



Sistem Cerdas Klasifikasi Gejala Awal COVID-19 dan Influenza Menggunakan Metode SVM

Bella Aprilia Sandra*¹, Hary Sabita²

^{1,2} Program Studi Teknik Informatika, Institut Informatika dan Bisnis Darmajaya, Bandar Lampung, Indonesia

*Email Penulis Korespondensi: bellaas.1711010048@mail.darmajaya.ac.id

Abstrak

Self-diagnosis merupakan tindakan menentukan sendiri penyakit berdasarkan informasi yang dimiliki, menjadi tantangan dalam mengklasifikasikan gejala awal COVID-19 dan Influenza. Tujuan penelitian ini adalah mengembangkan sistem cerdas menggunakan metode Support Vector Machine untuk membedakan antara COVID-19 dan Influenza berdasarkan gejala awal pasien RSUD Ragab Begawe Caram Mesuji. Metode pengumpulan data meliputi studi literatur dan dokumentasi dari berbagai sumber. Evaluasi dilakukan dengan menggunakan confusion matrix untuk mengukur akurasi model SVM dalam mengklasifikasikan gejala awal COVID-19 dan Influenza. Hasil penelitian menunjukkan bahwa model SVM mencapai akurasi sebesar 62%, dengan performa yang bervariasi di setiap kelas gejala. Evaluasi memperlihatkan bahwa sistem ini dapat membedakan antara gejala awal COVID-19 dan Influenza, memberikan diagnosa berdasarkan hasil klasifikasi. Rekomendasi bagi pengembangan selanjutnya adalah peningkatan jumlah Dataset untuk meningkatkan akurasi model secara keseluruhan, serta desain antarmuka yang dapat digunakan secara efektif oleh tenaga medis dan masyarakat.

Kata kunci— COVID-19, SVM, Influenza, Klasifikasi, Self-diagnosis

Abstract

Self-diagnosis is the act of determining one's own illness based on available information, poses a challenge in classifying early symptoms of COVID-19 and Influenza. The aim of this research is to develop an intelligent system using Support Vector Machine method to differentiate between COVID-19 and Influenza based on early symptoms of patients at RSUD Ragab Begawe Caram Mesuji. Data collection methods include literature studies and documentation from various sources. Evaluation is conducted using a Confusion Matrix to measure the accuracy of the SVM model in classifying early symptoms of COVID-19 and Influenza. The research findings indicate that the SVM model achieved an accuracy of 62%, with varying performance across symptom classes. The evaluation demonstrates that this system can distinguish between early symptoms of COVID-19 and Influenza, providing diagnoses based on classification outcomes. Recommendations for further development include increasing the Dataset size to enhance overall model accuracy, and designing an interface that can be effectively used by medical professionals and the public.

Keywords— COVID-19, SVM, Influenza, Classification, Self-Diagnosis

1. PENDAHULUAN

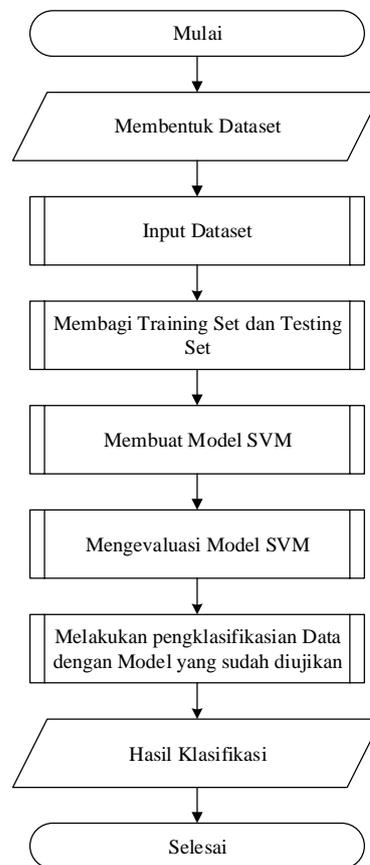
Pesatnya kemajuan teknologi informasi dan komunikasi telah membuka akses luas masyarakat terhadap informasi kesehatan melalui internet. Fenomena ini, meskipun memberikan keuntungan, juga memicu praktik *self-diagnosis*, di mana individu tanpa bimbingan medis cenderung menilai gejala kesehatan mereka sendiri. Seorang yang sedang dihadapi beberapa keluhan bisa langsung menanyakan keluhannya ataupun membaca informasi yang berkaitan dengan keluhannya di internet [1]. Praktik ini kerap berujung pada kesalahan diagnosis, kecemasan berkelanjutan, dan bahkan obsesi terhadap kondisi kesehatan yang belum diverifikasi secara medis. Konteks pandemi COVID-19 memberikan sorotan khusus terhadap pentingnya pemahaman yang lebih mendalam terkait gejala klinis, terutama dalam membedakannya dari penyakit umum seperti Influenza. Flu atau Influenza merupakan penyakit biasa dan banyak orang yang mengira penyakit ini akan sembuh dengan sendirinya [2]. Namun ketika wabah COVID-19 menjangkit, Influenza mendapat perhatian serius karena Influenza dan COVID-19 sama-sama menyerang saluran pernafasan hanya saja pada COVID-19 infeksi saluran pernafasan terjadi lebih cepat. Masyarakat perlu mengetahui bahwa kedua penyakit ini memiliki gejala awal yang sama, tetapi pada COVID-19, gejala ini akan semakin memburuk bahkan dapat mengakibatkan radang paru-paru [3].

