

# Implementasi Metode Support Vector Machine Pada Klasifikasi Diagnosis Penyakit Hipertensi

Hilda Zaqya Elnaz Putri\*, Hisyam Fahmi

Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia

hildazaqyaelnaz@gmail.com\*, hisyam.fahmi@uin-malang.ac.id

## Abstrak

Hipertensi merupakan salah satu penyakit yang menjadi penyebab utama kematian di dunia. Penyakit ini sering kali disebut *the silent killer* karena penyakit ini dapat menyebabkan kematian tanpa disadari yang mana orang yang mengidap hipertensi tidak memiliki keluhan. Oleh karena itu, deteksi dan penanganan dini hipertensi adalah hal yang sangat penting. Penelitian ini bertujuan untuk mendapatkan klasifikasi penyakit hipertensi menggunakan metode *Support Vector Machine* (SVM) dengan menggunakan berbagai atribut seperti usia, kebiasaan merokok, gaya hidup, tekanan darah, dan diagnosis hipertensi serta mengetahui tingkat akurasi dari hasil klasifikasi penyakit hipertensi dengan metode SVM. Metode SVM dilatih dengan berbagai parameter kernel dan *hyperparameter* untuk menemukan model terbaik. Hasil penelitian menunjukkan bahwa model terbaik untuk klasifikasi penyakit hipertensi dengan metode SVM menggunakan kernel RBF dengan parameter  $C = 100$  dan  $\gamma$  (*gamma*) = 0,1 dengan akurasi sebesar 97,15%. Hal ini menunjukkan bahwa metode SVM mampu mengklasifikasikan penyakit hipertensi dengan sangat baik serta mampu memberikan kontribusi yang signifikan dalam deteksi dini dan penanganan penyakit hipertensi.

**Kata kunci:** Hipertensi; *Support Vector Machine*; SVM; *Data Mining*; Klasifikasi;

---

## Abstract

Hypertension is one of the leading causes of death worldwide. This disease is often referred to as the silent killer because it can lead to death without noticeable symptoms, leaving those affected unaware of their condition. Therefore, early detection and management of hypertension are crucial. This research aims to obtain the classification of hypertension using the *Support Vector Machine* (SVM) method by utilizing various attributes such as age, smoking habits, lifestyle, blood pressure, and hypertension diagnosis, as well as determining the accuracy level of hypertension classification results using the SVM method. The SVM method is trained with various kernel parameters and hyperparameters to find the best model. The research findings indicate that the best model for classifying hypertension using the SVM method employs the RBF kernel with parameters  $C = 100$  and  $\gamma$  (*gamma*) = 0,1, achieving an accuracy of 97.15%. This demonstrates that the SVM method is capable of classifying hypertension very well and significantly contributes to the early detection and management of hypertension.

**Keywords:** Hypertension; *Support Vector Machine*; SVM; *Data Mining*; Classification;

---

## PENDAHULUAN

Kemajuan teknologi memiliki peranan yang penting dalam hampir seluruh aspek kehidupan. Pertumbuhan yang pesat ini tak terlepas dari adanya pengaruh dari komputer. Komputer telah

mengubah segala hal menjadi lebih efisien dan lebih cepat melalui sistemnya. Salah satu aspek penting dari pengolahan informasi ini adalah *data mining*, dimana komputer menggunakan teknik-teknik khusus untuk menggali wawasan berharga dari kumpulan data besar menjadi informasi yang berharga. *Data mining* dapat digunakan untuk mengekstrak informasi dari data yang berjumlah besar menjadi bentuk yang berguna dan dapat dipahami [8].

