

ANALISIS PREDIKSI KONSUMSI ENERGI LISTRIK BERBASIS WAKTU DENGAN METODE ARIMA

LAPORAN TUGAS AKHIR

Laporan ini disusun untuk memenuhi salah satu syarat menyelesaikan pendidikan
Diploma Empat Program Studi Teknik Informatika di Jurusan Teknik Komputer
dan Informatika

Oleh:

Muhammad Ridwan Fathin

NIM: 151524022



POLITEKNIK NEGERI BANDUNG

2018

ANALISIS PREDIKSI KONSUMSI ENERGI LISTRIK BERBASIS WAKTU DENGAN METODE ARIMA

Oleh:

Muhammad Ridwan Fathin

NIM: 151524022

Tugas Akhir ini telah disidangkan pada tanggal ... Juli 2018 dan
disahkan sesuai ketentuan.

Pembimbing II,

Pembimbing I,

Nurjannah Syakrani, Dra., M.T., Dr.

NIP 19631213 199201 2 001

Penguji II,

Yudi Widhiyasana, S.Si., M.T.

NIP 19740718 200112 1 002

Penguji I,

Jurusan Teknik Komputer dan Informatika

Ketua Jurusan,

Ketua Program Studi D-IV,

Bambang Wisnuadhi, S.Si., M.T.

NIP 19720106 199303 1 002

Santi Sundari, S.Si., M.T.

NIP 19710903 199903 2 001

KATA PENGANTAR

Puji dan syukur panjatkan kehadiran Allah SWT yang telah memberikan rahmat serta karunia-Nya sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul **Analisis Prediksi Konsumsi Energi Listrik Berbasis Waktu dengan Metode ARIMA**. Sholawat serta salam penulis curah limpahkan kepada junjungan alam Nabi Muhammad SAW.

Tugas Akhir ini dibuat dan diajukan sebagai salah satu syarat dalam menyelesaikan program Diploma IV di Jurusan Teknik Komputer dan Informatika Politeknik Negeri Bandung. Tugas Akhir ini tidak akan dapat dikerjakan tanpa bantuan dan bimbingan dari berbagai pihak, baik secara moral maupun spiritual yang tidak terkira nilainya. Oleh karena itu dengan segenap rasa hormat dan ketulusan hati, penulis mengucapkan terima kasih kepada:

1. Bambang Wisnuadhi, S.Si., M.T., selaku Ketua Jurusan Teknik Komputer dan Informatika atas arahan dan dukungannya;
2. Santi Sundari, S.Si., M.T., selaku Ketua Program Studi D-IV Teknik Informatika atas arahan dan dukungannya;
3. Bapak Yudi Widhiyasana, S.Si., M.T. dan Ibu Nurjannah Syakrani, Dra., M.T., Dr. selaku dosen pembimbing yang telah membimbing dan memberikan motivasi selama penelitian Tugas Akhir;
4. Bapak Ade Chandra Nugraha, S.Si., M.T., Ibu Ani Rahmani, S.Si., M.T. dan Bapak Suprihanto, BSEE., M.Sc selaku Koordinator Tugas Akhir yang telah memberikan arahan dan motivasi dalam penyusunan Tugas Akhir;
5. Ibu Transmissia Semiawan, BSCS., MIT., Ph.D. selaku dosen yang memberikan pencerahan penelitian pada mata kuliah Metodologi Penelitian;

6. Keluarga yang selalu memberikan dukungannya yang besar khususnya selama penulis melaksanakan praktik kerja lapangan;
7. Rekan-rekan mahasiswa/i D-IV Teknik Informatika angkatan 2014 yang telah memberikan dukungan dalam menyelesaikan penelitian dan penulisan laporan Tugas Akhir;
8. Serta seluruh pihak yang telah memberikan do'a dan dukungan demi kelancaran dan kemudahan saya selama mengerjakan penelitian dan penulisan laporan Tugas Akhir.

Mengingat akan kemampuan penulis yang sangat terbatas, penulis mengharapkan kritik dan sarannya dari semua pihak demi peyempurnaan pembuatan laporan in.

Demikian laporan ini penulis susun, semoga dapat bermanfaat khususnya bagi penulis dan umumnya untuk semua pihak. Akhir kata penulis ucapkan terima kasih.

Bandung, Mei 2019

Muhammad Ridwan Fathin

DAFTAR ISI

KATA PENGANTAR	iii
DAFTAR ISI.....	v
DAFTAR GAMBAR	viii
DAFTAR TABEL.....	ix
BAB I PENDAHULUAN.....	1
I.1 Latar Belakang.....	1
I.2 Perumusan/Identifikasi Masalah.....	2
I.3 <i>Research Question</i>	3
I.4 Hipotesis	3
I.5 Tujuan	3
I.6 Ruang Lingkup	3
I.7 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA.....	5
II.1 Karya Ilmiah Sejenis Sebelumnya.....	5
II.2 Dasar Teori	7
II.2.1 Energi Listrik.....	7
II.2.2 Prediksi	7
II.2.3 Metode Prediksi.....	7
II.2.3.1 Regresi.....	9
II.2.3.2 <i>Time series</i>	9
II.2.3.3 Autoregressive Integrated Moving Average (ARIMA)	10
II.2.3.3.1 Tahapan ARIMA.....	11
II.2.3.3.2 Uji Stationer	11
II.2.3.3.3 Penentuan Nilai AR dan MA	13
II.2.3.3.4 Penentuan Model Terbaik	15
II.2.4 Evaluasi Metode Prediksi	16

BAB III METODOLOGI PENELITIAN.....	17
III.1 Jenis Penelitian.....	17
III.2 Variabel Penelitian.....	17
III.2.1 Variabel Bebas	17
III.2.2 Variabel Terikat	18
III.3 Objek Penelitian.....	18
III.4 Perangkat Pendukung.....	18
III.5 Tahapan Penelitian.....	18
III.5.1 Studi Pustaka.....	20
III.5.2 Pengumpulan Data	20
III.5.3 Eksplorasi.....	21
III.5.4 Penyiapan Data Latih dan Data Uji.....	21
III.5.5 Uji Stasioner.....	22
III.5.6 <i>Differencing</i>	22
III.5.7 Penentuan Dugaan Awal AR(p) dan MA(q).....	22
III.5.8 Penentuan Kombinasi Order ARIMA	23
III.5.9 Penentuan Order Model Terbaik.....	23
III.5.10 Pembuatan Model ARIMA	23
III.5.11 <i>Fitting</i> Model	23
III.5.12 Menyusun Rancangan Eksperimen	24
III.5.12.1 Tujuan Eksperimen	24
III.5.12.2 Bahan, Alat, Perlengkapan, dan Unit Eksperimen.....	24
III.5.12.2.1 Bahan Eksperimen	24
III.5.12.2.2 Alat Eksperimen.....	24
III.5.12.2.3 Perlengkapan Eksperimen.....	25
III.5.12.2.4 Unit Eksperimen.....	25
III.5.12.3 Metode Eksperimen	26
III.5.12.4 Skenario Eksperimen	26

