Vol. 5 No. 3, 2024, pp. 262-274
©2024 Ilmu Komputer Unila Publishing Network all rights reserved
Jurnal Pepadun

Komparasi Peramalan ARIMA dan RNN pada Tunggakan Peserta BPJS Kesehatan Cabang Jambi

[*1]Ulfa Khaira dan [2]Bisma Aulia

[1,2]Jurusan Teknik Elektro dan Informatika, Fakultas Sains dan Teknologi, Universitas Jambi, Jl. Jambi –Muara Bulian No. KM. 15, Mendalo Darat, Kec. Jambi Luar Kota, Kabupaten Muaro Jambi, Provinsi Jambi, Indonesia
e-mail: *[1]**ulfakhaira@unja.ac.id**, [2]bismaaulia82@gmail.com

Abstract *— BPJS Kesehatan is a public legal entity that is directly responsible to the President who has the task of providing Social Security to the community and has the authority to hold dues every month for participants. BPJS Kesehatan participants who are independent participants have an obligation to pay dues every month. Pandemic conditions affect participants' incomes in paying dues as a result participants' arrears rise. Arrears participants also occurred in BPJS Kesehatan Jambi Branch Office. BPJS Kesehatan branch office has SIMANIS (Sistem Penagihan Iuran Terintegrasi) application. Through this system, BPJS Kesehatan can reduce the arrears participants, especially Jambi branch offices. Thus, it makes easier for each branch office to evaluate the right policy, through arrears data it can be processed into forecasting information using traditional statistic methods ARIMA and machine learning methods RNN (Recurrent Neural Network). Therefore, this study aims to compare ARIMA and RNN in forecasting on premium arrears data of BPJS Kesehatan Jambi Branch Office participants. The dataset used is data on arrears of premium participants BPJS Kesehatan branch office Jambi per week. Through this forecasting, the results RMSE (Root Mean Square Error) and MAPE (Mean Absolute Percentage Error) from RNN 1 hidden layer 1 neuron method have RMSE 2739803819 and MAPE 1.97%, which is better than ARIMA (3,1,3) with RMSE 3238308628 and MAPE 2.62%. Then, to make it easier for non-programmer users to forecast built a graphical user interface (GUI) based on the web with Streamlit library. This research is expected to validate the forecasting model that is suitable for the arrears data of BPJS Kesehatan Jambi Branch Office participants.*

Keywords**:** *ARIMA; BPJS; Forecasting; RNN.*

1. PENDAHULUAN

Setiap warga negara Indonesia berhak mendapatkan jaminan atas kesejahteraan hidupnya. Berdasarkan UU No. 11 Tahun 2009 tentang Kesejahteraan Sosial, kesejahteraan sosial diselenggarakan untuk memenuhi kebutuhan dasar warga negara melalui berbagai program, seperti rehabilitasi sosial, jaminan sosial, pemberdayaan sosial, dan perlindungan sosial. Untuk melaksanakan undang-undang tersebut, dibentuklah BPJS Kesehatan (Badan Penyelenggara Jaminan Sosial Kesehatan), sebuah badan hukum publik yang bertugas menyediakan jaminan kesehatan bagi seluruh warga negara [1]. BPJS Kesehatan bertanggung jawab langsung kepada Presiden dan merupakan hasil transformasi dari PT. Askes (Persero), yang mulai beroperasi pada 1 Januari 2014. BPJS Kesehatan memiliki wewenang untuk menetapkan iuran bulanan bagi peserta dan menagih pembayaran iuran sesuai dengan ketentuan dalam UU No. 24 Tahun 2011 tentang Badan Penyelenggara Jaminan Sosial, pasal 10. Peserta BPJS Kesehatan yang tergolong dalam kategori kepesertaan mandiri wajib membayar iuran setiap bulan. Pada 1 Januari 2020, iuran bagi peserta mandiri mengalami kenaikan seiring dengan defisit yang terjadi pada program JKN (Jaminan Kesehatan Nasional) [2]. Salah satu penyebab defisit ini adalah ketidakmampuan peserta dalam membayar iuran, yang sesuai dengan temuan dalam penelitian Latifah et al. [3], yang menyatakan bahwa pendapatan seseorang memengaruhi tingkat kepatuhannya dalam membayar iuran. Selain itu, kondisi pandemi pada tahun 2020 juga memengaruhi pendapatan peserta, yang berdampak pada meningkatnya tunggakan pembayaran BPJS Kesehatan.

©2024 Ilmu Komputer Unila Publishing Network all rights reserved

Untuk mengatasi masalah tunggakan pada peserta kepesertaan mandiri, BPJS Kesehatan telah mempermudah akses pembayaran dengan menyediakan berbagai kanal pembayaran seperti *mobile banking*, PT. POS, Indomaret, Tokopedia, Alfamart, dan berbagai bank yang mudah dijangkau oleh peserta. Selain itu, BPJS Kesehatan juga melakukan upaya penagihan melalui *tellecollecting*, yang mencakup penghubungan kepada peserta melalui pesan singkat, panggilan telepon, atau media WhatsApp. Dengan penagihan melalui *tellecollecting*, BPJS Kesehatan dapat lebih efisien dalam melakukan pendataan dan penagihan kepada peserta yang menunggak. Tunggakan pembayaran peserta juga terjadi di BPJS Kesehatan Kantor Cabang Jambi. Menurut mantan Kepala Cabang BPJS Kesehatan Jambi, Rizki Lestari, pada tahun 2016 tercatat ada tunggakan sebesar 11 miliar Rupiah di wilayah tersebut [4]. Kantor Cabang BPJS Kesehatan Jambi mencakup wilayah Kabupaten Muara Jambi, Kabupaten Batanghari, Kabupaten Tanjung Jabung Timur, dan Kabupaten Tanjung Jabung Barat. Untuk mendukung proses penagihan, BPJS Kesehatan Kantor Cabang Jambi juga menggunakan aplikasi SIMANIS, yang membantu staf dalam melakukan penagihan dan mempercepat proses pengumpulan tunggakan.

Aplikasi SIMANIS berfungsi untuk memudahkan staf BPJS Kesehatan dalam mengelola data tunggakan peserta. Melalui sistem ini, BPJS Kesehatan di Kantor Cabang Jambi dapat mengurangi tingkat tunggakan dengan lebih efektif. Selain itu, aplikasi ini juga membantu setiap kantor cabang untuk mengevaluasi kebijakan yang tepat sesuai dengan kondisi di masing-masing wilayah. Tanpa adanya evaluasi yang efektif, kemungkinan besar akan terjadi kecurangan, seperti peserta yang menunggak iuran setelah menerima layanan kesehatan, yang tentunya merugikan fasilitas kesehatan serta warga negara secara keseluruhan [5]. Berdasarkan fakta-fakta tersebut, diperlukan penelitian mengenai peramalan tunggakan pembayaran di BPJS Kesehatan Kantor Cabang Jambi. Dengan adanya peramalan tunggakan pembayaran, pemangku kepentingan dapat membuat keputusan yang lebih tepat dan mengevaluasi kebijakan-kebijakan yang ada untuk mengatasi masalah tunggakan. Selain itu, peramalan ini juga dapat mencegah terjadinya defisit dan membantu dalam merancang kebijakan kesehatan yang lebih efektif serta dalam pengalokasian anggaran di masa depan [5]. Pola dari data peramalan tunggakan pembayaran ini juga bisa menjadi indikator kinerja bagi staf-staf BPJS Kesehatan, yang memungkinkan untuk melakukan perbaikan dan penguatan kinerja di lapangan. Metode yang digunakan untuk meramalkan tunggakan pembayaran ini adalah analisis runtun waktu (*time series*), yang dapat memberikan gambaran yang lebih akurat mengenai tren dan pola tunggakan di masa depan.