*Data mining* bertujuan untuk mengidentifikasi pola yang dapat membantu dalam pengambilan keputusan, klasifikasi, dan prediksi. Klasifikasi merupakan teknik pengelompokan *data mining* ke dalam kelas-kelas tertentu atau kategori berdasarkan karakteristiknya [9]. Salah satu algoritma/metode yang terkenal dalam melakukan klasifikasi adalah *Support Vector Machine* (SVM). Metode SVM bertujuan untuk mencari *hyperplane* atau fungsi pemisah (*decision boundary*) terbaik yang dapat memisahkan dua atau lebih kelas data pada ruang *input*. Dalam melakukan klasifikasi, SVM memiliki konsep yang jelas secara sistematis dan konsisten jika dibandingkan dengan algoritma-algoritma klasifikasi lainnya. Metode SVM memiliki keunggulan, yaitu dapat menghasilkan nilai akurasi yang baik meskipun data yang dianalisis bersifat tidak seimbang (*unbalanced data*), dapat mengurangi kesalahan klasifikasi untuk sampel yang tidak diketahui dan meningkatkan kemampuan generalisasi dibandingkan dengan metode JST (Jaringan Syaraf Tiruan) dan kepekaan terhadap *outlier* [5].

*Data mining* dapat digunakan untuk melakukan klasifikasi dan prediksi suatu penyakit. Terdapat beberapa penelitian terdahulu yang menggunakan *data mining* dalam mengklasifikasikan atau memprediksi suatu penyakit. Pertama, penelitian yang dilakukan oleh Andriansyah, yaitu klasifikasi hipertensi menggunakan metode *Fuzzy Decision Tree Iterative Dichotomiser 3* (ID3). Penelitian tersebut menggunakan 12 atribut, yaitu umur, tekanan darah, lingkaran perut, tinggi badan, berat badan, berat massa indeks, merokok, konsumsi gula, konsumsi garam, olahraga dan konsumsi kafein dan diagnosis hipertensi. Dari penelitian tersebut dihasilkan akurasi sebesar 75,56% [1]. Kedua, penelitian yang dilakukan oleh Tino, yaitu membandingkan algoritma SVM dan *Neural Network* untuk mengklasifikasikan penyakit jantung. Dari percobaan terhadap kedua metode tersebut, metode SVM menghasilkan nilai akurasi 83% dan metode *Neural Network* menghasilkan nilai akurasi 82%. Hal ini mengindikasikan bahwa metode SVM memiliki model yang lebih akurat daripada *Neural Network* [11]. Maka dari itu, dapat disimpulkan bahwa metode *Support Vector Machine* mempunyai kemampuan yang baik dalam melakukan suatu klasifikasi atau prediksi penyakit. Salah satunya untuk pengklasifikasian diagnosis penyakit hipertensi.

Hipertensi adalah salah satu penyakit yang menjadi penyebab utama kematian di dunia. Penyakit ini sering kali disebut *the silent killer* karena penyakit ini dapat menyebabkan kematian tanpa disadari yang mana orang yang mengidap hipertensi tidak memiliki keluhan. Hipertensi adalah sebuah istilah medis dari penyakit tekanan darah tinggi. Umumnya, hipertensi merupakan kondisi asimtomatik (keadaan tanpa gejala) dimana tekanan di arteri tinggi secara tidak normal yang berpotensi dapat meningkatkan peluang terjadinya stroke, aneurisma, gagal jantung, serangan jantung, kerusakan ginjal, dan kematian [4].

Berdasarkan data Riset Kesehatan Dasar 2018, tingkat kematian masyarakat Indonesia akibat hipertensi sebanyak 427.218 kematian [7]. Lalu berdasarkan data Badan Pusat Statistik Kota Malang, hipertensi menduduki peringkat kedua sebagai salah satu penyakit dengan jumlah kasus terbanyak di Kota Malang pada tahun 2019 yaitu sebanyak 13.102 kasus [2]. Oleh karena itu, deteksi dan penanganan dini hipertensi adalah hal yang sangat penting. Kemajuan teknologi saat ini telah memudahkan pendiagnosaan penyakit, sehingga penanganan dapat dilakukan lebih cepat dan peluang sembuh pun meningkat. Salah satu cara untuk meningkatkan akurasi dan efisiensi para ahli medis dalam mengklasifikasikan suatu penyakit adalah dengan mempertimbangkan sistem klasifikasi yang dikembangkan menggunakan *data mining*.

