*, Elin Haerani, Teddie Darmizal

Fakultas Sains dan Teknologi, Prodi Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: ¹12250124499@students.uin-suska.ac.id, ^{2,*}siskakurniagusti@uin-suska.ac.id, ³elin.haerani@uin-suska.ac.id,

⁴teddie.darmizal@uin-suska.ac.id

Email Penulis Korespondensi: siskakurniagusti@uin-suska.ac.id

Abstrak—Stroke merupakan salah satu penyakit tidak menular dengan tingkat kematian dan kecacatan yang tinggi, sehingga diperlukan pendekatan klasifikasi yang mampu membantu proses deteksi secara lebih efektif. Ketidakseimbangan kelas pada dataset stroke menyebabkan model klasifikasi cenderung bias terhadap kelas mayoritas dan menghasilkan performa klasifikasi yang kurang optimal. Penelitian ini bertujuan untuk menganalisis kinerja *Recursive Feature Elimination* (RFE) pada model *Support Vector Machine* (SVM) dengan penanganan ketidakseimbangan data menggunakan *Adaptive Synthetic Sampling* (ADASYN) dalam klasifikasi penyakit stroke. Dataset yang digunakan merupakan dataset sekunder dari *Kaggle* yang terdiri dari 5109 data setelah tahap preprocessing. Proses pemodelan dilakukan melalui pengujian beberapa rasio pembagian data serta kombinasi kernel dan parameter SVM menggunakan pendekatan *5-fold cross validation*. Hasil penelitian menunjukkan bahwa model terbaik diperoleh pada rasio 80:20 dengan kernel polynomial dan parameter $C = 0.1$, yang menghasilkan akurasi sebesar 0.75, precision sebesar 0.14, recall sebesar 0.82, F1-score sebesar 0.24, serta nilai AUC sebesar 0.8245. Penerapan RFE menghasilkan peningkatan performa model dibandingkan tanpa RFE, meskipun besarnya peningkatan yang diperoleh relatif kecil. Nilai precision yang masih rendah menunjukkan bahwa model masih menghasilkan banyak prediksi positif palsu, sehingga tantangan klasifikasi pada dataset stroke belum sepenuhnya teratasi. Di sisi lain, nilai AUC sebesar 0.8245 menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam membedakan kedua kelas secara keseluruhan, meskipun penerapannya dalam konteks klinis masih memerlukan penyempurnaan lebih lanjut.

Kata Kunci: ADASYN; Imbalanced data; RFE; Stroke; SVM

Abstract—Stroke is a non-communicable disease with high mortality and disability rates, necessitating a classification approach that can facilitate more effective detection. Class imbalance in stroke datasets causes classification models to be biased toward the majority class, resulting in suboptimal classification performance. This study aims to analyze the performance of Recursive Feature Elimination (RFE) in a Support Vector Machine (SVM) model with data imbalance handling using Adaptive Synthetic Sampling (ADASYN) in stroke classification. The dataset used is a secondary dataset from Kaggle consisting of 5109 data points after the preprocessing stage. The modeling process was conducted by testing various data split ratios as well as combinations of kernels and SVM parameters using a 5-fold cross-validation approach. The results show that the best model was obtained with an 80:20 split ratio, a polynomial kernel, and a C parameter of 0.1, yielding an accuracy of 0.75, precision of 0.14, recall of 0.82, an F1-score of 0.24, and an AUC of 0.8245. The application of RFE resulted in improved model performance compared to without RFE, although the magnitude of the improvement was relatively small. The still low precision value indicates that the model still produces many false positives, so the classification challenge on the stroke dataset has not been fully resolved. On the other hand, an AUC value of 0.8245 indicates that the model performs reasonably well in distinguishing between the two classes overall, although its application in a clinical context still requires further refinement.

Keywords: ADASYN; Imbalanced Data; RFE; Stroke; SVM

1. PENDAHULUAN

Beban kesehatan akibat stroke masih tergolong tinggi karena penyakit ini berkontribusi besar terhadap angka kematian dan disabilitas di berbagai negara. Kondisi ini terjadi ketika aliran darah menuju otak terganggu akibat penyumbatan atau pecahnya pembuluh darah, sehingga jaringan otak mengalami kekurangan oksigen dan nutrisi. Kondisi tersebut dapat menyebabkan gangguan fungsi tubuh seperti gangguan motorik, kesulitan berbicara, hingga kematian [1]. Berdasarkan *World Stroke Organization Global Stroke Fact Sheet 2025*, prevalensi stroke global mencapai hampir 94 juta kasus orang yang hidup dengan riwayat stroke, dengan lebih dari 7 juta kematian setiap tahun. Antara tahun 1990 hingga 2021, jumlah kejadian stroke meningkat hingga 70%, kematian sebesar 44%, dan disabilitas meningkat 32%, di mana sebagian besar kasus tersebut ditemukan pada negara dengan tingkat pendapatan rendah dan menengah [2].

Kondisi serupa juga terjadi di Indonesia, dimana stroke masih termasuk penyakit dengan dampak kesehatan yang signifikan. Hasil Riset Kesehatan Dasar (Riskesdas) 2018 menunjukkan bahwa prevalensi stroke berdasarkan diagnosis dokter pada penduduk usia ≥ 15 tahun mencapai 10,9 per 1.000 penduduk, meningkat dibandingkan tahun 2013 yang berada pada angka 7 per 1.000 penduduk [3]. Selain itu, hasil Survei Kesehatan Indonesia (SKI) 2023 melaporkan bahwa prevalensi stroke nasional masih berada pada angka 8,3 per 1.000 penduduk [4]. Tingginya angka kejadian stroke menunjukkan perlunya upaya deteksi dini dan penanganan yang efektif. Namun, keterbatasan akses layanan kesehatan serta faktor sosial ekonomi sering menjadi hambatan dalam diagnosis dan penanganan yang tepat waktu [5]. Seiring dengan perkembangan teknologi, pemanfaatan *machine learning* menjadi salah satu pendekatan yang mulai banyak dikembangkan sebagai alternatif untuk membantu proses klasifikasi dan prediksi penyakit secara lebih efisien [6].

Sejumlah penelitian telah mencoba menerapkan pendekatan *machine learning* dalam klasifikasi penyakit stroke. Salah satunya adalah penelitian yang mengkombinasikan metode *Adaptive Synthetic Sampling* (ADASYN) dengan *Support Vector Machine* (SVM) untuk mengatasi ketidakseimbangan kelas pada dataset stroke. Hasilnya menunjukkan adanya peningkatan kemampuan model dalam mengenali kelas minoritas dengan akurasi sebesar 0.7266, namun nilai



precision dan F1-score masih relatif rendah, yang mengindikasikan adanya kesalahan klasifikasi yang cukup tinggi [7]. Rendahnya nilai precision dan F1-score menunjukkan bahwa model klasifikasi masih menghasilkan cukup banyak kesalahan prediksi pada kelas stroke. Selain dipengaruhi oleh ketidakseimbangan data, performa klasifikasi juga diduga dapat dipengaruhi oleh keberadaan fitur yang kurang relevan atau memiliki kontribusi yang rendah terhadap proses pembentukan model. Kondisi tersebut berpotensi meningkatkan kompleksitas data dan mempersulit model dalam menemukan pola yang benar-benar membedakan kelas stroke dan non-stroke. Berdasarkan kondisi tersebut, diperlukan mekanisme yang dapat menyaring atribut-atribut yang benar-benar memberikan informasi penting bagi pembentukan model klasifikasi [8].

Feature Selection merupakan proses pemilihan subset fitur yang paling relevan untuk meningkatkan kualitas dalam proses pemodelan *machine learning* [9]. Salah satu metode yang banyak digunakan adalah *Recursive Feature Elimination* (RFE), yang bekerja dengan mengeliminasi fitur secara bertahap berdasarkan kontribusinya terhadap model [10]. Berbagai penelitian menunjukkan bahwa penerapan RFE dapat memberikan dampak positif terhadap performa model klasifikasi. Namun, efektivitasnya sangat bergantung pada karakteristik dataset yang digunakan. Pada beberapa kasus, RFE mampu mengurangi jumlah fitur tanpa menurunkan performa, bahkan meningkatkan akurasi sekaligus efisiensi komputasi [11]. Seperti pada kasus *Gestational Diabetes Mellitus*, penerapan RFE menghasilkan akurasi sebesar 97,31% dan F1-score sebesar 96,6% [12]. Hasil serupa juga ditemukan pada klasifikasi kanker payudara menggunakan SVM, dengan penerapan RFE penelitian tersebut menghasilkan akurasi mencapai 97% [10].