Analisis runtun waktu atau *time series* merupakan metode analisis data kuantitatif yang digunakan untuk memprediksi pola data yang telah dikumpulkan dalam periode waktu tertentu [6]. Untuk melakukan peramalan data *time series*, berbagai model prediksi dapat diterapkan, baik dengan teknik statistik maupun teknik *machine learning*. Beberapa teknik statistik yang umum digunakan termasuk ARIMA (*Auto Regressive Integrated Moving Average*), ETS, Holt, dan lain-lain, sementara dalam kecerdasan buatan, teknik *machine learning* yang populer mencakup *neural network*, *decision tree*, dan *support vector machine* [7]. Model ARIMA adalah salah satu model statistik tradisional yang sering digunakan untuk peramalan jangka pendek, namun jika digunakan untuk peramalan jangka panjang, model ini cenderung menghasilkan grafik yang datar karena sifatnya yang tidak mempertimbangkan faktor perubahan jangka panjang.

Beberapa penelitian sebelumnya telah membahas penggunaan model ARIMA dan *Recurrent Neural Network* (RNN) dalam peramalan. Salah satunya adalah penelitian Suhartono dan A.J. Endharta pada 2016 yang membandingkan model ARIMA musiman ganda dan *Elman-Recurrent Neural Network* untuk meramalkan konsumsi listrik jangka pendek [8]. Selain itu, penelitian oleh Makridakis et. al. pada 2018 juga membandingkan metode statistik tradisional (seperti ARIMA) dengan metode *machine learning* (seperti RNN) dalam konteks peramalan [9]. Meskipun banyak penelitian membahas perbandingan model ARIMA dan RNN, belum ada studi yang membandingkan kedua model ini dalam konteks peramalan tunggakan pembayaran BPJS Kesehatan Kantor Cabang Jambi.

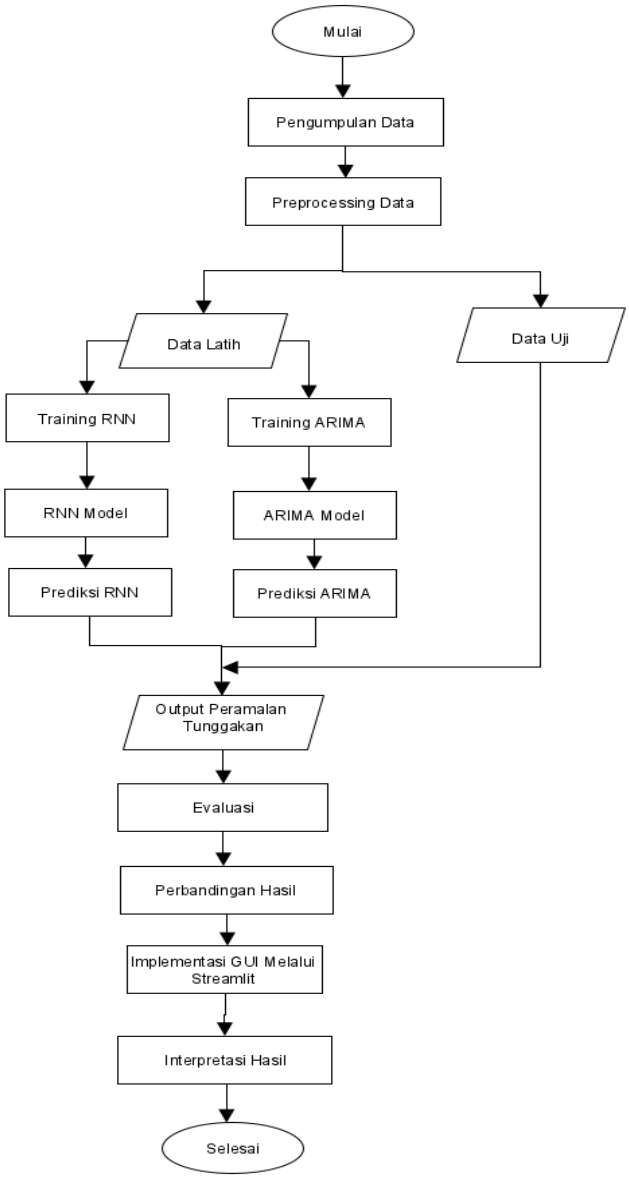
Untuk mendukung kelangsungan proses peramalan yang dapat dilakukan oleh pengguna terkait, penelitian ini menggunakan web interaktif berbasis library *Streamlit*. Streamlit adalah sebuah *library open-source* Python yang memungkinkan pembuatan antarmuka web secara mudah (docs.streamlit). Dengan menggunakan *Streamlit*, pengguna yang tidak memiliki latar belakang pengkodean tetap dapat menjalankan proses peramalan. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja model ARIMA yang berbasis teknik statistik tradisional dengan model *Recurrent Neural Network* (RNN) yang berbasis teknik *machine learning*, dengan menyediakan antarmuka pengguna grafis (GUI). Diharapkan hasil penelitian ini

©2024 Ilmu Komputer Unila Publishing Network all rights reserved

dapat memberikan kontribusi dalam pengambilan kebijakan kesehatan di BPJS Kesehatan Kantor Cabang Jambi, serta menyediakan alat yang berguna untuk meramalkan tunggakan pembayaran secara lebih akurat dan efisien.

2. METODOLOGI PENELITIAN

Pada penelitian ini, tahapan-tahapan yang dilakukan ditunjukkan pada Gambar 1.



Gambar 1. Alur penelitian.

2.1 Pengumpulan Data

Dalam melakukan penelitian ini, diperlukan pengumpulan *dataset* yang berupa kumpulan data *time series* tunggakan pembayaran peserta premi BPJS Kesehatan Kantor Cabang Jambi dari November 2019 hingga Oktober 2021. Data yang dikumpulkan berupa data numerik. Sampel data dari tunggakan pembayaran peserta dapat dilihat pada Tabel 1.

©2024 Ilmu Komputer Unila Publishing Network all rights reserved

Tabel 1. Sampel data.