Berdasarkan yang telah dipaparkan di atas, maka penulis memilih metode *Support Vector Machine* (SVM) untuk mengklasifikasikan diagnosis penyakit hipertensi. Penelitian ini akan memanfaatkan *tools Jupyter Notebook* dan bahasa pemrograman *python* sebagai alat bantu

penelitian. Penelitian ini diharapkan dapat membantu ahli kesehatan dalam menentukan diagnosis penyakit hipertensi.

## METODE PENELITIAN

### Data dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder. Data sekunder merupakan data yang diperoleh dari sumber yang telah tersedia, seperti buku, jurnal atau situs *web*. Penelitian ini menggunakan data sekunder yaitu berupa data *soft file* PTM (Penyakit Tidak Menular) hipertensi di Puskesmas Mojolangu Kota Malang. Jumlah data yang digunakan yaitu sebanyak 1373 data. Tabel 1 menyajikan variabel-variabel yang digunakan dalam penelitian ini.

Tabel 1 Variabel Penelitian

No.	Variabel	Keterangan
1.	Usia	Usia pasien pada saat penelitian
2.	Jenis Kelamin	Laki-laki/Perempuan
3.	Kebiasaan Merokok	Ya/Tidak
4.	Kurang Aktivitas Fisik (Min. 30 Menit/Hari)	Ya/Tidak
5.	Konsumsi Gula Berlebih (>4 Sendok Makan/Hari)	Ya/Tidak
6.	Konsumsi Garam Berlebih (>1 Sendok Teh/Hari)	Ya/Tidak
7.	Konsumsi Lemak Berlebih (>4 Potong Gorengan/Hari)	Ya/Tidak
8.	Kurang Makan Buah dan Sayur (<5 Porsi/Hari)	Ya/Tidak
9.	Konsumsi Alkohol	Ya/Tidak
10.	Tekanan Darah Sistolik	Tekanan darah saat jantung memompa darah Contoh : 140 mmHg
11.	Tekanan Darah Diastolik	Tekanan darah saat jantung beristirahat Contoh: 80 mmHg
12.	Diagnosis Hipertensi	0: Tidak Berpenyakit hipertensi 1: Berpenyakit Hipertensi

### Tahapan Penelitian

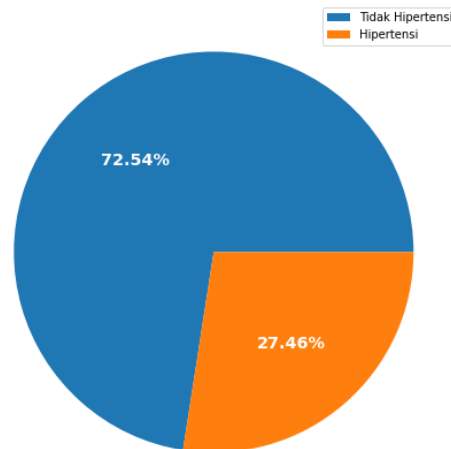
1. Pengumpulan data penyakit hipertensi yang terdiri dari variabel dependen dan independen.
2. Melakukan *Exploratory Data Analysis* (EDA) untuk mengetahui informasi-informasi serta *insight* yang ada di dalam data.
3. Melakukan *preprocessing data*  
Terdapat beberapa tahapan dalam proses *preprocessing data* yaitu:
  - a. *Data Cleaning*. *Data cleaning* adalah proses mendeteksi dan memperbaiki (atau menghapus) data yang salah, rusak, duplikat atau format yang tidak sesuai.
  - b. *Data Transformation*. Dalam hal ini, akan diubah atribut yang memiliki tipe data kategorikal menjadi numerik.

