

FORECASTING TOTAL PEMAKAIAN TENAGA LISTRIK MENGGUNAKAN MODEL HYBRID ARIMA-SVM

PEMODELAN DAN ANALITIKA PREDIKTIF



DOSEN PENGAMPU :

Dr. Wiwik Anggraeni, S.Si., M.Kom.

KELOMPOK 2 FINAL PROJECT - PAP Kelas A:

- | | | |
|----|-----------------------|--------------|
| 1. | Evanriza Safiq Ariadi | (5026211139) |
| 2. | Muhammad Rhakan | (5026211166) |
| 3. | Dewi Maharani | (5026221046) |

**DEPARTEMEN SISTEM INFORMASI
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2024**

DAFTAR ISI

DAFTAR ISI.....	1
KATA PENGANTAR.....	2
BAB I	
PENDAHULUAN.....	3
1.1. Latar Belakang.....	3
1.2. Rumusan Masalah.....	4
1.3 Tujuan dan Manfaat.....	4
BAB II	
METODOLOGI PENELITIAN.....	6
2.1 Data.....	6
2.2 Model Hybrid ARIMA-SVM.....	6
2.3 Tools untuk Mengelola Data.....	7
BAB III	
PEMBAHASAN.....	14
BAB IV.....	
PENUTUP.....	18
4.1. Simpulan.....	18
4.2. Saran.....	18
DAFTAR PUSTAKA.....	19

KATA PENGANTAR

Puji syukur kami panjatkan kepada Tuhan Yang Maha Esa atas berkat dan rahmat-Nya, sehingga kami dapat menyelesaikan laporan ini. Laporan ini berjudul "Forecasting Total Pemakaian Tenaga Listrik Menggunakan Model Arima-SVM ". Laporan penelitian ini disusun berdasarkan hasil pengolahan data yang telah dilakukan oleh penulis. Penelitian ini menggunakan data terkait jumlah pelanggan dan pemakaian listrik rumah tangga yang berada pada Kota Jember, Gresik dan Pamekasan. Penulis menyadari bahwa laporan forecasting ini masih memiliki kekurangan. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun dari pembaca. Demikian laporan ini, semoga bermanfaat.

Surabaya, 28 Juni 2024

BAB I

PENDAHULUAN

1.1. Latar Belakang

Listrik adalah salah satu kebutuhan pokok dalam kehidupan modern. Penggunaan listrik yang terus meningkat seiring dengan pertumbuhan penduduk dan perkembangan teknologi memerlukan manajemen yang baik agar distribusi dan pemanfaatannya dapat berjalan efisien dan berkelanjutan. Pemantauan dan peramalan pemakaian tenaga listrik menjadi sangat penting untuk memastikan ketersediaan listrik yang memadai dan untuk mendukung perencanaan operasional serta pengambilan keputusan strategis oleh perusahaan listrik.

Kota Jember, Gresik, dan Pamekasan merupakan daerah yang mengalami pertumbuhan penduduk dan ekonomi yang cukup pesat di Jawa Timur. Hal ini berdampak langsung pada peningkatan jumlah pelanggan listrik dan pemakaian tenaga listrik rumah tangga. Untuk itu, diperlukan suatu model yang dapat memprediksi total pemakaian tenaga listrik di masa mendatang dengan akurasi tinggi, sehingga perusahaan listrik dapat mengantisipasi kebutuhan dan meningkatkan kualitas layanan. Model ARIMA (AutoRegressive Integrated Moving Average) telah lama digunakan dalam peramalan deret waktu dan terbukti efektif dalam memprediksi data yang bersifat linier. Namun, pemakaian tenaga listrik seringkali dipengaruhi oleh banyak faktor yang bersifat non-linier, seperti perubahan cuaca, kebijakan tarif listrik, dan perubahan perilaku konsumen. Oleh karena itu, diperlukan pendekatan yang dapat menangkap pola-pola non-linier tersebut.

Salah satu pendekatan yang dapat digunakan adalah SVM (Support Vector Machine). SVM merupakan metode machine learning yang efektif dalam menangani data non-linier dengan menggunakan kernel trick. Kombinasi antara ARIMA dan SVM diharapkan dapat meningkatkan akurasi prediksi dengan memanfaatkan kelebihan masing-masing metode, yaitu kemampuan ARIMA dalam menangkap pola linier dan kemampuan SVM dalam menangkap pola non-linier. Dalam penelitian ini, data yang digunakan mencakup pemakaian listrik rumah tangga dan jumlah pelanggan di Kota Jember, Gresik, dan Pamekasan. Data tersebut akan dianalisis dan diprediksi menggunakan model ARIMA-SVM. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi nyata dalam bidang peramalan pemakaian listrik, serta membantu perusahaan listrik dalam mengoptimalkan distribusi dan perencanaan operasional.