No.	Minggu ke-	Data Tunggakan (Rupiah)
1.	I, November 2019	251.077.648
2.	II, November 2019	299.282.300
3.	III, November 2019	158.931.108
4.	IV, November 2019	169.742.316
5.	I, Desember 2019	207.817.855

2.2 *Preprocessing* Data

Tahap selanjutnya adalah *preprocessing* data. *Preprocessing* data pada penelitian ini dilakukan untuk memeriksa apakah data yang akan diolah terdapat *missing value* atau tidak. *Missing value* adalah hilangnya sebagian atau beberapa informasi dalam sebuah *dataset* [19]. Pada data ini, dilakukan pengecekan *missing value* menggunakan fungsi dari Python yaitu isna().

2.3 Data Latih dan Data Uji

Pada tahapan ini dilakukan proses pembagian *dataset* untuk dilakukan proses pelatihan dan data yang akan diuji. Data latih atau data *training* adalah kumpulan data yang memiliki label yang digunakan untuk mengenal karakteristik dari data tersebut oleh mesin agar dapat menghasilkan pola atau model dari data tersebut. Data latih digunakan untuk membuat permodelan, sedangkan data uji atau data *testing* adalah yang memiliki label untuk menguji performa dan kebenaran dari model yang telah dilatih [20]. Pada penelitian ini, dilakukan pembagian data latih sebesar 80% dan data uji sebesar 20% dari total 102 *record* data.

2.4 Model ARIMA dan RNN

Setelah dilakukan proses pelatihan untuk melatih model yang akurat dilakukan proses pelatihan, akan terbentuk sebuah model dari proses latihan *dataset* tersebut, kemudian akan dilakukan proses prediksi dari model yang telah didapatkan. Dalam melakukan pemodelan, penelitian ini menggunakan dua metode yakni model ARIMA dan *Reccurent Neural Network* (RNN).

2.5 Evaluasi

Proses selanjutnya yaitu evaluasi kinerja dari 2 model tersebut melalui *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE). MAPE dipilih karena ukuran variabel peramalan adalah faktor utama dalam mengevaluasi keakurasian peramalan yang dilakukan, selain itu MAPE membandingkan kesalahan antara nilai aktual dan peramalan [21]. RMSE adalah metode untuk mengevaluasi teknik peramalan dengan mengukur tingkat akurasinya. Persamaan 1 menunjukkan formula perhitungan RMSE [22].

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(A_t - F_t)^2}{n}} \tag{1}$$

Keterangan:
A_t : nilai data aktual
F_t : nilai hasil peramalan
n : banyak data

MAPE adalah rata-rata dari diferensiasi antara hasil sebenarnya dan nilai peramalan yang dinyatakan melalui persentase nilai aktual. Persamaan 2 menunjukkan formula perhitungan MAPE [23].

$$\text{MAPE} = \frac{\sum_{t=1}^{n} |\text{PEt}|}{n} \tag{2}$$

Keterangan:
PE_t : persentase kesalahan pada waktu ke-t
E_t : nilai kesalahan waktu ke-t
Z_t : data aktual waktu ke-t

©2024 Ilmu Komputer Unila Publishing Network all rights reserved

n : banyak data

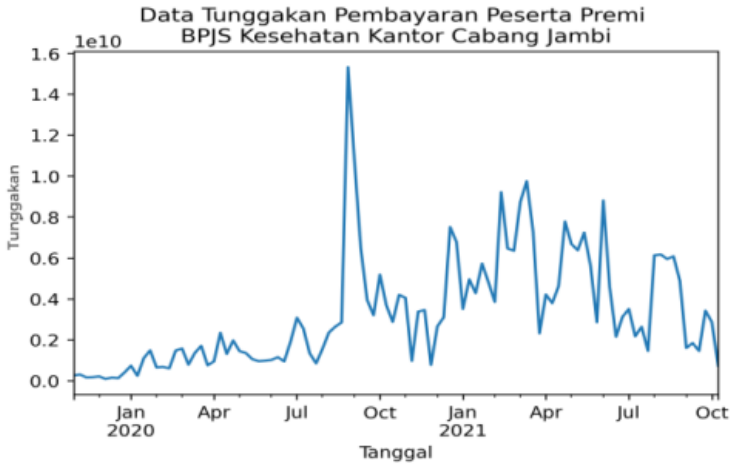
2.6 Implementasi GUI melalui Streamlit

Setelah dilakukan *preprocessing* data hingga evaluasi data secara manual untuk mengukur *dataset* tunggakan peserta premi BPJS Kesehatan Kantor Cabang Jambi, maka selanjutnya dapat dibangun GUI yang dapat digunakan oleh pengguna secara mudah tanpa memiliki dasar *coding*. Aplikasi ini disusun dengan dasar dari pengujian yang telah dilakukan pada tahap sebelumnya

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Tahap awal yang dilakukan adalah pengumpulan data tunggakan peserta premi BPJS Kesehatan Kantor Cabang Jambi yang dijadikan *dataset*. *Dataset* diperoleh dari BPJS Kesehatan Kantor Cabang Jambi. Bentuk *dataset* adalah data tunggakan pembayaran peserta premi BPJS Kesehatan Kantor Cabang Jambi dalam kurun waktu kurang lebih 2 tahun yaitu 2019-2021. *Dataset* tersebut disusun ke dalam *file excel* dan memiliki format ekstensi xlsx. *Dataset* yang diperoleh sebanyak 102 *records*, terhitung per-minggu dari November 2019 hingga Oktober 2021 yang disajikan pada Gambar 2.



Gambar 2. *Plot* data tunggakan peserta BPJS kesehatan November 2019 - Oktober 2021.

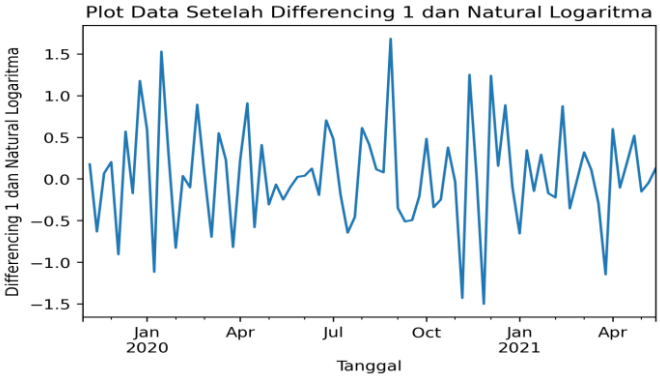
3.2 Praproses Data

Tahap selanjutnya adalah praproses data. Praproses data ini dilakukan untuk melihat apakah data yang akan diolah terdapat *missing value* atau tidak. Pada data ini dilakukan pengecekan menggunakan fungsi dari Python yaitu isna(). Hasil dari pengecekan pada data tidak ditemukan *missing value*. Dengan demikian, data dapat dilanjutkan untuk dibagi menjadi data latih dan data uji, yakni sebesar 80% data latih dan 20% data uji. Data latih dan data uji menggunakan *dataset* yang sama yaitu data tunggakan peserta BPJS Kesehatan Kantor Cabang Jambi.