- c. *Oversampling Data*. *Oversampling data* adalah teknik yang digunakan untuk meningkatkan jumlah sampel dari kelas minoritas dalam kumpulan data yang tidak seimbang. Teknik *oversampling* yang akan digunakan SMOTE (*Synthetic Minority Oversampling Technique*).
4. Membagi data menjadi dua bagian, yaitu data *training* dan data *testing* dengan rasio 80:20. Data *training* akan digunakan untuk model klasifikasi dan data *testing* akan digunakan untuk evaluasi.
5. Melakukan proses klasifikasi menggunakan metode SVM  
Proses klasifikasi menggunakan metode SVM terdiri dari beberapa langkah, yaitu sebagai berikut:
  - a. Memilih metode untuk menentukan *hyperplane* SVM yang akan digunakan pada tahap klasifikasi.
  - b. Memilih fungsi kernel untuk membentuk model SVM.
  - c. Menentukan nilai parameter yang akan digunakan untuk model SVM.
  - d. Membentuk model SVM menggunakan fungsi kernel.
6. Melakukan evaluasi terkait ketepatan model klasifikasi metode SVM menggunakan *k-fold cross validation*.
7. Menganalisis kinerja model klasifikasi metode SVM untuk mengetahui hasil prediksi klasifikasi diagnosis penyakit hipertensi menggunakan *confusion matrix*.

## HASIL DAN PEMBAHASAN

### *Exploratory Data Analysis*

Penelitian ini menggunakan 10 variabel independen dan 1 variabel dependen setelah menghapus variabel jenis kelamin pada tahapan *data cleaning*. Variabel independen yang digunakan yaitu tanggal lahir, kebiasaan merokok, kurang aktivitas fisik, konsumsi gula berlebih, konsumsi garam berlebih, konsumsi lemak berlebih, kurang makan buah dan sayur, konsumsi alkohol, tekanan darah sistolik, dan tekanan darah diastolik. Lalu untuk variabel dependen yang digunakan adalah diagnosis hipertensi. Berikut perbandingan pasien yang terdiagnosis berpenyakit hipertensi dan tidak hipertensi pada Gambar 1.



Gambar 1 *Pie Chart* Diagnosis Hipertensi

Berdasarkan *pie chart* pada Gambar 1, dapat diketahui bahwa kategori tidak hipertensi lebih banyak daripada hipertensi. Total pasien yang tidak berpenyakit hipertensi sebanyak 996 data, sedangkan yang berpenyakit hipertensi sebanyak 377 data. Hal ini menunjukkan adanya indikasi *imbalanced data* atau disebut juga sebagai data yang tidak seimbang.

**Data Cleaning**

Pada tahapan ini dilakukan penghapusan data terhadap atribut Jenis Kelamin karena seluruh nilai atribut memiliki nilai yang sama yaitu Perempuan. Langkah ini bertujuan untuk menyederhanakan data dan mengurangi kompleksitas yang tidak perlu sehingga berpotensi untuk memberikan kontribusi yang lebih signifikan terhadap analisis yang dilakukan.

**Data Transformation**

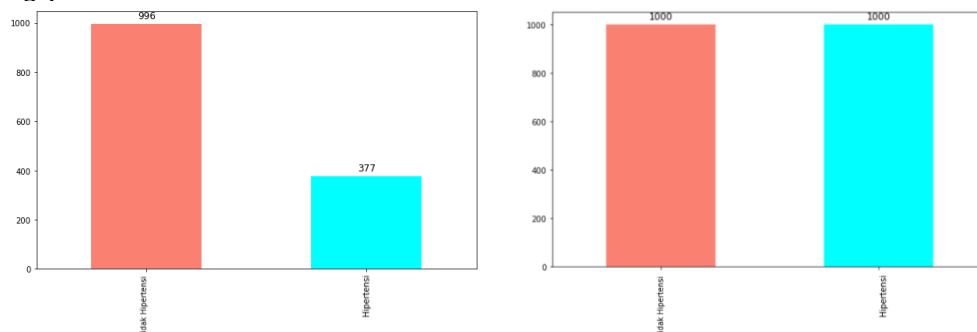
Proses transformasi data dilakukan untuk mengubah data pada *dataset* sesuai dengan format yang dapat diproses oleh *software* yang digunakan. Pada tahapan ini, terdapat proses transformasi data atribut-atribut yang bersifat kategorikal menjadi numerik. Berikut tampilan transformasi data kategorikal ke numerik pada Tabel 2.