III.5.13 Eksperimen.....	27
III.5.14 Evaluasi Hasil Eksperimen	27
BAB IV ANALISIS, PERANCANGAN, IMPLEMENTASI DAN PENGUJIAN ...	28
IV.1 Analisis	28
IV.1.1 Pembuatan Model Prediksi	28
IV.1.1.1 Data Latih.....	28
IV.1.1.2 Uji Stasioner.....	29
IV.1.1.3 Penentuan Dugaan Awal AR(p) dan MA(q).....	31
IV.1.1.4 Penentuan Kombinasi Order ARIMA	32
IV.1.1.5 Penentuan Order Model Terbaik	33
IV.1.2 Analisis Kebutuhan.....	34
IV.1.2.1 Analisis Kebutuhan Data.....	34
IV.1.2.2 Analisis Kebutuhan Aplikasi.....	34
IV.2 Perancangan	35
IV.3 Implementasi.....	35
IV.4 Pengujian	36
BAB V HASIL DAN PEMBAHASAN.....	37
BAB VI KESIMPULAN DAN SARAN	38
VI.1 Kesimpulan	38
VI.2 Saran	38
DAFTAR PUSTAKA	39

DAFTAR GAMBAR

Gambar II.1 Contoh plot grafik ACF dan PACF [8]	13
Gambar III.1 Variabel Penelitian	17
Gambar III.2 Metodologi Penelitian	19
Gambar IV.1 Plot grafik pada data latih per jam	29
Gambar IV.2 Plot grafik pada data latih setelah proses transformasi ke nilai logaritmanya dan proses <i>differencing</i>	30
Gambar IV.3 Grafik ACF data latih.....	31
Gambar IV.4 Grafik PACF data latih	32

DAFTAR TABEL

Tabel II.1 Hasil penelitian [3]	5
Tabel II.2 Hasil Penelitian [6]	6
Tabel II.3 Rentang waktu dalam prediksi	8
Tabel III.1 Contoh data pada dataset [9]	20
Tabel III.2 Contoh data pada data latih	21
Tabel III.3 Contoh kombinasi model ARIMA yang diduga	23
Tabel III.4 Alat eksperimen	24
Tabel III.5 Format catatan nilai RMSE hasil eksperimen	26
Tabel III.6 Skenario Eksperimen	27
Tabel IV.1 Contoh isi dari data latih	28
Tabel IV.2 Tabel perhitungan rata-rata dan variasi awal	30
Tabel IV.3 Tabel perhitungan rata-rata dan variasi setelah proses transformasi ke nilai logaritmanya dan proses <i>differencing</i>	31
Tabel IV.4 Kombinasi model ARIMA yang diduga	32
Tabel IV.5 Hasil AIC dari masing-masing kombinasi model ARIMA yang diduga ..	33
Tabel IV.6 Kebutuhan aplikasi eksperimen	34

BAB I

PENDAHULUAN

I.1 Latar Belakang

Perkembangan populasi manusia, pembangunan dan aplikasi teknologi menyebabkan peningkatan pesat dalam konsumsi energi listrik. Sehingga memprediksi konsumsi energi listrik diperlukan dalam pengambilan keputusan untuk manajemen energi listrik [1]. Untuk pemerintah negara berkembang, membuat prediksi konsumsi listrik jangka panjang sangat penting untuk meningkatkan efisiensi energi. Terlalu banyak penyediaan energi listrik akan menyebabkan kapasitas berlebih yang berarti pemborosan sumber daya, sedangkan penyediaan yang terlalu sedikit akan menyebabkan biaya operasi yang lebih tinggi untuk pemasok energi tambahan. Oleh karena itu, pemodelan prediksi konsumsi listrik dengan akurasi yang baik menjadi sangat penting untuk meminimalisir kerugian [2].

Terdapat berbagai cara untuk pemodelan prediksi diantaranya adalah model ekonometrik dan model *time series*. Model ekonometrik menganalisis korelasi antar variabel makroekonomi, dalam hal ini biasanya para peneliti menganalisis pendapatan perkapita, harga energi listrik dan populasi untuk mengestimasi nilai konsumsi energi listrik [3]. Artinya untuk menggunakan model ekonometrik membutuhkan data variabel makroekonomi lain juga pengetahuan terkait ekonomi untuk dapat menganalisis korelasi variabel tersebut dengan konsumsi energi listrik.

Model *time series* menganalisis nilai-nilai masa lalu dari variabel yang sama untuk mengestimasi nilai-nilai berikutnya [4]. Prediksi dilakukan dengan belajar dari nilai-nilai data masa lalu yang data historisnya diperoleh dan dianalisis untuk memberikan hasil nilai-nilai di masa datang [5]. Model ini hanya membutuhkan nilai konsumsi energi listrik di masa lalu sehingga model ini merupakan model yang paling sederhana untuk diterapkan [3].

Salah satu metode pada model *time series* yang cukup populer adalah metode *Autoregressive integrated moving average* (ARIMA). Seperti yang dilakukan oleh Yao Lu [3], Vitaliy Y. dan Oleksandr B. [6], dan Kanyaphorn S., dkk [7] yang membandingkan ARIMA dengan model *time-series* lain untuk prediksi energi terbarukan di amerika, *failure* pada software angular, dan resiko kebakaran secara berurutan. Ketiganya memperoleh hasil bahwa ARIMA memiliki akurasi yang lebih baik dibanding metode-metode *time-series* lain yang dibandingkan pada karya ilmiah tersebut yaitu metode exponential smoothing-additive model, exponential smoothing-multiplicative model dan seasonal dummy with GARCH.

Menghitung akurasi model prediksi dilakukan dengan membandingkan hasil prediksi dengan nilai sebenarnya. Salah satu caranya adalah dengan menghitung nilai *root mean squared error* (RMSE). Semakin kecil nilai RMSE (mendekati nilai nol) maka semakin besar akurasi hasil prediksi yang dilakukan oleh model prediksi, begitupun sebaliknya [8]. Akurasi dari suatu metode prediksi juga dipengaruhi oleh jangka waktu prediksinya. Penelitian [3] [6] [7] tidak menjadikan jangka waktu prediksi sebagai salah satu variabel eksperimen, sehingga tidak diperoleh pengetahuan mengenai jangka waktu prediksi optimal dari metode yang diteliti.

I.2 Perumusan/Identifikasi Masalah

Berdasarkan latar belakang yang sudah dijelaskan, terdapat permasalahan berikut: beberapa penelitian tidak menyatakan jangka waktu prediksi optimal dari ARIMA. Hal itu diperlukan untuk mengetahui apakah ARIMA paling optimal untuk mendukung keputusan operasional (jangka pendek, yaitu jam hingga hari), keputusan taktis (jangka menengah, yaitu minggu hingga bulan), atau keputusan strategis (jangka panjang, yaitu tahun hingga beberapa tahun) dalam kasus prediksi konsumsi energi listrik.

I.3 Research Question

Dari rumusan masalah yang ada, dibentuk *research question* berikut: Bagaimana jangka waktu prediksi optimal dari metode ARIMA dalam prediksi konsumsi energi listrik berdasarkan akurasinya ?

I.4 Hipotesis

Dari *research question*, diperoleh hipotesis berikut: ARIMA optimal untuk prediksi jangka menengah, yaitu prediksi satu bulan kedepan dalam kasus konsumsi energi listrik.