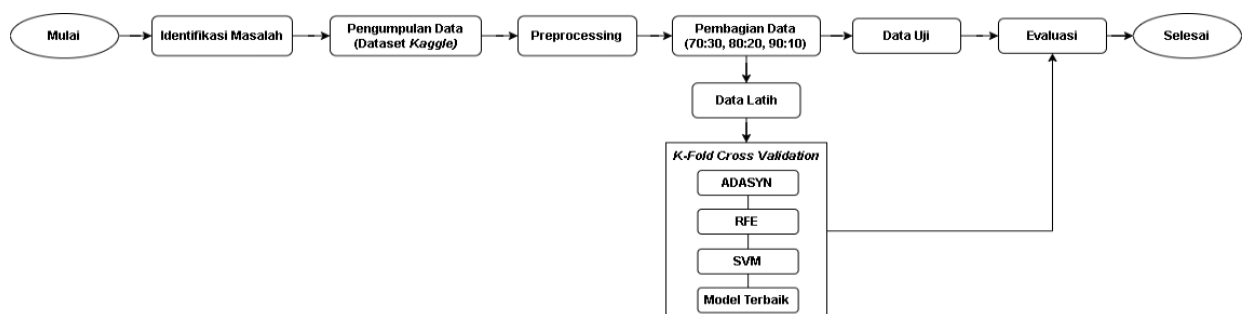
Efektivitas RFE juga telah diuji oleh penelitian pada data klinis yang mengkombinasikan teknik seleksi fitur dan penanganan ketidakseimbangan data. Pada klasifikasi penyakit dermatitis, penerapan RFE dan ADASYN mampu meningkatkan akurasi model secara signifikan dari 60,15% menjadi 90,89% [13]. Hasil yang konsisten juga ditunjukkan pada penelitian lain menggunakan algoritma Decision Tree, di mana akurasi meningkat dari 60,19% menjadi 92,44% setelah penerapan RFE [14]. Temuan tersebut menunjukkan bahwa RFE memiliki potensi yang kuat dalam meningkatkan kualitas model pada kondisi tertentu, khususnya ketika dataset mengandung fitur yang tidak relevan atau redundan. Selain itu, pendekatan hibrida SVM-RFE pada data biomedis yang kompleks juga menunjukkan kemampuan yang kuat dalam memilih fitur yang paling relevan untuk proses klasifikasi [15].

Di sisi lain, permasalahan ketidakseimbangan data juga menjadi tantangan utama dalam proses klasifikasi. Distribusi kelas yang tidak seimbang menyebabkan model lebih dominan mempelajari pola pada kelas mayoritas sehingga kemampuan dalam mengenali kelas minoritas menjadi menurun. Teknik oversampling seperti ADASYN telah banyak digunakan untuk mengatasi masalah ini dengan menghasilkan data sintesis secara adaptif pada kelas minoritas [16]. Sejumlah penelitian menunjukkan bahwa kombinasi ADASYN dan SVM mampu meningkatkan performa klasifikasi dibandingkan metode oversampling lainnya. Penelitian yang membandingkan teknik oversampling menunjukkan bahwa kombinasi ADASYN-SVM memberikan hasil yang lebih baik dibandingkan beberapa metode oversampling lain pada deteksi diabetes tipe 2 [17]. Lebih lanjut, kombinasi antara teknik *feature selection* dan penanganan ketidakseimbangan data terbukti memberikan hasil yang lebih optimal. Penerapan ADASYN dan RFE secara bersamaan mampu meningkatkan akurasi hingga 98,59% melalui penyaringan fitur yang relevan dan pengurangan kompleksitas data [18]. Hasil serupa juga ditunjukkan pada penelitian klasifikasi penyakit tiroid, di mana kombinasi ADASYN, RFE, dan SVM efektif meningkatkan akurasi hingga 98,39% dengan nilai precision, recall, dan F1-score yang tinggi [19].

Meskipun berbagai penelitian telah menunjukkan potensi RFE dalam meningkatkan performa model, kajian yang secara khusus menganalisis kinerja RFE pada klasifikasi penyakit stroke dengan kondisi data tidak seimbang masih relatif terbatas. Oleh karena itu, penelitian ini berfokus pada analisis kinerja *Recursive Feature Elimination* (RFE) terhadap performa klasifikasi penyakit stroke menggunakan *Support Vector Machine* (SVM) dengan penanganan ketidakseimbangan data menggunakan *Adaptive Synthetic Sampling* (ADASYN). Analisis ini diharapkan dapat memberikan pemahaman yang lebih jelas mengenai kontribusi seleksi fitur terhadap performa klasifikasi pada data stroke yang tidak seimbang.

2. METODOLOGI PENELITIAN

Metodologi penelitian ini disusun secara sistematis untuk mengembangkan model klasifikasi penyakit stroke menggunakan pendekatan *machine learning*. Berikut alur penelitian yang dilakukan pada Gambar 1.



Gambar 1. Tahapan Penelitian



2.1 Identifikasi Masalah

Penelitian ini berfokus pada permasalahan klasifikasi penyakit stroke yang dipengaruhi oleh ketidakseimbangan distribusi kelas pada dataset. Ketidakseimbangan distribusi data membuat model cenderung membentuk pola berdasarkan kelas yang jumlahnya lebih banyak, sehingga kemampuan identifikasi terhadap kelas minoritas menjadi berkurang. Selain itu, tidak seluruh atribut pada dataset memberikan informasi yang sama pentingnya bagi pembentukan model, sehingga keberadaan fitur yang kurang relevan berpotensi mempengaruhi performa model. Untuk mengatasi permasalahan tersebut, penelitian ini mengusulkan penerapan *Recursive Feature Elimination* (RFE) sebagai metode seleksi fitur dan *Adaptive Synthetic Sampling* (ADASYN) sebagai teknik penanganan data tidak seimbang pada algoritma *Support Vector Machine* (SVM).

2.2 Pengumpulan Data

Penelitian ini menggunakan dataset sekunder yang diperoleh dari platform *kaggle* dengan nama dataset *healthcare-dataset-stroke-data*, dataset tersedia secara publik melalui <https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset>. Dataset ini terdiri dari 5110 data dengan 11 atribut fitur dan 1 atribut target (stroke). Atribut tersebut meliputi *gender*, *age* (umur), *hypertension* (hipertensi), *heart disease* (riwayat jantung), *ever married* (status pernikahan), *work type* (tipe pekerjaan), *residence type* (tipe tempat tinggal), *avg glucose level* (kadar glukosa), BMI (*body mass index*), *smoking status* (status merokok), serta atribut stroke digunakan sebagai label klasifikasi untuk membedakan antara yang mengalami stroke dan yang tidak mengalami stroke.

2.3 Preprocessing Data

Sebelum memasuki tahap pemodelan, dataset terlebih dahulu melalui proses *preprocessing* untuk menyesuaikan kondisi data agar lebih optimal saat digunakan pada algoritma *machine learning* [20]. Tahapan ini mencakup pembersihan data, transformasi data, serta normalisasi fitur. Kolom id dihapus karena tidak memiliki pengaruh terhadap proses klasifikasi. Selain itu, dilakukan pembersihan data dengan menghapus data yang tidak valid, yaitu satu record dengan kategori "Other" karena tidak representatif. Selain itu, penanganan *missing value* dilakukan pada atribut tertentu menggunakan imputasi nilai median untuk menjaga kestabilan distribusi data tetap stabil. Selanjutnya, dilakukan transformasi data dengan mengubah atribut kategorikal menjadi bentuk numerik agar dapat diproses oleh algoritma *machine learning*. Tahap selanjutnya adalah normalisasi data menggunakan metode *Min-Max Scaling* untuk menyamakan rentang nilai antar fitur sehingga setiap atribut berada pada skala yang seragam.

2.4 Pembagian Data

Dalam penelitian ini, dilakukan pembagian data menggunakan rasio umum, yaitu 70:30, 80:20, dan 90:10, sebagai bagian dari proses untuk mengevaluasi pengaruh komposisi data terhadap performa model. Data latih digunakan untuk membangun serta menentukan konfigurasi model terbaik, sedangkan data uji disimpan dan tidak dilibatkan selama proses pelatihan, serta hanya dimanfaatkan pada tahap evaluasi akhir. Pada data latih diterapkan metode *Stratified K-Fold Cross Validation* untuk menjaga proporsi distribusi kelas pada setiap subset [21]. Dalam setiap iterasi, data dibagi menjadi beberapa subset dengan proporsi kelas yang seimbang, di mana sebagian subset digunakan sebagai data pelatihan dan sisanya digunakan sebagai data validasi secara bergantian [22].

