

Penerapan Metode SVM Untuk Deteksi Dini Penyakit Stroke (Studi Kasus : RSUD Dr. H. Ishak Umarella Maluku Tengah dan RS Sumber Hidup-GPM)

Berny Pebo Tomasouw^{1*}, Francis Yunito Rumlawang¹

¹Jurusan Matematika FMIPA Universitas Pattimura, Jl. Ir. M. Putuhena, Ambon. Indonesia.

*Email: bptomasouw@gmail.com

Manuscript submitted : March 2023

Accepted for publication : April 2023

doi : <https://doi.org/10.30598/tensorvol4iss1pp37-44>

Abstract: Stroke is a significant health problem in today's modern society. Early detection of stroke usually takes a long time. To prevent the risk of a significant disabling stroke, it is good to pay attention and recognize the symptoms of a stroke early on. In this study, the Support Vector Machine (SVM) method was used to detect stroke based on stroke risk factors, including blood pressure, age, LDL, and blood sugar. Based on the results obtained, the nonlinear SVM method has a better level of accuracy than the linear SVM. This is because of the two data-sharing schemes, the linear SVM only has an accuracy rate of 81.25%, while the nonlinear SVM has an accuracy rate of 84.38%. Especially for the nonlinear SVM, the RBF kernel has a better level of accuracy than the polynomial kernel. This can be seen from the results of testing the two data-sharing schemes, the RBF kernel has the best results, namely the highest accuracy rate of 84.38% and 84% respectively

2010 Mathematical Subject Classification : 9010

Keywords: early detection, stroke, SVM.

1. Pendahuluan

Stroke merupakan masalah kesehatan yang utama bagi masyarakat modern saat ini. Di samping masalah-masalah kesehatan lainnya, stroke semakin menjadi masalah serius yang dihadapi hampir diseluruh dunia. Hal tersebut dikarenakan serangan stroke yang mendadak dapat mengakibatkan kematian, kecacatan fisikdan mental baik pada usia produktif maupun usia lanjut [1]. Menurut World Health Organization (WHO), kematian akibat stroke sebesar 51% di seluruh dunia disebabkan oleh tekanan darah tinggi. Selain itu, diperkirakan sebesar 16% kematian stroke disebabkan tingginya kadar glukosa darah dalam tubuh. Tingginya kadar gula darah dalam tubuh secara patologis berperan dalam peningkatan konsentrasi glikoprotein, yang merupakan pencetus beberapa penyakit vaskuler. Kadar glukosa darah yang tinggi pada saat stroke akan memperbesar kemungkinan meluasnya area infark karena terbentuknya asam laktat akibat

metabolisme glukosa secara anaerobik yang merusak jaringan otak [2].

Pendeteksian awal penyakit stroke biasanya membutuhkan waktu yang cukup lama. Dengan kemajuan teknologi, stroke dapat dicegah dengan adanya pendeteksian resikonya sejak dini sehingga dapat mendapatkan penanganan dengan cepat dan memperbesar kemungkinan penyembuhan. Keuntungan lainnya dengan adanya pendeteksian secara cepat yaitu efisiensi untuk biaya perawatan, dan semakin banyak orang yang dapat mendiagnosa resiko tersebut secara cepat [3]. Untuk mencegah resiko stroke besar yang melumpuhkan, ada baiknya memperhatikan dan mengenali gejala-gejala stroke sejak dini.

Oleh karena itu, perlu dikaji suatu metode yang mampu mendeteksi secara dini penyakit stroke berdasarkan gejala yang dialami. Metode yang diusulkan untuk dipakai dalam penelitian ini adalah metode *Support Vector Machine* (SVM). SVM adalah salah satu algoritma dalam *machine learning* yang digunakan untuk pemodelan dan klasifikasi data. SVM termasuk metode pembelajaran yang terarah, yang berarti SVM menggunakan data yang telah dilabeli dengan benar untuk melatih modelnya. SVM sendiri memiliki beberapa kelebihan antara lain, dapat mengatasi masalah klasifikasi linear dan nonlinier dengan menggunakan fungsi kernel, sehingga mampu menemukan pemisah yang optimal, efektif dalam menangani data dengan dimensi tinggi, dan mampu menangani masalah overfitting karena mencari pemisah dengan margin maksimum sehingga dapat mengenali dengan baik pada data yang belum pernah dilihat sebelumnya.

