

IMPLEMENTASI METODE ARIMA DALAM FORECASTING JUMLAH KASUS PENDERITA PENYAKIT HIV/AIDS (STUDI KASUS : KABUPATEN/KOTA DI PROVINSI SUMATERA SELATAN)

Fathoni, Siti Hariza Marshella, Aliyananda Risyahputri,
Naila Raihana Putri, Eka Therina Lakeisyah, Ali Ibrahim

Sistem Informasi, Fakultas Ilmu Komputer, Universitas Sriwijaya
Jalan Palembang - Prabumulih No.KM 32, Indralaya Indah, Kec. Indralaya Kab.Ogan Ilir Sumatera Selatan
09031282227071@student.unsri.ac.id

ABSTRAK

HIV/AIDS senantiasa menjadi persoalan kesehatan masyarakat yang serius di Indonesia, utamanya di Provinsi Sumatera Selatan yang menunjukkan adanya tren peningkatan kasus dalam beberapa tahun terakhir. Pola penyebaran yang sulit diprediksi dengan metode konvensional menjadi tantangan bagi pemerintah dalam melakukan intervensi dini. Penelitian ini berfokus pada prediksi jumlah kasus HIV/AIDS di Provinsi Sumatera Selatan pada 3 tahun kedepan menggunakan metode time series ARIMA (Autoregressive Integrated Moving Average). Data yang digunakan merupakan data tahunan jumlah kasus HIV/AIDS di Provinsi Sumatera Selatan dari tahun 2018 hingga 2023. Setelah melalui proses identifikasi, estimasi, dan diagnostik model, diperoleh bahwa model ARIMA(1,1,1) merupakan model terbaik berdasarkan uji ACF dan PACF. Hasil ramalan menunjukkan adanya fluktuasi dengan lonjakan kasus, khususnya beberapa daerah seperti Lubuk Linggau, sementara Kota Palembang diprediksi tetap menjadi wilayah dengan jumlah kasus tertinggi, meskipun menunjukkan tren penurunan. Evaluasi model menggunakan MAPE dan RMSE menunjukkan bahwa ARIMA mampu memberikan hasil yang cukup baik pada wilayah dengan pola data stabil seperti Muara Enim, Ogan Ilir, dan Banyuasin. Penelitian ini diharapkan dapat menawarkan kontribusi untuk pengambilan kebijakan berbasis data serta meningkatkan kesadaran berbagai pihak terhadap pentingnya pemantauan kasus secara berkala guna menekan laju penyebaran HIV/AIDS di wilayah Sumatera Selatan.

Kata kunci: HIV/AIDS, ARIMA, *Forecasting*, *Time series*, Sumatera Selatan.

1. PENDAHULUAN

Human Immunodeficiency Virus (HIV) adalah virus yang menyerang sel darah putih dan mengakibatkan sistem imun melemah, alhasil penderitanya rentan terhadap berbagai penyakit lainnya[1]. Sebenarnya HIV dan AIDS adalah kedua kondisi yang berlainan. HIV merupakan infeksi pada tubuh yang ditimbulkan oleh virus, sedangkan AIDS merupakan tahap akhir dari infeksi HIV yang terjadi dalam jangka waktu lama [2].

HIV/AIDS senantiasa menjadi persoalan kesehatan masyarakat yang serius di Indonesia, utamanya di Provinsi Sumatera Selatan. Berdasarkan data [3], jumlah kasus HIV/AIDS yang dilaporkan di Sumatera Selatan pada tahun 2023 mencapai 846 kasus. Peningkatan ini terjadi tidak merata di seluruh kabupaten/kota, dengan konsentrasi tertinggi di Kota Palembang dengan 423 kasus, dan terendah pada Kota Pagaralam dengan 8 kasus. Data menunjukkan tren peningkatan kasus HIV/AIDS yang signifikan dalam 3 tahun terakhir di Sumatera Selatan. Pada tahun 2021, tercatat 321 kasus meningkat drastis menjadi 639 kasus pada tahun 2022 dan terus meningkat hingga 846 kasus di tahun 2023 [3]. Peningkatan ini menjadi perhatian serius mengingat dampaknya terhadap kesehatan masyarakat dan beban sistem kesehatan daerah.

Berdasarkan analisis terhadap tren peningkatan tersebut, dinamika kasus HIV/AIDS di Provinsi Sumatera Selatan menunjukkan pola yang sulit

diprediksi dengan metode konvensional, hal ini menjadi tantangan bagi pemerintah dalam melakukan intervensi atau pencegahan dini. Prediksi yang kurang akurat ini diperparah oleh adanya stigma sosial yang menganggap bahwa HIV/AIDS hanya menimpa kelompok tertentu seperti pekerja seks atau suntik narkoba yang akhirnya membuat penderita cenderung menutup diri dan enggan untuk mengakses pengobatan [4]. Untuk mengatasi keterbatasan ini, dibutuhkan sebuah pendekatan inovatif melalui penerapan teknologi yang mampu meminimalisir dampak negatif tersebut. Salah satu cara penerapannya adalah dengan penggunaan *machine learning* yang lebih canggih dan mampu menangkap pola kompleks sehingga dapat mencapai akurasi yang lebih baik serta memberikan wawasan untuk pemerintah kedepannya [5].

