

Model Hibrida ARIMA-Neural Network untuk Peramalan Kasus Tuberkulosis

Arriza Agung Setyabudi, Etik Zukhronah*, Isnandar Slamet

Fakultas Matematika dan Ilmu Pengetahuan Alam, Program Studi Statistika, Universitas Sebelas Maret, Surakarta, Indonesia

Email: ¹arrizaas03@student.uns.ac.id, ^{2,*}etikzukhronah@staff.uns.ac.id, ³isnandarlamet@staff.uns.ac.id

Email Penulis Korespondensi: etikzukhronah@staff.uns.ac.id

Abstrak—Tuberkulosis (TBC) masih menjadi tantangan kesehatan masyarakat yang signifikan di Kota Surakarta, sehingga memerlukan metode peramalan yang akurat untuk mendukung strategi pengendalian yang efektif dan terencana. Penelitian ini bertujuan mengevaluasi kinerja model hibrida *Autoregressive Integrated Moving Average-Neural Network* (ARIMA-NN) dalam upaya meramalkan jumlah kasus TBC bulanan di wilayah Surakarta. Kinerja model hibrida ini selanjutnya dibandingkan dengan model ARIMA. Data penelitian yang digunakan adalah data kasus TBC bulanan dari Januari 2019 hingga September 2024 yang diperoleh dari Dinas Kesehatan Kota Surakarta. Data dibagi menjadi dua yaitu data *training* periode Januari 2019 sampai dengan Desember 2023 dan data *testing* periode Januari 2024 sampai dengan September 2024. Model ARIMA(0,1,1) diidentifikasi sebagai model terbaik untuk menangkap komponen linear data, yang menghasilkan *Mean Absolute Percentage Error* (MAPE) sebesar 15,49 % pada data *training* dan 10,53 % pada data *testing*. Residu dari model ARIMA(0,1,1) ini kemudian dimodelkan lebih lanjut menggunakan *Neural Network* dengan arsitektur 5 neuron *hidden*, periode *lookback* 6, dan *learning rate* 0.1, untuk menangkap pola non-linear yang tersisa. Model hibrida ARIMA(0,1,1)-NN yang dikembangkan menunjukkan kinerja peramalan yang lebih baik, dengan nilai MAPE sebesar 14,34% pada data *training* dan 14,48% pada data *testing*. Hasil ini mengindikasikan bahwa pendekatan hibrida ARIMA-NN menawarkan potensi peningkatan akurasi dibandingkan model ARIMA dalam konteks peramalan kasus TBC di Surakarta.

Kata Kunci: ARIMA; Neural Network; Model Hibrida; Tuberkulosis

Abstract—Tuberculosis (TB) remains a significant public health challenge in Surakarta City, necessitating accurate forecasting methods to support effective and planned control strategies. This study aims to evaluate the performance of the Autoregressive Integrated Moving Average-Neural Network (ARIMA-NN) hybrid model in forecasting monthly TB cases in the Surakarta region. The performance of this hybrid model is further compared with the ARIMA model. The research data used consists of monthly TB case data from January 2019 to September 2024 obtained from the Surakarta City Health Department. The data is divided into two sets: training data from January 2019 to December 2023 and testing data from January 2024 to September 2024. The ARIMA(0,1,1) model was identified as the best model for capturing the linear component of the data, yielding a Mean Absolute Percentage Error (MAPE) of 14.52% on the training data and 16.55% on the testing data. The residuals from the ARIMA(0,1,1) model were then further modeled using a Neural Network with 5 hidden neuron architecture, period lookback 6, and a learning rate of 0.1, to capture the remaining non-linear patterns. The developed ARIMA(0,1,1)-NN hybrid model showed better forecasting performance, with a MAPE value of 14.34% on the training data and 14.48% on the testing data. These results indicate that the ARIMA-NN hybrid approach offers the potential for improved accuracy compared to the ARIMA model in the context of TB case forecasting in Surakarta.

Keywords: ARIMA; Neural Network; Hybrid Model; Tuberculosis

1. PENDAHULUAN

Penyakit Tuberkulosis (TBC), yang disebabkan oleh infeksi bakteri *Mycobacterium tuberculosis*, tetap menjadi salah satu masalah kesehatan masyarakat paling mendesak di tingkat global [1]. Indonesia, sebagai negara dengan populasi besar, termasuk dalam sepuluh negara dengan beban TBC tertinggi di dunia. Kondisi ini juga tercermin di tingkat regional, di mana beberapa daerah seperti Kota Surakarta menghadapi tingginya angka kasus TBC. Data dari Kementerian Kesehatan mencatat sebanyak 2.738 kasus TBC di Surakarta pada tahun 2023, yang menunjukkan urgensi penanganan dan pengendalian penyakit ini. Berbagai faktor kompleks seperti kepadatan penduduk yang tinggi, laju urbanisasi, tingkat kemiskinan, serta akses yang terbatas terhadap layanan kesehatan berkualitas memperburuk situasi penyebaran TBC di Surakarta. Mengingat tingginya tingkat penularan TBC, terutama melalui udara dari percikan dahak penderita, maka upaya penanganan yang cepat dan hati-hati menjadi sangat krusial. Oleh karena itu, ketersediaan data dan informasi yang akurat, serta kemampuan untuk memperamalan tren kasus di masa depan, menjadi fondasi penting dalam merancang strategi intervensi yang efektif dan efisien. Peramalan jumlah kasus TBC dapat memberikan panduan bagi otoritas kesehatan dalam mengalokasikan sumber daya, merencanakan program pencegahan, dan membuat keputusan berbasis bukti untuk menekan laju penyebaran penyakit [2].

