

# Analisis Klasifikasi Penyakit *Multiple Sclerosis* Menggunakan Algoritma *Logistic Regression* dan SVM

Ida Nur Laela<sup>1</sup>, Wiga Maulana Baihaqi<sup>2</sup>

<sup>1</sup>Sistem Informasi, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto

<sup>2</sup>Teknologi Informasi, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto

E-mail: [1laellaida719@gmail.com](mailto:laellaida719@gmail.com), [2wiga@amikompurwokerto.ac.id](mailto:wiga@amikompurwokerto.ac.id)

*Corresponden Author*: [laellaida719@gmail.com](mailto:laellaida719@gmail.com)

Diterima Redaksi: 07 Juli 2023 Revisi Akhir: 01 Desember 2023 Diterbitkan Online: 28 Januari 2024

**Abstrak** – Kesehatan merupakan aspek yang paling penting untuk menunjang kegiatan sehari-hari. Tentunya dengan mempunyai tubuh yang sehat setiap orang dapat melakukan beragam aktivitas dengan nyaman dan tenang. Setiap individu pastinya memiliki insting yang kuat untuk hidup sehat dan terbebas dari penyakit, salah satunya dengan meningkatkan kekebalan tubuh. *Multiple Sclerosis* (MS) adalah penyakit neurodegeneratif autoimun yang mempengaruhi sistem saraf pusat. Kelainan MS itu ditandai tanda-tanda termasuk inflamasi kronis, demielinisasi, gliosis, dan kematian neuron. Gejala yang dihadapi oleh pasien MS tidak dapat diprediksi sehingga perlu adanya penelitian klasifikasi terkait penyakit tersebut. Oleh karena itu dilakukan penelitian klasifikasi menggunakan algoritma *logistic regression* dan SVM. Metode yang digunakan dalam penelitian ini adalah studi kepustakaan dengan bahasa pemrograman Python. Hasil penelitian ini menunjukkan bahwa algoritma SVM memiliki tingkat akurasi yang tinggi sebanyak 88,33% dari algoritma *logistic regression*. Sehingga dapat disimpulkan dari penelitian ini adalah metode SVM memiliki performa kinerja yang baik untuk mengolah dataset *multiple sclerosis*.

**Kata Kunci** — Data Mining, Multiple Sclerosis, Python, Logistic Regression, SVM.

**Abstract** – Health is the most important aspect to support daily activities. Of course, by having a healthy body, everyone can carry out various activities comfortably and calmly. Every individual certainly has a strong instinct to live a healthy life and be free from disease, one of which is by increasing the body's immunity. *Multiple sclerosis* (multiple sclerosis/MS) is a neurodegenerative autoimmune disease that affects the central nervous system. The affliction of MS is characterized by chronic inflammation, demyelination, gliosis, and neuronal death. The symptoms faced by MS patients are unpredictable, so there is a need for a classification related to the disease. Therefore, a classification study was carried out using the *logistic regression* algorithm and SVM. The method used in this research is a literature study with the Python programming language. The results of this study indicate that the SVM algorithm has a high accuracy rate of 88.33% of the *logistic regression* algorithm. So it can be concluded from this study that the SVM method has good performance for processing *multiple sclerosis* datasets.

