

PROPOSAL PENELITIAN

PREDIKSI TINGKAT KEPARAHAN COVID 19 DENGAN MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DAN LINEAR DISCRIMINANT ANALYSIS



Anggota Kelompok A:

Bonifasius Ariesto Adrian Finantyo	00000042580
Frederich Elisafan Latuputty	00000044441
Charisma Christ Anugrah	00000043903
Geraldus Dwi Atmaka	00000047880

KELAS IF540-C

SEMESTER GASAL 2022-2023

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2022

JUDUL: PREDIKSI TINGKAT KEPARAHAN COVID 19 DENGAN MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DAN LINEAR DISCRIMINANT ANALYSIS

1. Latar Belakang

Pada akhir tahun 2019, dunia dehebohkan dengan penemuan virus jenis baru yang cukup berbahaya. Virus ini dianggap cukup berbahaya karena tingkat infeksi yang terjadi pada saluran pernapasan dan usus sangat parah. Virus ini pertama kali teridentifikasi di daerah Wuhan provinsi Hubei, Tiongkok. Setelah menginfeksi dan menyebabkan kematian ribuan orang di Cina, virus ini kemudian mulai menyebar ke seluruh dunia dan setiap harinya jumlah kasus yang terkonfirmasi positif terus meningkat. Hingga saat ini, tercatat telah terjadi 29 juta kasus positif COVID-19 yang tersebar di 216 negara [1]. Tentunya angka tersebut bukanlah nilai yang kecil dan menarik perhatian seluruh negara. Organisasi Kesehatan Dunia menamai penyakit ini sebagai Corona Virus Diseases 2019 (COVID-19) dan selanjutnya dinyatakan sebagai pandemi karena infektivitas yang menyebar luas dan tingkat penularan yang tinggi.

Covid-19 pertama dilaporkan masuk ke Indonesia pada tanggal 2 Maret 2020, yakni dengan jumlah 2 kasus terkonfirmasi positif [2]. Setiap harinya, kasus penyebaran Covid-19 di Indonesia terus meningkat. Pemerintah menghimbau masyarakat untuk melakukan social distancing demi memutus mata rantai penyebaran Covid-19 yang tersebar di berbagai wilayah di Indonesia [3]. Berdasarkan keputusan menteri kesehatan RI Nomor HK.01.07/MENKES/4641/2021 tentang Panduan Pelaksanaan Pemeriksaan, Pelacakan, Karantina dan Isolasi dalam Rangka Percepatan Pencegahan dan Pengendalian COVID-19 terdapat 5 derajat gejala COVID-19 yaitu tanpa gejala, gejala ringan, gejala sedang, gejala berat, dan kritis [4]. Penderita atau orang yang diduga terkena COVID-19 dengan gejala ringan hingga sedang dapat melakukan karantina mandiri di rumah selama 14 hari. Sedangkan bagi yang memiliki gejala berat hingga kritis disarankan untuk dilakukan perawatan di rumah sakit oleh tenaga medis. Tanda dan gejala umum yang dapat diamati dari seseorang yang terinfeksi antara lain batuk, sesak napas, dan demam. Pada kasus yang lebih berat, infeksi dapat menyebabkan pneumonia, sindrom pernapasan akut, gagal ginjal, hingga kematian [5]. Di Indonesia sendiri, angka kematian yang disebabkan oleh COVID-19 sejauh ini telah mencapai 158.380 orang. Lonjakan angka kematian ini tidak hanya dialami Indonesia saja, melainkan juga terjadi di seluruh dunia [6].

Lonjakan ini tentunya menjadi ancaman yang cukup berbahaya bagi Indone-

sia, terutama pada segmen usia produktif yang berdampak langsung pada berbagai bidang kehidupan masyarakat. Dengan demikian, deteksi awal menjadi kunci dalam menekan angka kematian yang disebabkan oleh COVID-19. Sejauh ini cara mendeteksi COVID-19 yang paling efektif dan akurat adalah dengan tes PCR atau Polymerase Chain Reaction, dimana metode ini memiliki tingkat keakuratan berkisar 71-98 persen [7]. Akan tetapi kenyataan di lapangan, tes PCR membutuhkan waktu dan biaya tambahan yang tidak sedikit.

Berangkat dari permasalahan ini, peneliti memutuskan untuk membuat sistem yang dapat memprediksi tingkat keparahan infeksi COVID-19 yang dialami oleh seseorang. Sistem ini dibangun guna membantu percepatan penanganan dan mitigasi COVID-19 di Indonesia. Melalui sistem yang hendak dibangun ini, diharapkan dapat mengatasi masalah yang terjadi dengan cara memungkinkan masyarakat untuk melakukan diagnosis dini secara mandiri terkait status COVID-19 dan tingkat keparahan infeksi berdasarkan gejala yang dialami. Selain itu, sistem ini juga dapat dimanfaatkan oleh tenaga kesehatan untuk melakukan diagnosis awal dalam rangka pemberian tindakan darurat sebelum dilakukannya tes PCR.