Dalam konteks peramalan data runtun waktu, seperti data kasus penyakit, berbagai metode statistik dan komputasi telah dikembangkan. Salah satu metode yang populer adalah *Autoregressive Integrated Moving Average* (ARIMA) [3]. ARIMA dikenal karena kemampuannya dalam mengenali pola historis, dependensi linear antar observasi, dan memberikan peramalan yang cukup akurat untuk jangka pendek [4]. Model ini bekerja dengan mengidentifikasi komponen *autoregressive* (AR) yang menunjukkan ketergantungan nilai saat ini pada nilai-nilai sebelumnya, komponen *integrated* (I) yang melibatkan proses *differencing* untuk mencapai stasioneritas data, dan komponen *moving average* (MA) yang memperhitungkan kesalahan peramalan masa lalu. Meskipun demikian, ARIMA memiliki keterbatasan fundamental, yaitu asumsi linearitas hubungan dalam data dan lebih cocok untuk data yang bersifat stasioner atau dapat distasionerkan [5]. Data epidemiologi seringkali menunjukkan pola yang kompleks, non-linear, dan dipengaruhi oleh berbagai faktor yang sulit ditangkap oleh model linear murni.

Untuk mengatasi keterbatasan model linear seperti ARIMA, metode *Neural Networks* (NN) atau Jaringan Saraf Tiruan (JST) menawarkan pendekatan alternatif [6]. NN adalah model komputasi yang terinspirasi dari cara kerja otak manusia, yang mampu mengenali pola-pola kompleks, non-linear, dan non-stasioner dalam data. Keunggulan NN terletak pada kemampuan generalisasinya yang kuat, menjadikannya efektif untuk berbagai jenis peramalan, termasuk penyebaran penyakit. Namun, NN juga memiliki tantangan tersendiri, seperti kebutuhan akan dataset yang relatif besar untuk pelatihan, kompleksitas dalam penentuan arsitektur dan parameter model, serta sering dianggap sebagai model "kotak hitam" karena hasilnya yang terkadang sulit diinterpretasikan secara langsung dibandingkan model statistik tradisional [7].

Mengingat kelebihan dan kekurangan dari masing-masing metode tersebut, pendekatan hibrida yang mengkombinasikan ARIMA dan NN yang disebut dengan hibrida ARIMA-NN telah banyak diusulkan dan dikembangkan. Konsep dasar model hibrida ARIMA-NN adalah memanfaatkan ARIMA untuk menangkap komponen linear dari data runtun waktu, sementara residu yang dihasilkan oleh model ARIMA yang diasumsikan mengandung pola non-linear yang belum terjelaskan kemudian dimodelkan lebih lanjut menggunakan NN [8]. Dengan cara ini, karakteristik linear dan non-linear dalam data diharapkan dapat diakomodasi secara lebih komprehensif, sehingga menghasilkan peramalan yang lebih akurat dan andal. Beberapa penelitian terdahulu telah menunjukkan keberhasilan penerapan model hibrida ARIMA-NN dalam berbagai bidang. Sebagai contoh, menerapkan tiga pendekatan model, yaitu ARIMA, NN, dan model Hibrida ARIMA-NN untuk meramalkan kasus Tuberkulosis di Kenya, dan menemukan bahwa model Hibrida ARIMA-NN memiliki performa yang lebih unggul dibandingkan model ARIMA dan NN secara terpisah [9], terutama dalam menangkap pola musiman dan tren jangka panjang. Demikian pula, dalam peramalan harga beras menggunakan model hibrida ARIMA-NN mendapatkan hasil akurasi yang lebih baik dibandingkan model ARIMA saja [10]. Kontribusi dari penelitian ini adalah penerapan dan evaluasi model hibrida ARIMA-NN secara spesifik pada data kasus TBC di Kota Surakarta, dengan harapan dapat memberikan metode peramalan yang lebih akurat untuk mendukung upaya pengendalian TBC yang lebih efektif dan berbasis data di wilayah tersebut. Penelitian ini juga bertujuan untuk membandingkan secara langsung kinerja model hibrida dengan model ARIMA berdasarkan nilai *Mean Absolute Percentage Error* (MAPE). Urgensi penelitian ini terletak pada kebutuhan akan alat bantu pengambilan keputusan yang lebih baik bagi Dinas Kesehatan Kota Surakarta dalam menghadapi tantangan TBC yang berkelanjutan, sekaligus meningkatkan kualitas kebijakan kesehatan melalui integrasi teknologi dan analisis data yang inovatif, yang dapat mendorong efektivitas intervensi kesehatan masyarakat secara menyeluruh dan berkelanjutan.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Dinas Kesehatan Kota Surakarta. Data tersebut mencakup informasi jumlah penderita penyakit TBC di Surakarta yang telah dikumpulkan dan tercatat oleh Dinas Kesehatan Kota Surakarta dalam waktu Januari 2019 – September 2024. Data tersebut dibagi menjadi dua yaitu data *training* dan data *testing* dengan data *training* terdiri dari 60 data, serta data *testing* terdiri dari 9 data. Berikut adalah langkah-langkah penelitian untuk mencapai tujuan penelitian ini.