Dengan diterapkannya *machine learning* dalam prediksi kasus HIV/AIDS di Sumatera Selatan, diharapkan dapat membantu Dinas Kesehatan dalam mendistribusikan sumber daya dan pengobatan untuk penderita penyakit HIV/AIDS di Sumatera Selatan. Selain itu, hasil prediksi ini dapat membantu pemerintah dalam mempersiapkan langkah mitigasi terhadap potensi lonjakan kasus dan meningkatkan kesadaran masyarakat terhadap penyakit HIV/AIDS khususnya di lingkup kabupaten dan kota.

Berbeda dari penelitian sebelumnya, penelitian ini akan mencakup data kasus HIV/AIDS yang terspesifikasi pada level kabupaten/kota, khususnya di Provinsi Sumatera Selatan. Mengacu pada hal tersebut,

diharapkan dapat memberikan analisis yang lebih terperinci serta solusi yang nyata untuk pengambilan keputusan di masa depan. Selain itu, diferensiasi yang paling menonjol dalam penelitian kali ini mengacu pada metode atau algoritma yang digunakan. Untuk kasus ini, peneliti memilih untuk menerapkan model *machine learning Autoregressive Integrated Moving Average* (ARIMA) yang diyakini dapat melakukan analisis pola data yang bersifat *time-series* dengan lebih akurat. Pernyataan ini diperkuat dengan adanya penelitian oleh [6] yang mengimplementasikan model ARIMA dan memberikan hasil prediksi kasus yang mendekati data sebenarnya. Dengan demikian, diharapkan penelitian dapat memberikan analisis yang lebih terperinci serta solusi yang nyata untuk pengambilan keputusan dalam penanganan HIV/AIDS Sumatera Selatan di masa depan.

2. TINJAUAN PUSTAKA

Prediksi kasus penyakit di Indonesia bukan hal yang baru. Penelitian ini [7] menyelidiki model ARIMA untuk meramalkan kejadian diare di masa depan di kota Kendari. Hasil penelitian menunjukkan berdasarkan uji diagnostik dengan Ljung-Box test, model ARIMA (1,1,1) ditemukan sebagai model terbaik karena memenuhi syarat white noise. Peramalan menggunakan model ARIMA (1,1,1) menunjukkan penurunan kasus diare untuk dua tahun ke depan. Adapun penelitian lain [8] menggunakan algoritma regresi linier dengan mempertimbangkan faktor iklim yang mempengaruhi kejadian DBD di BRSUD Kabupaten Tabanan. Evaluasi akurasi peramalan menggunakan dua metrik, yaitu MAPE dan RMSE. Nilai MAPE sebesar 19% (kategori baik) dan RMSE 0.156 pada data training, sedangkan pada data test didapatkan MAPE sebesar 35% (kategori wajar) dengan RMSE 0.252. Hasil ini mengindikasikan bahwa metode regresi linier mampu diterapkan dalam memprediksi jumlah kasus DBD.

Forecasting juga ditemui pada kasus yang lebih spesifik yakni penyakit HIV. Terdapat pada penelitian [8] yang membahas mengenai prediksi penyakit HIV dengan algoritma Random Forest. Hasil Pemodelan menggunakan algoritma Random Forest mengindikasikan akurasi 92,86%, recall 91,43%, presisi 94,12%, dan F1-Score 86,50%, menandakan efektivitas model dalam menganalisis prediksi HIV. Penelitian [1] tersebut juga membahas mengenai prediksi jumlah kasus baru HIV berdasarkan provinsi melalui pendekatan *machine learning* menggunakan teknik Levenberg-Marquardt. Teknik ini merupakan salah satu teknik dari algoritma ANN. Penelitian ini dianalisis menggunakan tiga model arsitektur jaringan, yaitu 3-15-1, 3-20-1, dan 3-25-1. Model 3-15-1 menunjukkan akurasi tertinggi sebesar 88%. Di lain sisi, penelitian oleh [9] menggunakan metode linear regression dengan performa 0.816 menghasilkan kasus HIV/AIDS yang terjadi di Kabupaten Sidoarjo diperkirakan akan terus meningkat setiap tahunnya

dan seringkali lebih banyak pada laki-laki dibandingkan perempuan

2.1. Machine Learning

Machine Learning adalah algoritma matematika atau program komputer yang dilatih dan diadopsi dari data yang ada untuk membuat sebuah perencanaan maupun prediksi suatu hal di masa depan [11]. Kategori dalam *machine learning* sendiri terbagi menjadi tiga, yakni *supervised learning*, *unsupervised machine learning*, serta *reinforcement learning*. Pada konteks penelitian ini, digunakan pendekatan *supervised machine learning* dengan metode *forecasting* yang bertujuan untuk melakukan *tracking* terhadap pola dari dataset jumlah kasus penderita penyakit HIV/AIDS di Sumatera Selatan untuk mendapatkan hasil prediksi di masa yang akan datang, khususnya tiga tahun setelahnya.

2.2. Forecasting

Forecasting atau analisis prediktif adalah bentuk kalkulasi untuk memprediksi peristiwa di masa depan menggunakan analisis dari basis data historis [10]. *Forecasting* dapat dilakukan untuk memprediksi penyebaran jumlah penyakit terkhusus penyakit HIV dalam beberapa tahun kedepan. Berbagai metode *forecasting* telah dikembangkan, mulai dari pendekatan statistik matematik sampai pemanfaatan teknologi kecerdasan buatan seperti *Machine Learning* [11]. *Forecasting* kerap digunakan sebagai bentuk pengendalian dan perencanaan di masa depan. Seperti dalam penelitian ini, *forecasting* digunakan untuk memprediksi lonjakan baik kenaikan atau penurunan jumlah kasus penderita HIV/AIDS di tiap kabupaten/kota di Provinsi Sumatera Selatan agar dapat merencanakan strategi penanggulangan dan pengendalian yang tepat untuk lonjakan atau perubahan jumlah kasus penderita HIV/AIDS di masing-masing daerah.