Sistem prediksi tingkat keparahan COVID-19 akan dibangun dengan menggunakan algoritma Support Vector Machine (SVM) dan Linear Discriminant Analysis (LDA). Algoritma Support Vector Machine dipilih karena dianggap algoritma terbaik dalam melakukan klasifikasi dan menghasilkan nilai akurasi yang sangat tinggi yakni 91,91% [8]. Sedangkan, algoritma Linear Discriminant Analysis dipilih karena algoritma ini merupakan teknik klasifikasi yang mudah diimplementasikan dan menghasilkan nilai akurasi yang cukup tinggi yakni 87,10% [9]. Dengan demikian peneliti menggunakan kedua algoritma ini untuk melakukan prediksi dan membandingkan algoritma mana yang terbaik untuk menyelesaikan permasalahan ini.

2. Rumusan Masalah

Berdasarkan latar belakang diatas, maka peneliti merumuskan beberapa rumusan masalah sebagai berikut :

- Apakah algoritma analisis diskriminan linier dapat digunakan untuk memprediksi tingkat keparahan COVID-19?
- Apakah algoritma *support vector machine* dapat digunakan untuk memprediksi tingkat keparahan COVID-19?
- Bagaimana ketepatan hasil prediksi algoritma analisis diskriminan linier dan *support vector machine* dalam memprediksi tingkat keparahan COVID-19?

- Manakah algoritma yang tepat dalam melakukan prediksi tingkat keparahan COVID-19 dari kedua algoritma yang digunakan?

3. Batasan Masalah

Berikut ini adalah batasan masalah yang digunakan pada penelitian ini, antara lain yaitu:

- Variabel yang akan digunakan dalam melakukan prediksi tingkat keparahan adalah *fever, tiredness, dry cough, difficulty in breathing, sore throat, none symptom, pains, nasal congestion, runny nose, diarrhea, experiencing with COVID 19, age, dan contact with COVID-19 patients*.
- Sampel dari penelitian ini menggunakan dataset berjudul *COVID-19 Symptoms Checker* yang diambil dari Kaggle.
- Prediksi tingkat keparahan COVID-19 dilakukan dengan menggunakan algoritma analisis diskriminan linier dan *support vector machine*.

4. Tujuan Penelitian

Dari latar belakang dan perumusan masalah yang telah ditetapkan, maka dapat ditentukan tujuan penelitian sebagai berikut :

- Mengetahui algoritma analisis diskriminan linier dapat digunakan untuk memprediksi tingkat keparahan COVID-19.
- Mengetahui algoritma *support vector machine* dapat digunakan untuk memprediksi tingkat keparahan COVID-19.
- Mengetahui ketepatan hasil prediksi yang dilakukan analisis diskriminan linear dan *support vector machine* dalam memprediksi tingkat keparahan COVID-19.
- Mengetahui algoritma yang tepat dalam melakukan prediksi tingkat keparahan COVID-19.

5. Manfaat Penelitian

Dari latar belakang, perumusan masalah dan tujuan penelitian yang telah diteliti oleh peneliti, peneliti dapat menentukan manfaat penelitian sebagai berikut :

- Penelitian ini diharapkan dapat menjadi referensi bagi pemerintahan Indonesia untuk merumuskan kebijakan pemerintah dalam rangka mengetahui tingkat keparahan COVID-19.

- Penelitian ini diharapkan dapat menambah pengetahuan dalam melihat keterkaitan seberapa parah gejala dari COVID-19.
- Hasil penelitian ini diharapkan dapat digunakan sebagai bahan rujukan penelitian lebih lanjut.

6. Telaah Literatur

Dalam menyusun penelitian ini, peneliti menggunakan beberapa literatur berikut sebagai dasar penelitian.