2.5 Pemodelan

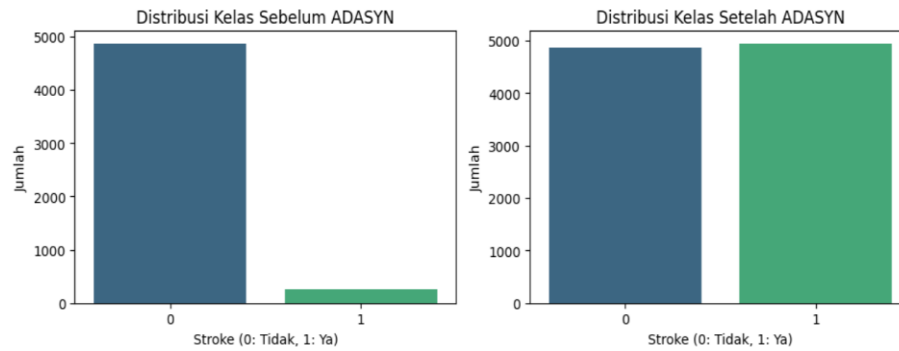
Pemodelan dilakukan pada data latih menggunakan pendekatan *K-Fold Cross Validation*. Pada setiap iterasi, diterapkan proses pipeline yang terdiri dari penyeimbangan data menggunakan ADASYN, seleksi fitur menggunakan RFE, dan proses klasifikasi menggunakan algoritma SVM. Seluruh tahapan tersebut dilakukan pada data latih di setiap fold untuk menghindari terjadinya data leakage selama proses pelatihan dan validasi model.

2.5.1 Penyeimbangan Data dengan ADASYN

Adaptive Synthetic Sampling (ADASYN) merupakan metode oversampling yang menghasilkan sampel sintesis secara adaptif dengan memfokuskan penambahan sampel pada data minoritas yang sulit dipelajari oleh model. Proses ADASYN dimulai dengan menghitung tingkat ketidakseimbangan antara kelas mayoritas dan minoritas. Selanjutnya, ditentukan jumlah sampel sintesis yang perlu ditambahkan sesuai dengan parameter keseimbangan yang diinginkan. Setiap data minoritas dievaluasi tingkat kesulitannya menggunakan pendekatan tetangga terdekatnya, data minoritas yang berada pada area dengan tingkat kesulitan tinggi akan memperoleh lebih banyak sampel sintesis [16]. Secara matematis, jumlah sampel sintesis yang dihasilkan dapat dituliskan sebagai:

$$G = (m_l - m_s) \times \beta \quad (1)$$

Sampel sintesis (G) dihitung berdasarkan selisih antara jumlah data kelas mayoritas (m_l) dan jumlah data kelas minoritas (m_s), kemudian dikalikan dengan parameter β yang mengontrol tingkat keseimbangan yang diinginkan [23]. Dalam penelitian ini, kasus stroke merupakan kelas minoritas. Distribusi data sebelum dan sesudah penerapan ADASYN dapat dilihat pada Gambar 2.



Gambar 2. Distribusi data saat proses ADASYN

Sebelum dilakukannya proses ADASYN terdapat ketimpangan yang signifikan antara kelas 0 dan kelas 1 dengan 4860 data pada kelas 0 (non-stroke) dan 249 data pada kelas 1 (stroke) dari total 5109 data. Setelah ADASYN diterapkan, jumlah data pada kelas minoritas meningkat secara signifikan, sehingga distribusi antar kelas menjadi jauh lebih seimbang. Sebagai gambaran, ketika suatu data dari kelas stroke dikelilingi oleh beberapa tetangga terdekat yang sebagian besar justru berasal dari kelas non-stroke, maka data tersebut cenderung sulit dikenali oleh model. Dalam kondisi seperti ini, ADASYN akan memfokuskan pembuatan data sintesis lebih banyak di area tersebut, dibandingkan pada titik-titik minoritas yang berada di lingkungan dengan dominasi kelas yang sama [7].

2.5.2 Recursive Feature Elimination (RFE)

Recursive Feature Elimination (RFE) merupakan metode seleksi fitur yang bekerja dengan cara menghilangkan fitur-fitur yang memiliki kontribusi paling rendah terhadap model klasifikasi. Metode ini termasuk dalam pendekatan *wrapper* karena proses pemilihan fitur dilakukan berdasarkan performa model pembelajaran yang digunakan [24]. Dalam penelitian ini, RFE dikombinasikan dengan SVM sebagai model dasar. Pada tahap seleksi fitur, RFE menggunakan SVM linear sebagai estimator untuk menentukan tingkat kontribusi masing-masing fitur. Pemilihan SVM linear dilakukan karena metode RFE memerlukan bobot fitur yang dapat digunakan untuk proses perankingan dan eliminasi fitur secara bertahap. SVM linear menyediakan nilai koefisien yang dapat diinterpretasikan secara langsung sebagai ukuran kontribusi fitur terhadap pembentukan hyperplane, sedangkan kernel non-linear tidak menghasilkan bobot fitur yang dapat digunakan secara langsung dalam proses seleksi fitur. Bobot fitur diperoleh dari parameter model SVM yang mempresentasikan kontribusi masing-masing fitur terhadap fungsi keputusan, di mana fitur dengan bobot terkecil dianggap memiliki kontribusi paling rendah dan akan dieliminasi pada setiap iterasi [11]. Meskipun RFE menggunakan SVM linear sebagai estimator untuk proses seleksi fitur, model klasifikasi akhir tetap dievaluasi menggunakan berbagai kernel SVM. Penggunaan SVM linear pada tahap RFE bertujuan untuk memperoleh peringkat fitur berdasarkan bobot koefisiennya, sedangkan pemilihan kernel terbaik dilakukan secara terpisah berdasarkan hasil evaluasi performa model. Bobot ini diperoleh dari parameter SVM, yaitu koefisien Lagrange (α_k), label kelas (y_k), dan vector fitur (x_k). Persamaan bobot ditulis sebagai:

$$w = \sum_1^k \alpha_k y_k x_k \quad (2)$$

Pada persamaan (2), w merupakan vektor bobot fitur yang digunakan oleh SVM. Selanjutnya, setiap fitur diberi peringkat berdasarkan bobotnya dengan mengkuadratkan komponen bobot yang sesuai:

$$c_k = w_k^2, k = 1, 2, \dots, |S| \quad (3)$$

Pada persamaan (3), c_k menunjukkan nilai kontribusi fitur ke- k sedangkan w_k merupakan bobot dari fitur ke- k . Proses RFE diawali dengan pelatihan model menggunakan seluruh fitur yang ada, kemudian setiap fitur dievaluasi berdasarkan tingkat kontribusinya terhadap kinerja model [25]. Fitur dengan kontribusi terendah kemudian dieliminasi, dan model dilatih kembali menggunakan subset fitur yang tersisa. Proses ini dilakukan secara iteratif hingga diperoleh jumlah fitur yang optimal sesuai dengan kebutuhan penelitian [19].