- a. Melakukan pemodelan ARIMA melalui tahapan berikut.
 1. Melakukan plot runtun waktu dari data *training*.
 2. Menguji stasioneritas data dalam variansi dan rata-rata. Jika data belum stasioner dalam variansi maka dilakukan transformasi *Box-Cox* dan jika data belum stasioner dalam rata-rata maka dilakukan *differencing*.
 3. Menentukan orde p dan q berdasarkan plot ACF dan PACF dari data yang sudah stasioner.
 4. Melakukan estimasi dan uji signifikansi parameter model ARIMA dengan nilai orde p dan q yang sudah diidentifikasi.
 5. Melakukan pemeriksaan diagnostik model ARIMA meliputi uji normalitas menggunakan uji Kolmogorov-Smirnov dan uji *white nois* menggunakan uji Ljung-Box
 6. Memilih model ARIMA yang memenuhi seluruh uji diagnostik.
 7. Melakukan peramalan dengan model ARIMA yang terbentuk.
 8. Menghitung nilai MAPE dari data *training* dan data *testing*.
- b. Pemodelan Neural Network (NN)
 1. Melakukan normalisasi data residu model ARIMA dengan menggunakan metode *Min-max scaling*.
 2. Membangun model *Neural Network* dengan menentukan jumlah lapisan, jumlah unit, fungsi aktivasi, ukuran *batch*, jumlah *epoch*, *optimizer*, dan *learning rate*.
 3. Melakukan hyperparameter tuning, yaitu mengumpulkan kombinasi hyperparameter terbaik untuk mendapatkan model Neural Network (NN) yang paling optimal
 4. Menghitung nilai MAPE dari data *training* dan data *testing*.
- c. Pemodelan Hibrida ARIMA
 1. Menggunakan residu dari model ARIMA terpilih sebagai input untuk model Neural Network. Menjumlahkan hasil peramalan ARIMA pada Langkah 1.h dengan hasil peramalan NN pada Langkah 2.d
 2. Menghitung nilai MAPE dari data *training* dan data *testing*.

2.2 Landasan Teori

2.2.1 Autoregressive Integrated Moving Average (ARIMA)

ARIMA adalah salah satu metode statistik untuk melakukan analisis peramalan pada data runtun waktu. Model ini terdiri dari 3 komponen yaitu *autoregressive* (AR), *differencing* (I), dan *moving average* (MA) [11]. Ketiga komponen tersebut digabungkan menjadi ARIMA(p,d,q) dimana p adalah orde dari komponen *autoregressive*, d adalah orde *differencing* yang diperlukan untuk membuat data stasioner, dan q adalah orde dari komponen *moving average* [12]. Model ARIMA(p,d,q) dapat dinyatakan secara umum sebagai:

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)a_t \quad (1)$$

dengan

$\phi_p(B)$ adalah parameter model AR(p),

$\theta_q(B)$ adalah parameter model MA(q),

B adalah operator *backshift*,

d adalah parameter *differencing*,

Z_t adalah nilai aktual pada periode t,

a_t adalah error pada periode t dengan asumsi *white noise* dan berdistribusi normal dengan rata-rata 0 dan variansi σ^2 .

2.2.2 Neural Network

Jaringan Syaraf Tiruan (JST) merupakan metode komputasi yang terinspirasi dari sistem jaringan syaraf biologis, khususnya otak manusia, dalam memproses informasi [13]. Secara umum, JST terdiri atas tiga lapisan utama, yaitu lapisan input, lapisan tersembunyi (hidden layer), dan lapisan output. Struktur setiap lapisan beserta hubungan antar lapisan dalam jaringan ini disebut sebagai arsitektur JST. Salah satu arsitektur yang paling sering digunakan adalah *multi-layer network*, karena umumnya memanfaatkan algoritma *backpropagation* dalam proses pembelajarannya.