**Keywords** — Data Mining, Multiple Sclerosis, Python, Logistic Regression, SVM.



## 1. PENDAHULUAN

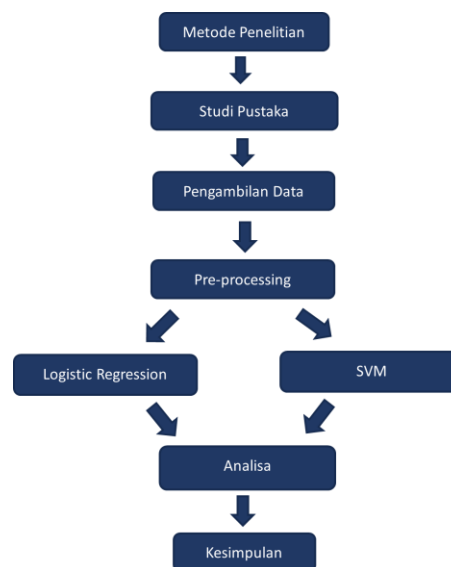
Kesehatan merupakan aspek yang paling penting untuk menunjang kegiatan sehari-hari. Tentunya dengan mempunyai tubuh yang sehat setiap orang dapat melakukan beragam aktivitas dengan nyaman dan tenang. Setiap individu pastinya memiliki insting yang kuat untuk hidup sehat dan terbebas dari penyakit, salah satunya dengan meningkatkan kekebalan tubuh. Kekebalan tubuh dan imunitas yang baik sangat berpengaruh untuk melawan segala jenis penyakit sehingga banyak masyarakat melakukan Program Perilaku Hidup Bersih dan Sehat (PHBS) [1]. Namun dalam kehidupan realitasnya berbagai penyakit masih menjadi tantangan yang serius bagi kesehatan masyarakat. *Multiple Sclerosis* merupakan salah satu penyakit yang mampu mempengaruhi kegiatan atau aktivitas sehari-hari, karena hal ini memberikan gangguan pada mata, tulang belakang, dan saraf otak [2]. Selain itu, Kejang otot, kelumpuhan, masalah dengan kandung kemih, usus, fungsi seksual, perubahan mental, depresi, dan epilepsi adalah beberapa dari banyak komplikasi yang dapat disebabkan oleh *multiple sclerosis* [3].

*Multiple sklerosis (MS)* adalah penyakit neurodegeneratif autoimun yang mempengaruhi sistem saraf pusat. Kelainan MS itu ditandai tanda-tanda termasuk *inflamasi kronis, demielinisasi, gliosis*, dan kematian *neuron* [4]. Penyakit ini biasanya diderita oleh orang berusia 20 hingga 50 tahun. MS berkembang secara bertahap dan dapat mengakibatkan kecacatan. Dalam 15 tahun setelah penyakit muncul, sekitar 50% penderita MS akan membutuhkan bantuan untuk berjalan. Sampai saat ini, penyebab Multiple Sclerosis masih belum diketahui. Ada banyak faktor yang dapat menyebabkan kerusakan saraf, yang pada gilirannya menyebabkan penyakit ini [5]. Dilansir dari CNN Indonesia gejala dari penyakit *Multiple Sklerosis (MS)* tidak dapat diperkirakan. Bisa saja suatu saat terlihat sehat, tapi detik berikutnya ambruk lalu tanpa aba-aba bisa kembali ke sediakala [6]. Maka dari itu penyakit *Multiple Sclerosis* harus dilakukan penanganan yang tepat sehingga tidak semakin memperburuk keadaan.

Atas dasar permasalahan yang telah disebutkan di atas, rumusan masalah yang ada pada penelitian ini adalah bagaimana melakukan analisis klasifikasi yang efektif untuk penyakit *Multiple Sclerosis*. Dalam dunia kesehatan data mining berperan dalam menghasilkan prediksi-prediksi mengenai diagnosa suatu penyakit. Penelitian sebelumnya mengenai penyakit multiple sclerosis telah dilakukan yaitu penggunaan algoritma decision tree untuk mengklasifikasi perkembangan sekunder pada penyakit multiple sclerosis, dari penelitian tersebut decision tree mendapatkan akurasi 82% [7]. Penelitian lain juga dilakukan yaitu penggunaan algoritma KNN menghasilkan akurasi 83% menggunakan divergensi Kullback-Leibler diikuti divergensi Jensen-shannon dengan ACC 81% [8].