Tabel 2 Tampilan Perubahan Atribut Data

Atribut	Sebelum Transformasi	Sesudah Transformasi
Kebiasaan Merokok	Ya	1
	Tidak	0
Kurang Aktivitas Fisik	Ya	1
	Tidak	0
Konsumsi Gula Berlebih	Ya	1
	Tidak	0
Konsumsi Garam Berlebih	Ya	1
	Tidak	0
Konsumsi Lemak Berlebih	Ya	1
	Tidak	0
Kurang Makan Buah dan Sayur	Ya	1
	Tidak	0
Konsumsi Alkohol	Ya	1
	Tidak	0
Diagnosis Hipertensi	Berpenyakit Hipertensi	1
	Tidak Berpenyakit Hipertensi	0

**Oversampling Data**

Berdasarkan Gambar 1, dapat disimpulkan bahwa data penelitian yang digunakan tidak seimbang. Hal ini dapat menyebabkan nilai performa yang tidak baik dari akurasi, presisi, *recall* dan *F1-Score*. Oleh karena itu, untuk membuat data tersebut *balance* digunakan teknik *oversampling*, yaitu dengan meratakan sampel data dari kedua nilai dengan memperbanyak data minoritas hingga setara nilai mayoritas. Berikut gambaran label dari kelas data sebelum dan setelah proses *oversampling* pada Gambar 2.



Gambar 2 Tampilan Data Sebelum Dan Sesudah *Oversampling*

### Klasifikasi dengan Metode SVM

Proses klasifikasi dimulai dengan membagi data menjadi dua, yaitu data *training* dan *testing* dengan rasio 80:20. Rasio tersebut dipilih berdasarkan hasil uji coba yang telah dilakukan, di mana terbukti bahwa model yang dilatih dengan rasio tersebut mencapai tingkat akurasi yang tinggi. Pada data *training* yang terdiri dari beberapa variabel dan kelas target dijadikan sebagai data sampel untuk membentuk model SVM. Sedangkan data *testing* digunakan untuk mengevaluasi ketepatan model klasifikasi berdasarkan *confusion matrix* yang terbentuk.

Klasifikasi SVM untuk keseluruhan data menggunakan bahasa pemrograman *python* dan *software Jupyter Notebook*. Penelitian ini menggunakan metode SVM linier dan non-linier untuk uji coba mencari akurasi terbaik. Fungsi kernel yang digunakan untuk klasifikasi SVM adalah kernel linier dengan parameter  $C = 0,01; 0,1; 1; 10$  dan  $100$ , kernel RBF dengan parameter  $C = 0,01; 0,1; 1; 10$  dan  $100$  serta parameter  $\gamma$  (*gamma*) =  $0,01; 0,1; 1; 10$  dan  $100$ . Berikut Tabel 3 hasil nilai akurasi SVM menggunakan kernel linier.

Tabel 3 Nilai Akurasi Parameter Model SVM Kernel Linier

<b>C (Cost)</b>	<b>Akurasi</b>	<b>Jumlah Support Vector</b>
0,01	92,75%	634
0,1	94,75%	332
1	95,25%	257
10	95,25%	245
100	95,25%	246