Algoritma *backpropagation* sendiri berperan penting dalam meminimalkan kesalahan output jaringan dengan cara melakukan penyesuaian terhadap bobot berdasarkan selisih antara output aktual dan output yang diharapkan [14]. Dalam proses ini, fungsi aktivasi juga memegang peranan krusial. fungsi aktivasi digunakan sebagai dasar dalam proses pembaruan bobot, sehingga dapat memperkuat kinerja jaringan. Fungsi aktivasi berperan dalam menentukan keluaran dari setiap neuron dan menjadi dasar dalam proses pembaruan bobot selama proses pelatihan. Adapun fungsi aktivasi yang sering diterapkan dalam pelatihan jaringan adalah fungsi sigmoid biner dan fungsi linier.

2.2.3 Hibrida ARIMA-NN

Pada Model hibrida ARIMA-NN, yang diperkenalkan oleh Zhang (2003), merupakan pendekatan gabungan yang mengintegrasikan metode ARIMA dan NN untuk mengatasi tantangan data runtun waktu yang sering memiliki karakteristik linier dan nonlinier [15]. Dalam pendekatan ini, model ARIMA digunakan untuk menangani komponen linier dari data, sementara residu yang dihasilkan, yang masih memuat informasi nonlinier, dimodelkan lebih lanjut menggunakan NN [16]. Dengan menggabungkan kedua pendekatan ini, karakteristik linier dan nonlinier dalam data runtun waktu dapat diakomodasi secara lebih efektif [17]. Zhang (2003) menjelaskan konsep kombinasi kedua peramalan tersebut dapat dihitung dengan rumus

$$\widehat{H}_t = \widehat{Z}_t + \widehat{N}_t \quad (2)$$

dengan

$\widehat{H}_t$ adalah hasil peramalan hibrida ARIMA-NN,

$\widehat{Z}_t$ adalah komponen ARIMA

$\widehat{N}_t$ adalah komponen NN

Dari residual ARIMA yang sebelumnya telah diperoleh, tahap selanjutnya adalah membangun model Neural Network (NN) untuk memodelkan pola nonlinier yang masih terkandung dalam data. Residual tersebut digunakan sebagai data latih pada model NN, dengan harapan jaringan saraf dapat menangkap hubungan nonlinier yang tidak terdeteksi oleh model ARIMA. Setelah model NN selesai dibangun dan menghasilkan peramalan, hasil akhir dari model hibrida diperoleh dengan menggabungkan peramalan dari ARIMA dan NN. Pendekatan ini memungkinkan karakteristik linier dan nonlinier dalam data runtun waktu dapat ditangani secara lebih efektif dan menyeluruh.

2.2.4 Evaluasi Model

Dalam evaluasi model, penelitian ini menggunakan Mean Absolute Percentage Error (MAPE) sebagai metrik utama untuk menilai akurasi peramalan. MAPE dipilih karena kemampuannya dalam mengukur kesalahan peramalan secara relatif terhadap nilai aktual, sehingga memberikan gambaran yang jelas mengenai performa setiap model dalam meramalkan data runtun waktu [18]. Dengan menggunakan MAPE, penelitian ini dapat menentukan model terbaik berdasarkan nilai kesalahan yang paling kecil, yang mencerminkan ketepatan dan keandalan hasil peramalan Lewis, C. D. (1982). Nilai MAPE dihitung menggunakan Persamaan (3).

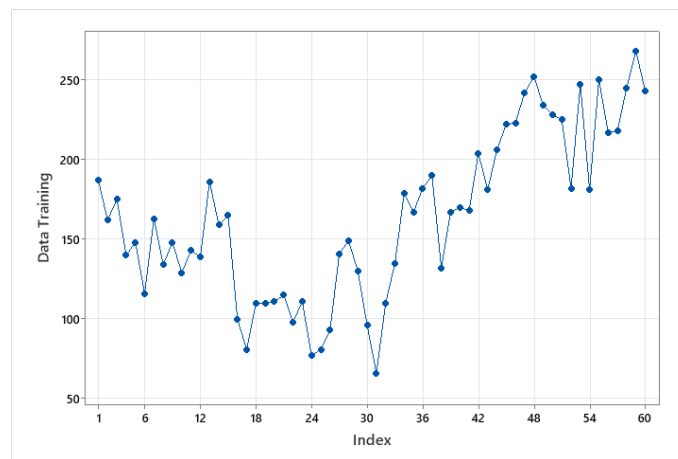
$$MAPE = \left(\frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \right) \times 100\% \quad (3)$$

Tabel 1. Kategori Nilai MAPE

Nilai MAPE	Kategori
$MAPE < 10\%$	Sangat baik
$10\% < MAPE \leq 20\%$	Baik
$20\% < MAPE \leq 50\%$	Cukup baik
$MAPE > 50\%$	Buruk

3. HASIL DAN PEMBAHASAN

Data yang digunakan penelitian ini merupakan data bulanan kasus TBC di Kota Surakarta. Data ini ditampilkan pada Gambar 1.