Penelitian ini akan menggunakan algoritma *logistic regression* dan *support vektor machine* untuk melakukan klasifikasi penyakit *multiple sclerosis*. Untuk memprediksi probabilitas atau kemungkinan suatu peristiwa, logistic regresi digunakan dalam statistik untuk memprediksi, yang biasanya disebut sebagai model logit. *Logistic Regression* merupakan regresi linear yang sederhana namun memiliki fungsi dengan kemampuan melibatkan variabel terikat biner [9] dan mempunyai performa yang tinggi. *Super Vektor Machine (SVM)* dipilih karena merupakan algoritma dengan tingkat akurasi tinggi dan memiliki kinerja yang baik. Dilansir dari penelitian sebelumnya oleh Hilda Amalia (2018) [10] melakukan penelitian perbandingan suport vektor machine dan nn pada klasifikasi penyakit ginjal kronis, algoritma svm memperoleh akurasi lebih tinggi yaitu sebanyak 95,16%. Penelitian ini diharapkan dapat mendukung pengambilan keputusan yang lebih akurat untuk menangani karakteristik dari pasien *Multiple Sclerosis* sehingga dapat menjadi pertimbangan pengembangan terapi yang lebih efektif.

## 2. METODE PENELITIAN



Gambar 1. Alur Metode Penelitian

### A. Studi Pustaka

Studi pustaka merupakan pendekatan yang dilakukan untuk mengumpulkan data dan merumuskan masalah terkait penyakit *Multiple Sclerosis*. Untuk melakukan penelitian, setiap peneliti perlu mempersiapkan bekal yang cukup dan terpercaya seperti referensi dari buku, jurnal, dan sumber informasi tambahan tentang subjek penelitian yang akan dilakukan [11]. Studi pustaka sangat penting dilakukan untuk mencari referensi dalam menulis jurnal sehingga jurnal dapat bersifat valid dan akurat.

### B. Pengambilan Data

Penelitian ini menggunakan data tentang penyakit *Multiple Sclerosis* yang terdiri dari 20 variabel yang bersumber dari [12] (open source) di [www.kaggle.com](http://www.kaggle.com). Adapun data yang digunakan dapat dilihat pada Gambar. 1 sebagai berikut.

Gambar 2. Dataset

Data tersebut merupakan dataset dari pasien Mexican mestizo yang baru didiagnosis dengan CIS yang dipresentasikan di Institut Nasional Neurologi dan Bedah Saraf (NINN) di Mexico City, Meksiko, antara 2006 dan 2010. Data umum terdiri dari *ID, Age, Schooling, Gender, Breastfeeding, Variacella, Initial\_Symptoms, Mono\_or\_Polysymptomatic, Oligoclonal\_Bands, LLSSEP, ULSSEP, VEP, BAEP, Periventricular\_MRI, Infratentorial\_MRI, Spinal\_Cord\_MRI, Initial\_EDSS, final\_EDSS, dan Group*. Untuk variabel dependen (Y) yang ada dalam dataset terdiri dari dua kategori yang dikategorikan dalam kolom *Group* yaitu :

Y = 1, untuk pasien dengan kategori terkena *CMDS*

Y = 2, untuk pasien dengan kategori *Non-CMDS*

Penelitian ini menggunakan variabel tersebut untuk mengetahui klasifikasi dari penyakit *Multiple Sclerosis*. Data tersebut dimasukkan dalam software anaconda dan *jupyter notebook* dengan bahasa pemrograman *python*.

### C. Pre-Processing

Merupakan tahap untuk mengolah data sehingga data siap untuk diolah. Tahap ini digunakan untuk memahami data lebih dalam dan memeriksa statistik deskriptifnya serta distribusinya untuk kepentingan modelling dengan algoritma dalam machine learning. Tujuan dilakukannya tahap pre-processing adalah data yang digunakan lebih presisi dan memudahkan dalam menjalankan klasifikasi dengan algoritma yang dipilih. Tahap *pre-processing* melakukan tahapan *cleaning* data dan cek duplikasi data.

```
*[195]: # Menghapus variabel yang tidak berpengaruh
df.drop(['Unnamed: 0', 'Initial_EDSS', 'Final_EDSS'], axis=1, inplace=True)
df.dropna(inplace=True)
df.shape
```