1.2. Rumusan Masalah

Dalam upaya mencapai tujuan pokok penelitian ini, kami mengajukan beberapa rumusan masalah yang berkaitan dengan *forecasting* total pemakaian tenaga listrik menggunakan model ARIMA-SVM. Berikut adalah rumusan masalah yang menjadi landasan utama dalam eksplorasi dan analisis kami:

1. Bagaimanakah ramalan total pemakaian tenaga listrik untuk Kota Jember, Gresik, dan Pamekasan untuk 5 tahun kedepan?
2. Seberapa baik performa model ARIMA-SVM dalam meramalkan pemakaian tenaga listrik berdasarkan 3 skenario pembagian dataset (80% train - 20% test, 70% train - 30% test, 60% train - 40% test)?

1.3 Tujuan dan Manfaat

Seiring dengan rumusan masalah yang telah diajukan, penelitian ini bertujuan untuk mencapai serangkaian tujuan untuk menghasilkan sebuah peramalan total pemakaian tenaga listrik. Berikut adalah tujuan yang ingin dicapai dalam penelitian ini:

1. Mengembangkan model peramalan untuk pemakaian tenaga listrik rumah tangga di Jember, Gresik, dan Pamekasan menggunakan metode ARIMA-SVM.
2. Mengevaluasi performa model ARIMA-SVM dalam meramalkan pemakaian tenaga listrik berdasarkan tiga skenario pembagian dataset (80% train - 20% test, 70% train - 30% test, 60% train - 40% test).
3. Membuat proyeksi pemakaian tenaga listrik rumah tangga di Jember, Gresik, dan Pamekasan untuk lima tahun ke depan.
4. Mengidentifikasi metode evaluasi yang paling tepat untuk mengukur akurasi peramalan model.

Selain tujuan yang telah dijabarkan, penelitian ini juga diharapkan dapat memberikan manfaat baik secara akademis maupun praktis dan memberikan kontribusi positif terhadap pemakaian tenaga listrik di Kota Jember, Gresik, dan Pamekasan. Adapun manfaat dari penelitian ini melibatkan beberapa aspek, antara lain:

1. Memberikan wawasan dan pengetahuan baru mengenai penerapan metode ARIMA-SVM dalam peramalan pemakaian tenaga listrik, serta evaluasi performa model berdasarkan skenario pembagian dataset yang berbeda.
2. Membantu dalam penyusunan kebijakan dan strategi pengelolaan energi listrik yang lebih baik berdasarkan proyeksi pemakaian tenaga listrik yang akurat.
3. Menjamin ketersediaan pasokan listrik yang lebih stabil dan efisien, serta mendukung kegiatan sehari-hari dan pertumbuhan ekonomi masyarakat.
4. Menambah literatur dan referensi dalam bidang peramalan pemakaian energi listrik, serta membuka peluang penelitian lebih lanjut mengenai kombinasi metode statistik dan machine learning dalam peramalan deret waktu.

BAB II

METODOLOGI PENELITIAN

2.1 Data

Penelitian ini menggunakan data yang mencatat jumlah total pemakaian tenaga listrik rumah tangga di daerah Jember, Gresik, dan Pamekasan untuk periode waktu tertentu. Data yang digunakan meliputi konsumsi listrik dari tanggal 1 hingga 12 Januari untuk setiap tahun dari 2004 hingga 2022, dan tanggal 1 hingga 6 Januari untuk tahun 2023, yang kami dapatkan melalui dosen mata kuliah Pemodelan dan Analitika Prediktif yaitu Dr. Wiwik Anggraeni, S.Si., M.Kom.

2.2 Model Hybrid ARIMA-SVM

Metodologi ARIMA, yang dikembangkan oleh Box dan Jenkins pada tahun 1976, adalah pendekatan yang digunakan untuk mengembangkan dan merancang model ARIMA (Autoregressive Integrated Moving Average) untuk peramalan deret waktu. Model ARIMA menangkap hubungan berurutan dalam data deret waktu melalui tiga komponen utama: autoregressive (AR), integrated (I), dan moving average (MA). Komponen AR (autoregressive) menggambarkan hubungan linier antara nilai saat ini dan nilai lag sebelumnya dalam deret waktu. Komponen I (integrated) menangani non-stasioneritas dalam data dengan melakukan differencing untuk mencapai stasioneritas, yaitu keadaan di mana rata-rata, varians, dan autokorelasi tetap konstan dari waktu ke waktu. Komponen MA (moving average) menggambarkan ketergantungan variabel dependen pada kesalahan sebelumnya. Model ARIMA biasanya digunakan untuk data stasioner, tetapi dapat diterapkan pada data non-stasioner setelah transformasi yang sesuai. Persamaan model ARIMA menggabungkan komponen AR dan MA untuk menghasilkan peramalan yang akurat dengan menguraikan data historis menjadi proses AR yang mempertahankan kejadian masa lalu, proses I yang menstasionerkan data, dan proses MA yang menangkap kesalahan.