2.5.3 Support Vector Machine (SVM)

Support Vector Machine (SVM) digunakan sebagai algoritma klasifikasi utama untuk membedakan data stroke dan non-stroke. Prinsip kerja SVM adalah menentukan hyperplane terbaik yang mampu memisahkan data dari masing-masing kelas secara optimal [26]. Hyperplane tersebut diperoleh melalui proses optimasi dengan memaksimalkan jarak antar kelas atau *margin*, sehingga model diharapkan mampu mempertahankan performa klasifikasi ketika dihadapkan pada data yang belum pernah dipelajari sebelumnya. Titik data yang memiliki jarak paling dekat terhadap *hyperplane* berperan penting dalam pembentukan batas keputusan dan dikenal sebagai *support vector*. Dalam ruang berdimensi dua, hyperplane dipresentasikan sebagai garis lurus, sedangkan pada dimensi yang lebih tinggi dapat berupa bidang atau permukaan pemisah [27]. Secara umum, hyperplane pada SVM dapat dinyatakan sebagai berikut:

$$w \cdot x + b = 0 \quad (4)$$



Pada persamaan tersebut, w menunjukkan vektor bobot yang menentukan arah hyperplane, x mempresentasikan fitur input, dan b adalah bias yang menggeser hyperplane dari titik asal. Ketika pola data tidak dapat dipisahkan secara linear, SVM menggunakan fungsi kernel untuk melakukan transformasi ruang fitur sehingga batas pemisahan yang lebih kompleks dapat dibentuk. Penelitian ini menggunakan tiga jenis kernel, yaitu *linear* kernel, *polynomial* kernel, dan *radial basis function* (RBF) kernel [27]. Secara matematis, masing-masing kernel dapat dinyatakan sebagai berikut:

Tabel 1. Persamaan kernel SVM

Jenis Kernel	Rumus
Linear	$K(x_i, x_j) = x_i \cdot x_j$
Polynomial	$K(x_i, x_j) = (x_i \cdot x_j + 1)^2$
RBF	$K(x_i, x_j) = \exp\left(-\frac{\ x_i - x_j\ ^2}{2\sigma^2}\right)$

Pada persamaan kernel tersebut, x_i dan x_j merepresentasikan dua vektor data yang akan dihitung tingkat kemiripannya. Kernel linear menghitung kemiripan berdasarkan hasil perkalian dot product antar vektor sehingga sesuai untuk data yang memiliki pola pemisahan linear. Kernel polynomial memetakan data ke ruang fitur yang lebih tinggi melalui fungsi polinomial sehingga mampu menangkap hubungan yang lebih kompleks. Sementara itu, kernel RBF mengukur kemiripan berdasarkan jarak antar data dan memungkinkan pemodelan pola non-linear yang lebih fleksibel. Penggunaan ketiga kernel ini bertujuan untuk membandingkan kemampuan SVM dalam menangani karakteristik data stroke yang belum diketahui secara pasti pola pemisahannya.

2.6 Evaluasi

Evaluasi model merupakan proses untuk mengetahui model dalam mengklasifikasikan data stroke data non-stroke berdasarkan hasil prediksi terhadap label sebenarnya [26]. Penilaian kinerja dilakukan menggunakan *confusion matrix*, yaitu tabel yang memuat jumlah prediksi benar maupun salah pada setiap kelas, komponen pada *confusion matrix* terdiri dari *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) [10]. Berdasarkan nilai-nilai tersebut, performa model dianalisis menggunakan beberapa metrik evaluasi sebagai berikut [28]: Akurasi adalah perhitungan untuk mengukur prediksi yang benar terhadap total keseluruhan data, pada persamaan 5.

$$Akurasi = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (5)$$

Presisi mengukur ketepatan prediksi positif dengan menghitung seberapa besar bagian prediksi positif yang benar-benar termasuk kelas positif, melalui persamaan 6.

$$Presisi = \frac{TP}{(TP+FP)} \quad (6)$$

Recall mengukur kemampuan model dalam mengenali seluruh data yang benar-benar termasuk kelas positif, melalui persamaan 7.

$$Recall = \frac{TP}{(TP+FN)} \quad (7)$$

F1-score adalah gabungan antara presisi dan recall yang digunakan untuk memberikan evaluasi kinerja model secara seimbang, melalui persamaan 8.

$$F1 - score = \frac{2(Presisi \times Recall)}{(Presisi + Recall)} \quad (8)$$

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Dataset yang digunakan dalam penelitian ini memiliki karakteristik data kesehatan individu yang mencakup berbagai atribut seperti usia, kondisi medis, serta gaya hidup yang berpotensi mempengaruhi risiko terjadinya stroke.

Tabel 2. Kumpulan Data Stroke

id	Gender	Age	Hyper tension	Heart disease	...	Avg glucose level	bmi	Smoking status	Stroke
34120	Male	75	1	1	...	221.29	25.8	Smokes	1
39373	Female	82	1	0	...	196.92	22.2	Never smoked	1
...
34138	Male	42	0	0	...	89	36.3	Formerly smoked	0
49894	Female	78	1	1	...	206.53	N/A	Never smoked	0



Tabel 2 menunjukkan bahwa setiap baris merepresentasikan satu individu dengan berbagai atribut yang digunakan sebagai variabel prediktor. Sedangkan kolom stroke berperan sebagai label kelas yang membedakan antara individu yang mengalami stroke (kelas 1) dan tidak mengalami stroke (kelas 0).

Tabel 3. Deskripsi Atribut

Atribut	Deskripsi
<i>Gender</i>	Jenis kelamin
<i>age</i>	Usia
<i>hypertension</i>	Riwayat hipertensi
<i>heart disease</i>	Riwayat penyakit jantung
<i>ever married</i>	Status pernikahan
<i>work type</i>	Jenis pekerjaan
<i>residence type</i>	Tipe tempat tinggal
<i>avg glucose level</i>	Rata-rata kadar glukosa darah
<i>body mass index</i>	Indeks massa tubuh
<i>smoking status</i>	Status merokok
<i>stroke</i>	Label stroke

Tabel 3 menunjukkan semua atribut yang digunakan dalam dataset bersama dengan deskripsinya. Atribut-atribut tersebut digunakan sebagai dasar dalam proses preprocessing dan pemodelan untuk mengidentifikasi pola yang berkaitan dengan risiko stroke. Selain itu dataset ini menunjukkan kondisi ketidakseimbangan kelas yang cukup tinggi karena jumlah data non-stroke jauh lebih banyak dibandingkan data stroke. Dari total 5110 data, terdapat 4861 data pada kelas non-stroke dan 249 data pada kelas stroke.

3.2 Preprocessing Data

Sebelum digunakan pada tahap pemodelan, dataset terlebih dahulu dipersiapkan melalui proses *preprocessing* untuk memastikan kualitas dan kesiapan data. Kolom identitas (id) dihapus karena tidak memiliki relevansi terhadap proses klasifikasi. Selain itu, terdapat satu record data dengan nilai "Other" pada atribut gender yang juga dihapus karena tidak representatif terhadap kebutuhan analisis, setelah proses tersebut, total data yang digunakan menjadi 5109 sampel. Pada atribut BMI ditemukan sejumlah *missing values* yang kemudian ditangani menggunakan imputasi nilai median, untuk menjaga kestabilan distribusi data. Selanjutnya, atribut kategorikal diubah ke dalam bentuk numerik agar dapat diproses oleh model *machine learning*. Atribut dengan dua kategori seperti *gender*, *ever married*, dan *residence type* ditransformasikan menggunakan *label encoding*, sedangkan atribut dengan lebih dari dua kategori seperti *work type* dan *smoking status* diubah menggunakan *one-hot encoding* sehingga setiap kategori direpresentasikan secara terpisah dalam bentuk numerik. Setelah proses transformasi, dilakukan normalisasi data menggunakan metode *Min-Max Scaling* untuk menyamakan rentang nilai antar fitur sehingga perbedaan skala antar atribut tidak mempengaruhi proses pelatihan model.

3.3 Hasil Pemodelan

Setelah tahap *preprocessing* kemudian dataset dibagi menjadi data latih dan data uji. Untuk menentukan rasio pembagian data yang optimal, pengujian dilakukan terhadap tiga rasio yaitu 70:30, 80:20, 90:10. Proses pemodelan dilakukan menggunakan metode *5-fold cross validation*. Pemilihan nilai $k=5$ didasarkan pada hasil perbandingan eksperimen antara $k=5$ dan $k=10$. Hasil pengujian menunjukkan bahwa penggunaan $k=10$ menghasilkan performa yang sebanding dengan $k=5$ pada rasio 80:20, namun memerlukan waktu komputasi yang jauh lebih lama. Selain itu, pada rasio 70:30 dan 90:10, penggunaan $k=10$ menghasilkan nilai recall kelas stroke yang lebih rendah dibandingkan $k=5$. Dengan mempertimbangkan keseimbangan antara performa model dan efisiensi komputasi, $k=5$ dipilih sebagai konfigurasi akhir karena menghasilkan sensitivitas model yang lebih baik dengan waktu komputasi yang lebih efisien. Pada setiap fold, diterapkan *pipeline* yang terdiri dari penyeimbangan data menggunakan ADASYN, seleksi fitur menggunakan RFE, serta proses klasifikasi menggunakan algoritma SVM. Pendekatan pipeline ini memastikan bahwa seluruh proses transformasi data hanya dilakukan berdasarkan informasi dari data latih pada setiap fold, sehingga tidak terjadi kebocoran informasi dari data validasi ke dalam proses pelatihan.