3.3 Pengolahan Data dengan ARIMA

Sebelum pemodelan dilakukan, dicek terlebih dahulu apakah data yang akan diolah apakah sudah stasioner atau belum. Dapat dilihat pada Gambar 3, bahwa grafik data tampak tidak rata dan cenderung naik. Maka, data dilakukan pengujian ADF. Setelah proses pengujian didapatkan bahwa *p-value* yang diperoleh adalah 0,0206 dengan *p-value* < 0,05 = α. Maka, dapat diambil kesimpulan bahwa data stasioner secara *mean*. Akan tetapi, karena secara variansi data tidak stasioner, maka dilakukan proses *differencing* orde 1 dan natural logaritma. Setelah proses uji hipotesis baru, didapatkan hasil dimana nilai *p-value* (0,000) < 0,05 = α. Maka, secara *mean* data telah stasioner. Secara varians hasil dapat dilihat pada Gambar 4 berikut ini.

©2024 Ilmu Komputer Unila Publishing Network all rights reserved



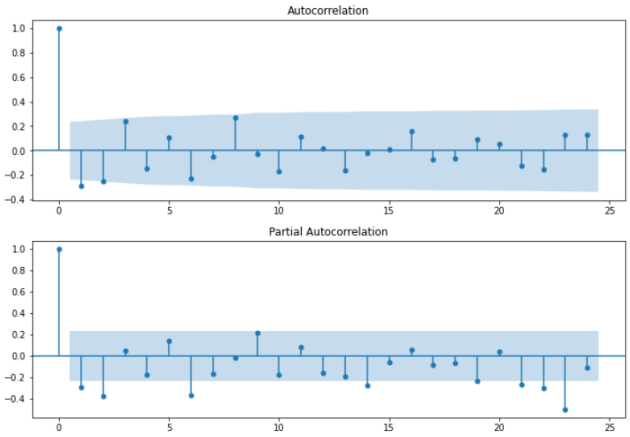
Gambar 3. *Plot* data setelah *differencing* dan natural logaritma.

Dari hasil uji hipotesis tersebut dapat dilihat bahwa data sudah stasioner, selain itu pola grafik juga menunjukkan variansi data relatif lebih stabil. Maka, data telah dianggap sudah stasioner *mean* dan variansi setelah proses *differencing* ke-1. Proses berhenti ke proses diferensiasi 1 karena data telah stasioner secara variansi dan rata-rata. Proses diferensiasi dapat dilakukan lebih dari 1 kali, akan tetapi umumnya diferensiasi dilakukan sebanyak 2 kali. Hal ini karena jika lebih dari 2 kali, maka biasanya akan terjadi *overdifferencing*. *Overdifferencing* adalah proses *differencing* secara berlebihan yang akan mengakibatkan data kehilangan efektivitasnya dan kecenderungan hilangnya informasi dari deret waktu [24]. Sampel perbandingan antara *dataset* asli dan data setelah proses diferensiasi 1 ditunjukkan pada Tabel 2.

Tabel 2. Sampel perbedaan data aktual dan data diferensiasi 1.

Minggu ke-	**Tanggal**	**Tunggakan Asli (Rupiah)**	**Tunggakan Setelah Diferensiasi 1**
I November 2019	03-11-19	251077648	0
II November 2019	10-11-19	299282300	48204652.0
III November 2019	17-11-19	158931108	-140351192.0

Tahap selanjutnya adalah mengidentifikasi data melalui *plot* ACF (*Autocorrelation Function)* dan PACF (*Partial Autocorrelation Function)* untuk mengestimasi model ARIMA dari data tunggakan peserta. Pada tahap sebelumnya, data telah stasioner menggunakan diferensiasi tingkat satu. Maka, telah didapatkan *order* dari (p, d, q) yakni *order* d, yaitu d=1. Maka, akan ditentukan *order* dari p dan q melalui *plot* ACF dan PACF. Berikut pada Gambar 4 adalah *plot* ACF dan PACF dari data yang telah diproses.



(a)

(b)

Gambar 4 (a). *Plot* ACF (b) *Plot* PACF.

Jumlah *lag* pada masing-masing *plot* adalah 24 *lag*. Pada penelitian ini, dilakukan estimasi parameter model ARIMA melalui pengamatan ACF (*Autocorrelation Function)* dan PACF (*Partial Autocorrelation Function)*. Dalam melakukan pengamatan, perlu dilihat dari pola ACF dan PACF yaitu pola *cut-off*. Pola *cut-off* terjadi

©2024 Ilmu Komputer Unila Publishing Network all rights reserved

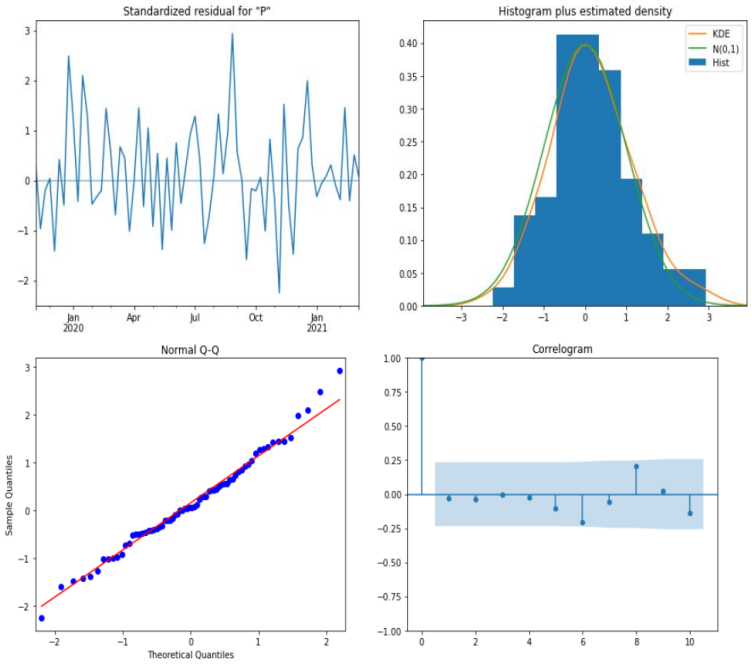
apabila data mendekati nol dan tidak melewati batas daerah biru. Dapat diketahui dari Gambar 4 bahwa grafik ACF *cut off* pada *lag* ke-3 dan grafik PACF tampak *cut-off* pada *lag* ke-3. Grafik ACF mengindikasikan model MA (*Moving Average*), sedangkan PACF mengindikasikan model AR (*AutoRegresi*). Maka dapat diketahui bahwa estimasi model dari data tersebut adalah ARIMA (3,1,3).

Setelah proses pemilihan model ARIMA (3,1,3), tahap selanjutnya adalah melakukan pemeriksaan diagnostik yang terdiri dari uji *white noise* dan uji normalitas. Dalam penelitian ini menggunakan Ljung-Box dalam menguji *white noise*. Jika nilai *p-value* lebih dari nilai α maka asumsi *white noise* terpenuhi. Pada hasil yang telah didapatkan diketahui bahwa model ARIMA (3,1,3) memenuhi syarat *white noise*, yakni *p-value* (0,8). Kemudian dilakukan pengujian normalitas untuk mengecek apakah model telah berdistribusi normal atau belum. Dalam pengujian normalitas, peneliti menggunakan uji Jarque-Bera. Dalam pengujian normalitas juga disimpulkan data telah berdistribusi normal apabila *p-value* > α = 0,05. Pada hasil yang telah diujikan didapatkan bahwa model ARIMA (3,1,3) telah berdistribusi normal karena memiliki *p-value* (0,58) > α = 0,05. Hasil dari model ARIMA (3,1,3) dapat dilihat pada Tabel 3.