Metodologi SVM (Support Vector Machine) adalah teknik pembelajaran mesin yang digunakan untuk klasifikasi dan regresi, yang dikembangkan berdasarkan teori pembelajaran statistik oleh Vapnik dan koleganya pada tahun 1990-an. SVM bekerja dengan menemukan

hyperplane optimal yang memisahkan data dalam ruang fitur dengan margin maksimum. Dalam konteks regresi, dikenal sebagai Support Vector Regression (SVR), SVM mencoba untuk menemukan fungsi yang paling sesuai dengan data sambil meminimalkan kesalahan dalam batas margin tertentu. SVM menggunakan kernel trik untuk menangani data yang tidak dapat dipisahkan secara linier dengan memetakan data ke dimensi yang lebih tinggi, memungkinkan pemisahan yang lebih mudah. Beberapa kernel yang umum digunakan termasuk linear, polynomial, radial basis function (RBF), dan sigmoid. Metodologi SVM memiliki keunggulan dalam menangani data berdimensi tinggi dan memiliki performa yang baik pada masalah non-linear, menjadikannya alat yang kuat dalam berbagai aplikasi seperti pengenalan pola, bioinformatika, dan peramalan deret waktu.

Metodologi hybrid adalah pendekatan yang menggabungkan dua atau lebih model peramalan untuk memanfaatkan kekuatan masing-masing model dalam menangkap karakteristik data yang kompleks. Dalam konteks peramalan deret waktu, metodologi hybrid yang umum digunakan adalah kombinasi dari model ARIMA (Autoregressive Integrated Moving Average) dan SVM (Support Vector Machine). Model ARIMA efektif dalam menangkap pola linier dalam data, seperti trend dan musiman, dengan menggunakan komponen autoregressive (AR), differencing (I), dan moving average (MA). Di sisi lain, SVM unggul dalam menangani pola non-linier berkat kemampuannya menggunakan kernel trik untuk memetakan data ke dimensi yang lebih tinggi. Dalam model hybrid ARIMA-SVM, pertama-tama, model ARIMA digunakan untuk menangkap dan menghilangkan komponen linier dari data. Residual atau kesalahan yang dihasilkan dari model ARIMA kemudian digunakan sebagai input untuk model SVM, yang bertugas menangani komponen non-linier yang tersisa. Pendekatan ini memungkinkan model hybrid untuk memberikan peramalan yang lebih akurat dan andal dibandingkan dengan menggunakan model ARIMA atau SVM secara terpisah, karena menggabungkan keunggulan dari kedua model dalam menangani berbagai karakteristik data deret waktu.

2.3 Tools untuk Mengelola Data

Dalam penelitian ini, pengelolaan data pemakaian tenaga listrik rumah tangga di Jember, Gresik, dan Pamekasan menggunakan dua tools utama, yaitu Microsoft Excel dan Google Colab dengan Python. Microsoft Excel digunakan untuk tahap awal preprocessing data, termasuk

pembersihan data dan analisis deskriptif sederhana seperti perhitungan rata-rata dan visualisasi awal menggunakan grafik. Google Colab, dengan dukungan Python, menjadi platform utama untuk analisis lebih mendalam. Di sini, data yang lebih kompleks dapat diolah menggunakan library seperti pandas dan numpy untuk manipulasi data serta scikit-learn untuk pemodelan statistik, termasuk implementasi dan evaluasi model ARIMA dan SVM untuk peramalan. Google Collab juga memfasilitasi visualisasi hasil peramalan dan analisis menggunakan library seperti matplotlib dan seaborn. Penggunaan kombinasi kedua tools ini memungkinkan peneliti untuk melakukan pengolahan data yang komprehensif, analisis yang mendalam, dan pengembangan model peramalan yang efektif sesuai dengan tujuan penelitian ini.

2.4. Langkah-langkah Model Hybrid ARIMA-SVM

Untuk melakukan *forecasting* total pemakaian tenaga listrik dengan menggunakan model hybrid ARIMA-SVM yang diimplementasikan dalam bahasa pemrograman Python, *order* dari model ARIMA perlu diketahui terlebih dahulu sebelum melakukan persilangan dengan model SVM. Maka dari itu, langkah awal yang harus dilakukan adalah melakukan import library yang relevan untuk tahap *forecasting*. Kemudian, data diimpor ke *google collab* dan diberlakukan uji ADF pada ketiga kolom yang terdapat pada data *time-series* yang kita miliki. Sebelum melakukan uji ADF, tiga kolom *time-series* tersebut perlu diketahui apabila nilai p (p-value) dari masing-masing kolom tersebut kurang dari tingkat signifikansi atau tidak. Apabila nilai p ketiga kolom tersebut kurang dari tingkat signifikansi yaitu 0.05, maka dapat dikatakan bahwa data yang dimiliki bersifat stasioner. Berikut merupakan nilai p yang dimiliki oleh masing-masing kolom sebelum diberlakukan *differencing*:

No	Kolom	P-value
1	JBR	0.7152406760495255
2	GSK	0.9964851902621912
3	PMK	0.08448770006630785

Tabel 2.4.1. P-value awal data

Berdasarkan tabel 2.4.1, data dapat dinyatakan tidak stasioner karena nilai p yang dimiliki belum kurang dari tingkat signifikansi. Maka dari itu, *differencing* akan dilakukan terhadap data dengan menggunakan fungsi 'adfuller'. *Differencing* pertama menghasilkan hasil yang berbeda untuk masing-masing kolom pada data. Berikut merupakan hasil nilai p pada data setelah melalui *differencing pertama*:

No	Kolom	P-value
1	JBR	0.0001465280170014774
2	GSK	0.9862691492943414
3	PMK	0.027769482976931468

Tabel 2.4.2. P-value setelah *differencing pertama*

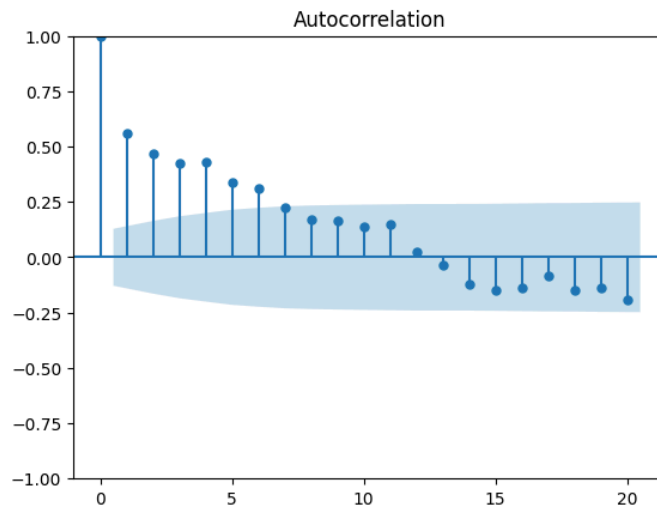
Dapat dilihat dari tabel 2.4.2 bahwa kolom JBR dan PMK telah memiliki nilai p kurang dari tingkat signifikansi. Hal tersebut menunjukkan bahwa data pada kolom JBR dan PMK bersifat stasioner setelah *differencing pertama*. Maka dari itu, model ARIMA yang akan diimplementasikan untuk kolom JBR dan PMK akan memiliki nilai $d = 1$. Di sisi lain, data yang terdapat pada kolom GSK belum bersifat stasioner karena masih memiliki nilai p yang belum kurang dari tingkat signifikansi. Maka dari itu, *differencing* lanjutan perlu dilakukan untuk kolom GSK. Berikut merupakan hasil *differencing* kedua untuk kolom GSK:

No	Kolom	P-value
1	GSK	8.583638769404253e-23

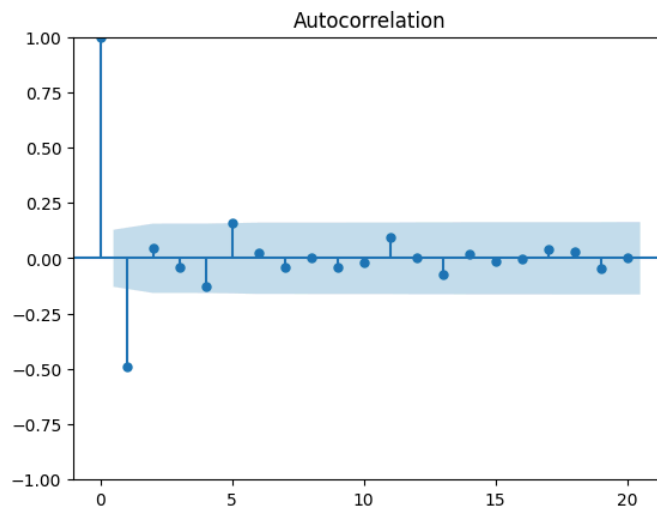
Tabel 2.4.2. P-value setelah *differencing kedua*

Setelah melalui *differencing* kedua, dapat dilihat bahwa data pada kolom GSK memiliki nilai p kurang dari tingkat signifikansi. Maka dengan itu, semua kolom data dapat dinyatakan stasioner. Setelah menentukan stasioneritas data, maka fungsi autokorelasi (ACF) dan fungsi autokorelasi parsial (PACF) perlu diketahui untuk menentukan nilai q dan p pada *order* ARIMA secara berurutan. Kedua fungsi tersebut dapat diketahui melalui data dari masing-masing kolom yang di plot di dalam sebuah grafik. Untuk plot kedua fungsi tersebut, fungsi "dropna()" juga

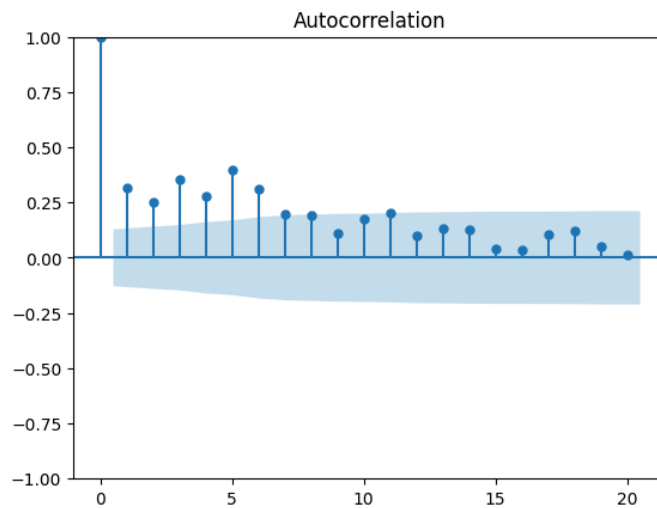
perlu dilakukan untuk menghilangkan null yang dimiliki oleh data hasil *differencing*. Berikut merupakan grafik plot untuk fungsi autokorelasi;