I.5 Tujuan

Tujuan dari penelitian ini adalah mengetahui jangka waktu prediksi optimal dari metode ARIMA dalam prediksi konsumsi energi listrik berdasarkan akurasinya.

I.6 Ruang Lingkup

Batasan dan ruang lingkup yang ada pada penelitian ini diantaranya:

- Data yang digunakan untuk pembuatan model dan simulasi menggunakan dataset Hourly Energy Consumption [9].
- Prediksi dilakukan untuk satu area distribusi, sehingga lokasi tidak menjadi parameter.

I.7 Sistematika Penulisan

Sistematika penulisan pada tugas akhir ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang penelitian ini dilakukan, perumusan masalah yang akan diselesaikan dalam penelitian, tujuan, ruang batasan pada penelitian, dan sistematika penulisan laporan ini;

BAB II TINJAUAN PUSTAKA

Bab ini berisi teori yang mendukung topik penelitian yang akan dikerjakan, yaitu tentang metode prediksi, ARIMA, dan evaluasi metode prediksi yang didapat dari jurnal, tesis, buku, dan internet;

BAB III METODOLOGI PENELITIAN

Bab ini membahas langkah-langkah yang dilakukan dalam proses penelitian serta cara yang akan dilakukan untuk mencapai hasil yang diharapkan dalam penelitian ini;

BAB IV ANALISIS, PERANCANGAN, IMPLEMENTASI DAN PENGUJIAN

Bab ini membahas analisis problem domain, analisis metode yang akan digunakan, analisis kebutuhan aplikasi yang akan dibangun untuk eksperimen, perancangan, dan implementasi aplikasi.

BAB V HASIL DAN PEMBAHASAN

Bab ini menuangkan hasil eksperimen yang dilakukan berdasarkan skenario eksperimen pada bab metodologi penelitian, dan pembahasan dari hasil eksperimen yang telah dilakukan;

BAB VI KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari penelitian yang telah dilakukan dan saran pengembangan kedepannya dalam rangka proses perbaikan.

BAB II TINJAUAN PUSTAKA

II.1 Karya Ilmiah Sejenis Sebelumnya

Yao Lu pada [3] melakukan analisis terhadap prediksi *time-series* pada konsumsi energi terbarukan di Amerika. Penelitian yang dilakukan pada thesis ini membuktikan bahwa exponential smoothing-additive model, exponential smoothing-multiplicative model, seasonal dummy with GARCH dan ARIMA merupakan pendekatan yang layak untuk prediksi konsumsi energi terbarukan. Ke empat metode tersebut digunakan untuk memprediksi jangka menengah yaitu satu bulan kedepan. Dalam paper ini digunakan Mean Absolute Percentage Error(MAPE), Mean Absolute Error(MAE), dan Root Mean Squared Error(RMSE) untuk evaluasi hasil prediksinya. Hasil dari eksperimen yang dilakukan pada penelitian ini dapat dilihat pada tabel II.1.

Tabel II.1 Hasil penelitian [3]

Model	MAPE	MAE	RMSE
Exponential smoothing-additive	0.04	33.19	42.31
Exponential smoothing-multiplicative	0.05	37.77	51.22
ARIMA	0.03	25.79	33.86
GARCH	0.04	29.97	40.35

Dari tabel tersebut dapat diketahui bahwa ARIMA memiliki akurasi paling tinggi dengan memperoleh nilai paling kecil di semua perhitungan evaluasi hasil prediksi.

Vitaliy Y. dan Oleksandr B. pada [6] melakukan perbandingan antar model Holt-Winters smoothing model dan ARIMA regression model untuk memprediksi *software failure* pada aplikasi berbasis Angular dalam kurun waktu mingguan. Data diperoleh dari Github bug-tracking system yang berisi data *failures* selama dua tahun. Kedua model tersebut digunakan untuk memprediksi dengan jangka waktu prediksi menengah yaitu satu minggu kedepan. Data yang digunakan untuk proses *training* adalah 90% dan 10% digunakan untuk proses *test*. Untuk mengukur akurasi dari hasil prediksi kedua model, digunakan root mean squared error(RMSE)

antara hasil prediksi dengan nilai sebenarnya. Hasil dari eksperimen yang dilakukan pada penelitian ini dapat dilihat pada tabel II.2.

Tabel II.2 Hasil Penelitian [6]

Model	RMSE
Holt-Winters	220.64
ARIMA	35.05

Kanyaphorn S., Chunchom P., dan Savitri G. pada [7] melakukan analisis pada prediksi *time-series* untuk memprediksi resiko kebakaran dengan menggunakan data meteorologi. Data yang digunakan adalah data faktor meteorologi pada tahun 2010-2014 dari departemen meteorologi Thailand. Metode yang digunakan adalah ARIMA dan Holt-Winters. Kedua model digunakan untuk memprediksi dengan jangka waktu prediksi menengah, yaitu satu bulan kedepan. Untuk mengukur akurasi, digunakan mean absolute percentage error(MAPE) antara hasil prediksi dengan nilai sebenarnya. Hasil penelitian menunjukkan bahwa model Holt-Winters memperoleh akurasi yang lebih baik untuk memprediksi curah hujan, tekanan dan kelembaban relatif. Sedangkan model ARIMA memperoleh akurasi yang lebih baik untuk memprediksi kecepatan angin rata-rata dan data suhu. Namun pada penelitian ini tidak disampaikan mengenai karakteristik data ataupun alasan mengapa data tersebut memperoleh akurasi yang lebih baik menggunakan model tertentu dibanding model lainnya.

Ketiga penelitian [3] [6] [7] tidak menjadikan jangka waktu prediksi sebagai salah satu variabel eksperimen, sehingga tidak diperoleh pengetahuan mengenai jangka waktu prediksi optimal dari metode yang diteliti. Hal itu diperlukan untuk mengetahui ARIMA paling optimal untuk mendukung keputusan operasional(jangka pendek, yaitu jam hingga hari), keputusan taktis(jangka menengah, yaitu minggu hingga bulan), dan keputusan strategis(jangka panjang, yaitu tahun hingga beberapa tahun) dalam kasus prediksi konsumsi energi listrik.

II.2 Dasar Teori

II.2.1 Energi Listrik

Energi adalah sesuatu yang bersifat abstrak yang sukar dibuktikan tetapi dapat dirasakan adanya. Energi adalah kemampuan untuk melakukan kerja. Energi dikategorikan menjadi beberapa macam yaitu energi mekanik, energi listrik, energi elektromagnetik, energi kimia, energi nuklir, dan energi termal (panas). Energi listrik adalah energi yang berkaitan dengan akumulasi arus elektron, dinyatakan dalam Watt-jam atau kilo Watt-jam. Bentuk transisinya adalah aliran elektron melalui konduktor jenis tertentu [10].