SVM sendiri telah berhasil diaplikasikan dalam problema dunia nyata (*real-world problems*), dan memberikan solusi yang lebih baik untuk kasus tertentu dibandingkan metode konvensional seperti misalnya jaringan syaraf tiruan. Dalam [4] dijelaskan bahwa Konsep dasar SVM bermula dari masalah klasifikasi dua kelas yang tergolong dalam masalah linier (*linear problem*). SVM berusaha menemukan *hyperplane* (pemisah) terbaik pada input space untuk memisahkan dua kelas tersebut. Selanjutnya SVM dikembangkan agar dapat bekerja pada masalah non-linear, dengan memasukkan konsep kernel pada ruang kerja berdimensi tinggi. Dalam mencari *hyperplane* terbaik, SVM akan dihadapkan pada proses mencari solusi dari masalah pemrograman kuadratik. Penelitian [5]-[13] memperlihatkan bahwa SVM memiliki hasil yang lebih baik dalam masalah klasifikasi.

2. Metode Penelitian

Data yang digunakan dalam penelitian ini adalah data rekam medis pasien yang memuat faktor-faktor penyebab resiko stroke yang terdiri dari tekanan darah, umur, LDL, dan gula darah. Sebelum data ini diproses menggunakan metode SVM, maka data harus distandarisasi ke dalam skala [-1, 1] sehingga semua nilai variabel input akan menjadi seragam dan juga akan mengoptimalkan hasil yang diperoleh dari metode SVM. Rumus yang digunakan untuk standarisasi data adalah sebagai berikut:

$$x_{new} = -1 + 2 \left(\frac{x - x_{min}}{x_{max} - x_{min}} \right)$$

Pada Tabel 1. diperlihatkan contoh hasil standarisasi beberapa data

Tabel 1. Standarisasi Data

No.	Data Sebelum Standarisasi				No.	Data Setelah Standarisasi			
1	130	43	184	107	1	-0.0329	-0.8498	0.4742	-0.2488
2	160	40	117	100	2	0.2488	-0.8779	-0.1549	-0.3146
3	140	49	100	115	3	0.0610	-0.7934	-0.3146	-0.1737

2.1. SVM Linier

Misalkan terdapat m data *pelatihan* $\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ dimana $x_i \in \mathbf{R}^n$ adalah sampel data dan $y_i \in \{1, -1\}$ adalah target atau kelas dari sampel data. Misalkan juga bahwa data untuk kedua kelas terpisah secara linier (*linearly separable*) maka ingin dicari fungsi pemisah

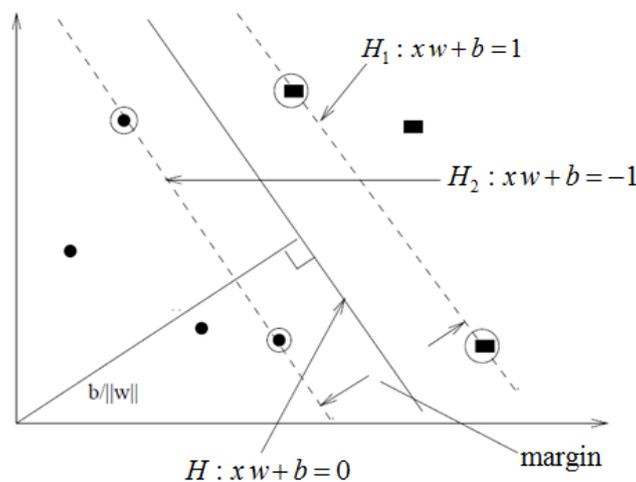
$$f(x) = xw + b = 0 \tag{1}$$

dimana $w \in \mathbf{R}^{n \times 1}$ adalah parameter bobot dan $b \in \mathbf{R}$ adalah parameter bias serta berlaku

$$\begin{aligned} x_i w + b &> 0 && \text{untuk } y_i = 1 \\ x_i w + b &< 0 && \text{untuk } y_i = -1 \end{aligned} \tag{2}$$