Tabel 3. Nilai uji *White Noise* dan Normalitas.

White Noise	Nilai	Normalitas	Nilai
Ljung-Box(Q)	0.07	Jarque-Bera (JB)	1.11
Prob(Q)	0.80	Prob (JB)	0.58
Heteroskedascity (H)	0.81	Skew	0.29
Prob(H) (two-sided)	0.62	Kurtosis	3.21

Maka, melalui uji normalitas dan uji *white noise* tersebut dapat disimpulkan bahwa model ARIMA (3,1,3) telah memenuhi untuk dijadikan model peramalan tunggakan peserta premi BPJS Kesehatan Kantor Cabang Jambi. *Plot* dari pengujian tersebut yaitu ditunjukkan pada Gambar 5. Gambar grafik *standardized residual* menunjukkan bahwa residu dari waktu ke waktu yang tidak menampilkan musim yang jelas dan tampak seperti *white noise*. *Plot histogram plus estimated density* menunjukkan bahwa garis berwarna oranye atau KDE yang mendekati garis N, maka menunjukkan bahwa data telah berdistribusi normal. *Plot* Normal Q-Q menunjukkan bahwa asumsi normalitas terpenuhi dilihat dari titik-titik yang telah mengikuti garis, sedangkan *plot* Correlogram menunjukkan bahwa asumsi *white noise* telah terpenuhi, dengan tidak ada garis *lag* yang melewati batas.



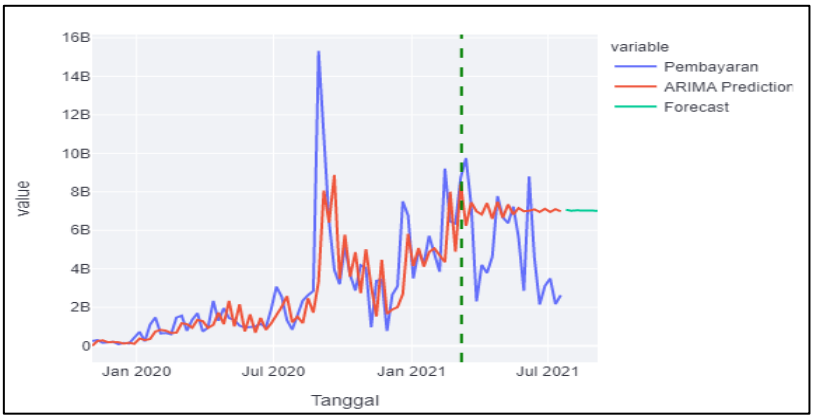
Gambar 5. *Plot* pengujian Normalitas dan *White Noise.*

©2024 Ilmu Komputer Unila Publishing Network all rights reserved

Tahap terakhir dari pengolahan data tunggakan peserta premi BPJS Kesehatan Kantor Cabang Jambi menggunakan metode ARIMA adalah evaluasi model yang telah dibuat yaitu ARIMA (3,1,3). Pada proses pengujian model ini dilakukan evaluasi menggunakan RMSE (*Root Mean Square Error*) yaitu 3.238.308.628 dan MAPE (*Mean Absolute Percentage Error*) sebesar 2,62%. Grafik perbandingan antara data aktual dan hasil peramalan ditunjukkan pada Gambar 6 dan nilai perbandingan ditunjukkan pada Tabel 4.

Tabel 4. Sampel perbedaan data aktual dan data prediksi.

Minggu Ke-	Tunggakan Aktual (Rupiah)	Prediksi (Rupiah)
I Mei 21	7.280.281.803	7.450.374.695
II Mei 21	2.311.292.184	6.987.865.237
III Mei 21	4.207.215.266	6.822.671.025
IV Mei 21	3.787.075.980	7.415.443.194
I Jun 21	4.625.594.771	6.605.671.820
II Jun 21	7.775.119.149	7.486.055.628
III Jun 21	6.677.584.866	6.657.059.613
IV Jun 21	6.370.596.920	7.345.018.973



Gambar 6. Perbandingan data aktual dan data prediksi ARIMA.

3.4 Pengolahan Data dengan RNN

Metode kedua yang digunakan adalah RNN. Pemrosesan data menggunakan RNN digunakan bantuan bahasa pemrograman Python dengan IDE *Jupiter*. Proses pemodelan menggunakan *neural network* membutuhkan penentuan *input* dan banyaknya lapisan tersembunyi (*hidden layer*). Jaringan ini menggunakan 1 *hidden layer* dengan *neuron* 1, 2, 3, 4, 5, 6, 7, 8, 9, dan 10. *Output* data dari penelitian RNN adalah 1. Untuk *epoch* dalam penelitian ini yang dipilih adalah 100.

Tabel 5. Uji coba RNN.

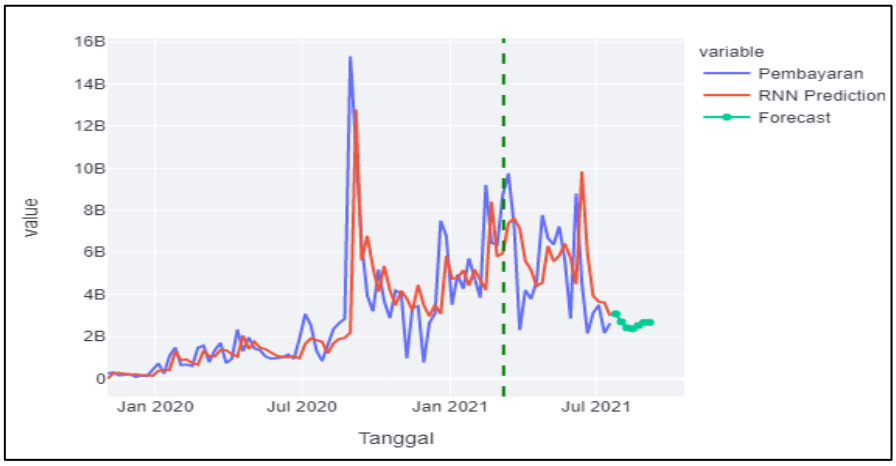
Neuron	RMSE	MAPE
1	2739803819	1,97%
2	3550350399	2,62%
3	2872294522	2,22%
4	5325247043	2,5%
5	4710586924	3,09%
6	12528099154	4,7%
7	13398089276	2,82%
8	5600628162	2,83%
9	5087021643	2,41%
10	278109330740	6,4%

©2024 Ilmu Komputer Unila Publishing Network all rights reserved

Sesuai dengan uji coba yang telah dilakukan pada Tabel 5, diketahui bahwa dari RMSE dan MAPE 1 *hidden layer* dengan 1 *neuron* memiliki nilai yang rendah. Oleh karena itu, dipilih 1 *unit neuron* sebagai *neuron hidden layer* yang optimal dibandingkan dengan *neuron hidden layer* lainnya. Perbandingan antara nilai aktual dan nilai prediksi atau peramalan dari percobaan pada 1 *hidden layer* dengan 1 *neuron* disajikan pada Tabel 6. Grafik perbandingan antara data aktual dan data prediksi ditunjukkan pada Gambar 7.