Gambar 3. Cleaning Data

Variabel dalam dataset tersebut memiliki missing value sehingga perlu adanya pembersihan data. Kolom yang dihapus dalam dataset yaitu kolom “*Unnamed, Initial\_EDSS, dan Final\_EDSS*. Selanjutnya dilakukan pengecekan duplikasi data dengan *query df.duplicated().sum()*. Setelah dilakukan pengecekan sudah tidak ada data terduplikat, selanjutnya akan dilakukan pengecekan banyaknya missing value dengan *query df.isnull().sum()*.

```
[257]: df.info()
<class 'pandas.core.frame.DataFrame'>
Int64Index: 271 entries, 0 to 272
Data columns (total 17 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Gender                 271 non-null    int64
1   Age                   271 non-null    int64
2   Schooling              271 non-null    float64
3   Breastfeeding          271 non-null    int64
4   Varicella              271 non-null    int64
5   Initial_Symptom        271 non-null    float64
6   Mono_or_Polysymptomatic 271 non-null    int64
7   Oligoclonal_Bands      271 non-null    int64
8   ILSSEP                 271 non-null    int64
9   ULSSEP                 271 non-null    int64
10  VEP                    271 non-null    int64
11  BAEP                   271 non-null    int64
12  Periventricular_MRI    271 non-null    int64
13  Cortical_MRI           271 non-null    int64
14  Infratentorial_MRI     271 non-null    int64
15  Spinal_Cord_MRI        271 non-null    int64
16  group                  271 non-null    int64
dtypes: float64(2), int64(15)
memory usage: 38.1 KB
```

Gambar 4. Data Info

Pada gambar 4 menunjukkan data setelah dihapus terdiri dari 17 kolom dimana data awal sebelum dihapus adalah 20 kolom. Tahap selanjutnya akan dilakukan modelling dan evaluasi dengan beberapa algoritma dalam machine learning,

D. Algoritma Logistic Regression

Metode analisis statistik yang dikenal sebagai Regresi Logistik digunakan untuk mengevaluasi hubungan antara variabel prediktor dan variabel respons dalam kasus di mana variabel respons termasuk dalam salah satu dari dua kategori [13]. Dalam penelitian ini *Logistic Regression* berperan dalam memodelkan variabel apa saja yang berpengaruh signifikan pada kejadian positif CMDS atau positif penyakit *Multiple Sclerosis*. Metode ini digunakan untuk *variabel predictor* atau variabel dependen (y), yang memiliki skala nominal atau kategorik yang terdiri dari dua biner atau lebih. Adapun berikut persamaan matematis model logistic regression.

$$P(y|x) = \pi(x) \tag{1}$$

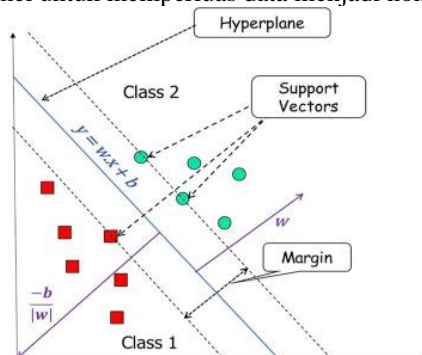
dengan nilai  $x$  diperoleh kemudian dipetakan menggunakan fungsi persamaan umum.

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)} \tag{2}$$

Dimana  $y$  bernilai 1 untuk pasien terkena CMDS dan 2 untuk Non-CMDS.

E. Algoritma SVM

*Support Vector Machine (SVM)* adalah metode prediksi untuk klasifikasi dan regresi yang merupakan salah satu algoritma dalam machine learning [14]. Klasifikasi dengan menggunakan SVM mengidentifikasi set yang terlibat dalam proses untuk membedakan kelasnya. Tujuan pemilihan SVM adalah untuk membedakan set yang terlibat dalam proses untuk membedakan kelasnya. Tujuan SVM adalah untuk memisahkan sebuah data dengan hyperplane dan menggunakan trik kernel untuk memperluas data menjadi nonlinier. [15].