Misalkan $H : xw + b = 0$ adalah pemisah yang ingin dicari sedangkan $H_1 : xw + b = 1$ dan $H_2 : xw + b = -1$ adalah pemisah dari klas 1 dan klas -1. Untuk mendapatkan H yang optimal maka jarak H_1 dan H_2 ke H haruslah sama dengan syarat bahwa tidak ada sampel data antara H_1 dan H_2 serta jarak H_1 ke H_2 adalah jarak yang maksimal [17].

Untuk memaksimalkan jarak H_1 dan H_2 maka digunakan sampel data positif yang terletak pada H_1 dan sampel data negatif yang terletak pada H_2 . Sampel data ini disebut *support vector* karena fungsinya sebagai penentu dalam mendapatkan pemisah yang optimal. Sedangkan sampel data yang lain dapat dibuang atau digerakkan menuju H_1 dan H_2 asalkan tidak melewati masing-masing pemisah.



Gambar 1. Pemisah optimal yang memisahkan dua kelas.

Misalkan $(x_0, y_0) \in \mathbb{R}^2$ sebarang titik maka jarak titik ini ke garis $Ax + by + C = 0$ adalah

$$\frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}} \tag{3}$$

sehingga jarak sampel data x yang terletak pada H_1 ke H adalah

$$\frac{|xw + b|}{\sqrt{w^T w}} = \frac{1}{\|w\|} \tag{4}$$

Karena jarak H_1 dan H_2 ke H adalah sama maka jarak H_1 ke H_2 adalah $\frac{2}{\|w\|}$.

Dengan demikian, masalah mencari parameter w dan b yang optimal agar diperoleh pemisah yang optimal merupakan masalah pemrograman kuadratik

$$\min_{w,b} \frac{1}{2} w^T w \tag{5}$$

dengan kendala

$$y_i (x_i w + b) \geq 1, \quad i = 1, \dots, m$$

Biasanya agak sulit menyelesaikan bentuk primal di atas, sehingga bentuk primal tersebut diubah ke bentuk dualnya dengan memperkenalkan pengali Lagrange.

Misalkan $\alpha \in \mathbb{R}^{m \times 1}$ adalah pengali Lagrange maka masalah pemrograman kuadratik (5) di atas berubah menjadi

$$L(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{i=1}^m \alpha_i (y_i (x_i w + b)) + \sum_{i=1}^m \alpha_i \tag{6}$$

Dengan menerapkan syarat Karush-Kuhn-Tucker (KKT) maka bentuk dual yang diperoleh adalah

$$\text{Max} \quad L(\alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m \alpha_i \alpha_j y_i y_j x_i x_j^T \tag{7}$$

dengan kendala

$$\sum_{i=1}^m \alpha_i y_i = 0 \quad \text{dan} \quad \alpha_i \geq 0 \quad \text{dimana} \quad i = 1, 2, \dots, m$$

Parameter bobot dan bias dapat dihitung dengan persamaan

$$w = \sum_{i=1}^{N_{sv}} \alpha_i y_i x_i \quad \text{dan} \quad b = \frac{1}{N_{sv}} \sum_{i=1}^{N_{sv}} (y_i - x_i w) \tag{8}$$