Gambar 2.4.1. ACF Data JBR

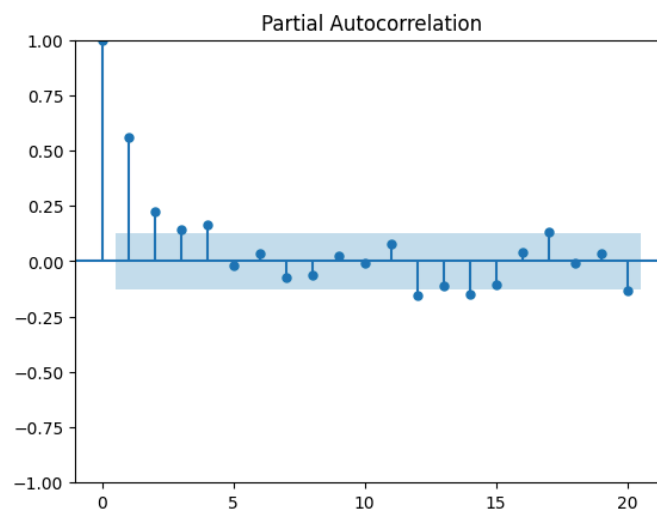


Gambar 2.4.2. ACF Data GSK

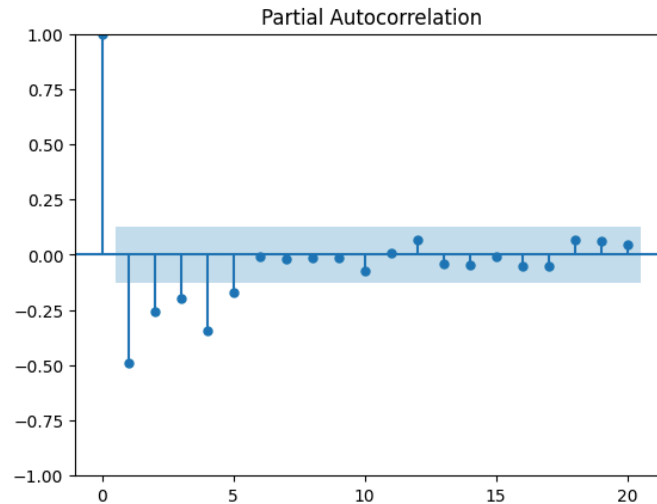


Gambar 2.4.3. ACF Data PMK

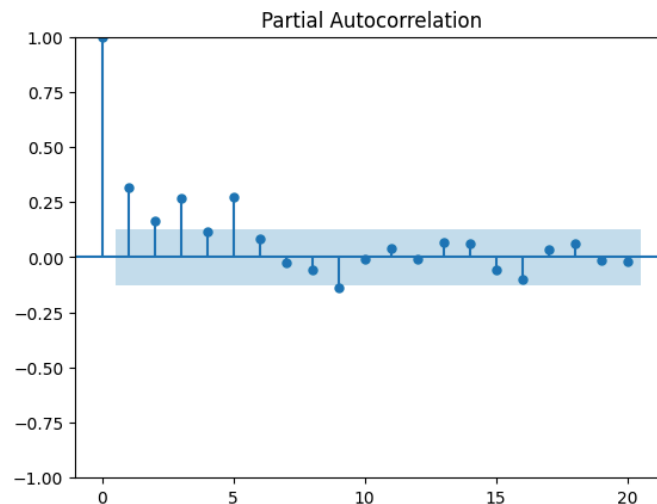
Gambar-gambar diatas menunjukkan fungsi autokorelasi yang dimiliki oleh data dengan tingkat *differencing* yang paling terakhir. Kolom JBR memiliki nilai tertinggi pada *lag* 1. Maka dari itu, q pada *order* ARIMA untuk kolom JBR akan memiliki nilai 1. Lalu, kolom PMK memiliki nilai tertinggi pada *lag* 5. Maka dari itu, q pada *order* ARIMA untuk kolom PMK akan memiliki nilai 5. Terakhir, kolom GSK memiliki nilai tertinggi pada *lag* 1. Maka dari itu, q pada *order* ARIMA untuk kolom GSK akan memiliki nilai 1. Setelah fungsi autokorelasi, proses akan dilanjutkan untuk fungsi autokorelasi parsial agar nilai p pada *order* ARIMA untuk masing-masing kolom data dapat ditentukan. Berikut merupakan grafik plot untuk fungsi autokorelasi parsial:



Gambar 2.4.4. PACF Data JBR



Gambar 2.4.5. PACF Data GSK



Gambar 2.4.6. PACF Data PMK

Berdasarkan grafik plot fungsi autokorelasi parsial, kolom JBR memiliki nilai tertinggi pada *lag* 1. Maka dari itu, p pada *order* ARIMA untuk kolom JBR akan memiliki nilai 1. Lalu, kolom PMK memiliki nilai tertinggi pada *lag* 1. Maka dari itu, p pada *order* ARIMA untuk kolom PMK akan memiliki nilai 1. Terakhir, kolom GSK memiliki nilai tertinggi pada *lag* 1. Maka dari itu, p pada *order* ARIMA untuk kolom GSK akan memiliki nilai 1. Setelah *differencing*, plot fungsi autokorelasi, dan plot fungsi autokorelasi parsial telah dilakukan, maka

model ARIMA untuk masing masing kolom dapat ditentukan. Berikut merupakan model ARIMA yang akan dipakai untuk model *hybrid* pada masing-masing kolom:

No	Kolom	Model
1	JBR	ARIMA(1,1,1)
2	GSK	ARIMA(1,2,1)
3	PMK	ARIMA(1,1,5)

Tabel 2.4.2. Model ARIMA Data

Setelah model ARIMA telah ditentukan untuk masing-masing kolom data, maka hal selanjutnya yang perlu dilakukan adalah membagi data menjadi beberapa skenario komposisi *train* dan *test*. Pada *forecasting* ini, tiga skenario komposisi telah ditentukan; 80% *train*-20% *test*, 70% *train*-30% *test*, dan 60% *train*-40% *test*. Setelah data dibagi sesuai dengan tiga skenario komposisi tersebut, maka model ARIMA dapat mulai dilatih dan dilakukan *fitting*. Setelah model ARIMA dilatih dan dilakukan *fitting*, *residual* model ARIMA akan digunakan sebagai sebuah *dataframe* untuk model SVM untuk menangkap pola non-linear yang ada dalam data. Lalu, hasil peramalan dari model ARIMA dan model SVM akan digabungkan sebagai sebuah hasil untuk model *hybrid* yang dikembangkan. Setelah hasil peramalan dari model *hybrid* telah diperhitungkan, maka peramalan untuk total pemakaian tenaga listrik ketiga kota 5 tahun yang akan datang dapat dimiliki. Tidak hanya itu, dengan semua model telah diperhitungkan, maka evaluasi terhadap masing-masing model dapat diperhitungkan juga untuk mengetahui performa dari masing-masing model peramalan.

BAB III PEMBAHASAN

Penelitian ini dilakukan dengan menggunakan tiga skenario yang berbeda pada proporsi data pelatihan (*training set*) dan data pengujian (*testing set*). Pada setiap skenario, data pelatihan digunakan untuk membangun model ARIMA yang kemudian di-*fit* terhadap data. Hasil fitting model ARIMA ini menghasilkan residu yang tidak dapat dijelaskan oleh komponen linear, sehingga model SVM digunakan untuk menangkap pola non-linear yang ada dalam data. Adapun hasil evaluasi model ARIMA dan SVM pada setiap skenario diuraikan sebagai berikut.

	ARIMA			SVM			Hybrid		
	0,6	0,7	0,8	0,6	0,7	0,8	0,6	0,7	0,8
MSE	2,3x10 ⁸	1,3x10 ⁸	4,7x10 ⁷	9,6x10 ¹¹	9,7x10 ¹¹	9,7x10 ¹¹	2,3x10 ⁸	1,3x10 ⁸	4,5x10 ⁷
MAE	13048,3	9941,4	5981,3	983859,4	984920,4	985661,2	12881,0	9751,6	5797,6
MAPE	1,324	1,008	0,606	100,01	100,01	100,01	1,307	0,989	0,588

Tabel 3.1. Evaluation Metrics Data JBR

	ARIMA			SVM			Hybrid		
	0,6	0,7	0,8	0,6	0,7	0,8	0,6	0,7	0,8
MSE	6,1x10 ¹⁰	2,6x10 ⁹	7,0x10 ⁸	2,9x10 ¹¹	3,5x10 ¹¹	4,2x10 ¹¹	6,1x10 ¹⁰	3,5x10 ₉	7,0x10 ⁸
MAE	205798,5	37016,4	19482,4	527861,9	581218,4	645306,2	205810,6	37035,2	19513,1
MAPE	34,718	5,48	2,722	100,002	100,003	100,01	34,720	5,484	2,727