Dalam penelitian ini, metode *Support Vector Machine* (SVM) dipilih untuk menganalisis pola klasifikasi data gejala awal COVID-19 dan Influenza. Pemilihan metode SVM didasari oleh temuan dari beberapa referensi yang membahas keunggulan SVM dalam menangani masalah klasifikasi pada berbagai konteks. Penelitian yang dilakukan oleh Puspitasari, dkk dalam 'Klasifikasi Penyakit Gigi Dan Mulut Menggunakan Metode *Support Vector Machine*' menunjukkan bahwa SVM mencapai akurasi klasifikasi rata-rata sebesar 94.442% dalam mengidentifikasi penyakit gigi dan mulut, menggunakan *Dataset* sebanyak 122 data dengan parameter-parameter tertentu [4]. Hasil ini menunjukkan keefektifan SVM dalam menangani masalah klasifikasi pada bidang kesehatan. Analisis terkait penggunaan SVM oleh Apriyani, H., & Kurniati dalam perbandingan dari dua algoritma yaitu algoritma naïve bayes dan algoritma *Support Vector Machine* untuk klasifikasi penyakit diabetes membuktikan hasil akurasi tertinggi yaitu algoritma *Support Vector Machine* dengan *kernel Polynomial* yang hasilnya 96.2704% dan tingkat *error* sebanyak 3.7296% dapat disimpulkan algoritma yang akurat yaitu algoritma *Support Vector Machine* [5].

Klasifikasi jenis penyakit malaria berdasarkan gejala yang ditimbulkan menggunakan SVM dengan *kernel RBF* berdasarkan 200 data yang diambil dari Dinas Kesehatan Kabupaten Nabire, Papua. Pengujian menggunakan *K-fold Cross Validation* dengan nilai *K-fold*=10. Hasil akurasi terbaik adalah 72% [6]. Mendiagnosa luka luar menggunakan algoritma *Support Vector Machine* (SVM) melalui tiga tahap—pengumpulan data, *preprocessing*, dan klasifikasi dengan menggunakan ekstraksi fitur GLCM, menghasilkan akurasi 96,39%, presisi 93,06%, *Recall* 92,85%, dan *F1-Score* 92,58%, menunjukkan bahwa SVM efektif untuk klasifikasi gambar luka luar [7]. Algoritma *Support Vector Machine* (SVM) dengan *kernel Polynomial* juga digunakan untuk mengklasifikasikan data anak balita di bawah 60 bulan di Kabupaten Enrekang, Sulawesi Selatan, guna menentukan apakah mereka mengalami stunting atau tidak, menunjukkan bahwa SVM dengan metode 10 *K-fold Cross Validation* menghasilkan rata-rata akurasi 96,98%, presisi 96,99%, *Recall* 96,98%, dan *F1-Score* 96,94%, dengan akurasi total 98% berdasarkan *Confusion Matrix* [8]. Referensi dari studi-studi ini memberikan landasan yang valid bagi penggunaan SVM dalam menganalisis pola klasifikasi gejala awal COVID-19 dan Influenza. Melalui penelitian yang melibatkan algoritma *Support Vector Machine*, diharapkan bahwa penggunaan rekam medis dari RSUD Ragab Begawe Caram Mesuji dapat memberikan kontribusi yang signifikan dalam mengurangi kesalahan *self-diagnosis*, meningkatkan pemahaman masyarakat, dan memberikan landasan yang lebih akurat untuk penanganan kesehatan, terutama dalam konteks penyakit menular.

2. METODE PENELITIAN

Metode *Support Vector Machine* (SVM) digunakan dalam klasifikasi data gejala awal COVID-19 dan Influenza. Proses pengembangan sistem ini mengikuti alur yang terdiri dari beberapa tahap seperti dilihat pada Gambar 1.



Gambar 1 Alur SVM

2.1 Membentuk Dataset

Untuk membentuk *Dataset* dari rekam medis Rumah Sakit Umum Daerah Ragab Begawe Caram Mesuji, data hasil diagnosis pasien dikumpulkan secara langsung dari rumah sakit, yang mencakup kelas Suspek Covid, *Probable Covid*, Influenza, dan Terkonfirmasi Covid, masing-masing menunjukkan tingkat kepastian dan hasil laboratorium yang berbeda. Data ini kemudian dimasukkan ke dalam *Microsoft Excel*, memastikan setiap entri memiliki kolom yang sesuai seperti gejala, hasil laboratorium, dan diagnosis akhir. Selanjutnya, *cleaning data* dengan mengatasi ketidakkonsistenan dalam penulisan label, mengubah label 'ya' dan 'tidak' menjadi '1' dan '0', serta memindahkan kolom Diagnosa Medis ke akhir tabel. Setelah data dibersihkan, file *Excel* tersebut disimpan dengan nama "RSUD (0, 1)".