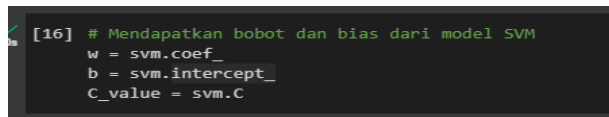
Gambar 5. Metode SVM

Sehingga dapat diperoleh persamaan suport vektor machinenya dimana  $w$  menyatakan bobot,  $x$  menyatakan data dan  $b$  menyatakan bias. Persamaan SVM sebagai berikut.

$$\begin{aligned} x_i \cdot w + b &= 0, \text{ untuk } y_i = 0 \text{ (hyperplane)} \\ x_i \cdot w + b &\geq +1, \text{ untuk } y_i = +1 \text{ (kelas positif)} \\ x_i \cdot w + b &\leq -1, \text{ untuk } y_i = -1 \text{ (kelas negatif)} \end{aligned} \quad (3)$$

Metode ini dilakukan splitting data training dan testing. Pada data training dilakukan pemetaan vektor input ke dalam feature space sesuai jenis kernel dan menyimpan nilai  $\alpha$  dan bias sehingga dapat dilakukan proses pengujian atau testing. Pada proses testing hasil perhitungan  $f(x) = w * xi + b$  lalu dikelompokkan berdasarkan nilai [16].

Dalam case ini, penulis mencoba untuk mengetahui nilai  $w$ , dan  $b$  yang digunakan dalam perhitungan klasifikasi dengan algoritma *support vektor machine*. Pengecekan dilakukan dengan *google colab* dengan bantuan library python. Dari hasil yang diperoleh nilai  $w$  yang menyatakan bobot dan  $b$  yang menyatakan bias sebagai berikut.



```
[16] # Mendapatkan bobot dan bias dari model SVM
w = svm.coef_
b = svm.intercept_
C_value = svm.c
```

Gambar 6. Bobot dan Bias

Dengan nilai  $w$  dan  $b$  yang diperoleh.

$$w = [[ 0.00330423 -0.11867812 -0.49315548 0.02970728 0.2354359 -0.97374493 0.71211144 0.32387113 -0.18430992 0.12964121 -0.1817647 0.13114955 -0.82460431 -0.32969564 -0.521532 0.00972095]]$$

$$b = [0.28921472]$$

Setelah itu dilakukan perhitungan dengan persamaan dasar dari algoritma menggunakan koefisien ( $w$ ), intersep ( $b$ ), dan nilai fitur dari sampel pertama untuk menghitung  $f(x_1)$ .

$$f(x_1) = 0.00330423 \cdot x_{11} - 0.11867812 \cdot x_{12} - 0.49315548 \cdot x_{13} + \dots - 0.521532 \cdot x_{116} + 0.28921472$$

Perhitungan ini memberikan nilai  $f(x_1)$ , yang dapat digunakan untuk membuat keputusan klasifikasi. Jika  $f(x_1)$  lebih besar dari 0, akan mengklasifikasikan sebagai kelas 1; jika  $f(x_1)$  kurang dari 0, akan mengklasifikasikan sebagai kelas 0.