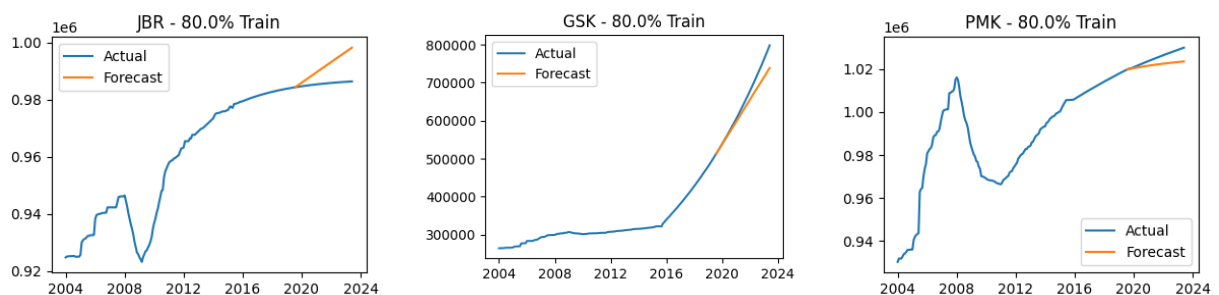
Tabel 3.2. Evaluation Metrics Data GSK

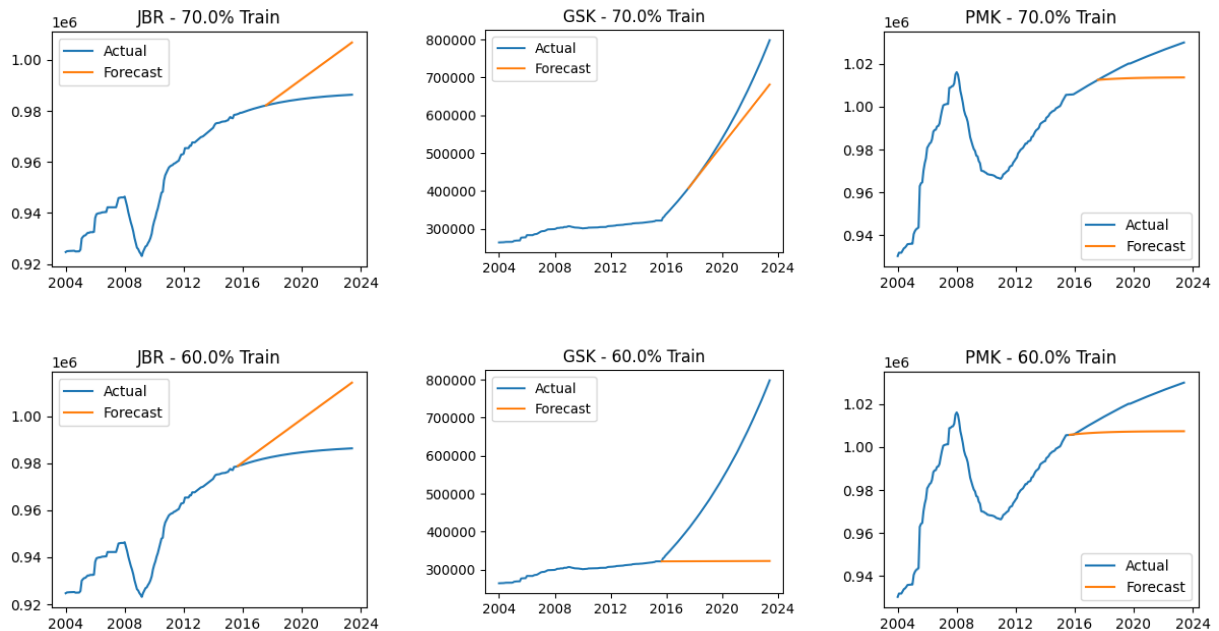
	ARIMA			SVM			Hybrid		
	0,6	0,7	0,8	0,6	0,7	0,8	0,6	0,7	0,8
MSE	$1,8 \times 10^{88}$	$1,0 \times 10^8$	$1,4 \times 10^7$	$1,0 \times 10^{12}$	$1,0 \times 10^{12}$	$1,0 \times 10^{12}$	$1,8 \times 10^8$	$9,7 \times 10^7$	$1,3 \times 10^7$
MAE	11784,9	8908,4	3222,6	$1,0 \times 10^6$	$1,0 \times 10^6$	$1,0 \times 10^6$	11772,0	8671,9	3070,2
MAPE	1,152	0,869	0,314	99,998	99,977	99,985	1,151	0,846	0,299

Tabel 3.3. Evaluation Metrics Data PMK

Tabel 3.1.1., 3.1.2., dan 3.1.3. menunjukkan hasil evaluasi kinerja model ARIMA, SVM, dan Hybrid menggunakan metrik MSE, MAE, dan MAPE. Nilai MSE dan MAE pada perhitungan *forecasting* setiap kota menunjukkan skala yang cukup tinggi. Akan tetapi, hal ini tidak secara absolut menyimpulkan bahwa model yang digunakan memiliki kinerja yang buruk. Hal ini dikarenakan skala data aktual berpengaruh terhadap nilai MSE dan MAE, sehingga data aktual yang bernilai besar akan cenderung menghasilkan perhitungan error dengan nilai yang besar pula. Oleh karena itu, nilai MAPE dapat dijadikan sebagai standar pengukuran kinerja model. Perhitungan MAPE dilakukan dengan membagi nilai error dengan data aktualnya, sehingga setiap data bernilai independen dan dapat dijadikan sebagai standar dari hasil pengukuran kinerja.

Setelah melakukan evaluasi kinerja pada data pelatihan, kita dapat melihat bagaimana model memprediksi data aktual dengan melakukan visualisasi hasil forecasting data pengujian.