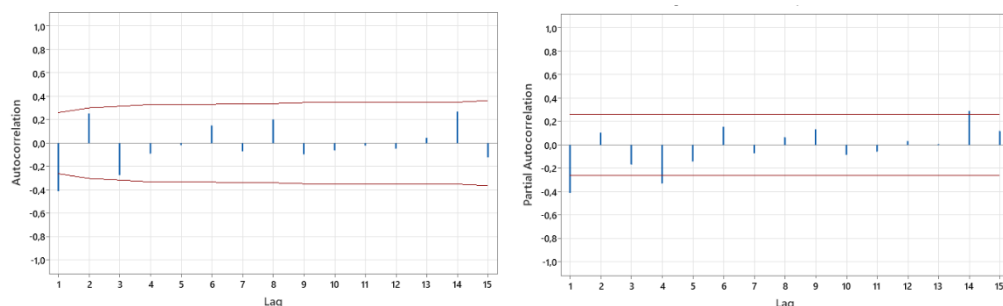


Gambar 1. Plot Runtun Data Kasus Penderita TBC di Surakarta

Gambar 1 menyajikan plot runtun waktu data *training* jumlah kasus penderita TBC di Kota Surakarta selama periode Januari 2019 hingga Desember 2023.. Berdasarkan pola yang terlihat pada Gambar 1, jumlah kasus TBC menunjukkan kecenderungan meningkat seiring waktu. Dengan demikian, dapat disimpulkan bahwa data memiliki pola tren naik yang cukup jelas dan konsisten sepanjang periode observasi.

3.1 Peramalan dengan Metode ARIMA

Pemodelan ARIMA mensyaratkan bahwa data yang digunakan harus stasioner baik terhadap variansi maupun rata-rata. Untuk menguji kestasioneran terhadap variansi, dilakukan transformasi Box-Cox pada data *training*. Hasil transformasi menunjukkan bahwa nilai parameter lambda (λ) sebesar 1, yang mengindikasikan bahwa data *training* telah stasioner terhadap variansi, sehingga tidak diperlukan transformasi tambahan. Selanjutnya, kestasioneran terhadap rata-rata diuji menggunakan plot Autocorrelation Function (ACF). Plot ACF awal menunjukkan pola autokorelasi yang menurun secara eksponensial, menandakan bahwa data belum stasioner terhadap rata-rata. Oleh karena itu, dilakukan proses *differencing* pada data *training*. Setelah dilakukan *differencing*, plot ACF terdapat pada Gambar 2a, sedangkan plot PACF terdapat pada Gambar 2b.



Gambar 2. (a). Plot ACF dan (b). Plot PACF

Gambar 2 menunjukkan plot ACF setelah dilakukan *differencing* satu kali, di mana lag ke-1 berada di luar batas kritis, yang mengindikasikan bahwa data telah mencapai kestasioneran. Kondisi ini menunjukkan bahwa orde q dapat dipertimbangkan pada nilai 0 atau 1. Sementara itu, pada plot PACF, lag ke-1, ke-4, dan ke-14 tampak keluar dari batas

kritis, sehingga orde p dapat dipertimbangkan pada rentang 0 hingga 3. Berdasarkan orde p , d dan q diperoleh kemungkinan sementara adalah ARIMA yang mungkin digunakan secara sementara antara lain adalah ARIMA(0,1,0), ARIMA(0,1,1), ARIMA(1,1,0), ARIMA(1,1,1), ARIMA(2,1,0), ARIMA(2,1,1), ARIMA(3,1,0), ARIMA(3,1,1), ARIMA(4,1,0), dan ARIMA(4,1,1). Model-model ini selanjutnya akan dievaluasi untuk menentukan model terbaik berdasarkan kriteria statistik dan diagnostik residu.

Berdasarkan hasil identifikasi awal, kandidat model ARIMA yang diuji meliputi ARIMA(0,1,0), ARIMA(0,1,1), ARIMA(1,1,0), ARIMA(1,1,1), ARIMA(2,1,0), ARIMA(2,1,1), ARIMA(3,1,0), ARIMA(3,1,1), ARIMA(4,1,0), dan ARIMA(4,1,1). Evaluasi terhadap masing-masing model dilakukan melalui uji signifikansi parameter, dengan memperhatikan nilai p -value dari setiap koefisien, serta uji diagnostik terhadap residual menggunakan uji Ljung-Box untuk menilai apakah residual bersifat white noise. Hasil evaluasi menunjukkan bahwa terdapat model yang memenuhi kedua kriteria tersebut, yaitu parameter yang signifikan dan residual yang bersifat white noise.