Tabel 6. Perbedaan data aktual dan data prediksi 1 *hidden layer* dan 1 *neuron*.

Minggu Ke-	Tunggakan Aktual (Rupiah)	Prediksi (Rupiah)
I Mei 21	7.280.281.803	7.333.916.662
II Mei 21	2.311.292.184	6.401.746.508
III Mei 21	4.207.215.266	5.038.203.332
IV Mei 21	3.787.075.980	4.856.551.942
I Jun 21	4.625.594.771	4.424.911.218
II Jun 21	7.775.119.149	4.564.543.416
III Jun 21	6.677.584.866	6.280.827.133
IV Jun 21	6.370.596.920	5.621.753.627



Gambar 7. Perbandingan data aktual dan data prediksi RNN.

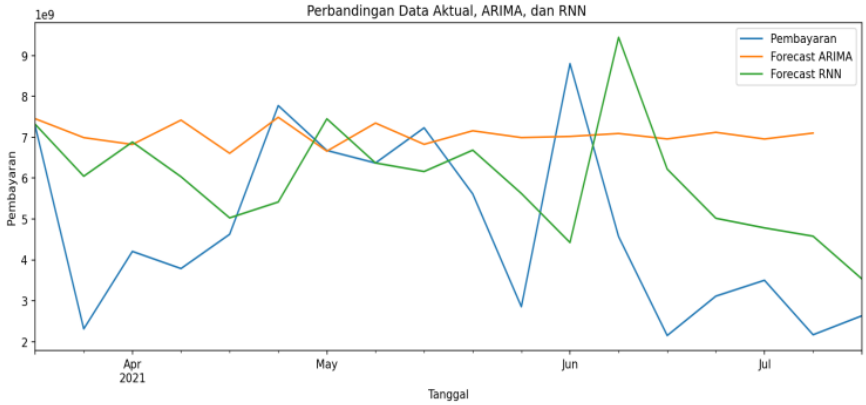
3.5 Perbandingan Model ARIMA dan RNN

Pada penelitian ini digunakan dua metode dalam melakukan pengolahan data tunggakan peserta premi BPJS Kesehatan Kantor Cabang Jambi. Pada hasil yang telah diperoleh, diketahui bahwa model ARIMA (3,1,3) dan model RNN dengan 1 *hidden layer* dengan 1 *neuron* merupakan yang terbaik. Dengan menggunakan model ARIMA (3,1,3) didapatkan nilai RMSE 3238308628 dan MAPE 2,62%, sedangkan menggunakan RNN didapatkan nilai RMSE 2739803819 dan MAPE 1,97%. Perbandingan secara visual antara data aktual dan metode ARIMA serta RNN dapat dilihat pada Gambar 8. Sampel perbandingan data aktual dan data peramalan antara ARIMA dan RNN disajikan pada Tabel 7.

Tabel 7. Perbandingan data aktual dan prediksi ARIMA dan RNN

Data Aktual	Data Prediksi ARIMA (Rupiah)	Data Prediksi RNN(Rupiah)
7.280.281.803	7.450.374.695	7.333.916.662
2.311.292.184	6.987.865.237	6.401.746.508
4.207.215.266	6.822.671.025	5.038.203.332
3.787.075.980	7.415.443.194	4.856.551.942
4.625.594.771	6.605.671.820	4.424.911.218

©2024 Ilmu Komputer Unila Publishing Network all rights reserved



Gambar 8. Perbandingan data aktual dan data prediksi ARIMA dan RNN.

Dari hasil yang telah didapatkan diketahui bahwa peramalan menggunakan RNN lebih optimal dibandingkan menggunakan metode ARIMA. Hasil dari penelitian ini sesuai dengan penelitian Wikantara [25], dimana dilakukan penelitian perbandingan antara ARIMA dan RNN dalam memprediksi jumlah wisatawan asing. Dalam penelitian tersebut didapatkan bahwa dalam meramalkan jumlah wisatawan asing lebih optimal menggunakan metode RNN dengan hasil uji MAPE sebesar 5.417% dan RMSE sebesar 21884 dibandingkan dengan metode ARIMA yang memiliki MAPE sebesar 6.965% dan RMSE 27181. Kemudian penelitian oleh Suhartono dan Endharta [8] dengan hasil perbandingan ketepatan ramalan model menunjukkan bahwa metode Elman-RNN, yaitu jaringan Elman-RNN adalah model terbaik untuk meramalkan konsumsi listrik per jam di Mengare Gresik. Akan tetapi, menurut penelitian yang dilakukan oleh Makridakis et. al. [9], metode ARIMA memilki nilai akurasi / MAPE sebesar 7,19%, sedangkan untuk RNN memiliki akurasi sebesar 9,48%. Dari hasil yang penelitian yang didapatkan dan dikaitkan dengan penelitian sebelumnya, metode ARIMA dan metode RNN sama-sama memiliki kelebihan dan kekurangan masing-masing, sesuai menurut Wiyanti dan Pulungan [6], bahwa tidak ada model peramalan tunggal yang terbaik pada setiap situasi. Pada model peramalan ARIMA baik untuk situasi peramalan jangka pendek, karena jika dilakukan proses peramalan secara jangka panjang akan mengakibatkan model yang akan mendatar [26]. Di sisi lain, RNN cocok untuk *timeseries* yang panjang karena dapat melatih modelnya agar lebih optimal [9].

3.6 Peramalan Jumlah Penunggakan Peserta Premi BPJS Kesehatan Kantor Cabang Jambi

Metode yang cocok digunakan dalam meramalkan jumlah penunggakan peserta premi BPJS Kesehatan Kantor Cabang Jambi adalah metode RNN. Pemilihan metode RNN karena hasil akurasi yang lebih baik dibandingkan peramalan menggunakan metode ARIMA. Berikut adalah hasil dari peramalan tunggakan peserta premi BPJS Kesehatan Kantor Cabang Jambi untuk 7 minggu selanjutnya ditunjukkan pada Tabel 8.