#### F. Analisa

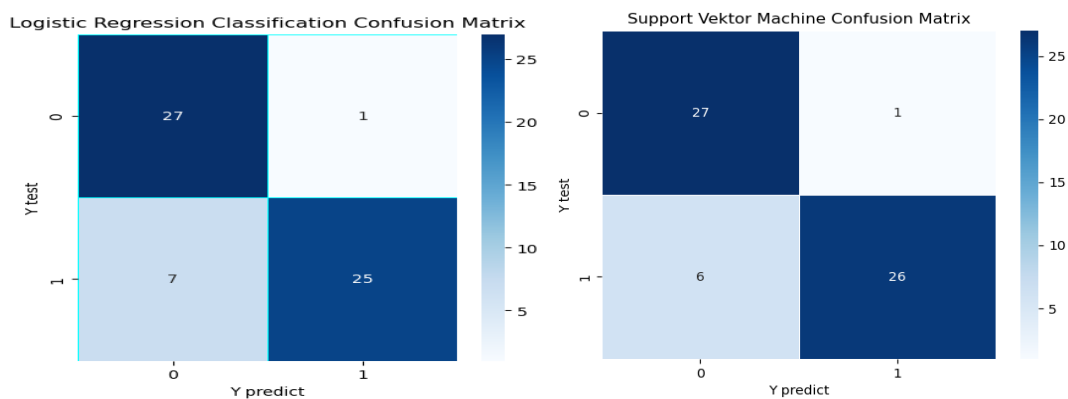
Dari proses klasifikasi akan diperoleh hasil yang akan disajikan dalam bentuk *confusion matrix*, *precision*, *recall*, dan *f1-score*. Penelitian ini melakukan analisa perbandingan untuk melihat algoritma yang lebih unggul antara algoritma *logistic regression* atau algoritma *svm*. Algoritma tersebut memiliki akurasi yang kuat sehingga analisa perbandingan ini sangat penting untuk melihat metode yang paling sesuai untuk klasifikasi penyakit *Multiple Sclerosis*.

#### G. Kesimpulan

Kesimpulan merupakan pernyataan yang didapatkan dari analisa yang diperoleh. Kesimpulan berisi jawaban dari rumusan masalah, tujuan penelitian, serta metode penelitian yang digunakan.

### 3. HASIL DAN PEMBAHASAN

Berdasarkan hasil analisa tentang akurasi perbandingan algoritma *Logistic Regression* dan *SVM* menghasilkan *confusion matrix*. *Confusion matrix* merupakan hasil dari proses klasifikasi, yang melibatkan menampilkan data yang benar atau salah diprediksi [17]. Menghitung jumlah hasil perhitungan benar dan salah adalah cara pembentukan matriks *confusion*. Setelah mendapatkan nilai matriks, nilai akurasi dapat dihitung. [18]. *Confusion matrix* yang dihasilkan sebagai berikut.



Gambar 7. Confusion Matrix Algoritma

Gambar 6 diatas menunjukkan *confusion matrix* dari masing-masing algoritma baik algoritma *logistic regression* dan algoritma *svm*. Kolom Y test dan Y predict menunjukkan data yang diprediksi oleh model sebagai kelas 0 dan kelas 1. Dalam gambar 5 terdapat 27 data yang diprediksi benar sebagai kelas 0 *True Negatif* (TN), 1 data yang diprediksi 1 *True Positif* (TP), 7 data yang diprediksi sebagai kelas 1 *False Positif* (FP), dan 25 data diprediksi sebagai kelas 0 *False Negatif* (FN). Sedangkan dari gambar 6 confusion matrix algoritma *svm* menunjukkan perbedaan dimana 6 data yang diprediksi sebagai kelas 1 *False Positif* (FP), dan 26 data diprediksi sebagai kelas 0 *False Negatif* (FN) dan sisanya menunjukkan kesamaan. Untuk dapat dilihat dengan jelas proses klasifikasi sebagai berikut.

Dari hasil *Confusion Matrix* diatas selanjutnya peneliti melakukan uji akurasi terhadap dua algoritma diatas. Hasil perbandingan setiap algoritma akan dilakukan perbandingan untuk melihat akurasi yang akurat. Hasil model akurasi dapat dilihat dalam tabel berikut.