Setelah melalui proses seleksi dan pengujian yang mencakup uji signifikansi serta diagnostik residu, diperoleh bahwa model ARIMA(0,1,1) merupakan model terbaik. Model ini memenuhi asumsi-asumsi penting dalam analisis runtun waktu, yaitu kestasioneran, signifikansi parameter, residu yang bersifat white noise, dan normalitas residual. Oleh karena itu, model ARIMA(0,1,1) dipilih sebagai model peramalan jumlah penderita kasus TBC di Kota Surakarta, dengan harapan dapat memberikan hasil peramalan yang akurat dan dapat diandalkan. Model ini memiliki parameter $p = 0$, $d = 1$, dan $q = 1$, dengan bentuk persamaan (4).

$$Z_t = Z_{t-1} + (0.421)a_{t-1} \quad (4)$$

Model ARIMA(0,1,1) dalam meramalkan jumlah penderita kasus TBC di Kota Surakarta dievaluasi menggunakan nilai kesalahan *Mean Absolute Percentage Error* (MAPE). Berdasarkan perhitungan menggunakan Persamaan (3), nilai MAPE yang diperoleh disajikan pada Tabel 2.

Tabel 2. Nilai MAPE ARIMA

Nilai MAPE	
Training	15,49 %
Testing	10,53 %

Tabel 2 menunjukkan bahwa nilai MAPE dari model ARIMA(0,1,1) berada di bawah 20%, sehingga model ini dapat dikategorikan memiliki tingkat akurasi yang baik dalam melakukan peramalan. Dengan nilai MAPE sebesar 15,49% untuk data training dan 10,53% untuk data testing, model ARIMA(0,1,1) dapat dinilai andal dalam memperkirakan jumlah kasus TBC di Kota Surakarta pada periode yang dianalisis.

3.2 Peramalan dengan Metode Hibrida ARIMA-NN

Residu data training yang diperoleh dari model ARIMA(0,1,1) digunakan sebagai input pada pemodelan jaringan saraf tiruan (*Neural Network*). Residu ini merepresentasikan komponen non-linier dari data yang belum sepenuhnya dijelaskan oleh model ARIMA. Oleh karena itu, pemodelan *Neural Network* digunakan untuk menangkap pola-pola kompleks dan non-linier dalam data tersebut sebagai bagian dari pendekatan hibrida ARIMA-NN. Sebelum digunakan dalam proses pelatihan, residu tersebut dinormalisasi menggunakan *metode Min-Max Scaling* agar nilainya berada dalam rentang 0 hingga 1, sesuai dengan kisaran fungsi aktivasi pada jaringan saraf tiruan. Normalisasi ini bertujuan untuk membuat proses pelatihan lebih stabil dan efisien serta mempercepat konvergensi model.

Setelah normalisasi, data residu siap digunakan sebagai input untuk melatih model *Neural Network*. Model ini bertugas memodelkan pola non-linier yang ada sehingga dapat meningkatkan akurasi peramalan apabila digabungkan dengan hasil peramalan model ARIMA. Untuk mendapatkan model *Neural Network* dengan nilai kesalahan minimum, dilakukan proses *hyperparameter tuning* guna mengoptimasi parameter-parameter jaringan. Residu yang telah diperoleh dari model ARIMA digunakan sebagai dasar dalam pemodelan *Neural Network* untuk menentukan komponen non-linier N_t dalam persamaan model hibrida ARIMA-NN.

Daftar parameter dan nilai yang digunakan dalam proses tuning model *Neural Network*, termasuk pengaturan *learning rate*, *lookback*, jumlah neuron pada *hidden layer*, fungsi aktivasi, serta algoritma optimasi, disajikan pada Tabel 3.

Tabel 3. Parameter Hibrida ARIMA-NN

Parameter	Jumlah
<i>Learning Rate</i>	0.1, 0.01, 0.001
<i>Lookback</i>	1, 2, 3, 4, 5, 6
<i>Batch Size</i>	25
<i>Neuron Hidden Layer</i>	1, 2, 3, 4, 5
<i>Optimizer</i>	Adam
<i>Fungsi Aktivasi</i>	ReLU
<i>Epoch</i>	100

Penentuan parameter pada model Neural Network tidak memiliki aturan baku, melainkan ditentukan berdasarkan hasil percobaan pada data yang digunakan. Proses *hyperparameter tuning* dilakukan dengan mencoba berbagai kombinasi parameter untuk menemukan model terbaik. Pemilihan model terbaik didasarkan pada nilai *Mean Squared Error (MSE)* terkecil yang diperoleh selama proses pelatihan. Dalam penelitian ini, metode *GridSearchCV* digunakan untuk melakukan pencarian parameter secara sistematis. Dari hasil *GridSearchCV*, diperoleh model Neural Network terbaik dengan parameter-parameter yang dirangkum pada Tabel 4.