2.3. Metode ARIMA

Metode ARIMA (*Autoregressive Integrated Moving Average*) merupakan metode peramalan yang didasarkan pada perilaku atribut yang ditinjau dan fenomena musiman tertentu dalam periode yang relatif sama berdasarkan data historis yang ada [10]. Metode ARIMA ini dipilih dalam perhitungan *forecasting* pada penelitian ini dikarenakan model ARIMA cocok digunakan dalam peramalan *timeseries*. Metode ARIMA memiliki rumus yang terdiri dari 3 gabungan komponen (p,d,q) yakni *autoregressive*, *integrated*, dan *moving average* yang dapat dituliskan secara sistematis seperti berikut [12] :

a. Autoregressive

Komponen *autoregressive* menunjukkan bahwa nilai saat ini yang dimiliki oleh deret waktu dipengaruhi oleh nilai-nilai sebelumnya, artinya nilai prediksi sekarang diperoleh dari nilai sebelumnya. Dinyatakan pada persamaan :

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} \quad (1)$$

Keterangan :

Y_t = nilai pada waktu ke- t , nilai pengamatan pada titik waktu tertentu.

$\phi_1, \phi_2, \dots, \phi_p$ = koefisien autoregresi yang menggambarkan seberapa besar nilai masa lalu mempengaruhi nilai sekarang dan ditentukan melalui estimasi.

p = jumlah periode waktu sebelumnya yang digunakan untuk memprediksi nilai sekarang pada model autoregresi.

b. Integrated

Menunjukkan seberapa banyak data yang dikurangi dengan data itu sendiri di waktu sebelumnya agar menjadi stabil dan memiliki rata-rata serta varians yang tetap. Dituliskan dalam persamaan :

$$\begin{aligned} Y_t^{(1)} &= Y_t - (Y_t - 1) \\ Y_t^{(2)} &= Y_t^{(1)} - (Y_t - 1)^{(1)} \\ Y_t^{(d)} &= Y_t^{(d-1)} - (Y_t - 1)^{(d-1)} \end{aligned} \quad (2)$$

Keterangan :

Y_t = nilai data pada waktu ke- t , yakni nilai atau variabel yang dicatat pada titik waktu tertentu.

d = nilai diferensiasi yakni pengurangan berulang terhadap nilai sebelumnya.

c. Moving Average

Komponen ini menunjukkan nilai saat ini dari data dipengaruhi oleh kesalahan (error) dari prediksi sebelumnya. Jadi, prediksi sekarang mempertimbangkan kesalahan yang terjadi di masa lalu. Dinyatakan dalam persamaan :

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} \quad (3)$$

Keterangan :

Y_t = data yang diamati dalam waktu ke- t

$\phi_1, \phi_2, \dots, \phi_p$ = koefisien model autoregresi yang merepresentasikan pengaruh nilai-nilai sebelumnya terhadap nilai saat ini

p = tingkat urutan dari model AR yang menunjukkan banyaknya nilai waktu sebelumnya yang digunakan untuk memprediksi nilai saat ini

2.4. HIV/AIDS

HIV (*Human Immunodeficiency Virus*) menjadi penyakit yang meningkat beberapa tahun terakhir dan ini menjadi tantangan yang besar untuk memprediksi perkembangan penyakit HIV guna mendeteksi HIV sedari dini dan merencanakan perawatan dan pengobatan yang dibutuhkan [13]. Faktor yang mendorong penyebaran virus ini ditandai dengan aktivitas seksual yang tidak aman hingga pengaplikasian jarum suntik lebih dari satu kali kepada orang yang berbeda. Penanganan yang cermat sangat diperlukan, sebab HIV dapat bertransformasi menjadi AIDS dan rentan menginfeksi tubuh manusia saat terjadi penurunan kondisi [14].

2.5. Provinsi Sumatera Selatan

Provinsi Sumatera Selatan telah dikenal berabad lalu dengan nama “Bumi Sriwijaya”. Secara

administratif, provinsi ini terbagi menjadi 13 kabupaten dan 4 kota dengan ibu kota provinsi adalah Kota Palembang [15].



Gambar 1. Peta Provinsi Sumatera Selatan

3. METODE PENELITIAN

3.1. Metode Pengumpulan Data

Fokus teknik yang diimplementasikan dalam penelitian kali ini adalah teknik kuantitatif dengan varian data sekunder. Kuantitatif dicetuskan sebagai sebuah pendekatan yang memiliki data bersifat angka, persentase, hingga frekuensi dalam rangka pengujian sebuah hipotesis penelitian [16]. Jenis data angka tersebut dapat tergolong menjadi data primer maupun sekunder berdasarkan cara memperolehnya. Data sekunder didapatkan melalui media yang tidak secara langsung bersentuhan dengan objek penelitian seperti jurnal, dataset, laporan, maupun sumber statistik terpercaya [17]. Bentuk varian data sekunder yang diolah adalah data Penderita Penyakit HIV/AIDS yang didapatkan dari dataset publik, yaitu Badan Pusat Statistik (BPS) Provinsi Sumatera Selatan. Pengambilan data dilakukan melalui situs resmi BPS Sumatera Selatan yakni sumsel.bps.go.id dan akan dianalisis secara lebih lanjut untuk didapatkan data yang relevan dengan tujuan penelitian [3].