Tabel 8. Hasil peramalan tunggakan premi peserta selama 7 minggu kedepan

Tanggal/Minggu Ke-	Prediksi (Rupiah)
25-Juli-2021	3.081.640.945
01-Agustus-2021	2.697.369.700
08- Agustus -2021	2.414.574.299
15- Agustus -2021	2.370.274.454
22- Agustus -2021	2.524.102.495
29- Agustus -2021	2.670.217.607
09- September -2021	2.675.204.621

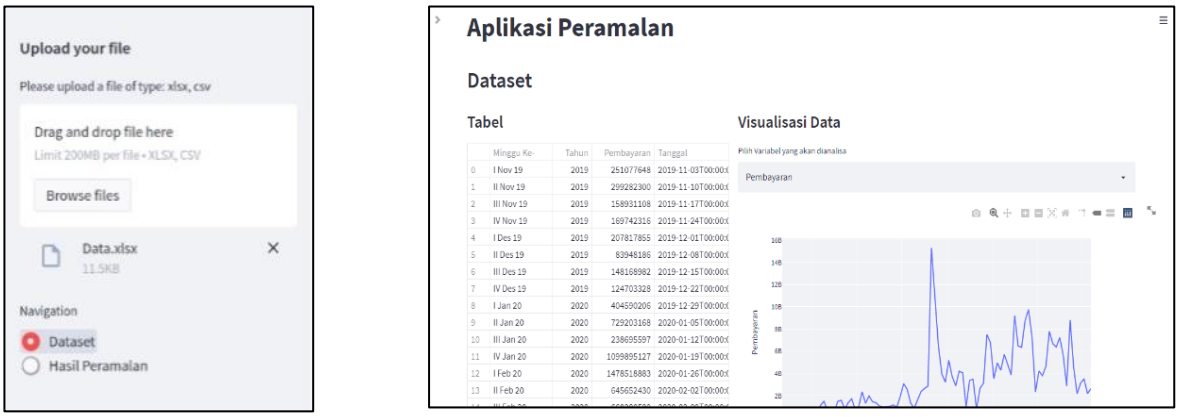
3.7 Implementasi Aplikasi Peramalan melalui GUI Streamlit

Pada penelitian ini, dirancang aplikasi peramalan yang dibangun melalui *library* Python Streamlit. Aplikasi ini digunakan untuk meramalkan tunggakan peserta premi BPJS Kesehatan Kantor Cabang Jambi. Aplikasi ini

©2024 Ilmu Komputer Unila Publishing Network all rights reserved

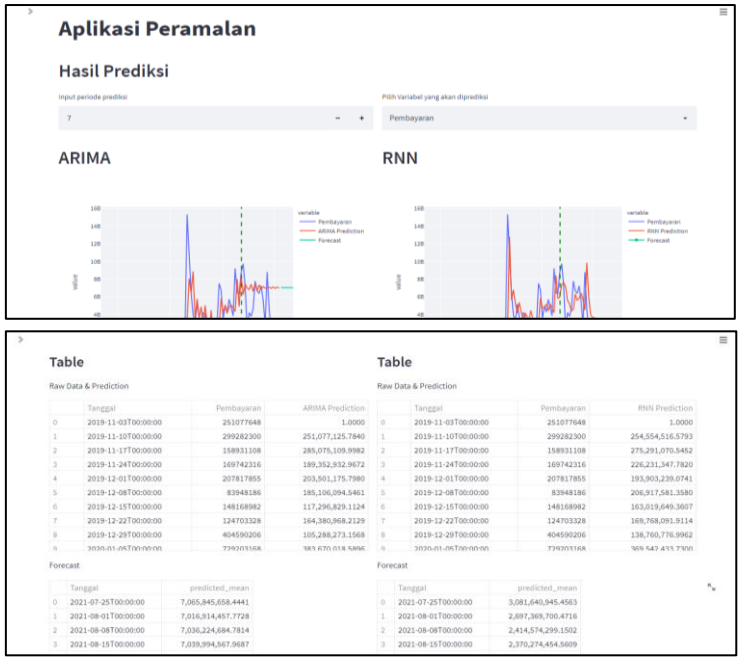
hanya digunakan untuk melakukan sebuah proses analisis *time series*, dengan *dataset* yang diolah hanya 1 variabel yakni *dataset* tunggakan. Aplikasi dikembangkan agar pengguna yang tidak memahami statistika atau *non-programmer* dapat melakukan peramalan terhadap *dataset* tunggakan peserta premi BPJS Kesehatan Kantor Cabang Jambi. Kebutuhan sistem yang disediakan dalam aplikasi ini yaitu *plot* dataset *time series* dan hasil peramalan ARIMA dan RNN. Aplikasi ini dikembangkan dari beberapa *library* yang ada di Python. Adapun *library-library* yang menunjang pengembangan aplikasi ini adalah *streamlit*, *numpy*, *pandas*, *matplotlib*, *pylab*, *statsmodels*, *sklearn*, *keras*, dan *math*. *Library-library* ini digunakan untuk metode atau model peramalan ARIMA dan RNN.

Antarmuka yang pertama adalah *sidebar*. Pada menu ini pengguna dapat memasukkan *dataset*, melihat nama *dataset*, dan memilih menu *dataset* atau hasil peramalan. Tampilan antarmuka *sidebar* ditunjukkan pada Gambar 9 (a). Antarmuka selanjutnya adalah menu *dataset*. Pada menu ini pengguna dapat melihat tabel *dataset* yang dimasukkan, variabel yang akan dipilih dan visualisasi *dataset* tersebut. Menu ini ditunjukkan pada Gambar 9 (b). Antarmuka yang terakhir adalah menu hasil peramalan. Pada menu ini dilakukan proses peramalan *dataset* yang telah dimasukkan sebelumnya. Pada menu ini ditampilkan hasil prediksi berupa visualisasi data perbandingan data uji dan data tes dari ARIMA dan RNN. Selain itu, ditampilkan juga nominal angka perbandingannya agar memudahkan pengguna. Kemudian, ringkasan dari model terbaik serta peramalan ke depannya. Berikut adalah tampilan dari menu hasil peramalan pada Gambar 10.



(a) (b)

Gambar 9 (a). Menu *sidebar* (b) Menu *dataset*.



Gambar 10. Menu hasil peramalan.

©2024 Ilmu Komputer Unila Publishing Network all rights reserved

4. KESIMPULAN

Berdasarkan penelitian yang dilakukan menggunakan *dataset* tunggakan peserta BPJS Kesehatan Kantor Cabang Jambi dengan membandingkan metode peramalan menggunakan ARIMA (3,1,3) dan RNN 1 *hidden layer* dengan 1 *neuron*, diperoleh hasil bahwa metode RNN lebih baik dalam melakukan peramalan dibandingkan ARIMA (3,1,3). Metode RNN memiliki MAPE sebesar 1,97% dan metode ARIMA (3,1,3) sebesar 2,62%, yang berarti bahwa kemampuan model peramalan sangat baik. Perbandingan RMSE RNN sebesar 2739803819 dan ARIMA (3,1,3) sebesar 3238308628. Pada penelitian ini dirancang pula antarmuka berbasis web agar memudahkan melakukan proses peramalan.

DAFTAR PUSTAKA

[1] (2020) Sejarah Perjalanan Jaminan Sosial di Indonesia. [Online]. Available: https://bpjs-kesehatan.go.id/bpjs (accessed Jun. 09, 2021).

[2] S. Syamsinar, A. S. Batara, & A. R. Amelia, "Gambaran Kemampuan Membayar Iuran BPJS Kesehatan pada Pedagang di Pasar Terong Kota Makassar," *Wind. Public Heal. J.*, vol. 1, no. 6, pp. 797–807, 2021, doi: 10.33096/woph.v1i6.333.