Tabel 4. Parameter Hibrida ARIMA-NN Terbaik

Parameter	Jumlah
<i>Learning Rate</i>	0.1
<i>Lookback</i>	6
<i>Batch Size</i>	30
<i>Neuron Hidden Layer</i>	5
<i>Optimizer</i>	Adam
<i>Fungsi Aktivasi</i>	ReLU
<i>Epoch</i>	10

Table 4 menunjukkan parameter terbaik dari model pada data training, yang kemudian digunakan dalam proses peramalan dengan model Neural Network. Proses pelatihan dilakukan sebanyak 10 epoch, menggunakan konfigurasi model terbaik berdasarkan kombinasi parameter yang menghasilkan nilai *Mean Squared Error (MSE)* terkecil. Berdasarkan hasil tersebut, diperoleh model terbaik hibrida ARIMA-NN dengan konfigurasi ARIMA(0,1,1) dan Neural Network menggunakan parameter *lookback* = 6, jumlah neuron tersembunyi = 5, serta *learning rate* = 0.1. Model ini menunjukkan performa yang baik dalam meramalkan jumlah kasus TBC di Kota Surakarta pada periode yang dianalisis.

Model hibrida ARIMA-NN(0,1,1)(6,5,0.1) yang digunakan untuk meramalkan jumlah penderita kasus TBC di Kota Surakarta dievaluasi menggunakan nilai *Mean Absolute Percentage Error (MAPE)*. Berdasarkan hasil perhitungan menggunakan Persamaan (3), nilai MAPE yang diperoleh disajikan pada Tabel 5.

Tabel 5. Nilai MAPE Hibrida ARIMA-NN

Nilai MAPE	
Training	4,14 %
Testing	6,43 %

Tabel 5 menunjukkan bahwa nilai MAPE yang dihasilkan oleh model hibrida ARIMA-NN adalah sebesar 4,14 % untuk data training dan 6,43 % untuk data testing. Dengan demikian, model hibrida ARIMA-NN dapat disimpulkan memiliki performa yang baik dan andal dalam meramalkan jumlah penderita kasus TBC di Kota Surakarta pada periode yang dianalisis.

4. KESIMPULAN

Berdasarkan hasil penelitian, model ARIMA(0,1,1) mampu menangkap pola linier pada data jumlah kasus TBC dengan baik, yang ditunjukkan oleh nilai Mean Absolute Percentage Error (MAPE) sebesar 15,49 % pada data training dan 10,53 % pada data testing. Namun, model ini memiliki keterbatasan dalam mengenali pola non-linier dan kompleksitas data yang sering muncul dalam deret waktu penyakit menular seperti TBC. Untuk mengatasi hal tersebut, penelitian ini menggabungkan model ARIMA dengan Neural Network menjadi metode hibrida ARIMA-NN. Model Neural Network yang digunakan sebagai komponen dalam metode hibrida berhasil menurunkan nilai MAPE menjadi 4,14 % pada data training dan 6,43 % pada data testing, menandakan peningkatan signifikan dalam kemampuan peramalan. Penurunan nilai error tersebut menunjukkan Neural Network efektif dalam menangkap pola non-linier yang tidak bisa dijelaskan oleh model ARIMA saja. Dengan menggabungkan kekuatan kedua metode, model hibrida ARIMA-NN terbukti lebih andal dan akurat dalam memperamalkan jumlah kasus TBC, khususnya pada data dengan pola yang kompleks dan bervariasi. Pendekatan ini memberikan solusi yang lebih baik untuk perencanaan dan pengendalian penyakit TBC secara efektif, sekaligus menjadi dasar yang kuat bagi pengembangan model peramalan penyakit lain dengan karakteristik data serupa. Dengan demikian, model ini dapat membantu pengambil kebijakan dalam merancang strategi kesehatan yang lebih tepat sasaran dan efisien dalam pemanfaatan sumber daya.

REFERENCES

- [1] H. Malcom Frank Palandeng and O. Waworuntu, "Holistic approach on pulmonary tuberculosis: a case study," *J Kedokt Kom Tropik*, vol. 11, no. 1, pp. 449–456, 2023.
- [2] Y. Zulaikha and R. A. Syakurah, "Manajemen Program Pencegahan dan Pengendalian Tuberculosis Saat Pandemi Covid-19 di Puskesmas Mangunharjo," *Jurnal Bidang Ilmu Kesehatan*, vol. 13, no. 2, pp. 130–144, 2023, doi: 10.52643/jbik.v13i2.2236.
- [3] S. A. Sinaga, "Implementasi Metode Arima (Autoregressive Moving Average) Untuk Prediksi Penjualan Mobil," *Journal Global Technology Computer*, vol. 2, no. 3, pp. 102–109, 2023, doi: 10.47065/jogtc.v2i3.4013.