Pengujian dilakukan terhadap beberapa kombinasi konfigurasi model, yang terdiri dari tiga jenis kernel SVM, yaitu linear, RBF, dan polynomial, Serta tiga nilai parameter C, yaitu 0.1, 1, dan 10, sehingga menghasilkan 9 kombinasi konfigurasi per rasionya. RFE digunakan untuk memilih 10 fitur yang paling relevan dalam proses klasifikasi dari semua fitur yang tersedia setelah proses preprocessing. Dalam penelitian ini pemilihan konfigurasi terbaik didasarkan pada nilai rata-rata F1-score kelas stroke dari seluruh fold.

Tabel 4. Performa Terbaik K-fold per Rasio

Rasio	Kernel	C	Akurasi	Precision (stroke)	Recall (stroke)	F1-score (stroke)
	Linear	0.1	0.6999	0.1214	0.8276	0.2118
	Linear	1	0.7120	0.1234	0.8047	0.2139
	Linear	10	0.7134	0.1233	0.7988	0.2136
	RBF	0.1	0.6882	0.1136	0.7933	0.1986



Rasio	Kernel	C	Akurasi	Precision (stroke)	Recall (stroke)	F1-score (stroke)
70:30	RBF	1	0.7449	0.1181	0.6497	0.1997
	RBF	10	0.7754	0.1224	0.5926	0.2024
	Polynomial	0.1	0.7422	0.1245	0.7128	0.2118
	Polynomial	1	0.7615	0.1245	0.6497	0.2086
	Polynomial	10	0.7760	0.1220	0.5864	0.2014
	Linear	0.1	0.7076	0.1269	0.8494	0.2208
	Linear	1	0.7147	0.1263	0.8191	0.2188
	Linear	10	0.7176	0.1280	0.8241	0.2216
	RBF	0.1	0.7034	0.1207	0.8094	0.2101
80:20	RBF	1	0.7490	0.1225	0.6732	0.2069
	RBF	10	0.7734	0.1285	0.6281	0.2126
	Polynomial	0.1	0.7595	0.1335	0.7138	0.2247
	Polynomial	1	0.7702	0.1300	0.6533	0.2165
	Polynomial	10	0.7773	0.1316	0.6329	0.2172
	Linear	0.1	0.7103	0.1275	0.8438	0.2215
	Linear	1	0.7212	0.1303	0.8305	0.2252
	Linear	10	0.7238	0.1314	0.8305	0.2268
	RBF	0.1	0.7029	0.1185	0.7901	0.2061
90:10	RBF	1	0.7447	0.1282	0.7277	0.2179
	RBF	10	0.7621	0.1298	0.6743	0.2174
	Polynomial	0.1	0.7527	0.1353	0.7544	0.2293
	Polynomial	1	0.7638	0.1351	0.7099	0.2269
	Polynomial	10	0.7610	0.1298	0.6787	0.2177

Berdasarkan Tabel 4, terlihat bahwa setiap rasio menghasilkan konfigurasi model terbaik yang berbeda. Rasio 70:30 cenderung optimal pada kernel linear dengan $C = 1$, sedangkan rasio 80:20 dan 90:10 menunjukkan performa terbaik pada kernel polynomial dengan parameter $C = 0.1$. Pemilihan konfigurasi ini didasarkan pada nilai F1-score karena F1-score dinilai lebih representatif dalam kasus klasifikasi data tidak seimbang, dibandingkan dengan akurasi yang cenderung bias terhadap kelas mayoritas. Penggunaan kernel polynomial dengan $C = 0.1$ yang menghasilkan performa terbaik menunjukkan bahwa pemetaan data ke ruang fitur yang lebih tinggi memberikan keuntungan dibandingkan pemisahan linear pada dataset yang digunakan. Ketiga rasio menghasilkan pola performa yang serupa, yaitu nilai precision yang rendah pada seluruh konfigurasi dan nilai recall yang cukup tinggi, kondisi ini menunjukkan bahwa model cenderung lebih sensitif dalam mendeteksi kelas stroke (recall tinggi), namun masih menghasilkan cukup banyak kesalahan prediksi pada kelas non-stroke (precision rendah). Rasio 90:10 menghasilkan F1-score tertinggi pada proses K-fold sebesar 0.2293, diikuti rasio 80:20 sebesar 0.2247, dan rasio 70:30 sebesar 0.2139. Perbedaan F1-score antar rasio yang relatif kecil menunjukkan bahwa performa model lebih dipengaruhi oleh karakteristik distribusi data dan ketidakseimbangan kelas dibandingkan oleh variasi pembagian data itu sendiri. Kondisi ini menyebabkan nilai F1-score tetap rendah meskipun recall tinggi, karena F1-score merupakan keseimbangan antara precision dan recall.

3.4 Seleksi Fitur RFE

Seleksi fitur menggunakan RFE dilakukan di dalam pipeline pada setiap fold dengan menggunakan SVM linear sebagai estimator, menghasilkan 10 fitur terpilih dari semua fitur yang tersedia. RFE bekerja dengan mengeliminasi fitur secara bertahap berdasarkan kontribusinya terhadap pembentukan hyperplane SVM melalui perhitungan bobot koefisien, hingga diperoleh fitur-fitur yang paling berkontribusi dalam proses klasifikasi pada setiap kondisi distribusi data latih. Pemilihan fitur oleh RFE didasarkan pada kontribusi matematis fitur terhadap pembentukan hyperplane SVM, bukan pada relevansi klinisnya sebagai faktor risiko stroke. Hasil seleksi fitur menunjukkan bahwa terdapat sejumlah fitur yang secara konsisten terpilih pada seluruh rasio pembagian data, yaitu *age*, *avg glucose level*, *bmi*, *work type Govt job*, *work type private*, *work type self employed*, *work type children*, *smoking status never smoked*. Konsistensi ini menunjukkan bahwa fitur-fitur tersebut memiliki kontribusi yang relatif kuat dan stabil dalam membantu SVM membedakan kelas stroke dan non-stroke, terlepas dari variasi distribusi data latih.

Terdapat beberapa fitur yang terpilih secara berbeda pada masing-masing rasio. Pada rasio 70:30 dan 80:20, fitur *gender* termasuk dalam fitur terpilih, sementara pada rasio 90:10 fitur tersebut digantikan oleh *hypertension*. Pada kategori *smoking status*, rasio 80:20 dan 90:10 memilih *smoking status smokes* sebagai salah satu fitur terpilih, sedangkan pada rasio 70:30 fitur tersebut tidak terpilih. Perbedaan fitur yang terpilih antar rasio ini menunjukkan bahwa komposisi data latih yang berbeda pada setiap rasio mempengaruhi penilaian kontribusi fitur oleh RFE terhadap pembentukan hyperplane SVM. Di sisi lain, fitur *hypertension* dan *heart disease* yang secara klinis dikenal sebagai faktor risiko utama stroke justru tidak terpilih secara konsisten oleh RFE. Kondisi ini menunjukkan bahwa kontribusi suatu fitur dalam model SVM tidak selalu sejalan dengan tingkat kepentingan klinisnya, melainkan ditentukan oleh kemampuannya dalam membentuk batas keputusan yang optimal pada ruang fitur yang digunakan. Dengan demikian, fitur yang penting secara klinis belum tentu memberikan kontribusi diskriminatif yang paling kuat terhadap proses klasifikasi pada dataset yang digunakan dalam penelitian ini.