Tabel 1 *Dataset*

Nama	Usia	Demam	Sesak Nafas	Mual	Muntah	Pusing	Nyeri Telan	Tidak Nafsu Makan	Batuk Kering	Lemas	Anosmia	Lidah Pahit	Nyeri Otot	Batuk Berdahak	Diagnosa Medis
Ny Yuni Astuti	40	0	0	0	0	0	0	0	0	0	0	0	0	0	Influenza
Tn Zaenudin	28	0	1	1	1	0	0	0	1	0	0	0	0	0	Influenza
Tn Tauhid	65	0	0	1	1	0	1	1	0	0	0	0	0	0	Influenza
Tn Untung Dwi Hartono	57	0	0	0	0	0	1	0	1	0	0	0	0	0	Influenza

Ny Nuryani	38	1	1	1	1	0	1	1	0	1	0	0	0	1	Probable Covid
Ny Katini	59	0	1	1	1	0	0	1	0	1	0	0	0	0	Suspek Covid
Tn Samsudin	57	1	0	1	1	1	0	1	0	1	0	0	0	0	Suspek Covid
Ny Nurlela	30	1	1	1	1	1	0	1	1	1	0	0	0	0	Terkonfirmasi Covid
By Ny Nurlela	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Terkonfirmasi Covid
Tn Regan Tormando	25	1	1	1	1	1	1	1	1	1	0	0	0	0	Probable Covid
Ny Maryani	48	0	1	1	0	0	1	0	0	0	0	0	0	0	Probable Covid
Tn Supono	56	1	1	1	1	1	0	1	1	1	0	0	0	0	Terkonfirmasi Covid
Tn Hendri AZ	50	1	1	0	0	0	1	0	1	0	0	0	0	0	Terkonfirmasi Covid
Ny Suyatnuk	55	1	1	0	0	1	1	1	0	1	0	0	1	1	Terkonfirmasi Covid
Tn Kholik Mawardi	50	1	0	1	1	0	0	1	1	1	0	0	1	0	Terkonfirmasi Covid
Ny Endang Suharsih	53	1	1	0	0	1	1	1	1	1	1	0	1	0	Terkonfirmasi Covid
Ny Dwi Rahayu Ningsih	38	1	0	0	1	1	1	1	0	0	1	1	0	1	Terkonfirmasi Covid
Nn Irma Yunira	23	0	1	0	0	0	0	1	0	1	0	0	0	0	Influenza
Nn Ganis Rengganis	36	1	1	1	1	1	1	1	0	1	1	0	0	1	Terkonfirmasi Covid
Ny Ustati	41	1	0	1	1	0	0	0	0	0	1	0	1	0	Influenza
Tn Yatimin	52	1	1	1	1	1	1	1	0	0	0	1	0	1	Terkonfirmasi Covid
By Ny Ustati	0	1	0	0	0	0	0	0	0	0	0	0	0	0	Influenza
Ny Kusnida Sari	31	1	1	1	0	0	0	0	1	1	1	1	0	0	Terkonfirmasi Covid
An Al Mutawakil Allallah	6	1	1	1	1	0	0	0	1	0	1	0	0	0	Probable Covid
Tn Takrip	65	1	1	1	1	1	0	0	1	1	1	0	0	0	Terkonfirmasi Covid
Tn Mat Juan	61	1	1	1	1	1	0	1	1	0	0	0	0	0	Probable Covid
Tn Pujo Yuhono	52	1	1	1	1	1	1	0	0	1	1	0	1	1	Terkonfirmasi Covid
Ny Neni Musawiyah	45	1	1	1	1	1	1	1	0	0	0	1	1	1	Terkonfirmasi Covid
Ny Susilowati	36	1	0	1	1	1	1	1	1	0	1	1	0	0	Terkonfirmasi Covid
An Regina	6	1	1	0	0	0	0	0	0	0	0	0	0	0	Probable Covid
Tn Sunadji	66	1	1	0	0	0	0	0	1	1	0	0	0	0	Suspek Covid
Tn Suyadi	51	0	0	1	1	1	1	1	0	1	1	1	1	1	Terkonfirmasi Covid
Tn Nur Huda	52	1	1	0	0	1	0	0	0	1	1	1	0	0	Terkonfirmasi Covid
Ny Misropah	35	1	1	1	0	0	1	1	1	0	0	0	0	0	Terkonfirmasi Covid
Tn Exo Darsanto	46	0	0	0	1	1	0	0	0	1	0	0	0	0	Terkonfirmasi Covid
Ny Sugiarti	25	0	0	1	1	0	0	0	0	1	0	0	0	0	Terkonfirmasi Covid
By Ny Sugiarti	0	1	0	1	1	0	0	0	0	0	0	0	0	0	Influenza
Ny Siti Rofiah	46	1	1	0	1	1	1	1	1	1	1	1	0	0	Terkonfirmasi Covid
Tn Agus Sudrajat	37	0	1	1	1	1	1	0	0	0	1	1	1	1	Terkonfirmasi Covid

Ny Sukamtingi	61	1	1	1	1	1	1	0	0	0	0	1	1	1	Terkonfirmasi Covid
Tn Mujianto	22	1	1	1	1	0	1	0	0	1	1	1	0	1	Terkonfirmasi Covid
Tn Misdi Udin	58	1	1	1	1	1	1	0	1	0	1	1	1	0	Terkonfirmasi Covid
Ny Siti Patonah	49	0	1	1	1	1	0	0	0	0	0	0	0	0	Influenza
Ny Dwi Astuti	25	1	1	1	1	1	1	1	1	0	1	1	0	0	Suspek Covid
Tn Sumono	45	1	1	1	1	1	1	0	0	1	1	1	1	1	Terkonfirmasi Covid
Tn Rismanto	69	1	1	1	1	1	0	1	1	1	1	0	0	0	Influenza
Tn Sutrisno	63	1	1	0	0	1	1	0	0	0	1	1	1	1	Influenza
Ny Usrek	62	1	1	1	1	1	0	0	1	0	1	1	1	0	Terkonfirmasi Covid
Ny Siti Markamah	53	1	1	1	0	1	1	1	1	0	1	1	0	0	Terkonfirmasi Covid
Ny Eka Melya	34	1	0	0	0	0	1	0	1	0	1	1	0	0	Terkonfirmasi Covid
Ny Hartatik	22	1	1	1	0	1	0	1	0	0	0	0	0	0	Influenza
Ny Tri Wahyuni	46	1	1	1	0	0	0	1	1	0	1	0	1	0	Terkonfirmasi Covid