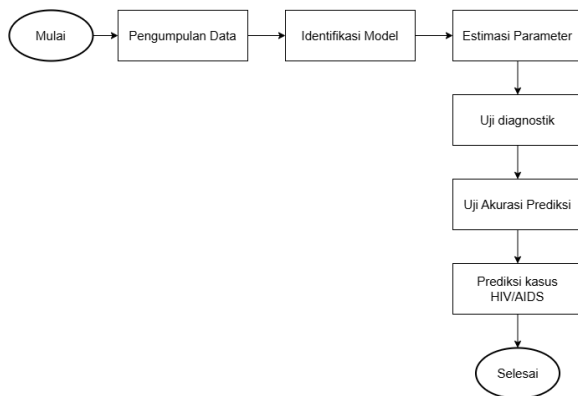
3.2. Metode Penentuan Sampel

Data yang digunakan sebagai sampel penelitian merupakan data penderita penyakit HIV/AIDS di Sumatera Selatan dengan rentang waktu yang cukup panjang yakni dari tahun 2015 hingga tahun 2023. Data ini diperoleh dari Badan Statistik Pusat provinsi Sumatera Selatan dalam bidang kesehatan. Sampel data terdiri dari jumlah kasus penderita penyakit HIV/AIDS yang tercatat setiap tahun di berbagai kabupaten/kota di Sumatera Selatan. Data tersebut termasuk ke dalam bentuk time series (deret waktu) dan akan digunakan untuk menganalisis tren kasus HIV/AIDS pada 3 tahun ke depan.

3.3. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode time series forecasting untuk menganalisis tren kasus HIV/AIDS di Sumatera

Selatan. Teknik yang digunakan dalam penelitian ini adalah ARIMA (AutoRegressive Integrated Moving Average), yang merupakan kombinasi dari algoritma moving average dan model autoregressive untuk memprediksi suatu fenomena berbasis data time-series yang telah ada sebelumnya [18]. Model kombinasi ini diperkirakan memiliki akurasi yang baik sebab memiliki proses yang detail dan dapat menyesuaikan dengan pola data yang tersedia [18]. Adapun tahapan penelitiannya sebagai berikut.



Gambar 2. Tahapan penelitian

- Mengumpulkan data jumlah penderita penyakit HIV/AIDS pada setiap kota/kabupaten di provinsi Sumatera Selatan dengan rentang tahun 2015 hingga 2023.
- Identifikasi model, mengevaluasi konsistensi data apakah terbilang stasioner menggunakan uji ADF (*Augmented Dickey-Fuller*). Apabila hasil menunjukkan tidak stasioner, dilakukan diferensiasi untuk menghilangkan tren. Dalam menentukan model tentatif ARIMA (p, d, q) dapat dilakukan analisis ACF (*Autocorrelation Function*) dan PACF (*Partial Autocorrelation Function*). Parameter p merupakan jumlah lag yang melampaui batas signifikan pada plot PACF, d menunjukkan jumlah diferensiasi yang diterapkan mencapai kestasioneran data, sedangkan q merepresentasikan banyaknya lag yang melampaui ambang signifikan pada grafik ACF [19].
- Estimasi parameter model menggunakan ACF dan PACF.
- Uji akurasi prediksi menggunakan MAPE (*Mean Absolute Percentage Error*) dan RMSE (*Rate Mean Square Error*) untuk melihat model ARIMA yang memiliki nilai error terkecil dan akurasi terbaik.
- Prediksi kasus HIV/AIDS 3 tahun ke depan menggunakan model ARIMA terbaik

3.4. Validasi Model

Dalam melakukan evaluasi akurasi model dengan bahasa pemrograman Python pada penelitian kali ini, digunakan dua buah parameter yakni MAPE (*Mean Absolute Percentage Error*) dan RMSE (*Root Mean Square Error*). Nilai MAPE diyakini sebagai mean atau rerata nilai error mutlak dari hasil prediksi yang disajikan dalam bentuk persentase [20]. MAPE dan RMSE sendiri lumrah digunakan sebagai model standar untuk validasi dalam *forecasting*, dengan membandingkan hasil prediksi dan nilai sebenarnya [21]. Menurut [21] bentuk persamaan yang digunakan untuk kedua pengukuran tersebut adalah sebagai berikut.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Z_t - \hat{Z}_t)^2} \quad (4)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|Z_t - \hat{Z}_t|}{Z_t} \quad (5)$$

Untuk mengetahui bahwa suatu prediksi termasuk ke dalam kategori baik atau buruk dapat dilihat melalui tabel rentang skor MAPE dijelaskan pada Tabel 1 [19].

Tabel 1. Kategori skor MAPE

Hasil	Keterangan
<10%	Hasil Prediksi sangat baik
10-20%	Hasil prediksi baik
20-50%	Hasil prediksi cukup baik
>50%	Hasil prediksi buruk

4. HASIL DAN PEMBAHASAN

Tahapan ini menjabarkan hasil analisis yakni data selection, data pre-processing, data transformation, data mining, dan data evaluation dari percobaan forecasting menerapkan metode ARIMA dengan Python Google Collaboratory untuk mendapatkan kesimpulan terkait penyebaran kasus HIV/AIDS di masing-masing kota di Provinsi Sumatera Selatan guna sebagai saran untuk penerapan kebijakan dalam mengatasi penyebaran HIV/AIDS pada tiap daerah di tiap tahunnya.