3.5 Evaluasi

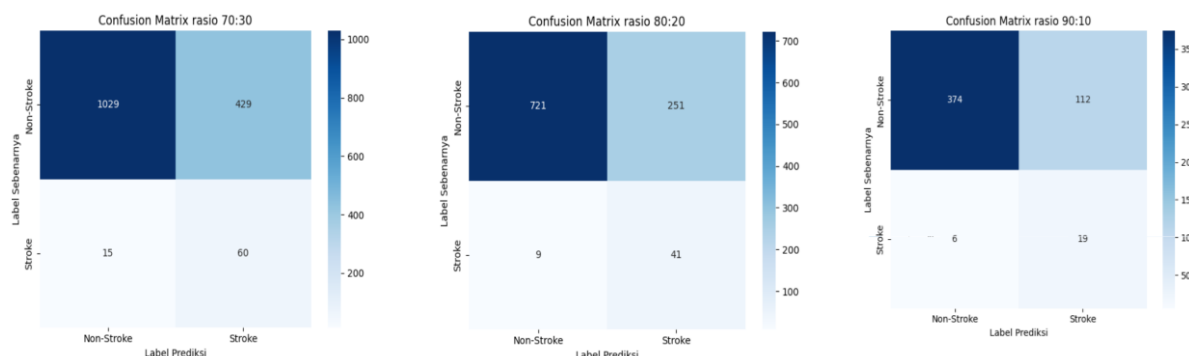
Evaluasi akhir dilakukan dengan menguji model terbaik dari masing-masing rasio pembagian data pada data uji yang sebelumnya tidak dilibatkan dalam proses pelatihan maupun pemilihan konfigurasi model. Model yang dievaluasi merupakan model final yang telah dilatih ulang menggunakan seluruh data latih pada masing-masing rasio dengan konfigurasi terbaik yang diperoleh dari proses 5-Fold. Evaluasi dilakukan pada ketiga rasio untuk memperoleh perbandingan yang komprehensif sebagai dasar pemilihan rasio terbaik.

Tabel 5. Evaluasi ketiga Rasio

Rasio	Kernel	C	Akurasi (stroke)	Precision (stroke)	Recall (stroke)	F1-Score (stroke)
70:30	Linear	1	0.71	0.12	0.80	0.21
80:20	Polynomial	0.1	0.75	0.14	0.82	0.24
90:10	Polynomial	0.1	0.77	0.15	0.76	0.24

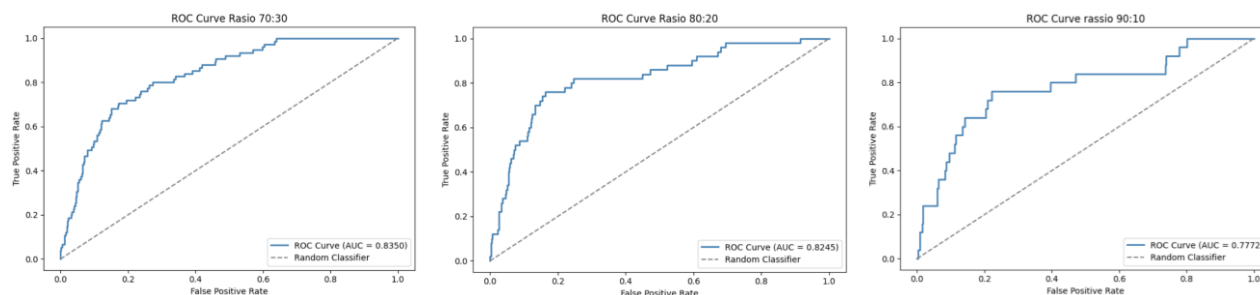
Berdasarkan hasil pada Tabel 5, terdapat perbedaan performa model pada masing-masing rasio pembagian data. Dari sisi akurasi, rasio 90:10 memperoleh nilai tertinggi sebesar 0.77, sedangkan rasio 80:20 dan 70:30 masing-masing menghasilkan akurasi sebesar 0.75 dan 0.71. Namun, untuk kasus data tidak seimbang, evaluasi tidak hanya bergantung pada akurasi. Jika ditinjau berdasarkan F1-score kelas stroke, rasio 80:20 dan 90:10 menunjukkan performa yang setara dengan nilai 0.24, serta lebih baik dibandingkan rasio 70:30 yang memperoleh nilai 0.21. Kemampuan model dalam mengenali kasus stroke paling baik ditunjukkan oleh rasio 80:20 dengan nilai recall sebesar 0.82. Nilai tersebut lebih tinggi dibandingkan rasio 70:30 yang mencapai 0.80 maupun rasio 90:10 sebesar 0.76.

Sementara itu, rasio 70:30 menghasilkan recall sebesar 0.80 dan rasio 90:10 sebesar 0.76. Hal ini menunjukkan bahwa meskipun rasio 90:10 memiliki akurasi lebih tinggi, sensitivitas model terhadap kelas minoritas tidak selalu lebih baik. Ketepatan prediksi terhadap kelas stroke masih relatif rendah pada seluruh rasio pengujian, yaitu berada pada rentang 0.12 hingga 0.15. Kondisi ini menunjukkan bahwa model masih menghasilkan cukup banyak prediksi positif yang tidak tepat (*false positive*), yang merupakan konsekuensi dari ketidakseimbangan kelas pada dataset stroke.



Gambar 3. Confusion Matrix Seluruh Rasio

Berdasarkan confusion matrix yang ada pada Gambar 3, terlihat bahwa model pada ketiga rasio mampu mengklasifikasikan sebagian besar data non-stroke dengan benar, namun menghasilkan jumlah *False Positive* yang cukup tinggi. Pada rasio 70:30, model menghasilkan True Negative (TN) sebesar 1029, False Positive (FP) sebesar 429, True Positive (TP) sebesar 60, dan False Negative (FN) sebesar 15. Pada rasio 80:20, model menghasilkan TN sebesar 721, FP sebesar 251, TP sebesar 41, dan FN sebesar 9. Sementara pada rasio 90:10, model menghasilkan TN sebesar 374, FP sebesar 112, TP sebesar 19, dan FN sebesar 6. Meskipun menggunakan rasio pembagian data yang berbeda, ketiga model memperlihatkan kecenderungan yang serupa, yaitu banyak menghasilkan *false positive* namun relatif sedikit *false negative*. Jumlah FN yang kecil pada seluruh rasio mengindikasikan bahwa model cukup sensitif dalam mendeteksi kasus stroke, namun tingginya FP menunjukkan bahwa model masih cenderung memprediksi banyak data non-stroke sebagai stroke. Kondisi ini menunjukkan bahwa model lebih cenderung memberikan peringatan palsu, namun berhasil meminimalkan kasus stroke yang terlewat.



Gambar 4. Kurva ROC-AUC Seluruh Rasio

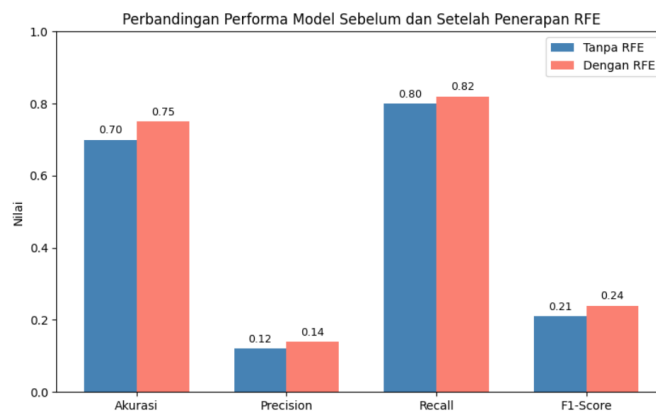


Gambar 4 merupakan kurva kinerja model yang dianalisis menggunakan kurva *Receiver Operating Characteristic* (ROC) dan nilai *Area Under Curve* (AUC). Kurva ROC digunakan untuk melihat kemampuan model dalam membedakan kelas stroke dan non-stroke melalui perbandingan antara *True Positive Rate* (TPR) dan *False Positive Rate* (FPR) pada berbagai nilai ambang. Berdasarkan hasil pengujian terhadap ketiga rasio, kurva ROC seluruh rasio berada di atas garis diagonal, menunjukkan bahwa model memiliki performa yang lebih baik dibandingkan tebakan acak. Rasio 70:30 menghasilkan AUC tertinggi sebesar 0.8350, diikuti rasio 80:20 sebesar 0.8245, dan rasio 90:10 sebesar 0.7772. Hasil AUC yang relatif tinggi pada seluruh rasio menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam membedakan kelas stroke dan non-stroke secara keseluruhan. Namun demikian, nilai precision yang masih rendah menunjukkan bahwa kemampuan diskriminasi tersebut belum sepenuhnya diikuti oleh ketepatan prediksi pada kelas stroke, sehingga performa model masih memiliki keterbatasan pada kondisi data yang tidak seimbang. Dengan mempertimbangkan keseimbangan antara F1-score dan AUC secara keseluruhan, rasio 80:20 dipilih sebagai konfigurasi terbaik karena menghasilkan performa yang paling seimbang antara kemampuan deteksi kelas stroke dan kemampuan diskriminasi antar kelas. Selanjutnya, untuk memperkuat analisis perbandingan performa model sebelum dan setelah penerapan RFE, evaluasi tambahan dilakukan pada konfigurasi terbaik tanpa menggunakan RFE.