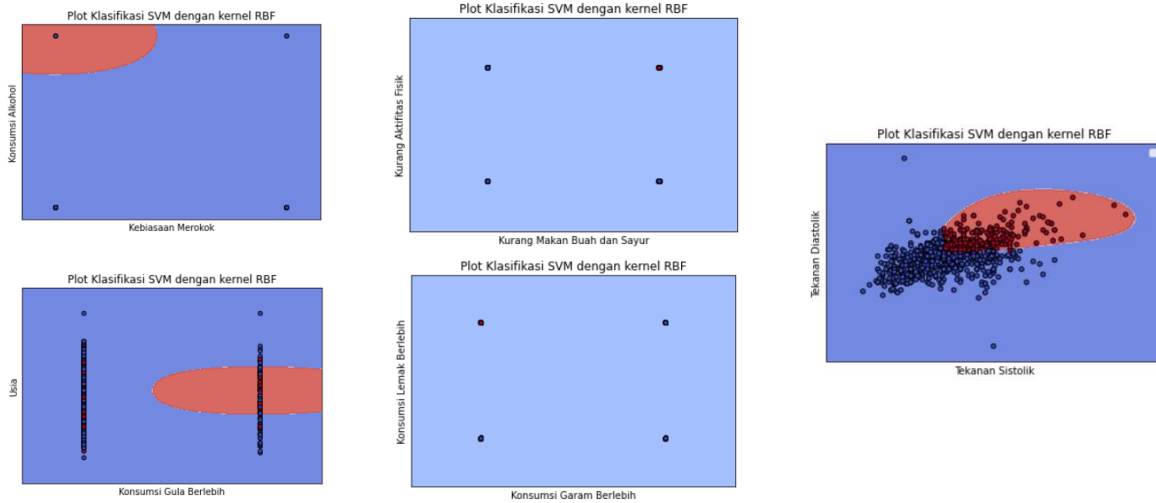
Sedangkan untuk parameter RBF diperoleh sebagai berikut pada Tabel 4.

Tabel 4 Nilai Akurasi Parameter Model SVM Kernel RBF

<b>Parameter</b>	<b>Akurasi</b>			
	$\gamma = 0,01$	$\gamma = 0,1$	$\gamma = 1$	$\gamma = 10$
$C = 0,01$	76,25%	89,75%	69,75%	49,75%
$C = 0,1$	90,75%	93,25%	89,25%	71,50%
$C = 1$	94,00%	95,25%	95,75%	90,25%
$C = 10$	95,25%	96,50%	96,75%	90,75%
$C = 100$	96,50%	98,00%	97,50%	90,75%

Berdasarkan *trial and error* yang dilakukan, diperoleh beberapa nilai akurasi dari beberapa percobaan menggunakan kernel linier dan RBF. Dari Tabel 3 dan 4, didapatkan akurasi tertinggi pada kernel RBF dengan parameter  $C = 100$  dan  $\gamma$  (*gamma*) =  $0,1$  dengan nilai akurasi sebesar 98%. Dari model tersebut, diperoleh jumlah *support vector* sebanyak 133 titik. Model tersebut dianggap sudah sangat baik, karena nilai *cost* ( $C$ ) yang kecil cenderung membuat margin lebih lebar dengan mengabaikan data yang tidak sesuai pada data *training*, sedangkan nilai *cost* ( $C$ ) yang besar cenderung menyesuaikan data *training* sehingga mengurangi kesalahan klasifikasi.

Model klasifikasi SVM yang telah terbentuk dapat divisualisasikan dalam bentuk plot. Pembuatan plot dibantu menggunakan *library matplotlib*. Plot dibentuk dengan kombinasi dua variabel independen ( $X$ ). Dari kombinasi dua variabel independen secara keseluruhan diperoleh plot-plot sebagai berikut pada Gambar 3.



Gambar 3 Plot Klasifikasi SVM Kernel RBF dengan  $\gamma = 0,1$  dan  $C = 100$

Berdasarkan Gambar 3, kombinasi dua variabel independen berhasil membentuk plot klasifikasi diagnosis penyakit hipertensi menggunakan SVM dengan kernel RBF. Pada plot tersebut, area berwarna merah ditandai kategori hipertensi, sedangkan area berwarna biru ditandai kategori tidak hipertensi. Titik berwarna merah menunjukkan data yang diklasifikasikan sebagai hipertensi, sedangkan titik berwarna biru menunjukkan data yang diklasifikasikan sebagai tidak hipertensi.

**Evaluasi**

Evaluasi terkait ketepatan model klasifikasi yang diperoleh dapat ditentukan menggunakan *K-Fold Cross Validation* dengan *k* sebanyak 5 *fold*. Setiap bagian data akan digunakan sebagai data uji sebanyak 1 kali, sedangkan data lainnya akan digunakan sebagai data latih. Dengan demikian, model klasifikasi akan dilatih sebanyak 5 kali dan diuji sebanyak 5 kali. Berikut hasil akurasi, presisi dan *recall* yang didapatkan menggunakan *5-Fold Cross Validation* pada Tabel 5.