4.1. Data Selection

a. Pemilihan Dataset

Dataset terdiri dari 17 baris dan 9 kolom yang dipetakan dalam 5 atribut (Kabupaten/Kota, data HIV/AIDS tahun 2015, data HIV/AIDS tahun 2016, data HIV/AIDS tahun 2017, data HIV/AIDS tahun 2018, data HIV/AIDS tahun 2019, data HIV/AIDS tahun 2020, data HIV/AIDS tahun 2021, data HIV/AIDS tahun 2022, dan data HIV/AIDS tahun 2023). Dataset terdiri dari 153 record data.

Tabel 2. Dataset selection

Kabupaten/Kota	2015	2016	2017	2018	2019	2020	2021	2022	2023
Ogan Komering Ulu	1	1	4	24	3	20	8	16	22
Ogan Komering Ilir	35	25	22	30	29	29	15	43	40
Muara Enim	3	9	4	49	16	18	28	49	46
Lahat	3	4	8	27	8	3	6	5	24

Kabupaten/Kota	2015	2016	2017	2018	2019	2020	2021	2022	2023
Musi Rawas	0	3	15	8	2	7	5	6	14
Musi Banyuasin	5	7	5	31	32	52	16	32	46
Banyu Asin	6	7	30	45	28	24	19	41	38
Ogan Komering Ulu Selatan	0	0	2	4	0	2	1	2	9
Ogan Komering Ulu Timur	1	8	5	29	24	13	22	38	53
Ogan Ilir	15	8	5	12	13	12	12	18	22
Empat Lawang	1	1	4	8	0	4	1	5	9
Pali	2	3	2	6	8	6	11	5	17
Musi Rawas Utara	0	0	1	2	7	0	5	1	9
Kota Palembang	152	187	208	343	231	189	142	353	423
Kota Prabumulih	8	8	10	18	5	12	20	15	35
Kota Pagaram	3	0	1	9	3	6	1	6	8
Kota Lubuk Linggau	9	3	19	9	5	9	9	4	31

b. Pemilihan Fitur

Fitur yang dipilih untuk data testing meliputi data Kabupaten/Kota, dan data time series 2015-2023 jumlah kasus HIV/AIDS di masing-masing daerah di provinsi Sumatera Selatan.

4.2. Data Pre-Processing

Melakukan pembersihan dan pemformatan data sebelum dilakukan uji forecasting dengan menggunakan metode ARIMA.

```
df.set_index("Kabupaten/Kota", inplace=True)
```

Gambar 3. Data pre-processing

Tahap ini dilakukan dengan mengubah kolom Kabupaten/Kota menjadi index utama dari data frame agar dapat mengakses, mengolah, dan

mentransformasi data berdasarkan nama daerah dengan mudah dan rapi.

4.3. Data Transform

Pada tahapan ini dilakukan normalisasi data dan tabel, setelah mengubah kolom kabupaten/kota menjadi index utama selanjutnya adalah mentranspose (mengubah baris dan kolom data) dikarenakan untuk memudahkan forecasting dengan format time series. Kemudian dilakukan juga perubahan tipe data index tahun dari string menjadi integer agar hasil format data lebih stabil dan akurat.

Data kolom sebelum dilakukan transpose masih berupa waktu (time series) dan data barisnya adalah Kabupaten/Kota seperti terlihat pada Tabel 1. Setelah dilakukan dataset transpose menjadi seperti berikut ini.

Tabel 3. Data Setelah Dilakukan Dataset Transpose

Tahun	OKU	OKI	Muara Enim	Lahat	Musi Rawas	Musi Banyuasin	Banyu Asin	OKUS	OKUT
2015	1	35	3	3	0	5	6	0	1
2016	1	25	9	4	3	7	7	0	8
2017	4	22	4	8	15	5	30	2	5
2018	24	30	49	27	8	31	45	4	29
2019	3	29	16	8	2	32	28	0	24
2020	20	29	18	3	7	52	24	2	13
2021	8	15	28	6	5	16	19	1	22
2022	16	43	49	5	6	32	41	2	38
2023	22	40	46	24	14	46	38	9	53

Tabel 4. Lanjutan Data Setelah Dilakukan Dataset Transpose

Tahun	Ogan Ilir	Empat Lawang	Pali	Musi Rawas Utara	Palembang	Prabumulih	Pagaralam	Lubuk Linggau
2015	15	1	2	0	152	8	3	9
2016	8	1	3	0	187	8	0	3
2017	5	4	2	1	208	10	1	19
2018	12	8	6	2	343	18	9	9
2019	13	0	8	7	231	5	3	5
2020	12	4	6	0	189	12	6	9
2021	12	1	11	5	142	20	1	9
2022	18	5	5	1	353	15	6	4
2023	22	9	17	9	423	35	8	31

Data kolom berubah menjadi data Kabupaten/Kota dan data barisnya adalah data tahun (time series) yang lebih ideal untuk melakukan

pengujian forecasting dengan melatih satu per satu data Kabupaten/Kota per tahunnya.