Tabel 6. Perbandingan Sebelum dan Sesudah RFE

RFE	Kernel	C	Akurasi (stroke)	Precision (stroke)	Recall (stroke)	F1-Score (stroke)
Sebelum	Linear	1	0.70	0.12	0.80	0.21
Sesudah	Polynomial	0.1	0.75	0.14	0.82	0.24

Berdasarkan Tabel 6, penerapan RFE memberikan peningkatan performa pada seluruh metrik. Nilai akurasi meningkat dari 0.70 menjadi 0.75, sementara precision, recall, dan F1-score juga mengalami peningkatan. Meskipun demikian, peningkatan yang diperoleh relatif kecil, yang menunjukkan bahwa sebagian besar informasi yang diperlukan untuk proses klasifikasi telah terwakili oleh fitur awal pada dataset sehingga eliminasi fitur tidak menghasilkan perubahan performa yang besar. Nilai recall meningkat dari 0.80 menjadi 0.82, yang menunjukkan adanya peningkatan sensitivitas model dalam mendeteksi kasus stroke. Peningkatan paling terlihat pada F1-score yang naik dari 0.21 menjadi 0.24, ini menunjukkan bahwa keseimbangan antara kemampuan model dalam mendeteksi kelas positif dan ketepatan prediksi menjadi sedikit lebih baik setelah penerapan RFE. Hasil perbandingan evaluasi divisualisasikan pada Gambar 5 berikut.



Gambar 5. Perbandingan hasil evaluasi

Pada Gambar 5, terlihat bahwa seluruh metrik evaluasi mengalami peningkatan setelah penerapan RFE. Peningkatan terhadap semua performa yang dihasilkan terhadap kelas stroke relatif kecil dan berada dalam rentang yang tidak terlalu signifikan, yang menunjukkan bahwa sebagian besar fitur awal mungkin telah cukup representatif dalam membentuk model klasifikasi, sehingga proses eliminasi fitur tidak menghasilkan perubahan performa yang besar. Selain itu, terlihat adanya perubahan konfigurasi model terbaik antara sebelum dan setelah penerapan RFE. Sebelum RFE, model terbaik menggunakan kernel linear, sedangkan setelah penerapan RFE model terbaik menggunakan kernel polynomial. Hal ini menunjukkan bahwa seleksi fitur tidak hanya mempengaruhi jumlah fitur yang digunakan, tetapi juga mempengaruhi karakteristik pemisahan data sehingga berdampak pada konfigurasi model yang optimal.

3.6 Pembahasan

Penelitian ini bertujuan untuk menganalisis kinerja RFE pada model SVM dengan penanganan ketidakseimbangan data menggunakan ADASYN dalam klasifikasi penyakit stroke. Berdasarkan hasil eksperimen yang telah dilakukan, penerapan RFE memberikan peningkatan performa pada model dibandingkan tanpa RFE. Dari ketiga rasio pembagian data yang diuji, rasio 80:20 dipilih sebagai konfigurasi terbaik karena menghasilkan keseimbangan performa yang paling baik berdasarkan nilai F1-score dan AUC. Hasil ini menunjukkan bahwa RFE membantu menyederhanakan ruang fitur sehingga model mampu membentuk hyperplane yang lebih stabil dan representatif terhadap pola data stroke. Nilai recall yang lebih tinggi dibandingkan precision menunjukkan bahwa model lebih sensitif dalam mendeteksi kasus stroke dibandingkan dalam memastikan ketepatan prediksi positifnya. Kondisi ini menyebabkan jumlah kasus stroke yang



terlewat relatif rendah, namun di sisi lain menghasilkan false positive yang cukup tinggi. Oleh karena itu, meskipun model menunjukkan kemampuan deteksi yang baik terhadap kelas stroke, nilai precision yang masih rendah menunjukkan bahwa performa model belum optimal dan masih memerlukan pengembangan lebih lanjut sebelum dapat diterapkan pada lingkungan klinis. Selain itu, nilai AUC yang diperoleh menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam membedakan kelas stroke dan non-stroke secara keseluruhan. Namun demikian, kemampuan diskriminasi tersebut belum sepenuhnya diikuti oleh ketepatan prediksi pada kelas stroke, sebagaimana ditunjukkan oleh nilai precision yang masih rendah. Dari aspek seleksi fitur, penerapan RFE menunjukkan bahwa tidak semua fitur awal memberikan kontribusi yang sama terhadap proses klasifikasi. Beberapa fitur yang berkaitan dengan karakteristik fisiologis dan gaya hidup menjadi lebih dominan dalam membentuk keputusan model. Konsistensi pemilihan fitur-fitur ini pada berbagai rasio pembagian data mengindikasikan bahwa fitur-fitur tersebut memang memiliki kontribusi informatif yang cukup stabil dalam membantu proses klasifikasi. Namun demikian, pengurangan jumlah fitur yang dilakukan tidak diikuti oleh peningkatan performa yang signifikan. Hal ini menunjukkan bahwa informasi yang dibutuhkan untuk proses klasifikasi sebagian besar telah terwakili dengan baik oleh fitur awal, sehingga eliminasi fitur tidak memberikan dampak yang signifikan terhadap kinerja model.

Hasil penelitian ini juga menunjukkan performa yang lebih baik dibandingkan penelitian [7] yang menggunakan kombinasi ADASYN dan SVM tanpa penerapan RFE pada dataset stroke yang sama. Peningkatan tersebut menunjukkan bahwa penerapan RFE membantu model dalam memilih fitur yang lebih relevan sehingga performa klasifikasi mengalami perbaikan, meskipun peningkatan yang diperoleh relatif terbatas. Namun, jika dibandingkan dengan penelitian [29] dan [19] yang mengombinasikan ADASYN, RFE, dan algoritma klasifikasi pada dataset yang berbeda, Performa pada penelitian ini masih berada di bawah hasil yang diperoleh pada penelitian kanker payudara dan penyakit tiroid. Perbedaan tersebut diduga dipengaruhi oleh karakteristik dataset yang berbeda, terutama tingkat ketidakseimbangan kelas pada dataset stroke yang lebih tinggi, sehingga proses klasifikasi menjadi lebih kompleks.

4. KESIMPULAN

Berdasarkan hasil penelitian, penerapan *Recursive Feature Elimination* (RFE) pada model *Support Vector Machine* (SVM) dengan penanganan ketidakseimbangan data menggunakan ADASYN memberikan perbaikan performa yang terbatas dibandingkan model tanpa RFE. Konfigurasi terbaik diperoleh pada rasio pembagian data 80:20 dengan kernel polynomial dan parameter $C = 0.1$, menghasilkan akurasi sebesar 0.75, precision 0.14, recall kelas stroke 0.82, F1-score 0.24, dan AUC 0.8245. Hasil tersebut menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam membedakan kelas stroke dan non-stroke serta mendeteksi sebagian besar kasus stroke, namun masih menghasilkan jumlah false positive yang cukup tinggi sebagaimana tercermin dari nilai precision yang rendah. Hasil seleksi fitur menunjukkan bahwa fitur *age*, *avg glucose level*, *bmi*, *work type*, dan *smoking status* terpilih secara konsisten, sedangkan *hypertension* dan *heart disease* tidak selalu terpilih karena proses seleksi didasarkan pada kontribusi fitur terhadap pembentukan model, bukan pada tingkat kepentingan klinisnya. Peningkatan performa yang relatif kecil mengindikasikan bahwa sebagian besar informasi penting dalam dataset telah terwakili oleh fitur awal, sehingga kontribusi utama RFE pada penelitian ini lebih terlihat pada penyederhanaan ruang fitur dibandingkan peningkatan performa klasifikasi yang substansial. Keterbatasan penelitian terletak pada kondisi ketidakseimbangan kelas yang masih tinggi serta rendahnya precision model, sehingga penelitian selanjutnya dapat difokuskan pada pengembangan pendekatan yang mampu meningkatkan precision dan mengurangi false positive melalui perbandingan metode seleksi fitur lain, penggunaan teknik penanganan ketidakseimbangan data yang berbeda maupun eksplorasi algoritma klasifikasi yang lebih sesuai untuk karakteristik dataset.