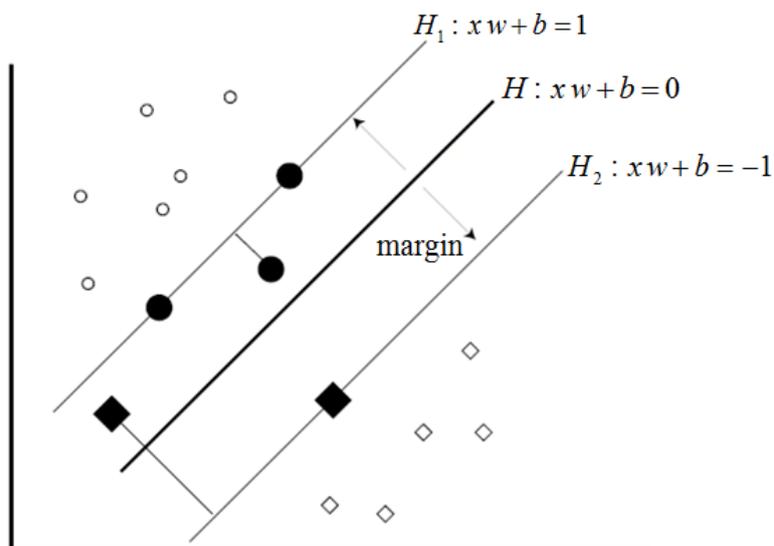
SV adalah himpunan *support vector* dan $i \in SV$ jika $\alpha_i \neq 0$. N_{sv} adalah jumlah *support vector*. Dengan menggunakan persamaan

$$f(x) = x w + b$$

maka data input $x \in \mathbb{R}^n$ yang baru diklasifikasikan menjadi

$$\begin{cases} \text{kelas} +1, & \text{jika } f(x) > 0 \\ \text{kelas} -1, & \text{jika } f(x) < 0 \end{cases} \tag{9}$$

Selanjutnya jika terjadi kasus pemisahan yang tidak sempurna, dalam hal ini terdapat data antara H_1 dan H_2 seperti tampak pada Gambar di bawah ini.



Gambar 2. Contoh kasus pemisahan data yang tidak sempurna

Untuk mengatasi masalah ini, akan diperkenalkan variabel *slack* μ yang tak negatif ($\mu \geq 0$) dan dimasukkan pada fungsi kendala (5) sehingga menjadi

$$y_i (x_i w + b) \geq 1 - \mu_i \quad , \quad i = 1, \dots, m$$

Sedangkan pada fungsi objektifnya ditambahkan parameter positif C sehingga menjadi

$$\frac{1}{2} w^T w + C \sum_{i=1}^m \mu_i$$

Secara lengkap, masalah pemrograman kuadrat (5) berubah menjadi

$$\min_{w,b} \frac{1}{2} w^T w + C \sum_{i=1}^m \mu_i$$

dengan kendala

(10)

$$y_i (x_i w + b) \geq 1 - \mu_i, \quad \mu_i \geq 0, \quad i = 1, \dots, m$$

Dengan menggunakan pengali Lagrange $\alpha \in \mathbb{R}^{m \times 1}$ maka bentuk primal (10) dapat diubah menjadi bentuk dual sebagai berikut

$$\text{Max} \quad L(\alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m \alpha_i \alpha_j y_i y_j x_i x_j^T$$

dengan kendala

(11)

$$\sum_{i=1}^m \alpha_i y_i = 0 \quad \text{dan} \quad 0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, m$$

2.2. SVM Nonlinier

Dalam masalah klasifikasi kebanyakan sampel data tidak terpisah secara linier sehingga jika digunakan SVM linier maka hasil yang diperoleh tidak optimal dan mengakibatkan hasil klasifikasi yang buruk. Yang menjadi salah satu keunggulan SVM terletak pada bagian ini, yakni SVM dapat diperluas untuk menyelesaikan masalah non-linier. SVM linier dapat diubah menjadi SVM non-linier dengan menggunakan metode kernel [14]. Metode ini bekerja dengan cara memetakan data input ke ruang *feature* yang dimensinya lebih tinggi menggunakan fungsi sebuah ϕ . Sebagai contoh, misalkan $u = (u_1, u_2)$ adalah data input pada \mathbb{R}^2 dan $\phi(u) = (1, \sqrt{2}u_1, \sqrt{2}u_2, u_1^2, u_2^2, \sqrt{u_1} u_2)$ adalah data input pada ruang *feature* yang berdimensi lebih tinggi yakni \mathbb{R}^5 . Diharapkan data input hasil pemetaan ke ruang *feature* akan terpisah secara linier sehingga dapat dicari pemisah yang optimal.