6.1. Support Vector Machine

SVM atau Support Vector Machine dikenalkan pertama kali oleh Vapnik pada tahun 1992 sebagai salah satu metode *learning machine* yang bekerja dengan prinsip untuk menemukan *hyperplane* terbaik. SVM memiliki fitur mendasar dari pemisahan data berlabel biner yang berpusat pada garis yang mencapai data berlabel jarak maksimum. Untuk membantu memisahkan data berlabel, SVM menggunakan *hyperplane* yang membagi bidang menjadi kelas dan mengukur margin maksimum di mana di kelas terletak di kedua sisi. Tujuan dari SVM adalah untuk menemukan *hyperplane* terbaik seperti dibawah ini [10].

$$w.x + b = 0 \quad (1)$$

fungsi keputusan dapat dinyatakan sebagai berikut:

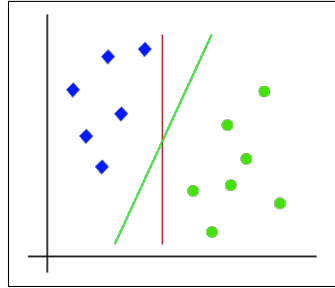
$$f(x) = \text{sign}(w.x + b) \quad (2)$$

Hyperplane sendiri adalah sebuah fungsi digunakan untuk pemisah antar kelas. Dalam 2-D fungsi yang digunakan untuk klasifikasi antar kelas disebut sebagai *line whereas*, fungsi yang digunakan untuk klasifikasi antar kelas dalam 3-D disebut *plane similarly*, sedangkan fungsi yang digunakan untuk klasifikasi di dalam ruang kelas dimensi yang lebih tinggi disebut *hyperplane* [11].

Terdapat 2 jenis SVM [12], yaitu:

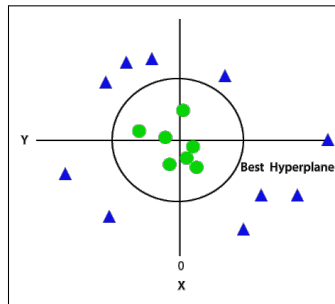
1. SVM Linear

SVM linear digunakan untuk data yang dapat dipisahkan secara linear, yang berarti jika sebuah dataset dapat diklasifikasi menjadi dua kelas dengan menggunakan sebuah garis lurus tunggal, maka data tersebut disebut sebagai data yang dapat dipisahkan secara linear, dan classifier yang digunakan disebut sebagai Linear SVM classifier.



Gambar 1. SVM Linear
source: [12]

2. SVM Non Linear SVM non-linear digunakan untuk data yang dapat dipisahkan secara non-linear, yang berarti jika sebuah dataset tidak dapat diklasifikasi menggunakan garis lurus, maka data tersebut disebut data non-linear dan classifier yang digunakan disebut sebagai Non-linear SVM classifier.



Gambar 2. SVM Non-Linear
source: [12]

6.2. Linear Discriminant Analysis

LDA merupakan salah satu supervised learning yang dikembangkan oleh R. A. Fisher pada tahun 1936. Classifier ini menentukan kombinasi karakteristik linear yang menggambarkan dua kelas atau lebih. LDA dikembangkan untuk mentransformasikan suatu parameter ke ruang dimensi rendah dengan memaksimalkan rasio varian antar kelas dan meminimalkan rasio varian dalam kelas, sehingga menjamin pemisahan kelas yang dilakukan secara maksimum [13].

Langkah 1. Untuk melakukan transformasi dengan LDA adalah menghitung separabilitas antar beda kelas (varian antar kelas) menggunakan Persamaan 1.

$$S_B = \sum_{i=1}^c n_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (3)$$

Dimana c adalah kelas, n_i adalah jumlah dari kelas i , μ_i adalah rata-rata kelas i , dan μ adalah rata-rata keseluruhan kelas.

Langkah 2. Menghitung jarak antara rata-rata dengan sampel dari tiap kelas (varian dalam kelas) menggunakan Persamaan 2.

$$S_W = \sum_{i=1}^c \sum_{j=1}^{n_i} (x_j^i - \mu_i)(x_j^i - \mu_i)^T \quad (4)$$

Dimana x_j^i adalah sampel j yang berada di kelas i .

Langkah 3. Merekonstruksi ruang dimensi rendah dengan menggunakan *Fisher's criterion* pada Persamaan 3.

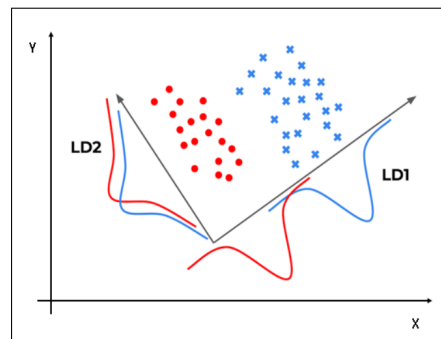
$$J(W) = \arg \max \frac{W^T S_B W}{W^T S_W W} \quad (5)$$

Dimana S_W adalah matriks non-singular dengan menghitung nilai eigen dan eigen vektornya. Hasil formulasi ini dapat dilihat pada Persamaan 4.

$$S_W W = \lambda_B W \quad (6)$$

Dengan W adalah transformasi matriks dan λ adalah nilai eigen dari transformasi matriks.