II.2.2 Prediksi

Prediksi adalah penggunaan data masa lalu dari sebuah variabel atau kumpulan variabel untuk mengestimasi nilainya di masa yang akan datang. Asumsi dasar dalam penerapan teknik prediksi adalah jika kita dapat memprediksi apa yang terjadi di masa depan maka kita dapat mengubah kebiasaan kita saat ini agar masa yang akan datang menjadi lebih baik. Prediksi dibutuhkan karena semua institusi/industri beroperasi dalam lingkungan yang tidak pasti dan keputusan yang dibuat hari ini akan mempengaruhi masa depan institusi/industri. Prediksi yang efektif sangat dibutuhkan untuk mencapai tujuan strategis dan operasional dari semua institusi/industri. Untuk penyedia energi listrik, prediksi diperlukan untuk mengatur sistem kendali produksi (pembangkitan) dan pendistribusian sesuai kebutuhan. Untuk sektor publik, prediksi merupakan bagian yang tidak terpisahkan dari perancangan kebijakan dan program, baik dalam bidang ekonomi, pendidikan, maupun kesehatan masyarakat [11].

II.2.3 Metode Prediksi

Pada umumnya metode prediksi dapat dibedakan menjadi beberapa jenis tergantung dari cara melihatnya, yaitu:

1. Dilihat dari sifat penyusunannya [11]

- a. Prediksi yang subjektif, yaitu prediksi yang didasarkan atas perasaan atau intuisi dari orang yang menyusunnya. Dalam hal ini pandangan orang yang menyusunnya sangat menentukan baik tidaknya hasil prediksi tersebut.
 - b. Prediksi yang objektif, yaitu prediksi yang didasarkan atas data yang relevan pada masa lalu, dengan menggunakan teknik-teknik dan metode-metode dalam penganalisaannya.
2. Dilihat dari jangka waktu prediksinya [11]

Tabel II.3 Rentang waktu dalam prediksi

Rentan Waktu	Tipe Keputusan	Contoh
Jangka Pendek	Operasional	Perencanaan Produksi, Distribusi
Jangka Menengah	Taktis	Penyewaan lokasi dan peralatan
Jangka Panjang	Strategis	Penelitian dan pengembangan akuisisi dan merger

- a. Prediksi jangka pendek (*short term forecasting*), yaitu prediksi yang dilakukan dengan jangka waktu harian hingga setiap jam. Biasa digunakan untuk studi perbandingan beban listrik prediksi dengan aktual (*realtime*).
 - b. Prediksi jangka menengah (*mid term forecasting*), yaitu prediksi yang dilakukan dengan jangka waktu mingguan hingga bulanan. Biasa digunakan untuk mempersiapkan jadwal persiapan dan operasional sisi pembangkit.
 - c. Prediksi jangka panjang (*long term forecasting*), yaitu prediksi yang dilakukan dengan jangka waktu tahunan atau beberapa tahun kedepan. Biasanya dapat digunakan untuk mempersiapkan ketersediaan unit pembangkitan, sistem transmisi, serta distribusi.
3. Dilihat dari sifat prediksi yang telah disusun [11]
- a. Prediksi kualitatif, yaitu prediksi yang didasarkan atas kualitatif pada masa lalu. Hasil prediksi yang dibuat sangat tergantung pada orang yang menyusunnya. Hal ini penting karena hasil prediksi tersebut ditentukan berdasarkan pemikiran yang bersifat intuisi, *judgement* atau pendapat dan pengetahuan serta pengalaman dari penyusunnya.

- b. Prediksi kuantitatif, yaitu prediksi yang didasarkan atas data kuantitatif pada masa lalu. Hasil prediksi yang dibuat sangat tergantung pada metode yang digunakan dalam prediksi tersebut. Dengan metode yang berbeda akan diperoleh hasil prediksi yang berbeda, adapun yang perlu diperhatikan dari penggunaan metode tersebut, adalah baik tidaknya metode yang digunakan, sangat ditentukan oleh perbedaan atau penyimpangan antara hasil prediksi dengan kenyataan yang terjadi. Metode yang baik adalah metode yang memberikan nilai-nilai perbedaan atau penyimpangan yang mungkin paling sedikit. Prediksi kuantitatif hanya dapat digunakan apabila adanya informasi tentang keadaan lain, informasi tersebut dapat dituliskan dalam bentuk data, dan dapat diasumsikan bahwa pola yang lalu akan berkelanjutan pada masa yang akan datang.

II.2.3.1 Regresi

Prediksi menggunakan regresi digunakan untuk mencari hubungan antara konsumsi energi listrik dengan faktor lain. Regresi dapat dilakukan pada data linear maupun non linear [12]. Pada ARIMA terdapat unsur regresi yaitu dengan mencari hubungan antara satu nilai dengan nilai itu sendiri di masa lalu (*Autoregressive*).

II.2.3.2 Time series

Time Series atau deret waktu adalah serangkaian pengamatan yang diambil berdasarkan urutan waktu dan antar pengamatan saling berkorelasi sehingga dikatakan bahwa tiap pengamatan yang diambil berkorelasi dengan variabel itu sendiri pada waktu sebelumnya. Metode ini berdasarkan pada asumsi data yang memiliki struktur didalamnya, seperti autokorelasi, *trend* ataupun variasi musiman. *Time series* telah digunakan dalam beberapa dekade untuk bidang ekonomi, *digital signal processing* (DSP), seperti halnya prediksi beban listrik. contoh metode yang sering digunakan diantaranya AR (*auto regressive*), MA (*moving average*), lalu

dikembangkan menjadi ARMA (*auto regressive moving average*), dan ARIMA (*auto regressive integrated moving average*) [11].

II.2.3.3 Autoregressive Integrated Moving Average (ARIMA)

Autoregressive integrated moving average (ARIMA) merupakan model yang dikembangkan oleh George Box dan Gwilyn Jenkinson yang diterapkan untuk analisis dan peramalan data deret waktu atau dikenal sebagai model Box-Jenkins. Langkah dalam ARIMA meliputi identifikasi model, pendugaan parameter, dan diagnosis model [8].

Model ARIMA merupakan gabungan antara model *autoregressive* (AR) dan model *moving average* (MA) dengan data telah mengalami proses *differencing* sebanyak n kali. Pada model AR regresi dilakukan terhadap dirinya sendiri atau menghubungkan nilai-nilai sebelumnya pada variabel itu sendiri. Model MA merupakan salah satu metode analisis teknikal sederhana dengan mencari rata-rata bergerak dari suatu variabel selama beberapa periode yang dipengaruhi oleh kesalahan atau residual pada saat ini dan masa lalu. Model AR secara umum seperti pada persamaan (8)

$$y_t = \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + e_t \quad (1)$$

Dengan keterangan:

- y_t : nilai variabel pada waktu ke- t
- α_i : koefisien *autoregressive*; $i : 1, 2, \dots, p$
- e_t : nilai *error* pada waktu ke- t
- p : order AR

Pada persamaan (1) model AR bergantung atau dipengaruhi oleh nilai observasi sebelumnya.

Model MA secara umum seperti pada persamaan (2)

$$y_t = e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (2)$$

Dengan keterangan:

- y_t : nilai variabel pada waktu ke- t
- θ_i : koefisien *moving average*; $i : 1, 2, \dots, q$
- e_t : nilai *error* pada waktu ke- t
- q : order MA

Pada persamaan (2) model MA bergantung atau dipengaruhi oleh nilai *error* saat ini dan nilai *error* dengan bobot tertentu pada masa lalu.