Tabel 1. Hasil Akurasi

No	Algoritma	Akurasi	Presisi	Recall	F1 Score
1	Logistic Regression	86,67%	79,41%	96,42%	87,89%
2	SVM	88,33%	96,42%	96,42%	88,52%

Dari tabel diatas menunjukkan hasil perbandingan akurasi dua algoritma. Algoritma *Logistic Regression* menghasilkan akurasi 86,67%, *presisi* sebesar 79,41%, *recall* sebesar 96,42% dan *F1 Score* sebesar 87,89%. Sedangkan algoritma SVM menghasilkan akurasi sebesar 88,33%, *presisi* sebesar 81,81%, *recall* sebesar 96,42% dan *F1 Score* sebesar 88,52%. Selisih dari hasil akurasi kedua algoritma adalah 1,66% dimana algoritma SVM menunjukkan tingkat akurasi yang lebih tinggi dari algoritma *Logistic Regression*.

#### 4. SIMPULAN

Penelitian ini menganalisis dataset penyakit *Multiple Sclerosis* yang diambil dari Kaggle. Peneliti melakukan analisis perbandingan untuk menentukan metode mana yang memiliki nilai akurasi yang lebih tinggi untuk dataset tersebut. Hasil penelitian menunjukkan bahwa metode Super Vektor Mesin (SVM) memiliki nilai akurasi yang lebih tinggi daripada metode Logistic Regression, dengan nilai akurasi masing-masing 86,67% dan 88,33%. Kesimpulan dari penelitian ini adalah bahwa metode SVM lebih baik dalam pengolahan dataset *Multiple Sclerosis* dan menghasilkan nilai akurasi yang lebih tinggi daripada metode *Logistic Regression*.

#### 5. SARAN

Hasil penelitian ini diharapkan dapat bermanfaat bagi pembaca dan berguna untuk penelitian selanjutnya. Penelitian ini masih belum menggunakan kombinasi dari beberapa algoritma dalam *machine learning*. Penggunaan kombinasi diharapkan akan meningkatkan hasil akurasi yang tinggi sehingga dapat membuat klasifikasi lebih akurat dan baik. Penelitian ini juga dapat menerapkan hasil diagnosis klasifikasi sehingga dapat dilakukan pembuatan aplikasi mobile atau website sehingga memudahkan para tenaga medis.