Misalkan $x \rightarrow \phi(x)$ maka Persamaan (11) dapat ditulis menjadi

$$\text{Max} \quad \psi(\alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m \alpha_i \alpha_j y_i y_j \phi^T(x_i) \phi(x_j)$$

dengan kendala

(12)

$$\sum_{i=1}^m \alpha_i y_i = 0 \quad \text{dan} \quad 0 \leq \alpha_i \leq C \quad \text{dimana} \quad i = 1, 2, \dots, m$$

Parameter bobot dan bias dapat dihitung dengan persamaan

$$w = \sum_{i=1}^{N_{SV}} \alpha_i y_i \phi(x_i) \quad \text{dan} \quad b = \frac{1}{N_{SV}} \sum_{i=1}^{N_{SV}} (y_i - w^T \phi(x_i)) \quad (13)$$

Sedangkan *pemisah* optimal pada Persamaan (2.1) berubah menjadi

$$f(x) = w^T \phi(x) + b = 0 \quad (14)$$

Permasalahan yang muncul adalah jika sampel data input untuk tahap *training* dalam jumlah besar maka perhitungan hasil kali titik $\phi^T(x_i) \phi(x_j)$ pada (2.16) akan membuat waktu perhitungan semakin lama.

Oleh karena itu diperlukan cara untuk menghitung $\phi^T(x_i) \phi(x_j)$ tanpa mengetahui fungsi ϕ .

Misalkan K adalah sebuah fungsi dengan sifat

$$K(u, v) = \phi^T(u) \phi(v)$$

dimana $u, v \in \mathbb{R}^n$ dan $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $n < m$. Fungsi K ini disebut sebagai fungsi kernel. Menurut [14], fungsi kernel yang sering digunakan adalah sebagai berikut :

- kernel linier : $K(u, v) = u^T v$.
- kernel polinomial : $K(u, v) = (1 + u^T v)^d$, $d \geq 2$.
- kernel RBF (Radial Basis Function) : $K(u, v) = \exp(-\gamma \|u - v\|^2)$, $\gamma > 0$.

Dengan menggunakan konsep fungsi kernel di atas maka Persamaan (12) berubah menjadi

$$\text{Max } \psi(\alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

dengan kendala

$$\sum_{i=1}^m \alpha_i y_i = 0 \quad \text{dan} \quad 0 \leq \alpha_i \leq C \quad \text{dimana} \quad i = 1, 2, \dots, m$$

parameter bias dapat dihitung dengan persamaan

$$b = \frac{1}{N_{SV}} \sum_{i=1}^{N_{SV}} \left(y_i - \sum_{i=1}^{N_{SV}} \alpha_i y_i K(x_i, x_j) \right)$$

sedangkan pemisah optimalnya

$$f(x) = \sum_{i=1}^{N_{SV}} \alpha_i y_i K(x_i, x) + b$$

Algoritma 1. Algoritma Klasifikasi dengan SVM [14]

- Normalisasi data *pelatihan* $\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ dimana $x_i \in \mathbf{R}^n$ adalah sampel data dan $y_i \in \{1, -1\}$.
- Tetapkan parameter C dan kernel yang akan digunakan.
- Selesaikan masalah pemrograman kuadratik (15) untuk mendapatkan nilai α_i .
- Dapatkan parameter bias dengan (16) dan pemisah optimal menggunakan (17).
- Data input $x \in \mathbf{R}^n$ yang baru tetap diklasifikasikan berdasarkan (9).

3. Hasil dan Pembahasan

Keseluruhan data yang telah distandarisasi akan dibagi ke dalam dua kelompok, yakni data pelatihan dan data pengujian. Dalam penelitian ini akan dibuat dua skema pembagian data yakni 60/40 dan 70/30 sehingga bisa diketahui tingkat akurasi untuk setiap skema. Sebagai contoh, skema 60/40 berarti data pelatihan merupakan 60% data dari total keseluruhan data sedangkan sisa 40% merupakan data pengujian.

3.1. Hasil Menggunakan Metode SVM Linier

Proses pelatihan untuk kedua skema pembagian data dilakukan dengan menggunakan metode SVM linier dimana parameter $C = 1$. Sedangkan tingkat akurasi dari hasil pengujian untuk kedua skema pembagian data dapat dilihat pada Tabel 2.