Tabel 5 Hasil Evaluasi Menggunakan *5-Fold Cross Validation*

Parameter <i>k</i>	Presisi	Recall	Akurasi
<i>k</i> = 1	0.957983	0.942149	0.970000
<i>k</i> = 2	0.975904	0.975904	0.990000
<i>k</i> = 3	0.972477	0.946429	0.977500
<i>k</i> = 4	1.000000	0.936620	0.955000
<i>k</i> = 5	1.000000	0.965000	0.965000
Rata-rata	0.981273	0.953220	0.971500

Berdasarkan Tabel 5, diperoleh rata-rata presisi sebesar 0.981273 atau 98,12%, rata-rata *recall* sebesar 0.953220 atau 95,32% dan rata-rata akurasi sebesar 0.971500 atau 97,15%. Hasil ini menunjukkan bahwa model SVM dengan kernel RBF memiliki performa yang baik dalam mengklasifikasikan data.

**Analisis Performa Model Klasifikasi**

Perhitungan ketepatan klasifikasi diagnosis penyakit hipertensi menggunakan metode SVM kernel RBF dengan parameter  $C = 100$  dan  $\gamma = 0.1$  dapat dianalisis menggunakan *confusion matrix* pada Tabel 6.

Tabel 6 *Confusion Matrix* Hasil Penelitian

Data	Aktual	
	Non Hipertensi	Hipertensi

<b>Prediksi</b>	<b>Non Hipertensi</b>	192	7
	<b>Hipertensi</b>	1	200

Berdasarkan Tabel 6 dari *confusion matrix* menunjukkan bahwa 192 pasien non hipertensi diprediksi benar sebagai non hipertensi, 7 pasien hipertensi diprediksi salah sebagai non hipertensi, 200 pasien hipertensi diprediksi benar sebagai hipertensi, dan 1 pasien non-hipertensi diprediksi salah sebagai hipertensi. Kemudian didapatkan perhitungan nilai akurasi, presisi, *recall* dan *F1-Score* sebagai berikut.

$$\begin{aligned} \text{Akurasi} &= \frac{192+200}{192+200+7+1} \times 100\% = 98\% \\ \text{Presisi} &= \frac{192}{192+7} \times 100\% = 96.48\% \\ \text{Recall} &= \frac{192}{192+1} \times 100\% = 99.48\% \\ \text{F1-Score} &= 2 \times \frac{96.48\% \times 99.48\%}{96.48\% + 99.48\%} \times 100\% = 97.95\% \end{aligned}$$

Berdasarkan hasil perhitungan tersebut, dapat disimpulkan bahwa klasifikasi diagnosis hipertensi menggunakan SVM memiliki akurasi sebesar 98%. Hal ini membuktikan bahwa metode SVM mampu mengklasifikasikan diagnosis hipertensi dengan sangat baik. Kemudian untuk nilai presisi didapatkan sebesar 96.48%, yang artinya dari 199 pasien non hipertensi yang diprediksi oleh model SVM, sebanyak 192 pasien di antaranya benar-benar non hipertensi. Kemudian untuk nilai *recall* didapatkan sebesar 99.48%, yang artinya dari 193 pasien non hipertensi yang ada, sebanyak 192 pasien berhasil diidentifikasi dengan benar oleh model SVM. Terakhir nilai *F1-Score* diperoleh sebesar 97.95% yang artinya model SVM mampu menyeimbangkan antara presisi dan *recall*.