DAFTAR PUSTAKA

- [1] A. Raksanagara and A. Raksanagara, "Perilaku Hidup Bersih Dan Sehat Sebagai Determinan Kesehatan Yang Penting Pada Tatanan Rumah Tangga Di Kota Bandung," *J. Sist. Kesehat.*, vol. 1, no. 1, pp. 30–34, 2016, doi: 10.24198/jsk.v1i1.10340.
- [2] R. A. Rosdiana, Destya Sabrina Pujian, H. Angelia, N. Zulpa, Elhapidi, and F. Rachmatullah, "Analisis Gejala Multiple Sclerosis Sejak Dini Dan Penangannya," vol. 5, pp. 1–23, 2023.
- [3] R. Estiasari, N. Devianca, K. Maharani, and D. Imran, "Multipel Sklerosis Progresif Sekunder, Gejala Klinis, Diagnosis, Dan Tata Laksana," *Maj. Kedokt. Neurosains Perhimpun. Dr. Spes. Saraf Indones.*, vol. 37, no. 3, 2020, doi: 10.52386/neurona.v37i3.166.
- [4] C. J. Suryo, "Sklerosis Multipel: Diagnosis dan Tatalaksana," *Cermin Dunia Kedokt.*, vol. 48, no. 8, pp. 296–303, 2021, doi: 10.55175/cdk.v48i8.111.
- [5] dr. Rashmeeta; Dr. dr. Stevent Sumantri; DAA; Sp.PD; K-AI, "Mengenal Multiple Sclerosis," 2022. <https://dokterimun.id/2022/07/23/mengenal-multiple-sclerosis/> (accessed Jun. 26, 2023).
- [6] S. Galikano, "Kisah Kanya Puspokusumo Berdamai dengan Multipel Sklerosis," 2015. <https://www.cnnindonesia.com/gaya-hidup/20151228095923-255-100666/kisah-kanya-puspokusumo-berdamai-dengan-multipel-sklerosis> (accessed Jun. 27, 2023).
- [7] R. Ramanujam *et al.*, "Accurate classification of secondary progression in multiple sclerosis using a decision tree," *Mult. Scler. J.*, vol. 27, no. 8, pp. 1240–1249, 2021, doi: 10.1177/1352458520975323.
- [8] S. Rebbah *et al.*, "Classification of Multiple Sclerosis patients using a histogram-based K-Nearest Neighbors algorithm To cite this version : HAL Id : hal-02156448 Classification of Multiple Sclerosis patients using a histogram- based KNN algorithm," 2019.
- [9] M. V. Babo and I. A. Dwiatmoko, "Analisis Faktor-Faktor Yang Berhubungan Dengan Terjangkitnya Penyakit Malaria Dengan Menggunakan Model Regresi Logistik Binomial," *Asimtot J. Kependidikan Mat.*, vol. 2, no. 1, pp. 27–37, 2020, doi: 10.30822/asimtot.v2i1.498.
- [10] H. Amalia, "Perbandingan Metode Data Mining Svm Dan Nn Untuk Klasifikasi Penyakit Ginjal Kronis," *Maret*, vol. 14, no. 1, p. 1, 2018, [Online]. Available: [www.bsi.ac.id](http://www.bsi.ac.id)
- [11] N. J. A. Kisma and C. R. S. Ajeng, "Analisis Aplikasi Di Playstore Berdasarkan Rating Dan Type Menggunakan Naive Bayes Dan Logistik Regresi," vol. 10, no. 2, pp. 174–184, 2023.
- [12] M. National Institute of Neurology and Neurosurgery (NINN) in Mexico City, "Multiple Sclerosis Disease," 2010. <https://www.kaggle.com/datasets/desalegngeb/conversion-predictors-of-cis-to-multiple-sclerosis> (accessed Jun. 30, 2023).
- [13] N. A. N. Roosyidah and P. K. Supriyatna, "Pemodelan Regresi Logistik untuk Diagnosis Dini Infeksi Covid-19 di Indonesia," *Jambura J. Math.*, vol. 4, no. 2, pp. 232–246, 2022, doi: 10.34312/jjom.v4i2.12653.
- [14] K. W. Patunduk and A. Pratiwi, "Pemodelan Pasien Covid-19 Di Kota Palopo Dengan Regresi Logistik ( Studi Perbandingan Regresi Logistik dan Analisis Survival )," *J. Penelit. Mat. dan Pendidik. Mat.*, vol. 5, pp. 260–269, 2022.
- [15] L. I. Permatasari, M. . Astri Novianty, S.T, and M. Pm. Dr. Tito Waluyo Purboyo, S.Si, S.T, "Heart Disorder Detection Based on Computerized Iridology Using SVM Classification Method," *e- Proceeding Eng.*, vol. 3, no. 1, pp. 786–792, 2016.
- [16] H. Syah and A. Witanti, "Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (Svm)," *J. Sist. Inf. dan Inform.*, vol. 5, no. 1, pp. 59–67, 2022, doi: 10.47080/simika.v5i1.1411.
- [17] A. Amrin and H. Saiyar, "Aplikasi Diagnosa Penyakit Tuberculosis Menggunakan Algoritma Data Mining," *J. Ris. Komput.*, vol. 5, no. 5, pp. 498–502, 2018, [Online]. Available: <http://ejournal.bsi.ac.id/ejurnal/index.php/paradigma>
- [18] D. C. rini Novitasari, M. F. Rozi, and R. Veriani, "Klasifikasi Kelainan Pada Jantung Melalui Citra Iris Mata Menggunakan Fuzzy C-Means Sebagai Pengambil Fitur Iris Dan Klasifikasi Menggunakan Support Vector Machine," *INTEGER J. Inf. Technol.*, vol. 4, no. 1, pp. 1–10, 2019, doi: 10.31284/j.integer.2019.v4i1.489.