Dari model AR pada persamaan (1) dan model MA pada persamaan (2) didapat model ARIMA secara umum seperti pada persamaan (3)

$$y_t = \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (3)$$

Dengan keterangan:

- y_t : nilai variabel pada waktu ke- t
- α_i : koefisien *autoregressive*; $i : 1, 2, \dots, p$
- θ_i : koefisien *moving average*; $i : 1, 2, \dots, q$
- p : order MA
- q : order MA
- e_t : nilai *error* pada waktu ke- t

Sehingga dari persamaan (3) pada model ARIMA, nilai saat ini bergantung atau dipengaruhi dari beberapa nilai sebelumnya, nilai kesalahan saat ini dan beberapa nilai kesalahan sebelumnya.

II.2.3.3.1 Tahapan ARIMA

II.2.3.3.2 Uji Stationer

Salah satu syarat menggunakan metode ARIMA, data *time series* yang dimiliki harus sudah dalam kondisi stasioner. Stasioner artinya tidak terjadi pertumbuhan dan penurunan data. Suatu kelompok data dikatakan stasioner apabila nilai rata-rata

dan variansinya konstan selama waktu tertentu. Data *time series* dikatakan stasioner apabila tidak ada unsur *trend* dalam data, tidak ada unsur musiman atau rata-rata dan variannya tetap. Terdapat beberapa cara untuk mengetahui suatu data *time series* berada dalam kondisi stasioner [8], diantaranya:

- Melihat grafik data *time series*, apabila terdapat unsur *trend* atau musiman dari plot grafik maka data tersebut tidak stasioner.
- *Summary Statistics*, dengan melihat *summary statistic* seperti rata-rata dan varian dari data *time series*, apakah terjadi perubahan nilai yang signifikan atau tidak dari nilai rata-rata dan varian pada periode waktu tertentu.
- *Statistical Test*, dengan menggunakan uji statistik untuk mengetahui kondisi stasioner pada data *time series*.

Apabila data tidak dalam kondisi stasioner, maka dapat dilakukan proses *differencing* untuk menjadikan data *time series* menjadi dalam bentuk stasioner. Proses *differencing* dilakukan seperti pada persamaan (4)

$$\nabla^n y_t = y_t - y_{t-1} \quad (4)$$

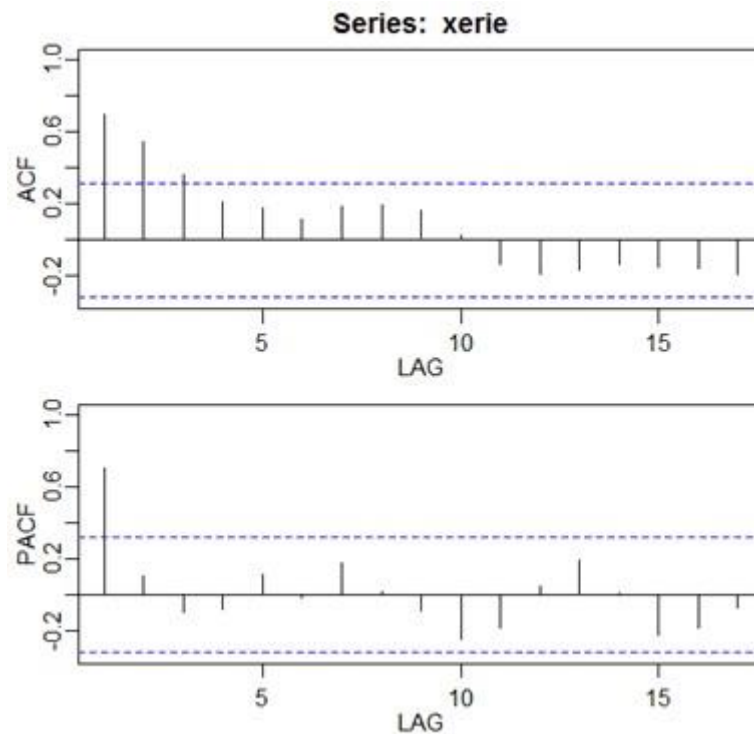
Dengan keterangan:

- y_t : nilai saat ini
- y_{t-1} : nilai sebelum
- ∇y_t : hasil *differencing*
- n : nilai order *differencing*

Setiap nilai pada data *time series* dilakukan proses *differencing* seperti pada persamaan (4). Proses *differencing* dapat dilakukan beberapa kali, hingga data *time series* menjadi dalam bentuk stasioner. Jumlah proses *differencing* yang dilakukan menunjukkan nilai order *differencing*. Biasanya digunakan variabel d untuk menunjukkan order *differencing*.

II.2.3.3 Penentuan Nilai AR dan MA

Setelah data *time series* sudah dalam bentuk stasioner, selanjutnya nilai dugaan awal AR dan MA dapat diketahui. Nilai AR dan MA dapat diketahui dengan menggunakan nilai *autocorrelation function* (ACF) dan *partial autocorrelation function* (PACF). Secara visual dapat dibuat grafik ACF dan PACF untuk mempermudah menentukan nilai AR dan MA [8]. Gambar 1 menunjukkan ilustrasi dari grafik ACF dan PACF.



Gambar II.1 Contoh plot grafik ACF dan PACF [8]

Garis vertikal dari grafik ACF dan PACF pada Gambar 1 merepresentasikan nilai koefisien korelasi yang disebut sebagai *lag*. Nilai *lag* ACF dan PACF didapat melalui persamaan (5) dan persamaan (6)

$$r_k = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2} \quad (5)$$

Dengan keterangan:

- r_k : koefisien autokorelasi pada $lag-k$
- k : selisih waktu
- n : jumlah observasi
- \bar{x} : rata-rata pengamatan
- x_t : pengamatan pada waktu ke- t
- x_{t+k} : pengamatan pada waktu ke $t+k$, $k = 1, 2, 3, \dots$

$$\pi_k = \begin{cases} 1, k = 0 \\ r_1, k = 1 \\ \frac{r_k - \sum_{t=1}^{k-1} \pi_{k-1,t} \times r_{k-t}}{1 - \sum_{t=1}^{k-1} \pi_{k-1,t} \times r_{k-t}} \end{cases} \quad (6)$$

Dengan keterangan:

- π_k : koefisien autokorelasi parsial pada $lag-k$
- r_k : koefisien autokorelasi pada $lag-k$

Dengan menggunakan persamaan rekursi durbin, nilai $\pi_{k,t}$ didapat melalui persamaan (7)

$$\pi_{k,t} = \pi_{k-1,t} - \pi_k \pi_{k-1,k-t} \quad (7)$$

Dengan keterangan:

- $\pi_{k,k} = \pi_k$

Garis horizontal dari grafik ACF dan PACF pada Gambar 1 merepresentasikan batas signifikan yang menunjukkan koefisien autokorelasi atau suatu lag signifikan atau tidak. Nilai batas signifikan didapat dari persamaan (8).

$$v = \pm \frac{1.96}{\sqrt{N}} \quad (8)$$

Dengan keterangan:

- v : nilai *critical value* atau nilai *significant*

- N : jumlah data observasi yang digunakan

Nilai AR dan MA didapat dari jumlah *lag* yang berada diluar batas signifikan. Sebagai contoh pada Gambar II.1 dari grafik ACF terdapat tiga *lag* yang signifikan (berada di luar batas signifikan), sehingga nilai MA awal bernilai 3, sedangkan 1 untuk nilai awal AR dari grafik PACF.