- [4] R. M. Faris, K. Kurniaji, D. Budiman, Y. Yoedani, M. W. Kusuma, and F. Lestari, "Analisis Peramalan Produksi Tanaman Kelapa Sawit Menggunakan Metode Arima pada PTPN Kebun Sukamaju," *Jurnal Bisnis dan Manajemen West Science*, vol. 3, no. 03, pp. 275–290, 2024, doi: 10.58812/jbmws.v3i03.1537.
- [5] S. M. Saragih and P. Sembiring, "Analisis Perbandingan Metode Arima Dan Double Exponential Smoothing Dari Brown Pada Peramalan Inflasi Di Indonesia," *Journal of Fundamental Mathematics and Applications (JFMA)*, vol. 5, no. 2, pp. 176–191, 2022, doi: 10.14710/jfma.v5i2.15312.
- [6] Listy Oktaviani, Sandy Erlangga, Bintang Aufa Sultan, Agus Perdana Windarto, and Putrama Alkhairi, "Penerapan Metode Jaringan Saraf Tiruan Dalam Memprediksi Produksi Daging Domba Menurut Provinsi," *Journal of Computing and Informatics Research*, vol. 3, no. 2, pp. 199–207, 2024, doi: 10.47065/comforch.v3i2.992.
- [7] L. Alzubaidi *et al.*, "A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications," *J Big Data*, vol. 10, no. 1, 2023, doi: 10.1186/s40537-023-00727-2.
- [8] C. Vista Magdalena Sihombing, S. Martha, and ainul Miftahul Huda, "Analisis Metode Hybrid Arima-Svr Pada Indeks Harga Saham Gabungan," *Buletin Ilmiah Math. Stat. dan Terapannya (Bimaster)*, vol. 11, no. 3, pp. 413–422, 2022.
- [9] I. K. Hasan and Ismail Djakaria, "Perbandingan Model Hybrid ARIMA-NN dan Hybrid ARIMA-GARCH untuk Peramalan Data Nilai Tukar Petani di Provinsi Gorontalo," *Jurnal Statistika dan Aplikasinya*, vol. 5, no. 2, pp. 155–165, 2021, doi: 10.21009/jsa.05204.
- [10] E. Lastinawati, A. Mulyana, I. Zahri, and S. Sriati, "Model ARIMA untuk Peramalan Harga Beras di Kabupaten Ogan Komering Ilir Provinsi Sumatera Selatan," *Prosiding Seminar Nasional Lahan Suboptimal*, vol. 5, no. 5, pp. 192–200, 2019.
- [11] A. A. Abidin, P. F. Buiney, and D. A. Nohe, "Peramalan Data Ekspor Kalimantan Barat Dengan Metode Autoregressive Integrated Moving Average (Arima)," ... *Nasional Matematika dan ...*, pp. 96–107, 2022, [Online]. Available: <http://jurnal.fmipa.unmul.ac.id/index.php/SNMSA/article/view/900%0Ahttp://jurnal.fmipa.unmul.ac.id/index.php/SNMSA/article/download/900/380>
- [12] M. A. Firmansyah and C. R. Oktarina, "Kinerja Peramalan Autoregressive Integrated Moving Average dan Seasonal Autoregressive Integrated Moving Average dalam Memprediksi Kejadian Gempa Bumi Sumatra," vol. 4, no. 1, pp. 1–16, 2025.
- [13] N. Kristianti and W. Widiatry, "Penggunaan Algoritma Hebb Dalam Pola Pengenalan Huruf," *Jurnal Teknologi Informasi: Jurnal Keilmuan dan Aplikasi Bidang Teknik Informatika*, vol. 18, no. 1, pp. 52–60, 2024, doi: 10.47111/jti.v18i1.12561.
- [14] Dwira Azi Pragana, D. W. Manurung, and Agus Perdana Windarto, "Analisa Metode Backpropagation Pada Prediksi Rata-rata Harga Beras Bulanan Di Tingkat Penggilingan Menurut Kualitas," *Journal of Computing and Informatics Research*, vol. 2, no. 3, pp. 77–84, 2023, doi: 10.47065/comforch.v2i3.855.
- [15] R. A. E. V. Targa Sapanji, S. Lestari, M. Murnawan, and R. Samiharjo, "Prediksi Indeks Bursa Efek Indonesia 2023 Pendekatan ARIMA, Machine Learning dengan R Programming," *Jurnal Manajemen Informatika (JAMIKA)*, vol. 13, no. 2, pp. 163–177, 2023, doi: 10.34010/jamika.v13i2.10777.
- [16] Zulhamidi and R. Hardianto, "PERAMALAN PENJUALAN TEH HIJAU DENGAN METODE ARIMA (STUDI KASUS PADA PT. MK)," *Jurnal PASTI*, vol. XI, no. 3, pp. 231–244, 2017.
- [17] M. F. M. Arsyada and R. Tyasnurita, "Evaluasi Model Jaringan Saraf Tiruan untuk Prediksi Konsentrasi Nitrogen Oksida (NO_x)," *Jurnal Sistem Informasi*, vol. 14, no. 3, pp. 2540–9719, 2025, [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [18] D. Arisandi, S. Salamun, and A. R. Putra, "Prediksi Penerimaan Siswa Baru dengan Metode Single Exponential Smoothing Melalui Metrik Evaluasi MAD, MSE dan MAPE," *Journal of Information System Research (JOSH)*, vol. 4, no. 4, pp. 1197–1204, 2023, doi: 10.47065/josh.v4i4.3658.