4.4. Data Mining

Pada tahapan ini dilakukan pengujian dataset dengan metode ARIMA untuk melihat forecasting jumlah penderita penyakit HIV/AIDS di Provinsi Sumatera Selatan. Prediksi dilakukan untuk 3 tahun kedepan

```
predictions = {}
actual_values = {}
future_years = [2024, 2025, 2026]
```

Gambar 4. Analisis prediksi

Menggunakan Python dilakukan pemodelan prediksi dengan arima dengan sampel sesuai dengan dataset testing, nilai sebenarnya sesuai dengan dataset testing dan untuk peramalan yang akan dilakukan adalah untuk kasus tahun 2024,2025, dan 2026.

```
for city in df.columns:
    try:
        series = df[city].astype(float)

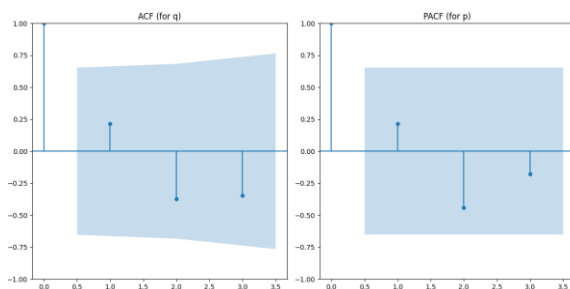
        model = sm.tsa.ARIMA(series, order=(1, 1, 1))
        model_fit = model.fit()

        forecast = model_fit.forecast(steps=3)

        predictions[city] = forecast.values
        actual_values[city] = series[-3:].values # Gu
    except Exception as e:
        print(f"Gagal memproses {city}: {e}")
```

Gambar 5. ARIMA model

Kemudian dilakukan penentuan model tentatif ARIMA (p,d,q) , dalam hal ini ditentukan nilai p,d,q adalah (1,1,1) hal ini merujuk pada test ACF dan PACF yang menunjukkan nilai atau lag yang melewati batas biru adalah lag = 1 maka nilai p dan 1 adalah 1, sehingga penentuan p,d,q adalah (1,1,1)



Gambar 6. ACF dan PACF

Selanjutnya adalah menentukan berapa banyak hasil model prediksi yang akan ditampilkan, dalam hal ini nilai yang dimasukkan adalah steps =3, artinya diperlihatkan hasil model forecasting dalam 3 tahun kedepan, yakni 2024-2026.

Setelah dilakukan penentuan ACF dan PACF, maka dapat merumuskan metode prediksi dan didapatkan hasil prediksi jumlah kasus HIV/AIDS di masing-

masing daerah di Provinsi Sumatera Selatan dengan rincian berikut :

Tabel 5. Hasil Prediksi Tahun 2024-2026

Tahun	2024	2025	2026
Ogan Komering Ulu	16,36	19,44	17,76
Ogan Komering Ilir	30,57	29,94	29,9
Muara Enim	42,31	42,61	42,59
Lahat	11,61	10,1	9,91
Musi Rawas	16,06	14,97	15,55
Musi Banyuasin	39,1	39,62	39,58
Banyu Asin	37,33	37,53	37,47
Ogan Komering Ulu Selatan	6,57	6,05	5,94
Ogan Komering Ulu Timur	52,2	52,8	52,35
Ogan Ilir	21,65	21,76	21,73
Empat Lawang	4,52	4,43	4,43
Pali	6,99	15,7	8,12
Musi Rawas Utara	2,95	8,29	3,58
Kota Palembang	383,21	362,35	351,42
Kota Prabumulih	23,46	28,93	26,34
Kota Pagaralam	4,81	5,19	5,14
Kota Lubuk Linggau	0,37	14,82	8

Dari hasil analisis yang telah dilakukan, hasil prediksi tahun 2024-2026 menunjukkan mayoritas Kabupaten/Kota di Provinsi Sumatera Selatan memiliki tren naik turun terhadap jumlah kasus HIV/AIDS namun tidak terlalu signifikan. Kota Palembang diprediksi menjadi daerah dengan kasus tertinggi setiap tahunnya, namun dapat dilihat terdapat tren penurunan pula pada setiap tahunnya. Sementara itu daerah lainnya rata-rata memiliki pola tren menaik kemudian menurun, seperti salah satu contoh kasus adalah Kota Lubuk Linggau yang diprediksi pada tahun 2024 memiliki sangat minim kasus atau 0 kasus, namun di tahun 2025 diprediksi memiliki 15 kasus namun di tahun 2026 terdapat penurunan kasus menjadi 8 kasus. Pola semacam ini menunjukkan adanya potensi lonjakan mendadak yang mungkin berkaitan dengan faktor lokal seperti peningkatan tes dan pelaporan kasus, perubahan perilaku populasi, atau terbatasnya akses terhadap pelayanan kesehatan preventif.

Perbedaan beberapa lonjakan kenaikan maupun penurunan kasus di masing-masing daerah di Provinsi Sumatera Selatan menjadi sinyal bahwa pentingnya strategi penanggulangan HIV/AIDS yang bersifat terlokalisasi dan kontekstual. Misalnya, perlu dilakukan strategi intervensi di Kota Palembang seperti difokuskan pada upaya sustainabilitas layanan dan edukasi, sementara di daerah seperti Lubuk Linggau, strategi responsif terhadap lonjakan kasus perlu diantisipasi, seperti peningkatan skrining rutin dan edukasi berbasis komunitas.

Prediksi tren ini memberikan gambaran awal yang penting untuk perencanaan dan alokasi sumber daya kesehatan secara lebih tepat sasaran. Pemerintah daerah dan instansi terkait disarankan untuk menggunakan data ini sebagai dasar dalam merancang

kebijakan berbasis bukti yang adaptif terhadap kondisi lokal masing-masing daerah di Sumatera Selatan.