2.2 Input Dataset

Mengimpor *library* yang diperlukan, seperti *pandas*, yang akan membantu dalam memanipulasi dan menganalisis data. File Excel yang berisi *Dataset* rekam medis "RSUD (0, 1)" kemudian diimpor ke dalam Python sebagai *DataFrame* menggunakan perintah:

```
import pandas as pd
data = pd.read_excel('RSUD (0,1).xlsx')
```

Langkah berikutnya adalah mengubah *DataFrame* tersebut ke dalam bentuk *array* dengan kode:

```
ini_data = data.iloc[:,1:15].to_numpy()
ini_label = data['Diagnosa Medis'].to_list()
```

Di mana '*ini_data*' memuat data dari kolom Demam hingga Batuk Berdahak dalam bentuk *array* dan '*ini_label*' memuat kolom Diagnosa Medis dalam bentuk *list*, yang diperlukan untuk pembuatan model *Support Vector Machine* (SVM) pada tahap selanjutnya.

2.3 Membagi Training Set dan Testing Set

Memisahkan data menjadi *training set* dan *testing set*, digunakan function '*train_test_split()*' dari *Scikit-learn*. *DataFrame* yang sebelumnya telah diubah ke dalam bentuk struktur data *array*, yaitu '*ini_data*' dan '*ini_label*', diperlukan. Aturan pembagian *Dataset* yang digunakan adalah 70% untuk *training set* dan 30% untuk *testing set*.

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(ini_data, ini_label, test_size=0.3,
                                                  random_state=0)
```

Pada kode program di atas, kedua *array* '*ini_data*' dan '*ini_label*' dimasukkan ke dalam function '*train_test_split()*', yang menghasilkan empat buah *array*: '*X_train*' dan '*y_train*' untuk *training set*, serta '*X_test*' dan '*y_test*' untuk *testing set*. Parameter '*test_size=0.3*' menunjukkan bahwa 30% dari seluruh isi *DataFrame* akan dijadikan *testing set*, sedangkan sisanya akan menjadi *training set*. Dalam bidang *Data Science*, huruf X (huruf besar) untuk mewakili *feature* dan huruf y (huruf kecil) untuk mewakili target *feature*. Oleh karena itu, fitur

di *training set* diwakili oleh 'X_train' dan fitur di *testing set* diwakili oleh 'X_test', sementara 'y_train' dan 'y_test' mewakili target untuk *training set* dan *testing set* [9].

2.4 Membuat Model SVM

Dalam pemilihan parameter *gamma* dan C untuk model SVM, telah dilakukan eksperimen dengan berbagai nilai yang berbeda untuk mengoptimalkan akurasi model. Rentang nilai *gamma* dan C yang biasanya digunakan dalam eksperimen ini sangat bervariasi tergantung pada data dan masalah yang dihadapi.

Tabel 2 Parameter SVM

Parameter	Value
Kernel Type	Poly
Gamma	0,001
C	5

Setelah mencoba banyak variasi, hasil terbaik untuk akurasi model tercapai dengan menggunakan *kernel Polynomial* dengan *gamma*=0.001 dan C=5, seperti yang tertera pada Tabel 2.

```

from sklearn import svm

# Membuat model SVM dengan kernel Poly dan parameter yang ditentukan
classifier2 = svm.SVC(kernel='poly', gamma=0.001, C=5)

# Melatih model menggunakan data training
classifier2.fit(X_train, y_train)

# Memprediksi target menggunakan data testing
y_predict2 = classifier2.predict(X_test)

```

'svm.SVC()' digunakan untuk inialisasi model SVM dengan *kernel Polynomial*. Parameter yang digunakan adalah *kernel='poly'* untuk menggunakan *kernel Polynomial*, *gamma*=0.001 untuk mengatur nilai *gamma*, dan C=5 untuk mengatur parameter C. Hasilnya menunjukkan bahwa kombinasi ini memberikan performa yang optimal untuk kasus ini.

2.5 Mengevaluasi Model SVM

Setelah model SVM dibuat, selanjutnya melakukan validasi keakuratan menggunakan metode *K-fold Cross Validation*. Metode ini digunakan untuk memastikan bahwa model memiliki performa yang konsisten dan dapat dipercaya. Variabel 'classifier2' berisi model SVC dengan parameter yang telah ditentukan sebelumnya seperti yang tercantum dalam Tabel 2. Model dilatih dengan menggunakan data training yang telah dipisahkan sebelumnya, 'X_train' dan 'y_train', menggunakan *method 'fit()'*. Selanjutnya, model diuji dengan menggunakan *method 'predict()'* pada data *testing 'X_test'*, dan hasilnya disimpan dalam variabel 'y_predict2'.

Setelah uji coba model, dilakukan validasi keakuratan menggunakan metode *K-fold Cross Validation*. *Cross Validation* adalah teknik yang digunakan untuk memastikan keakuratan model berdasarkan *Dataset* yang digunakan [10]. Pada kode program di atas, 'scores2' menyimpan hasil dari fungsi '*cross_val_score(classifier2, X, y, cv=3)*', yang menghitung skor akurasi model dengan melakukan validasi silang *K-fold* dengan 3 *fold*. Hasilnya ditampilkan dengan menggunakan '*print('Nilai scores 1:', scores2)*', yang menghasilkan nilai-nilai skor untuk setiap *fold*.

Selain itu, 'scores02' menghitung akurasi model dengan menggunakan 'metrics.accuracy_score(y_test, y_predict2)', dan hasilnya ditampilkan dengan 'print('Nilai scores 2:', scores02)'. Untuk evaluasi lebih rinci, hasil dari 'classification_report(y_test, y_predict2)' akan menampilkan laporan evaluasi model berdasarkan perhitungan *Confusion Matrix*. Berikut adalah contoh *output* yang ditampilkan:

```
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score
from sklearn.model_selection import cross_val_score

# Evaluasi model dengan K-fold Cross Validation
scores2 = cross_val_score(classifier2, X, y, cv=3)
scores02 = accuracy_score(y_test, y_predict2)
print(classification_report(y_test, y_predict2))
print('Nilai scores 1:', scores2)
print('Nilai scores 2:', scores02)
```

'classification_report(y_test, y_predict2)' akan menampilkan laporan evaluasi model dalam bentuk *Confusion Matrix*.

Tabel 1 *Confusion Matrix*

		Actual Values	
		Postive	Negative
Predicted Values	Postive	TP	FP
	Negative	FN	TN

- 1) TP (*True Positive*) adalah jumlah kasus di mana model SVM dengan benar mengidentifikasi data positif sebagai positif. Ini berarti bahwa model berhasil mendeteksi dengan benar adanya kondisi yang diharapkan.
- 2) FP (*False Positive*) adalah jumlah kasus di mana model SVM salah mengidentifikasi data negatif sebagai positif. Ini berarti bahwa model memberikan hasil positif padahal seharusnya negatif, sehingga menyebabkan alarm yang salah.
- 3) FN (*False Negative*) adalah jumlah kasus di mana model SVM salah mengidentifikasi data positif sebagai negatif. Ini berarti bahwa model gagal mendeteksi kondisi yang seharusnya terdeteksi, sehingga hasil negatif ini adalah kesalahan yang signifikan terutama jika kondisi tersebut kritis.
- 4) TN (*True Negative*) adalah jumlah kasus di mana model SVM dengan benar mengidentifikasi data negatif sebagai negatif. Ini berarti bahwa model berhasil mengenali dengan benar bahwa kondisi yang diuji memang tidak ada, yang menunjukkan kemampuan model untuk tidak memberikan alarm palsu. Adapun nilai-nilai dalam *Confusion Matrix* mencakup *Actual Values* yang merupakan nilai asli dari label atau kelas yang diperoleh dari rekam medis di RSUD Ragab Begawe Caram Mesuji, serta *Predicted Values* yang merupakan nilai prediksi dari model SVM berdasarkan data pengujian.

Confusion Matrix merupakan metode yang digunakan untuk mengukur atau melakukan perhitungan akurasi pada konsep data mining. *Confusion Matrix* terdiri dari baris data uji yang diprediksi benar dan tidak benar suatu data oleh model klasifikasi[11]. Nilai yang dapat dihasilkan dari *Confusion Matrix*, yaitu *Precision*, *Recall*, *F1-Score*, *Support*, *Accuracy*, *Macro Average*, dan *Weighted Average*.

1) *Precision*

Precision menggambarkan tingkat keakuratan antara data yang diminta dengan hasil prediksi yang diberikan oleh model maka, rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif seperti dapat dilihat pada formula (1).

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

2) Recall

Recall merupakan gambaran keberhasilan model dalam menemukan kembali sebuah informasi, maka rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif seperti pada formula (2).

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

3) F1-Score

F1-Score merupakan perbandingan rata-rata dari *Precision* dan *Recall* yang dibobotkan seperti pada formula (3).

$$F1 - Score = \frac{2 \times (Recall \times Precision)}{(Recall + Precision)} \quad (3)$$

4) Support

Support merupakan jumlah sampel dari setiap label yang dihitung.

5) Accuracy

Accuracy memberikan gambaran seberapa akurat model SVM dapat mengklasifikasikan data dengan benar. *Accuracy* merupakan rasio prediksi benar (TP dan TN) dengan keseluruhan data maka, dapat dikatakan bahwa *Accuracy* juga merupakan tingkat kedekatan nilai prediksi dengan nilai *Actual* (yang sebenarnya) dengan formula (4).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

6) Macro Average

Macro Average merupakan nilai rata-rata dari menjumlahkan seluruh nilai masing-masing *Precision*, *Recall*, dan *F1-Score* kemudian dibagi dengan jumlah label seperti disajikan pada formula (5).

$$Macro Average = \frac{Jumlah\ seluruh\ Precision/Recall/F1-Score}{Jumlah\ label} \quad (5)$$

7) Weighted Average