Jumlah *lag* yang signifikan untuk nilai order AR biasanya digunakan variabel p dan q untuk menunjukkan order MA. Sehingga pada tahap ini akan didapat dugaan awal nilai p dan q , pada Gambar II.1 p bernilai 1 dan q bernilai 3.

II.2.3.3.4 Penentuan Model Terbaik

Setelah nilai dugaan awal p dan q ditentukan, nilai tersebut dapat digunakan untuk membuat beberapa dugaan order model ARIMA yang mungkin. Nilai p yang didapat menunjukkan jumlah maksimal *lag* yang signifikan untuk digunakan pada model *autoregressive*, begitu juga dengan nilai q menunjukkan jumlah maksimal *lag* yang signifikan untuk digunakan pada model *moving average*. Oleh sebab itu dapat dibuat beberapa kombinasi order model ARIMA yang mungkin, sehingga setiap order model dapat dibandingkan, untuk menentukan order model ARIMA terbaik. Nilai p dan q digunakan sebagai batas maksimal order model ARIMA [8].

Untuk menentukan order model terbaik dapat menggunakan perhitungan *Akaike Information Criterion* (AIC). Perhitungan AIC didapat dari persamaan (9).

$$AIC = 2k - 2 \ln(\hat{L}) \quad (9)$$

Dengan keterangan:

- k : jumlah estimasi parameter yang digunakan, dalam hal ini adalah nilai p dan q , sehingga $k = p + q$
- L : nilai maksimum *likelihood*

Order model yang dianggap terbaik merupakan order model yang memiliki nilai AIC terkecil [8].

II.2.4 Evaluasi Metode Prediksi

Evaluasi metode prediksi digunakan untuk mengukur akurasi dari suatu metode prediksi. Salah satu caranya adalah dengan menghitung nilai RMSE. Formula untuk menentukan nilai RMSE ditunjukkan pada persamaan (10).

$$RMSE = \sqrt{\frac{\sum_{t=1}^N (y_t - \hat{y}_t)^2}{N}} \quad (10)$$

Dengan keterangan:

- y_t : nilai aktual *time series* pada waktu ke- t
- \hat{y}_t : nilai yang di prediksi pada waktu ke- t
- N : jumlah data yang digunakan untuk evaluasi prediksi

Nilai RMSE menunjukkan tingkat akurasi dari model yang dibuat untuk melakukan prediksi terhadap nilai yang sebenarnya. Semakin kecil nilai RMSE (mendekati nilai nol) maka semakin besar akurasi hasil prediksi yang dilakukan oleh model prediksi, begitupun sebaliknya [8].

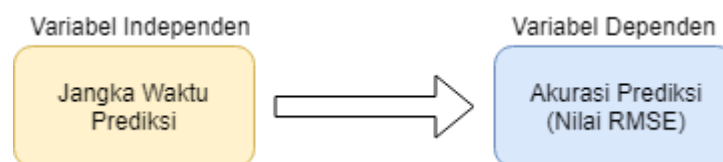
BAB III METODOLOGI PENELITIAN

III.1 Jenis Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan teknik penelitian eksperimental. Eksperimen dilakukan untuk menghasilkan data konsumsi energi listrik hasil prediksi menggunakan model ARIMA untuk setiap jangka waktu prediksi (1 hari, 1 minggu, 1 bulan, 3 bulan, 6 bulan, 1 tahun, 2 tahun, dan 4 tahun).

III.2 Variabel Penelitian

Variabel yang terdapat pada penelitian ini adalah variabel bebas (*independent variable*) dan variabel terikat (*dependent variable*). Nilai RMSE tersebut mewakili akurasi dari model. Pada pelaksanaan eksperimen ini, keterkaitan variabel ditunjukkan pada Gambar III.1.



Gambar III.1 Variabel Penelitian

Berikut adalah penjelasan dari masing-masing variabel yang terdapat pada **Error! Reference source not found.**

III.2.1 Variabel Bebas

Variabel bebas merupakan variabel yang dapat dimanipulasi dan nilainya akan mempengaruhi nilai dari variabel terikat [12]. Pada penelitian ini variabel independennya adalah jangka waktu prediksi.

III.2.2 Variabel Terikat

Variabel terikat merupakan variabel yang diakibatkan atau dipengaruhi oleh variabel bebas [12]. Variabel dependennya adalah nilai RMSE dari hasil prediksi.

III.3 Objek Penelitian

Objek penelitian ini adalah data konsumsi energi listrik dari American Electric Power(AEP) [9].