Tabel 2. Tingkat Akurasi Dengan Metode SVM Linier

No.	Skema Pembagian	Tingkat Akurasi
1	60/40	81.25%
2	70/30	76%

Pada tabel di atas terlihat bahwa tingkat akurasi tertinggi adalah 81.25% pada skema pembagian data 60/40.

3.2. Hasil Menggunakan Metode SVM Nonlinier

Untuk metode SVM nonlinier, akan digunakan dua buah kernel sebagai pembanding yakni kernel *Polynomial* dan kernel RBF (*Radial Basis Function*). Nilai parameter yang akan dibuat bervariasi pada kernel polinomial adalah nilai d (*degree*) sedangkan pada kernel RBF adalah nilai γ (*gama*). Sedangkan parameter C yang digunakan dalam proses pelatihan sama dengan nilai sebelumnya yakni 1. Hasil pengujian data menggunakan SVM nonlinier untuk kedua skema pembagian data dapat dilihat pada Tabel 3 dan Tabel 4.

Tabel 3. Tingkat Akurasi dengan Skema Pembagian Data 60/40

Jenis Kernel : Kernel polinomial								
Parameter Kernel	$d = 1$	$d = 2$	$d = 3$	$d = 4$	$d = 5$	$d = 6$	$d = 7$	$d = 8$
Tingkat Akurasi	81.25%	78.13%	78.13%	65.63%	65.63%	78.13%	71.88%	68.75%
Jenis Kernel : Kernel RBF (<i>Radial Basis Function</i>)								
Parameter Kernel	$\gamma = 1.5$	$\gamma = 2$	$\gamma = 2.5$	$\gamma = 2.6$	$\gamma = 2.7$	$\gamma = 2.8$	$\gamma = 2.9$	$\gamma = 3$
Tingkat Akurasi	81.25%	78.13%	75%	81.25%	81.25%	78.13%	84.38%	75%

Dari Tabel 3 terlihat bahwa hasil pengujian dengan kernel *Polynomial* memperoleh hasil terbaik pada nilai parameter $d = 1$, dengan tingkat akurasi sebesar 81.25% , kemudian selanjutnya dengan menggunakan kernel RBF diperoleh tingkat akurasi sebesar 84.38% untuk $\gamma = 2.9$.

Tabel 4.4. Tingkat Akurasi dengan Skema Pembagian Data 70/30

Jenis Kernel : Kernel polinomial								
Parameter Kernel	$d=1$	$d = 2$	$d = 3$	$d = 4$	$d = 5$	$d = 6$	$d = 7$	$d = 8$
Tingkat Akurasi	76%	76%	76%	72%	72%	72%	72%	72%
Jenis Kernel : Kernel RBF (<i>Radial Basis Function</i>)								
Parameter Kernel	$\gamma = 1.5$	$\gamma = 2$	$\gamma = 2.5$	$\gamma = 2.6$	$\gamma = 2.7$	$\gamma = 2.8$	$\gamma = 2.9$	$\gamma = 3$
Tingkat Akurasi	76%	76%	72%	72%	72%	76%	84%	76%

Dari Tabel 4 terlihat bahwa hasil pengujian dengan kernel polinomial memperoleh hasil terbaik pada nilai parameter $d = 1$, $d = 2$, dan $d = 3$ dengan tingkat akurasi sebesar 76%, kemudian selanjutnya dengan menggunakan kernel RBF diperoleh tingkat akurasi sebesar 84% untuk $\gamma = 2.9$.

Berdasarkan hasil yang diperoleh, metode SVM nonlinier memiliki tingkat akurasi yang lebih baik dibandingkan SVM linier. Hal ini dikarenakan dari kedua skema pembagian data, SVM linier hanya memiliki tingkat akurasi sebesar 81.25% sedangkan SVM nonlinier memiliki tingkat akurasi sebesar 84.38%. Khusus untuk SVM nonlinier, kernel RBF memiliki tingkat akurasi yang lebih baik dibandingkan kernel polinomial. Hal ini terlihat dari hasil pengujian kedua skema pembagian data, kernel RBF memiliki hasil terbaik yakni tingkat akurasi tertinggi masing-masing sebesar 84.38% dan 84% untuk nilai parameter $\gamma = 2.9$.