Weighted Average adalah nilai rata-rata yang didapat dari mengalikan *Precision*, *Recall*, dan *F1-Score* dengan *Support* dan dibagi total *Support* atau jumlah keseluruhan sampel yang diujikan seperti pada formula (6).

$$Weighted Average = \frac{(Precision \times Support) + (Recall \times Support) + (F1-Score \times Support)}{Total Support} \quad (6)$$

2.6 Melakukan Pengklasifikasian Data dengan Model yang Sudah Diujikan

$$hasil = classifier2.predict(ini_data)$$

Model SVM yang telah diinisialisasi sebelumnya dengan nama '*classifier2*' digunakan untuk memprediksi seluruh gejala awal yang telah diubah ke dalam bentuk struktur data *array* '*ini_data*'. Penggunaan *method* '*predict()*' pada model ini memungkinkan untuk menghasilkan prediksi berdasarkan *input* data yang diberikan, yang pada kasus ini adalah rekam medis dari RSUD Ragab Begawe Caram Mesuji.

2.7 Hasil Klasifikasi

Tujuan utama dari klasifikasi ini adalah untuk memberikan hasil klasifikasi yang nantinya akan dianalisis dan dievaluasi kinerjanya pada BAB Hasil dan Pembahasan. Berikut ini merupakan hasil klasifikasi diagnosis 52 pasien RSUD Ragab Begawe Caram Mesuji menggunakan model SVM yang telah dibuat sebelumnya:

Tabel 4 Hasil Klasifikasi Diagnosis Medis oleh Model SVM

No.	Nama	Hasil Klasifikasi
1.	Ny Yuni Astuti	Terkonfirmasi Covid
2.	Tn Zaenudin	Terkonfirmasi Covid
3.	Tn Tauhid	Influenza
4.	Tn Untung Dwi Hartono	Terkonfirmasi Covid
5.	Ny Nuryani	Terkonfirmasi Covid
6.	Ny Katini	Terkonfirmasi Covid
7.	Tn Samsudin	Terkonfirmasi Covid
8.	Ny Nurlela	Terkonfirmasi Covid
9.	By Ny Nurlela	Terkonfirmasi Covid
10.	Tn Regan Tormando	Terkonfirmasi Covid
11.	Ny Maryani	Terkonfirmasi Covid
12.	Tn Supono	Terkonfirmasi Covid
13.	Tn Hendri AZ	Terkonfirmasi Covid
14.	Ny Suyatnuk	Terkonfirmasi Covid
15.	Tn Kholik Mawardi	Terkonfirmasi Covid
16.	Ny Endang Suharsih	Terkonfirmasi Covid

Hasil klasifikasi ini akan digunakan untuk mengevaluasi kinerja model menggunakan *Confusion Matrix* seperti pada Tabel 3. Kemudian, hasil klasifikasi ini akan dianalisis lebih lanjut untuk membandingkan hasil klasifikasi model *SVM* dengan diagnosis asli pasien.

3. HASIL DAN PEMBAHASAN

Pada hasil penelitian ini dapat diketahui pasien dengan hasil diagnosa medis *Influenza* dapat dipastikan hanya mengalami *Influenza* saja sedangkan pasien dengan keterangan Suspek *Covid* dan *Probable Covid* memiliki hasil yang beragam serta pasien dengan hasil diagnosa medis Terkonfirmasi *Covid* dapat dipastikan positif terjangkit virus *COVID-19*.

Tabel 5 Perbandingan Diagnosa Medis Asli dan Hasil Prediksi Model SVM

No.	Nama	Diagnosa Medis Asli	Diagnosa Medis Hasil Prediksi
1.	Ny Yuni Astuti	Influenza	Terkonfirmasi Covid
2.	Tn Zaenudin	Influenza	Terkonfirmasi Covid
3.	Tn Tauhid	Influenza	Influenza
4.	Tn Untung Dwi Hartono	Influenza	Terkonfirmasi Covid
5.	Ny Nuryani	Probable Covid	Terkonfirmasi Covid
6.	Ny Katini	Suspek Covid	Terkonfirmasi Covid
7.	Tn Samsudin	Suspek Covid	Terkonfirmasi Covid
8.	Ny Nurlela	Terkonfirmasi Covid	Terkonfirmasi Covid
9.	By Ny Nurlela	Terkonfirmasi Covid	Terkonfirmasi Covid
10.	Tn Regan Tormando	Probable Covid	Terkonfirmasi Covid
11.	Ny Maryani	Probable Covid	Terkonfirmasi Covid
12.	Tn Supono	Terkonfirmasi Covid	Terkonfirmasi Covid
13.	Tn Hendri AZ	Terkonfirmasi Covid	Terkonfirmasi Covid
14.	Ny Suyatnuk	Terkonfirmasi Covid	Terkonfirmasi Covid

15.	Tn Kholik Mawardi	Terkonfirmasi Covid	Terkonfirmasi Covid
16.	Ny Endang Suharsih	Terkonfirmasi Covid	Terkonfirmasi Covid

Tabel 6 Evaluasi Hasil Klasifikasi dengan *Confusion Matrix*

		<i>Actual Values</i>			
		<i>Probable Covid</i>	Suspek Covid	Terkonfirmasi Covid	Influenza
<i>Predicted Values</i>	<i>Probable Covid</i>	0	0	0	0
	Suspek Covid	0	0	0	0
	Terkonfirmasi Covid	3	2	7	3
	Influenza	0	0	0	1

1) *Probable Covid*

TP = 0, yaitu pasien *Probable Covid* diprediksi *Probable Covid* berjumlah 0. FP = 0 + 0 + 0 = 0, yaitu pasien bukan *Probable Covid* diprediski *Probable Covid* berjumlah 0. FN = 0 + 3 + 0 = 3, yaitu pasien *Probable Covid* diprediksi bukan *Probable Covid* berjumlah 3. TN = 0 + 0 + 0 + 2 + 7 + 3 + 0 + 0 + 1 = 13, yaitu pasien bukan *Probable Covid* diprediksi bukan *Probable Covid* berjumlah 13.