REFERENCES

- [1] World Health Organization, *Integrating stroke services in health-care systems : A practical approach*. 2020. [Online]. Available: <https://apps.who.int/iris/handle/10665/336231>
- [2] V. L. Feigin *et al.*, "World Stroke Organization : Global Stroke Fact Sheet 2025," *Int. J. Stroke*, vol. 20, no. 2, pp. 132–144, 2025, doi: 10.1177/17474930241308142.
- [3] Badan Penelitian dan Pengembangan Kesehatan, "Laporan Nasional Riskesdas 2018," 2019, *Jakarta*. [Online]. Available: <https://repository.kemkes.go.id/book/1323>
- [4] Badan Kebijakan Pembangunan Kesehatan, "Survei Kesehatan Indonesia (SKI) 2023 Dalam Angka," Jakarta, 2023. [Online]. Available: <https://www.badankebijakan.kemkes.go.id/ski-2023-dalam-angka/>
- [5] E. Wiyarta, R. Hidayat, M. Kurniawan, G. H. P. Sinaga, and R. A. Budiman, "Factors associated with prehospital and in-hospital delays in acute ischaemic stroke care in Indonesia: a systematic review," *BMJ Open*, vol. 15, no. 3, 2025, doi: 10.1136/bmjopen-2024-095845.
- [6] T. Vu, Y. Kokubo, M. Inoue, M. Yamamoto, and A. Mohsen, "Machine Learning Approaches for Stroke Risk Prediction : Findings from the Suita Study," *J. Cardiovasc. Dev. Dis.*, 2024, doi: doi.org/10.3390/jcdd11070207.
- [7] Alwalyanto, S. K. Gusti, I. Afrianty, and F. Syafria, "Penerapan Metode ADASYN Dalam Mengatasi Imbalanced Data Untuk Klasifikasi Penyakit Stroke Menggunakan Support Vector Machine," *Bull. Comput. Sci. Res.*, vol. 5, no. 4, pp. 532–541, 2025, doi: 10.47065/bulletincsr.v5i4.612.
- [8] M. S. Hamdard and H. Lodin, "Effect of Feature Selection on the Accuracy of Machine Learning Model," *Int. J. Multidiscip. Res. Anal.*, vol. 06, no. 09, pp. 4460–4466, 2023, doi: 10.47191/ijmra/v6-i9-66.
- [9] I. Guyon, "An Introduction to Variable and Feature Selection 1 Introduction," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, 2003.



- [10] E. S. Septiany, H. H. Handayani, T. Al Mudzakir, A. Fitri, and N. Masruriyah, "Optimasi Metode Support Vector Machine Menggunakan Seleksi Fitur Recursive Feature Elimination dan Forward Selection untuk Klasifikasi Kanker Payudara TIN : Terapan Informatika Nusantara," *TIN Terap. Inform. Nusant.*, vol. 5, no. 2, pp. 144–154, 2024, doi: 10.47065/tin.v5i2.5324.
- [11] H. Sundari, M. A. Amrustian, A. Dwi, and P. Wicaksono, "Penerapan Recursive Feature Elimination pada Support Vector Machine untuk Klasifikasi Kanker Payudara," *LEDGER J. Inform. Inf. Technol.*, vol. 3, no. 2, pp. 60–65, 2024, doi: 10.20895/ledger.v3i2.1591.
- [12] N. D. M. Zulkifli and M. A. As'ari, "Optimized Supervised Learning Framework for Gestational Diabetes Mellitus (GDM) Detection based on Recursive Feature Elimination (RFE)," *Malaysian J. Fundam. Appl. Sci.*, vol. 21, pp. 2967–2976, 2025, doi: doi.org/10.11113/mjfas.v21n6.4462.
- [13] W. Hidayat, T. A. Y. Siswa, and R. Hasudungan, "Penerapan Algoritma Naive Bayes , Recursive Feature Elimination , dan Adaptive Synthetic Sampling Pada Klasifikasi Penyakit Dermatitis," *J. INSTEK*, vol. 10, no. 2, pp. 424–436, 2025, doi: <https://doi.org/10.24252/instek.v10i2.59737>.
- [14] L. B. E. Susilo, T. A. Y. Siswa, and N. A. Verdikha, "Penerapan Algoritma Decision Tree, Recursive Elimination Dan ADASYN Pada Data Klasifikasi Penyakit Dermatitis," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 5, pp. 8791–8798, 2025, doi: 10.36040/jati.v9i5.15212.
- [15] H. Hamla, K. Ghanem, and M. Constantine, "A Hybrid Feature Selection Based on Fisher Score and SVM-RFE for Microarray Data," *Informatica*, vol. 48, pp. 57–68, 2024, doi: 10.31449/inf.v48i1.4759.
- [16] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning," *Proc. Int. Jt. Conf. Neural Networks*, no. 3, pp. 1322–1328, 2008, doi: 10.1109/IJCNN.2008.4633969.
- [17] N. G. Ramadhan, "Comparative Analysis of ADASYN-SVM and SMOTE-SVM Methods on the Detection of Type 2 Diabetes Mellitus," *Sci. J. Informatics*, vol. 8, no. 2, pp. 276–282, 2021, doi: 10.15294/sji.v8i2.32484.
- [18] H. N. Jasim, W. Mohammed Jasim, and M. S. Ibrahim, "Optimized Machine Learning Approach for Breast Cancer Detection Using ADASYN with Feature Selection," in *2025 5th International Conference on Emerging Smart Technologies and Applications (eSmarTA)*, 2025, pp. 1–7. doi: 10.1109/eSmarTA66764.2025.11132282.
- [19] C. Wulandari, I. Afrianty, E. Budianita, and S. K. Gusti, "Thyroid Disease Classification Using Support Vector Machine and Recursive Feature Elimination Method," *bit-Tech (Binary Digit. - Technol.)*, vol. 8, no. 2, 2025, doi: 10.32877/bt.v8i2.3454.
- [20] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "The KDD Process for Extracting Useful Knowledge from Volumes of Data," *Commun. ACM*, vol. 39, no. 11, pp. 27–34, 1996, doi: 10.1145/240455.240464.
- [21] J. Allgaier and R. Pryss, "Practical approaches in evaluating validation and biases of machine learning applied to mobile health studies," *Commun. Med.*, pp. 1–11, 2024, doi: 10.1038/s43856-024-00468-0.
- [22] Wijiyanto, A. I. Pradana, Sopingi, and V. Atina, "Teknik K-Fold Cross Validation untuk Mengevaluasi Kinerja Mahasiswa," *J. Algoritma. Inst. Teknol. Garut*, pp. 239–248, 2024, doi: 10.33364/algoritma/v.21-1.1618.
- [23] M. Tiara *et al.*, "Pemanfaatan Algoritma ADASYN Dan Support Vector Machine Dalam Meningkatkan Akurasi Prediksi Kanker Paru-Paru," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 5, pp. 8773–8778, 2024, doi: 10.36040/jati.v8i5.10752.
- [24] A. M. Priyatno, T. Widiyaningtyas, I. Engineering, and U. N. Malang, "A Systematic Literatur Review: Recursive Feature Elimination Algorithms," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 9, no. 2, pp. 196–207, 2024, doi: 10.33480/jitk.v9i2.5015.
- [25] N. Sulistianingsih, S. A. A. Yusuf, A. Anggreni, F. N. Sari, M. P. Juniarta, and J. Permatasari, "Analisis Pengaruh Recursive Feature Elimination Terhadap Kinerja Model Prediksi Dini Diabetes Mellitus di RS PKU Muhammadiyah Bima," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 7, no. 3, pp. 548–560, 2025, doi: 10.35746/jtim.v7i3.774.
- [26] M. A. Muslim *et al.*, *Data Mining Algoritma C4.5*, Pertama. Semarang, 2019.
- [27] A. S. Nugroho, A. B. Witarto, and D. Handoko, "Support Vector Machine," *Proceeding Indones. Sci. Meeting Cent. Japan*, 2003, doi: 10.18411/trmio-12-2023-769.
- [28] D. R. Nurqotimah, A. N. Khudori, and R. S. Pradini, "Implementasi Algoritma Support Vector Machine (SVM) Untuk Klasifikasi Penyakit Stroke," *J. Appl. Comput. Sci. Technol. (JACOST)*, vol. 5, no. 2, pp. 179–185, 2024, doi: 10.52158/jacost.v5i2.817.
- [29] H. N. Jasim, W. M. Jasim, and M. S. Ibrahim, "Optimized Machine Learning Approach for Breast Cancer Detection Using ADASYN with Feature Selection," pp. 1–7, 2025, doi: 10.1109/eSmarTA66764.2025.11132282.