4. Kesimpulan

Berdasarkan proses pelatihan dan hasil pengujian dengan menggunakan metode SVM maka nilai parameter C yang terbaik adalah $C = 1$. Sedangkan parameter yang terbaik untuk fungsi kernel RBF adalah $\gamma = 2.9$. Hal ini terlihat dari tingkat akurasi yang dihasilkan sebesar 84.38% dan 84%.

References

- [1] Purwaningtiyas, D. P., Kusumawati, Y., & Nugroho, F. S. (2015). Hubungan antara gaya hidup dengan kejadian Stroke usia dewasa muda Di RSUD Dr. Moewardi Surakarta. *Prosiding Seminar Nasional Fakultas Ilmu Kesehatan*.
- [2] Rico, J., Suharyo, H., & Endang, K. (2008). Faktor-faktor yang mempengaruhi kejadian Stroke pada usia muda kurang dari 40 tahun. *Jurnal Epidemiologi*, 1-13.
- [3] Adelina, V., Ratnawati, D E., & Fauzi, M. A. (2018). Klasifikasi tingkat risiko penyakit stroke menggunakan metode GA-Fuzzy Tsukamoto. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(9), 3015-3021.
- [4] Tomasouw, B. P. & Irawan, M. I. (2012). Multiclass Twin Bounded Support Vector Machine untuk pengenalan ucapan. *Seminar Nasional Penelitian, Pendidikan dan Penerapan MIPA*, pp.49-58.
- [5] Leleury, Z. A. & Tomasouw, B. P. (2015). Diagnosa penyakit saluran pernapasan dengan menggunakan Support Vector Machine (SVM). *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 9(2), 109-119.
- [6] Damasela, R., Tomasouw, B. P., & Leleury, Z. A. (2021). Penerapan metode Support Vector Machine (SVM) untuk mendeteksi penyalahgunaan narkoba. *PARAMETER: Jurnal Matematika, Statistika dan Terapannya*, 1(2).
- [7] Hmeidi, I., Hawashin, B. & El-Qawasmeh, E. (2014). Performance of KNN and SVM classifiers on full word Arabic articles", *Advanced Engineering Informatics*, 22(1), 106-111.
- [8] Qiu, Y. (2014). Towards prediction of pancreatic cancer using SVM study model. *JSM Clinical Oncology and Research*, 2(4), 1031, 1-6.
- [9] Utami, I. T. (2018). Perbandingan Kinerja Klasifikasi Support Vector Machine (SVM) Dan Regresi Logistik Biner dalam Mengklasifikasikan Ketepatan Waktu Kelulusan Mahasiswa FMIPA UNTAD. *Jurnal Ilmiah Matematika dan Terapan*, 15 (2), 256-267.
- [10] Wahyudi, R. & Kusumawardhana, G. (2021). Analisis Sentimen pada review Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine. *JURNAL INFORMATIKA*, 8 (2), 200-207.
- [11] Anjasmos, M. T., Istiadi, I. & Marisa, F. (2020). Analisis Sentimen Aplikasi Go-Jek Menggunakan Metode SVM Dan NBC (Studi Kasus: Komentar Pada Play Store). *The 3rd Conference on Innovation and Application of Science and Technology (CIASTECH 2020)*. pp. 489-498.
- [12] Dong, S. (2021). Multi class SVM algorithm with active learning for network traffic classification. *Journal Expert Systems with Applications*, 176, 114885.
- [13] Onel, M., Kieslich, C. A., & Pistikopoulos, E. N. (2018). A nonlinear support vector machine-based feature selection approach for fault detection and diagnosis: Application to the Tennessee Eastman process. *AIChE Journal*, 65 (3), 992-1005.
- [14] Abe, S. (2010). *Support Vector Machines for Pattern Classification* (2nd Ed.). London: Springer-Verlag.