[3] N. Latifah, Y. Riza, & H. K. Anam, "Hubungan Pendapatan dan Pengetahuan Tentang Kewajiban Membayar Iuran dengan Kepatuhan Membayar Iuran Peserta Bpjs Non PBI di BPJS Kesehatan Kantor Kabupaten Banjar Tahun 2020," Doctoral dissertation, Universitas Islam Kalimantan, Indonesia, 2020.

[4] E. Maharani. (2016) Tunggakan Iuran BPJS di Jambi Capai Rp 11 Miliar. [Online]. Available: https://news.republika.co.id/berita/o7es4x335/tunggakan-iuran-bpjs-di-jambi-capai-rp-11-miliar (accessed Aug. 25, 2021).

[5] R. Annisa, S. Winda, E. Dwisaputro, & K. N. Isnaini, "Mengatasi Defisit Dana Jaminan Sosial Kesehatan Melalui Perbaikan Tata Kelola," *INTEGRITAS J. Antikorups*, vol. 6, no. 2, pp. 209–224, 2020, doi: 10.32697/integritas.v6i2.664.

[6] D. Wijayanti & R. Pulungan, "Peramalan Deret Waktu Menggunakan Model Fungsi Basis Radial (Rbf) Dan Auto Regressive Integrated Moving Average (Arima)," *J. MIPA Unnes*, vol. 35, no. 2, p. 114402, 2012.

[7] J. Liu, L. Ding, X. Guan, J. Gui, & J. Xu, "Comparative analysis of forecasting for air cargo volume: Statistical techniques vs. machine learning," *J. Data, Inf. Manag.*, vol. 2, no. 4, pp. 243–255, 2020, doi: 10.1007/s42488-020-00031-1.

[8] Suhartono & A. Endharta, "Suhartono, Peramalan Konsumsi Listrik Jangka Pendek dengan Arima Musiman Ganda dan Elman-Recurrent Neural Network," *Jurnal Ilmiah Teknologi Informasi*, 7(4), 183-190, 2009.

[9] S. Makridakis, E. Spiliotis, & V. Assimakopoulos, "Statistical and Machine Learning forecasting methods: Concerns and ways forward," *PLoS One*, vol. 13, no. 3, 2018, doi: 10.1371/journal.pone.0194889.

[10] H. T. Handoko, *Dasar-Dasar Manajemen Produksi dan Operasional*, Yogyakarta: BPFE, 2000.

[11] H. Diana & C. D. Raharjo, "Sistem Pendukung Keputusan Untuk Forecasting Penjualan di Toko Sumber Saudara," *Prosiding SNATIF*, pp. 275–280, 2015.

[12] J. Christy, A. P. Hintarsyah, & H. L. H. S. Warnars, "Forecasting Sebagai Decision Support Systems Aplikasi dan Penerapannya Untuk Mendukung Proses Pengambilan Keputusan.," *J. Sist. Komput.*, vol. 8, no. 1, pp. 19–27, 2018.

[13] D. C. Montgomery & L. A. Johnson, *Forecasting and Time Series Analysis*, New York: McGraw-Hill, 1998.

[14] R. A. Pitaloka, S. Sugito, & R. Rahmawati, "Perbandingan Metode Arima Box-Jenkins Dengan Arima

©2024 Ilmu Komputer Unila Publishing Network all rights reserved

Ensemble Pada Peramalan Nilai Impor Provinsi Jawa Tengah," *J. Gaussian*, vol. 8, no. 2, pp. 194–207, 2019, doi: 10.14710/j.gauss.v8i2.26648.

[15] D. E. Tarkus, S. R. U. A. Sompie, & A. Jacobus, "Implementasi Metode Recurrent Neural Network pada Pengklasifikasian Kualitas Telur Puyuh," *J. Tek. Inform.*, vol. 15, no. 2, pp. 137–144, 2020.

[16] A. M. Siregar, J. H. Jaman, & A. Mufti, "Analisa Prediksi Kesehatan Masyarakat Indonesia Menggunakan Recurrent Neural Network," *INTERNAL (Information System Journal)*, vol. 4, no. 1, pp. 28–34, 2021.

[17] A. N. Habowski, T. J. Habowski, & M. L. Waterman, "GECO: gene expression clustering optimization app for non-linear data visualization of patterns," *BMC Bioinformatics*, vol. 22, no. 1, 2021, doi: 10.1186/s12859-020-03951-2.

[18] M. C. Hidayatullah & A. Fathan, (2021). Membuat Aplikasi Web Sains Data dengan Mudah Menggunakan Streamlit. [Online]. Available: https://informatics.uii.ac.id/2021/03/15/streamlit-membuat-aplikasi-web-sains-data/.

[19] I. Eldiyana, E. Nurlaelah, & N. Herrhyanto, "Estimasi Missing Data dengan Metode Multivariate Imputation by Chained Equations (MICE) Untuk Membentuk Persamaan Regresi Linear Berganda," *J. EurekaMatika*, vol. 8, no. 1, 2020.

[20] W. Musu & A. Ibrahim, "Pengaruh Komposisi Data Training dan Testing terhadap Akurasi Algoritma C4.5," *SISITI: Seminar Ilmiah Sistem Informasi dan Teknologi Informasi*, Vol. 10, No. 1, pp. 186-195, 2021.

[21] M. A. Maricar, "Analisa Perbandingan Nilai Akurasi Moving Average Dan Exponential Smoothing Untuk Sistem Peramalan Pendapatan Pada Perusahaan XYZ," *J. Sist. dan Inform.*, vol. 13, no. 2, 2019.

[22] M. Hudzaifah, A. A. Rismayadi, & S. Topiq, "Implementasi Metode Time Series Decompostion Untuk Peramalan Lalu Lintas Jangka Pendek Berdasarkan Waktu Tempuh dan Cuaca," *J. Inform.*, vol. 6, no. 1, pp. 1–8, 2019.

[23] I. Mardiyah, W. D. Utami, D. C. R. Novitasari, M. Hafiyusholeh, & D. Sulistiyawati, "Analisis Prediksi Jumlah Penduduk di Kota Pasuruan Menggunakan Metode Arima," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 15, no. 3, pp. 525–534, Sep. 2021.

[24] A. A. Karia, I. Bujang, & I. Ahmad, "Fractionally integrated ARMA for crude palm oil prices prediction: Case of potentially overdifference," *J. Appl. Stat.*, vol. 40, no. 12, 2013.

[25] I. P. P. Wikantara, "Perbandingan Model Arima Dengan Model Recurrent Artificial Neural Network Dalam Memprediksi Jumlah Wisatawan Asing Yang Datang Ke Provinsi Bali," Undergraduate Thesis, Institut Teknologi Sepuluh November, Surabaya, Indonesia, 2016.

[26] (2021) Direktorat Analisis dan Pengembangan Statistik BPS. ARIMA (Autoregressive Integrated Moving Average). [Online]. Available: https://daps.bps.go.id/file_artikel/77/arima.pdf.