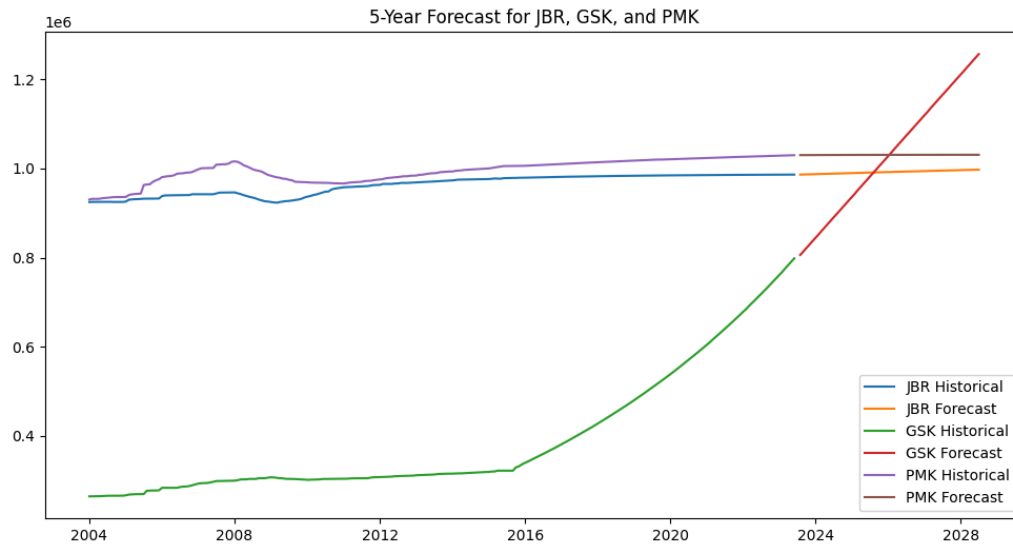




Gambar 3.1. Perbandingan data forecasting dengan data aktual pada tiga skenario pengujian

Melalui Gambar 3.1, skenario data pengujian 80% menghasilkan visualisasi perbandingan yang paling baik. Sebaliknya, data pengujian 60% menunjukkan hasil *forecasting* dengan margin error yang paling besar. Dalam hal ini, model terbaik didapatkan dengan menggunakan 80% dari sebagai *training set* dan 20% sebagai *testing set*.

Setelah melalui proses pembelajaran untuk mengetahui pola dan karakteristik dari dataset, model *Hybrid* digunakan untuk melakukan *forecasting* selama lima tahun ke depan yang dilakukan dengan 60 *steps* (menandakan periode 60 bulan ke depan). Model *forecasting* tersebut merupakan gabungan dari prediksi data aktual dengan disertai residu data tersebut. Hasil *forecasting* dengan menggunakan model *Hybrid* selama lima tahun ke depan dapat dilihat melalui gambar 3.2.



Gambar 3.2. Visualisasi Hasil *Hybrid Forecasting* Lima Tahun

BAB IV

PENUTUP

4.1. Simpulan

Dari hasil analisis forecasting menggunakan model *hybrid* ARIMA-SVM kami mendapati bahwa dari tiga skenario pembagian dataset yang diuji (80% *train* - 20% *test*, 70% *train* - 30% *test*, dan 60% *train* - 40% *test*), skenario dengan komposisi 80% *training* dan 20% *testing* memberikan hasil yang optimal dengan *margin error* yang paling kecil. Selain itu hasil evaluasi kinerja model menunjukkan bahwa nilai MSE dan MAE cukup tinggi, tetapi MAPE memberikan standar pengukuran kinerja yang lebih independen dari skala data aktual. Sehingga dapat disimpulkan bahwa model *hybrid* ARIMA-SVM yang dikembangkan mampu memberikan peramalan yang relatif akurat untuk pemakaian tenaga listrik rumah tangga di Kota Jember, Gresik, dan Pamekasan, terutama dengan menggunakan proporsi data 80% untuk pelatihan dan 20% untuk pengujian. Model ini dapat diandalkan untuk perencanaan distribusi dan pengelolaan energi listrik yang lebih efektif.

4.2. Saran

Berdasarkan penelitian yang telah dilakukan, dapat disarankan untuk memperpanjang periode pengumpulan data agar hasil peramalan lebih representatif dan dapat diaplikasikan secara lebih luas. Selain itu mempertimbangkan penggunaan model lain seperti GRU dapat memberikan hasil yang lebih kompleks pada data time-series. Penelitian lanjutan juga dapat mengeksplorasi kombinasi model *hybrid* lainnya untuk meningkatkan kinerja model. Selain itu, pemilihan kota untuk peramalan dapat menggunakan kota-kota yang lebih padat oleh penduduk dikarenakan tentunya memiliki penggunaan total pemakaian tenaga listrik yang relatif lebih banyak. Tidak hanya itu, saran lain yang dapat diberikan adalah untuk para instansi pengelola tenaga listrik dapat mempertimbangkan untuk memanfaatkan model peramalan agar dapat memenuhi permintaan tenaga listrik yang akan datang.

DAFTAR PUSTAKA

Shah, D. N., & Thaker, M. (2024). Forecasting Stock Index Using A Hybrid ARIMA-SVM Model. *IOSR Journal of Economics and Finance (IOSR-JEF)*, 15(2), <https://doi.org/10.9790/5933-1502011521>