2) *Suspek Covid*

TP = 0, yaitu pasien *Suspek Covid* diprediksi *Suspek Covid* berjumlah 0. FP = 0 + 0 + 0 = 0, yaitu pasien bukan *Suspek Covid* diprediksi *Suspek Covid* berjumlah 0. FN = 0 + 2 + 0 = 2, yaitu pasien *Suspek Covid* diprediksi bukan *Suspek Covid* berjumlah 2. TN = 0 + 0 + 0 + 3 + 7 + 3 + 0 + 0 + 1 = 14, yaitu pasien bukan *Suspek Covid* diprediksi bukan *Suspek Covid* berjumlah 14.

3) *Terkonfirmasi Covid*

TP = 7, yaitu pasien *Terkonfirmasi Covid* diprediksi *Terkonfirmasi Covid* berjumlah 7. FP = 3 + 2 + 3 = 8, yaitu pasien bukan *Terkonfirmasi Covid* diprediksi *Terkonfirmasi Covid* berjumlah 8. FN = 0 + 0 + 0 = 0, yaitu pasien *Terkonfirmasi Covid* diprediksi bukan *Terkonfirmasi Covid* berjumlah 0. TN = 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 1 = 1, yaitu pasien bukan *Terkonfirmasi Covid* diprediksi bukan *Terkonfirmasi Covid* berjumlah 1.

4) *Influenza*

TP = 1, yaitu pasien *Influenza* diprediksi *Influenza* berjumlah 1. FP = 0 + 0 + 0 = 0, yaitu pasien bukan *Influenza* diprediksi *Influenza* berjumlah 0. FN = 0 + 0 + 3 = 3, yaitu pasien *Influenza* diprediksi bukan *Influenza* berjumlah 3. TN = 0 + 0 + 0 + 0 + 0 + 0 + 0 + 3 + 2 + 7 = 12, yaitu pasien bukan *Influenza* diprediksi bukan *Influenza* berjumlah 12.

Tabel 7 Hasil Evaluasi Model SVM

	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
<i>Probable Covid</i>	0,00	0,00	0,00	3
<i>Suspek Covid</i>	0,00	0,00	0,00	2
<i>Terkonfirmasi Covid</i>	0,60	1,00	0,75	9
<i>Influenza</i>	1,00	0,50	0,67	2
<i>Accuracy</i>			0,62	16
<i>Macro Avg</i>	0,40	0,38	0,35	16
<i>Weighted Avg</i>	0,46	0,62	0,51	16

Berdasarkan Tabel 7, nilai *Precision* dari masing masing label, yaitu *Probable Covid* = 0,00, *Suspek Covid* = 0,00, *Terkonfirmasi Covid* = 0,60, dan *Influenza* = 1,00. Nilai *Recall* dari masing masing label, yaitu *Probable Covid* = 0,00, *Suspek Covid* = 0,00, *Terkonfirmasi Covid* = 1,00, dan *Influenza* = 0,50. Nilai *F1-Score* dari masing masing label, yaitu *Probable Covid* = 0,00, *Suspek Covid* = 0,00, *Terkonfirmasi Covid* = 0,75, dan *Influenza* = 0,67. *Support* adalah 16, terdapat 3 data pasien yang dijadikan sampel dengan label *Probable Covid*, 2 sampel dengan label *Suspek Covid*, 9 sampel dengan label *Terkonfirmasi Covid*, dan 2 sampel dengan label *Influenza* dengan total keseluruhan sampel yang digunakan yaitu 16 sampel. *Accuracy* dari hasil evaluasi model SVM yaitu 0,62 atau 62,00%. *Precision* mendapatkan *Macro Average* sebesar 0,40, *Recall* 0,38, dan *F1-Score* 0,35. *Precision* memiliki nilai *Weighted Average* yaitu 0,46, *Recall* 0,62 dan *F1-Score* 0,51.

Hasil evaluasi ini memberikan gambaran seberapa baik model SVM dapat mengklasifikasikan gejala awal COVID-19 berdasarkan data rekam medis. Meskipun akurasi model mencapai 62,00%, perlu diperhatikan bahwa performa model bervariasi di setiap kelas, dengan penekanan khusus pada kelas *Terkonfirmasi Covid* dan *Influenza*.

4. KESIMPULAN

Penelitian ini mengimplementasikan Metode *Support Vector Machine* (SVM) dalam Sistem Cerdas Klasifikasi Gejala Awal COVID-19 dan Influenza. Beberapa kesimpulan dapat diambil:

1) Kemampuan Klasifikasi Model SVM

Dari evaluasi model SVM, dapat disimpulkan bahwa nilai *Precision* menunjukkan variasi signifikan antar kelas, dengan *Terkonfirmasi Covid* mencapai 0,60 dan *Influenza* mencapai 1,00. Namun, nilai *Precision* untuk *Probable Covid* dan *Suspek Covid* adalah 0,00, menunjukkan kesulitan dalam mengenali kedua kelas ini dengan tepat. Sementara itu, *Recall* untuk *Terkonfirmasi Covid* mencapai 1,00, menunjukkan kemampuan model dalam mendeteksi kasus positif COVID-19 yang dikonfirmasi. Namun, *Recall* untuk *Probable Covid* dan *Suspek Covid* tetap 0,00, menunjukkan model SVM kesulitan dalam mengidentifikasi kelas-kelas tersebut. *F1-Score* untuk *Terkonfirmasi Covid* adalah 0,75 dan *Influenza* adalah 0,67, mencerminkan keseimbangan antara *Precision* dan *Recall*. Di sisi lain, *F1-Score* untuk *Probable Covid* dan *Suspek Covid* adalah 0,00, menunjukkan kesulitan dalam mencapai keseimbangan yang baik antara *Precision* dan *Recall* untuk kedua kelas tersebut. Total *accuracy* dari model SVM ini adalah 62,00%.