4.5. Data Evaluation

Berdasarkan analisis data forecasting yang telah dilakukan pada tahap sebelumnya, dibutuhkan evaluasi model guna memvalidasi hasil prediksi tersebut, dalam hal ini digunakan evaluasi menggunakan MAE dan RMSE yang merupakan bentuk evaluasi model dalam satuan asli data, didapatkan hasil sebagai berikut :

Tabel 6. Hasil Evaluasi Model Prediksi ARIMA

Kabupaten/Kota	MAPE (%)	RMSE
Ogan Komering Ulu	48,45	5,76
Ogan Komering Ilir	53,13	13,10
Muara Enim	23,86	9,26
Lahat	84,69	9,23
Musi Rawas	127,23	8,26
Musi Banyuasin	60,71	14,52
Banyu Asin	35,43	10,77
Ogan Komering Ulu Selatan	264,41	4,35
Ogan Komering Ulu Timur	59,15	19,42
Ogan Ilir	34,19	5,98
Empat Lawang	138,09	3,34
Pali	100,90	8,35
Musi Rawas Utara	276,88	5,38
Kota Palembang	63,14	145,36
Kota Prabumulih	44,96	9,67
Kota Pagaralam	143,29	2,78
Kota Lubuk Linggau	146,87	15,49
Total	1.705,41	291,092
Mean	100,31	17,12

Dari evaluasi tersebut, dapat diinterpretasikan sebagai berikut :

Tabel 7. Hasil Evaluasi Tergolong Baik

Kabupaten/Kota	MAPE (%)	RMSE
Muara Enim	23,86	9,26
Ogan Ilir	34,19	5,98
Banyuasin	35,43	10,77

Kabupaten Muara Enim memiliki hasil MAPE dan RMSE dengan akurasi tergolong baik dan kesalahan relatif kecil. Sementara itu, Kabupaten Ogan Ilir dan Banyuasin memiliki akurasi yang masih tergolong cukup baik dan masih dalam batas wajar (model masih dapat menangkap tren).

Tabel 8. Hasil Evaluasi Tergolong Kurang Baik

Kabupaten/Kota	MAPE (%)	RMSE
Musi Banyuasin	60,71	14,52
OKU Timur	59,15	19,42
Kota Palembang	63,14	145,36

Hasil evaluasi model Musi Banyuasin dan OKU Timur menunjukkan model tidak terlalu presisi dan memiliki kesalahan yang relatif besar. Selain itu untuk Palembang model sangat kesulitan memprediksi dikarenakan kota Palembang memiliki populasi terbesar dan kasus tertinggi, sehingga model mungkin

kesulitan untuk menangkap fluktuasi yang besar yang mungkin dianggap model sebagai outlier.

Selain evaluasi tersebut, Kabupaten/Kota memiliki kesalahan yang tergolong buruk (MAPE >100%) seperti Musi Rawas menyebabkan model gagal memprediksi dengan baik, hal tersebut mungkin disebabkan karena data sangat fluktuatif, beberapa jumlah kasus sangat rendah sehingga satu kesalahan sangat berdampak pada persentase evaluasi.

5. KESIMPULAN DAN SARAN

Sebagian besar Kabupaten/Kota di Sumatera Selatan berdasarkan hasil prediksi tahun 2024-2026 mengalami tren kasus HIV/AIDS yang fluktuatif, meskipun tidak terlalu signifikan. Kota Palembang diprediksi tetap memiliki jumlah kasus tertinggi namun menunjukkan tren penurunan. Sementara itu, beberapa daerah seperti Lubuk Linggau menunjukkan potensi lonjakan kasus secara tiba-tiba yang dapat disebabkan oleh peningkatan pelaporan, perubahan perilaku masyarakat, atau keterbatasan layanan kesehatan. Dari segi evaluasi, hasil pengujian yang tergolong baik terdapat pada 3 Kabupaten/Kota yakni Muara Enim, Ogan Ilir, dan Banyuasin yang artinya model ARIMA masih dapat menangkap tren yang ada, berbeda dengan Kabupaten/Kota yang memiliki evaluasi yang kurang baik sampai buruk, ARIMA kesulitan untuk menangkap tren yang ada disebabkan terdapat fluktuasi besar, maupun hasil prediksi jumlah kasus yang sangat rendah yang menyebabkan satu kesalahan dapat berdampak pada persentase evaluasi.

Variasi pola ini menunjukkan perlu diadakan evaluasi pengujian model yang mengkombinasikan ARIMA dengan metode lainnya sehingga evaluasi dapat lebih maksimal, kemudian adanya perbedaan hasil prediksi juga menegaskan perlunya strategi penanggulangan HIV/AIDS yang kontekstual dan disesuaikan dengan kondisi lokal. Palembang memerlukan pendekatan berkelanjutan dalam edukasi dan layanan, sedangkan daerah dengan lonjakan perlu strategi responsif seperti peningkatan skrining dan edukasi komunitas. Prediksi ini menjadi dasar penting untuk kebijakan berbasis data dan alokasi sumber daya yang lebih efektif di tiap daerah.