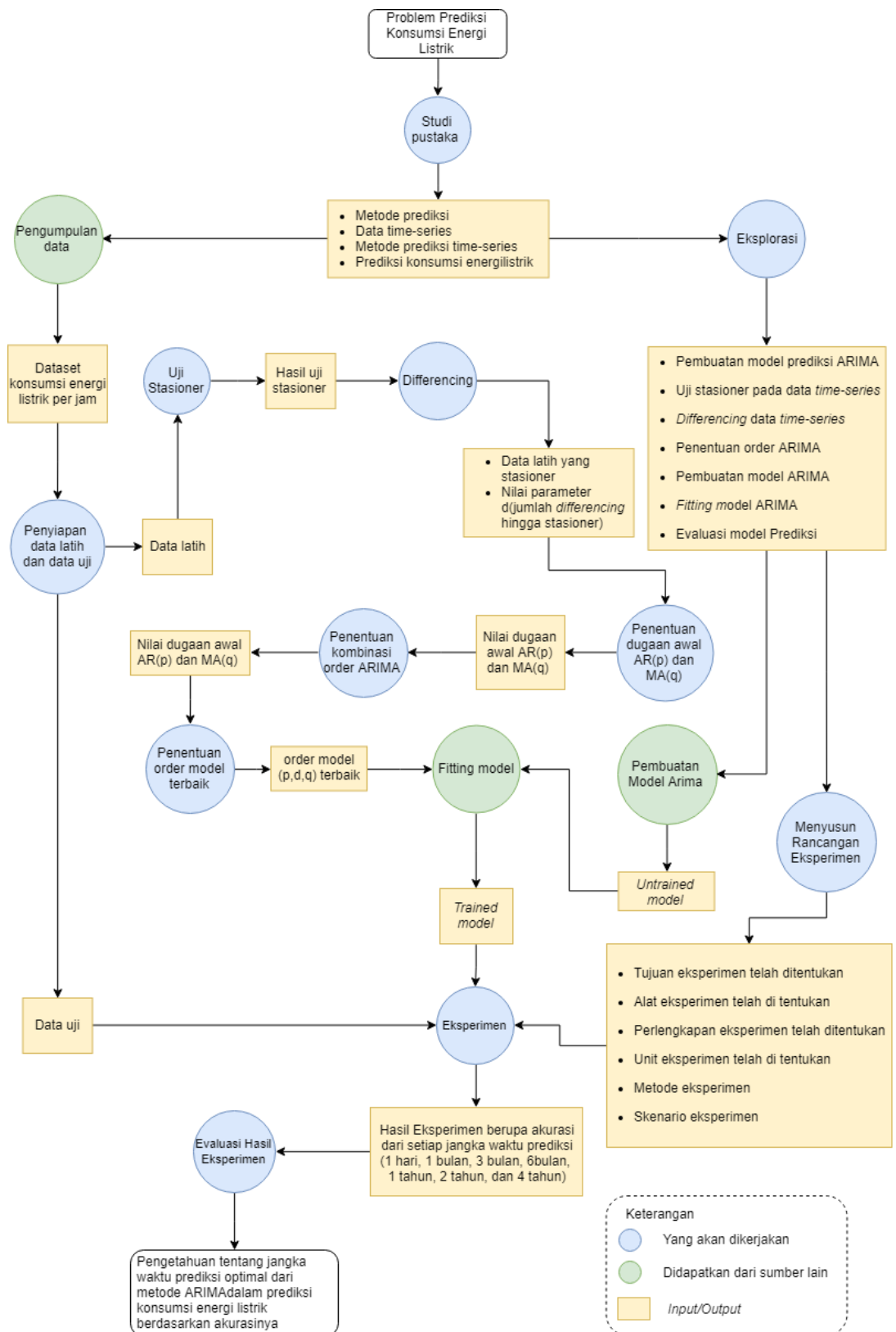
III.4 Perangkat Pendukung

Berikut merupakan beberapa alat pendukung untuk menunjang penelitian:

- Python : Bahasa pemrograman untuk eksplorasi metode prediksi time-series dan simulasi sistem prediksi penggunaan energi listrik;
- *Library* Statistika pada python: *Library* untuk persamaan statistika dan pembuatan grafik dari data;
- Terminal: untuk menjalankan perintah dari python, dan melihat hasil dari perintah yang dijalankan.

III.5 Tahapan Penelitian

Tahapan penelitian diilustrasikan pada Gambar III.2 berikut.



Gambar III.2 Metodologi Penelitian

Adapun rincian dari masing-masing proses dalam metodologi tersebut dijelaskan dalam uraian tahapan dibawah ini:

III.5.1 Studi Pustaka

Tahap studi pustaka merupakan tahapan untuk mencari, dan melakukan studi dari referensi terkait pendekatan *time-series* untuk prediksi penggunaan energi listrik dan memahami hal-hal dasar yang dibutuhkan untuk mengerjakan penelitian ini. Beberapa referensi yang dijadikan sebagai studi pustaka diantaranya:

- Metode prediksi
- Data time-series
- Metode prediksi time-series
- Prediksi konsumsi energi listrik

III.5.2 Pengumpulan Data

Untuk mendukung penelitian ini, dibutuhkan data untuk membuat model prediksi, evaluasi model, dan simulasi pada sistem prediksi. Sumber data yang digunakan pada penelitian ini didapat dari dataset [9]. Data diperoleh dari American Electric Power(AEP). Data tersebut berupa data konsumsi energi listrik perjam dalam megawatt(MW) dari pukul 01:00 tanggal 1 Oktober 2004 sampai pukul 00:00 tanggal 3 Februari 2018. Data Tabel III.1 merupakan contoh data dari dataset yang didapat.

Tabel III.1 Contoh data pada dataset [9]

Datetime	AEP_MW
8/1/2006 0:00	19926
8/1/2006 1:00	18397
8/1/2006 2:00	17347
8/1/2006 3:00	16510
8/1/2006 4:00	16102
...	...
7/31/2014 21:00	16491
7/31/2014 22:00	16418
7/31/2014 23:00	15323

III.5.3 Eksplorasi

Untuk melanjutkan studi yang lebih dalam dari hasil studi pustaka dan pembuatan rancangan eksperimen, dilakukan eksplorasi untuk memahami perilaku beberapa hal yang akan diimplementasikan / digunakan pada penelitian ini, eksplorasi yang dilakukan diantaranya:

- Pembuatan model prediksi ARIMA
- Uji stasioner pada data *time-series*
- *Differencing* data *time-series*
- Penentuan order ARIMA
- Pembuatan model ARIMA
- *Fitting model* ARIMA
- Evaluasi model Prediksi

III.5.4 Penyiapan Data Latih dan Data Uji

Setelah dataset diperoleh, akan dilakukan penyiapan data latih dan data uji. Pembagian data latih dan data uji dilakukan dengan memaksimalkan data yang tersedia pada dataset [9] setelah diambil data uji paling banyak yaitu 4 tahun. Data latih yang diambil adalah data dari dataset [9] dari tanggal 1 Agustus 2006 hingga tanggal 31 Juli 2014 dengan jumlah 70109 data. Contoh isi dari data latih terdapat pada tabel III.2.

Tabel III.2 Contoh data pada data latih

Datetime	AEP_MW
8/1/2006 0:00	19926
8/1/2006 1:00	18397
8/1/2006 2:00	17347
...	...
7/31/2014 21:00	16491
7/31/2014 22:00	16418
7/31/2014 23:00	15323

Untuk data uji, diambil data dari dataset [9] dimulai dari tanggal 1 Agustus 2014 dengan jumlah sesuai jangka waktu prediksi (1 hari, 1 minggu, 1 bulan, 3 bulan, 6 bulan, 1 tahun, 2 tahun, dan 4 tahun).

III.5.5 Uji Stasioner

Uji stasioner dilakukan dengan cara melihat *summary statistic* yaitu rata-rata dan varian dari data latih, apakah terjadi perubahan nilai yang signifikan atau tidak dari nilai rata-rata dan varian pada periode waktu tertentu. Data latih dibagi menjadi dua bagian untuk melihat perubahan nilai rata-rata dan varian dari bagian satu dengan bagian kedua.

III.5.6 Differencing

Apabila data tidak dalam kondisi stasioner, maka dapat dilakukan proses *differencing* untuk menjadikan data *time series* menjadi dalam bentuk stasioner. Proses *differencing* dilakukan seperti pada persamaan (4). Proses *differencing* dapat dilakukan beberapa kali, hingga data *time series* menjadi dalam bentuk stasioner. Jumlah proses *differencing* yang dilakukan menunjukkan nilai order *differencing*. Biasanya digunakan variabel d untuk menunjukkan order *differencing*.

III.5.7 Penentuan Dugaan Awal AR(p) dan MA(q)

Setelah data *time series* sudah dalam bentuk stasioner, selanjutnya nilai dugaan awal AR dan MA dapat diketahui. Nilai AR(p) dan MA(q) dapat diketahui dengan menggunakan nilai *autocorrelation function* (ACF) dan *partial autocorrelation function* (PACF). Nilai dugaan awal itu dapat dibuat beberapa kombinasi order model ARIMA yang mungkin, sehingga setiap order model dapat dibandingkan, untuk menentukan order model ARIMA terbaik. Nilai p dan q digunakan sebagai batas maksimal order model ARIMA.

III.5.8 Penentuan Kombinasi Order ARIMA

Dari dugaan awal nilai p dan q pada tahapan sebelumnya, dapat dibuat kombinasi beberapa model yang mungkin seperti pada tabel III.3.