## KESIMPULAN

Berdasarkan penjelasan sebelumnya, dapat disimpulkan bahwa hasil klasifikasi menggunakan metode SVM dengan kernel RBF memberikan model terbaik untuk diagnosis penyakit hipertensi. Parameter yang digunakan adalah  $C = 100$  dan  $\gamma$  (gamma) = 0,1. Hasil analisis klasifikasi menggunakan *confusion matrix* menunjukkan bahwa 192 pasien non-hipertensi diprediksi benar sebagai non-hipertensi, 7 pasien non-hipertensi diprediksi salah sebagai hipertensi, 200 pasien hipertensi diprediksi benar sebagai hipertensi, dan 1 pasien hipertensi diprediksi salah sebagai non-hipertensi. Akurasi klasifikasi diagnosis penyakit hipertensi menggunakan metode SVM dengan kernel RBF dan parameter  $C = 100$  serta  $\gamma$  (gamma) = 0,1 adalah sebesar 97,15%. Hal ini menunjukkan bahwa metode SVM mampu mengklasifikasikan diagnosis penyakit hipertensi dengan sangat baik. Berdasarkan hasil penelitian ini, metode SVM dapat menjadi alternatif metode klasifikasi yang andal untuk diagnosis penyakit hipertensi.

## DAFTAR PUSTAKA

- [1] Andriansyah, Mochamad Rafli, Edy Santoso, and Sutrisno. (2018). "Klasifikasi Risiko Hipertensi Menggunakan Fuzzy Decision Tree Iterative Dichotomiser 3 (ID3)." *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer* 2(Vol 2 No 12 (2018)):7088-96.
- [2] BPS Kota Malang. (2019). Jumlah Kasus Penyakit Terbanyak di Kota Malang (Jiwa), 2019. <https://malangkota.bps.go.id/indicator/30/371/1/jumlah-kasus-penyakit-terbanyak-di-kota-malang.html>. Diakses pada tanggal 5 Oktober 2023.
- [3] Don, D. P. W. R. (2018). Multiclass Classification Using Support Vector Machines [Tesis, Georgia Southern University]. <https://digitalcommons.georgiasouthern.edu/etd/1845/>. Diakses pada tanggal 5 Januari 2024.
- [4] Khalim, K. A., Hayati, U., & Bahtiar, A. (2023). Perbandingan Prediksi Penyakit Hipertensi Menggunakan Metode Random Forest Dan Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 498-504. doi:10.36040/jati.v7i1.6376



- [5] Maiyanti, S. I., Zayanti, D. A., Andriani, Y., Suprihatin, B., Desiani, A., Salsabila, A., & Marselina, N. C. (2023). Perbandingan Klasifikasi Penyakit Kanker Paru-Paru Menggunakan Support Vector Machine Dan K-Nearest Neighbor. *Jurnal Ilmiah Teknologi Kesehatan*, 18(1), 54-62. doi: 10.33998/processor.2023.18.1.700
- [6] P2TPM Kemenkes Kesehatan RI. (2016). Tekanan darah tinggi (hipertensi). Diakses pada 25 Oktober 2023, dari <https://p2ptm.kemkes.go.id/uploads/2016/10/Tekanan-Darah-Tinggi-Hipertensi.pdf>
- [7] P2TPM Kemenkes Kesehatan RI. (2019). Hari Hipertensi Dunia 2019 : “Know Your Number, Kendalikan Tekanan Darahmu dengan CERDIK.” <https://p2ptm.kemkes.go.id/tag/hari-hipertensi-dunia-2019-know-your-number-kendalikan-tekanan-darahmu-dengan-cerdik>. Diakses pada tanggal 15 Oktober 2023.
- [8] Septiani, W. D. (2017). Komparasi Metode Klasifikasi Data Mining Algoritma C4.5 Dan Naive Bayes Untuk Prediksi Penyakit Hepatitis. *Journal of Computing and Information System*, 13(1), 76-84.
- [9] Setiawati, I., Wibowo, A. P., & Hermawan, A. (2019). Pendahuluan tinjauan pustaka penelitian sebelumnya klasifikasi. *Jurnal of Information System Management*, 1(1), 13-17.
- [10] Tan, P.-N., Steinbach, M., & Kumar, V. (2016). Introduction to data mining (3rd ed.). Harlow, England: Pearson Education Limited.
- [11] Tino, M. D. F., Hasanah, H., & Santosa, T. D. (2023). Perbandingan Algoritma Support Vector Machines (Svm) Dan Neural Network Untuk Klasifikasi Penyakit Jantung. *INFOTECH journal*, 9(1), 232-235. doi: 10.31949/infotech.v9i1.5432