Nilai eigen merupakan kekuatan dari vektor eigen. Kekuatan vektor eigen tersebut mencerminkan kemampuan transformasi LDA untuk membedakan antar kelas. Dengan demikian, vektor eigen dengan nilai eigen tertinggi yang digunakan untuk membangun ruang dimensi rendah, sedangkan vektor eigen lainnya bisa diabaikan. Gambar 1 merupakan ilustrasi LDA dalam mengklasifikasi suatu kelas.



Gambar 3. Ilustrasi LDA
source: [14]

7. Daftar Pustaka

- [1] W. R. Hidayani, P. S. K. Masyarakat, and S. Respati, “Faktor faktor risiko yang berhubungan dengan covid 19 : Literature review,” *Jurnal Untuk Masyarakat Sehat (JUKMAS)*, vol. 4, pp. 120–134, 10 2020. [Online]. Available: <http://ejournal.urindo.ac.id/index.php/jukmas/article/view/1015>
- [2] N. Dwitri, J. A. Tampubolon, S. Prayoga, F. I. R. Zer, and D. Hartama, “Penerapan algoritma k-means dalam menentukan tingkat penyebaran pandemi covid-19 di indonesia,” *Jurnal Teknologi Informasi*, 2020.
- [3] S. Sindi, W. R. O. Ningse, I. A. Sihombing, F. I. R. Zer, and D. Hartama, “Analisis algoritma k-medoids clustering dalam pengelompokan penyebaran covid-19 di indonesia,” *Jurnal Teknologi Informasi*, 2020.
- [4] “Pasien Positif COVID-19 Tanpa Gejala Cukup Isoman.” [Online]. Available: <https://sehatnegeriku.kemkes.go.id/baca/rilis-media/20220204/1139250/pasien-positif-covid-19-tanpa-gejala-cukup-isoman/>
- [5] N. Wuryani, S. Agustiani, I. Komputer, and N. Mandiri, “Random forest classifier untuk deteksi penderita covid-19 berbasis citra ct scan,” *Jurnal Teknik Komputer AMIK BSI*, vol. 7, 2021. [Online]. Available: <http://ejournal.bsi.ac.id/ejurnal/index.php/jtk>
- [6] S. Sukesih, U. Usman, S. Budi, and D. N. A. Sari, “Pengetahuan dan sikap mahasiswa kesehatan tentang pencegahan covid-19 di indonesia,” *Jurnal Ilmu Keperawatan dan Kebidanan*, vol. 11, 2020.
- [7] V. Shah, R. Keniya, A. Shridharani, M. Punjabi, J. Shah, and N. Mehendale, “Diagnosis of covid-19 using ct scan images and deep learning techniques,” *Emergency Radiology*, vol. 28, pp. 497–505, 6 2021. [Online]. Available: <https://link.springer.com/article/10.1007/s10140-020-01886-y>
- [8] B. Shi, H. Ye, J. Zheng, Y. Zhu, A. A. Heidari, L. Zheng, H. Chen, L. Wang, and P. Wu, “Early recognition and discrimination of covid-19 severity using slime mould support vector machine for medical decision-making,” *IEEE Access*, vol. 9, pp. 121 996–122 015, 2021.
- [9] W. Astuti and A. Adiwijaya, “Principal component analysis sebagai ekstraksi fitur data microarray untuk deteksi kanker berbasis linear discriminant analysis,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 3, pp.

72–77, 4 2019. [Online]. Available: <http://ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/1161>

- [10] Z. Rustam, Y. Amalia, S. Hartini, and G. S. Saragih, “Linear discriminant analysis and support vector machines for classifying breast cancer,” *IAES International Journal of Artificial Intelligence*, vol. 10, 2021.
- [11] A. S. Nugroho, A. B. Witarto, and D. Handoko, “Support vector machine,” *Proceeding Indones. Sci. Meeting Cent. Japan*, 2003.
- [12] “Jenis-jenis SVM.” [Online]. Available: <https://www.trivusi.web.id/2022/04/algoritma-svm.html>
- [13] H. Ahmed and A. K. Nandi, *Condition Monitoring with Vibration Signals: Compressive Sampling and Learning Algorithms for Rotating Machines*, 12 2019.
- [14] B. P. Kamiel, Y. Ahmad, J. T. Mesin, F. Teknik, and U. M. Yogyakarta, “Deteksi kavitasi menggunakan linear discriminant analysis pada pompa sentrifugal,” *Turbo : Jurnal Program Studi Teknik Mesin*, vol. 9, 12 2020. [Online]. Available: <https://ojs.ummetro.ac.id/index.php/turbo/article/view/1326>