DAFTAR PUSTAKA

- [1] I. Zuhriyillah, F. Anggraini, and R. Dewantara, "Peramalan jumlah kasus baru HIV menurut provinsi menggunakan machine learning dengan teknik Levenberg-Marquardt," *J. Comput. Syst. Informatics*, vol. 3, no. 4, pp. 212–221, 2022.
- [2] L. A. Sari and D. Dasuki, "Respon Sosial Penderita HIV/AIDS di Yayasan Kanti Sehati Sejati Terhadap Dukungan Keluarga," *Jurnal Endurance: Kajian Ilmiah Problema Kesehatan*, vol. 5, no. 2, pp. 284–293, 2020.
- [3] Badan Pusat Statistik Sumatera Selatan, "Jumlah Kasus Penderita Penyakit (Kasus)," Badan Pusat Statistik Sumatera Selatan. Accessed: Feb. 13, 2025. [Online]. Available:

- <https://sumsel.bps.go.id/id/statistics-table/2/Mzc1IzI=/jumlah-kasus-penderita-penyakit-kasus-.html>
- [4] D. P. Asyari, A. Wahyuni, and E. L. Harmen, "Stigma Sosial dan Dampaknya pada Akses Layanan Kesehatan bagi Penderita HIV/AIDS Di Indonesia," *Appicare Journal*, vol. 1, no. 4, pp. 50–56, 2024.
 - [5] I. D. Id, *Machine Learning: Teori, Studi Kasus dan Implementasi Menggunakan Python*, vol. 1. Unri Press, 2021.
 - [6] L. Ainiyah and M. Bansori, "PREDIKSI JUMLAH KASUS COVID-19 MENGGUNAKAN METODE AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)(STUDI KASUS KABUPATEN SIDOARJO) FORECASTING COVID-19 CASES USING AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) METHOD (CASE FOR SIDOARJO REGENCY)," 2021.
 - [7] R. Tosepu and N. Y. Ningsi, "Forecasting of diarrhea disease using ARIMA model in Kendari City, Southeast Sulawesi Province, Indonesia," *Heliyon*, vol. 10, no. 22, 2024.
 - [8] M. W. Aditya, I. N. Sukajaya, and I. G. A. Gunadi, "Forecasting Jumlah Pasien DBD di BRSUD Kabupaten Tabanan Menggunakan Metode Regresi Linier," *Bali Medika Jurnal*, vol. 10, no. 1, pp. 1–12, 2023.
 - [9] A. Roihan, P. A. Sunarya, and A. S. Rafika, "Pemanfaatan machine learning dalam berbagai bidang," *Jurnal Khatulistiwa Informatika*, vol. 5, no. 1, p. 490845, 2020.
 - [10] T. D. Rimadhani, A. A. Arifiyanti, and R. Hadiwiyan, "Peramalan Jumlah Penderita Jenis Penyakit Utama Di Kota Surabaya Menggunakan Metode Autoregressive Integrated Moving Average (ARIMA)," *Student Research Journal*, vol. 1, no. 5, pp. 32–43, 2023.
 - [11] R. G. Wardhana, G. Wang, and F. Sibuea, "Penerapan Machine Learning Dalam Prediksi Tingkat Kasus Penyakit Di Indonesia," *Journal of Information System Management (JOISM)*, vol. 5, no. 1, pp. 40–45, 2023.
 - [12] M. Z. Rusminto, S. A. Wibowo, and F. S. Wahyuni, "Peramalan Harga Saham Menggunakan Metode Arima (Autoregressive Integrated Moving Average) Time Series," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 2, pp. 1263–1270, 2024.
 - [13] M. N. Fatorohman, K. Indriani, and M. N. Winnarto, "Analisa Prediksi Penyakit Hiv Menggunakan Random Forest," *Jurnal Infortech*, vol. 6, no. 2, pp. 150–155, 2024.
 - [14] S. K. M. Kristiani, S. Arifin, and S. K. M. Adi Nugroho, *Dukungan dan harapan: Kunci pencegahan HIV-AIDS*. Uwais Inspirasi Indonesia, 2025.
 - [15] Dinas Komunikasi dan Informatika Provinsi Sumatera Selatan, "Sekilas Sumatera Selatan," Pemerintah Provinsi Sumatera Selatan. Accessed: Apr. 10, 2025. [Online]. Available: <https://sumselprov.go.id/page/content/sekilas-sumatera-selatan>
 - [16] M. I. Syahroni, "Prosedur penelitian kuantitatif," *EJurnal Al Musthafa*, vol. 2, no. 3, pp. 43–56, 2022.
 - [17] A. R. Fadilla and P. A. Wulandari, "Literature review analisis data kualitatif: tahap pengumpulan data," *Mitita Jurnal Penelitian*, vol. 1, no. 3, pp. 34–46, 2023.
 - [18] A. K. Rachmawati, "Peramalan Penyebaran Jumlah Kasus Covid19 Provinsi Jawa Tengah dengan Metode ARIMA," *Zeta-Math Journal*, vol. 6, no. 1, pp. 11–16, 2021.
 - [19] I. Mardiyah, W. D. Utami, D. C. R. Novitasari, M. Hafiyusholeh, and D. Sulistiyawati, "Analisis prediksi jumlah penduduk di kota Pasuruan menggunakan metode arima," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 15, no. 3, pp. 525–534, 2021.
 - [20] A. Sulaiman and A. Juarna, "Peramalan Tingkat Pengangguran Di Indonesia Menggunakan Metode Time Series Dengan Model Arima Dan Holt-Winters," *Jurnal Ilmiah Informatika Komputer*, vol. 26, no. 1, pp. 13–28, 2021.
 - [21] A. Swaraj, K. Verma, A. Kaur, G. Singh, A. Kumar, and L. M. de Sales, "Implementation of stacking based ARIMA model for prediction of Covid-19 cases in India," *J Biomed Inform*, vol. 121, p. 103887, 2021.