2) Kesulitan Pembentukan Model SVM

Model SVM sulit dibentuk dengan optimal karena nilai data gejala, kecuali usia, cenderung bersifat biner (0 dan 1). Hal ini menyulitkan perhitungan akurat dan berkontribusi pada hasil prediksi yang kurang baik. Ketidakseimbangan data dan sifat biner pada nilai gejala menjadi hambatan utama dalam meningkatkan performa model.

5. SARAN

Berikut beberapa saran yang dapat diberikan untuk pengembangan lebih lanjut dari sistem cerdas klasifikasi gejala awal COVID-19 dan Influenza:

- 1) Melakukan upaya untuk mengatasi ketidakseimbangan data dan mempertimbangkan penggunaan fitur yang dapat memiliki variasi nilai, sehingga model SVM dapat membentuk hasil yang lebih akurat.
- 2) Membuat desain *interface* pada sistem agar bisa digunakan baik oleh tenaga medis maupun masyarakat luas.
- 3) Menambahkan varian dari COVID-19.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Tim Redaksi Jurnal Teknik Politeknik Negeri Sriwijaya yang telah memberi kesempatan, sehingga artikel ilmiah ini dapat diterbitkan.

DAFTAR PUSTAKA

- [1] Akbar, M. F. (2019). *Analisis Pasien Self-diagnosis Berdasarkan Internet Pada Fasilitas Kesehatan Tingkat Pertama*. INA-Rxiv. June, 25.
- [2] Librianty, N. (2015). *Panduan Mandiri Melacak Penyakit*. Jakarta Selatan: Lintas Kata.
- [3] Syam, A. F. (2020, Maret 8). Terserang Virus Corona Atau Hanya Sakit Flu? Bedanya ... (tvOne, Interviewer)
- [4] Puspitasari, A. M., Ratnawati, D. E., & Widodo, A. W. (2018). *Klasifikasi Penyakit Gigi Dan Mulut Menggunakan Metode Support Vector Machine*. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(2), 802-810.
- [5] Apriyani, H., & Kurniati. (2020). *Perbandingan Metode Naïve Bayes Dan Support Vector Machine Dalam Klasifikasi Penyakit Diabetes Melitus*. *Journal of Information Technology Ampera*, 1(3), 133-143.
- [6] Biantong, T. R., Furqon, M. T., & Soebroto, A. A. (2019). *Implementasi Metode Support Vector Machine Untuk Klasifikasi Jenis Penyakit Malaria*. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 3(2), 1215-1224.
- [7] Murinto, M., & Sunardi, S. (2023). *Medical External Wound Image Classification Using Support Vector Machine Technique*. *Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika*, 9(2), 98-103.
- [8] Gaffar, A. W. M., Halis, A. M., & Jabir, S. R. (2024). *Penerapan Algoritma Support Vector Machine untuk Klasifikasi Stunting pada Balita di Kabupaten Enrekang*. *Jurnal Minfo Polgan*, 13(1), 286-292.
- [9] Kurniawan, D. (2020). *Pengenalan Machine Learning dengan Python Solusi untuk Permasalahan Big Data*. Jakarta: PT Elex Media Komputindo.
- [10] Supartini, I. A. M., Sukarsa, I. K. G., & Srinadi, I. G. A. M. (2017). *Analisis Diskriminan Pada Klasifikasi Desa Di Kabupaten Tabanan Menggunakan Metode K-fold Cross Validation*. *E-Jurnal Mat*, 6(2), 106.
- [11] S. Andini, "Klasifikasi Dokumen Teks Menggunakan Algoritma Naïve Bayes dengan Bahasa Pemrograman Java," *Tekno. Inf. Pendidik.*, vol. 6, no. 2, pp. 140–147, 2013.
- [12] Sabita, H., & Herwanto, R. (2020). *Pantauan Prediktif COVID-19 Dengan Menggunakan Metode SIR dan Model Statistik Di Indonesia*. *TEKNIKA*, 14(2), 145 – 150.
- [13] Aditya, D. M. (2020). *Anosmia pada COVID-19: Studi Neurobiologi*. *KELUWIH: Jurnal Kesehatan dan Kedokteran*, 50-55.
- [14] WHO. (2020). *Clinical management of COVID-19*. *World Health Organization*.
- [15] Nareza, M. (2020, Juni 12). *Kenali Gejala Orang Terinfeksi Virus Corona di Minggu Pertama*. Retrieved November 13, 2020, from ALODOKTER: <https://www.alodokter.com>
- [16] KPCPEN. (2020). *Peta Sebaran COVID-19*. Retrieved November 13, 2020, from Satgas Penanganan COVID-19: <https://Covid19.go.id>