Tabel III.3 Contoh kombinasi model ARIMA yang diduga

No.	Order			Model ARIMA	Model AR/MA
	p	d	q		
1	0	1	1	ARIMA(0,1,1)	MA(1)
2	1	1	0	ARIMA(1,1,0)	AR(1)
3	1	1	1	ARIMA(1,1,1)	-
4	2	1	0	ARIMA(2,1,0)	AR(2)
5	2	1	1	ARIMA(2,1,1)	-

III.5.9 Penentuan Order Model Terbaik

Untuk menentukan order model terbaik dapat menggunakan perhitungan *Akaike Information Criterion* (AIC). Perhitungan AIC didapat dari persamaan (9). Order model yang dianggap terbaik merupakan order model yang memiliki nilai AIC terkecil.

III.5.10 Pembuatan Model ARIMA

Untuk pembuatan model ARIMA, akan digunakan library statsmodel yang diimplementasi dalam bahasa pemrograman python. Model ARIMA dibuat sesuai dengan persamaan (3).

III.5.11 *Fitting Model*

Model yang telah dibuat pada tahap pembuatan model ARIMA selanjutnya dilakukan *fitting model*. Untuk *fitting model*, digunakan data latih yang telah diproses hingga stasioner pada tahap sebelumnya. Selain itu, digunakan juga order ARIMA terbaik yang telah diperoleh pada tahap penentuan order model terbaik.

III.5.12 Menyusun Rancangan Eksperimen

Untuk mencapai tujuan penelitian, dibutuhkan rancangan eksperimen yang terstruktur sebelum melakukan eksperimen. Berikut merupakan penjelasan lebih lanjut mengenai rancangan eksperimen untuk keperluan melakukan eksperimen pada penelitian ini.

III.5.12.1 Tujuan Eksperimen

Tujuan dari eksperimen pada penelitian ini adalah menghasilkan data konsumsi energi listrik hasil prediksi menggunakan model ARIMA untuk setiap jangka waktu prediksi (1 hari, 2 minggu, 1 bulan, 3 bulan, 6 bulan, 1 tahun, 2 tahun, dan 4 tahun).

III.5.12.2 Bahan, Alat, Perlengkapan, dan Unit Eksperimen

III.5.12.2.1 Bahan Eksperimen

Data yang digunakan untuk eksperimen pada penelitian ini adalah data yang berasal dari *dataset* [9]. Akan terdapat beberapa skenario eksperimen dengan menggunakan beberapa jangka waktu prediksi. Data yang akan digunakan pada tahapan eksperimen untuk bagian *training* adalah data pada dataset [9] dari 1 Agustus 2006 hingga tanggal 31 Juli 2014 dengan jumlah 70109 data. Data yang digunakan untuk pengujian adalah data pada dataset [9] dimulai dari tanggal 1 Juli 2018. Prediksi akan dilakukan untuk jangka waktu 1 hari, 1 minggu, 1 bulan, 3 bulan, 6 bulan, 1 tahun, 2 tahun, dan 4 tahun.

III.5.12.2.2 Alat Eksperimen

Alat yang digunakan pada eksperimen pada penelitian ini berupa *server local* dengan spesifikasi seperti pada Tabel III.7.

Tabel III.4 Alat eksperimen

<i>Hardware</i>	
RAM	8 GB
<i>Hardisk</i>	500 GB
<i>Processor</i>	AMD Ryzen 3 2200G 3.5GHz

<i>Hardware</i>	
CPU Cores	4
<i>Software</i>	
Operating System	Windows 10 Pro 64-bit
Bahasa Pemrograman	Python
Library	Statsmodel
	Pandas
	Matplotlib
	Sklearn
	Math
	Numpy

III.5.12.2.3 Perlengkapan Eksperimen

Penelitian ini menggunakan perlengkapan eksperimen berupa beberapa perangkat pendukung untuk membantu dalam membangun, dan menjalankan program. Berikut adalah beberapa perangkat pendukung yang digunakan:

1. Visual Studio Code

Visual Studio Code merupakan teks editor yang digunakan untuk pembuatan program pada penelitian ini.

2. Terminal

Terminal digunakan untuk menjalankan program yang sudah dibuat untuk melakukan kegiatan eksperimen.

III.5.12.2.4 Unit Eksperimen

Unit yang digunakan untuk eksperimen pada penelitian ini adalah program untuk prediksi konsumsi energi listrik dengan metode ARIMA. Program yang dibuat akan menghasilkan file csv hasil prediksi konsumsi energi listrik sesuai jangka waktu prediksi yang ditentukan.

III.5.12.3 Metode Eksperimen

Metode eksperimen merupakan langkah-langkah yang harus dilakukan untuk menjalankan eksperimen. Langkah-langkah yang dilakukan untuk eksperimen adalah sebagai berikut:

1. Persiapkan alat dan bahan untuk eksperimen
2. Jalankan program yang telah dibuat
3. Masukkan jangka waktu prediksi

Program akan menghasilkan csv berisi hasil prediksi sesuai jangka waktu prediksi yang dimasukkan.

4. *Load* csv hasil prediksi
5. *Load* csv data uji sesuai jangka waktu prediksi yang ditentukan
6. Hitung nilai RMSE dari hasil 4 dan 5
7. Catat nilai RMSE yang dihasilkan sesuai tabel III.5

Tabel III.5 Format catatan nilai RMSE hasil eksperimen

Kode Eksperimen	Jangka Waktu Prediksi	RMSE			
		R1	R2	R3	Rata-rata

8. Langkah 1-7 dilakukan untuk setiap skenario eksperimen

III.5.12.4 Skenario Eksperimen

Skenario perhitungan akurasi akan dilakukan pada setiap jangka waktu prediksi yang ditentukan. Setiap skenario yang dijalankan akan diperoleh nilai konsumsi energi listrik hasil prediksi dan hasil evaluasi akurasi prediksinya berupa nilai RMSE. Setiap skenario dilakukan 3 kali dan diambil nilai rata-ratanya. Skenario eksperimen dapat dilihat pada Tabel III.6.

Tabel III.6 Skenario Eksperimen

Kode Eksperimen	Jangka Waktu Prediksi	RMSE			
		R1	R2	R3	Rata-rata
E01	1 Hari				
E02	1 Minggu				
E03	1 Bulan				
E04	3 bulan				
E05	6 bulan				
E06	1 tahun				
E07	2 tahun				
E08	4 tahun				

III.5.13 Eksperimen

Eksperimen dilakukan menggunakan *trained model* dari proses sebelumnya. Dengan alat, bahan dan proses eksperimen sesuai dengan rancangan eksperimen yang telah ditentukan sebelumnya. Hasil dari eksperimen berupa file csv hasil prediksi sesuai skenario yang telah ditentukan dan hasil evaluasi berupa nilai RMSE.

III.5.14 Evaluasi Hasil Eksperimen

Untuk melakukan evaluasi terhadap hasil eksperimen, semua hasil eksperimen tersebut akan diurutkan sesuai dengan nilai rmse terkecil. Nilai RMSE terkecil menunjukkan bahwa skenario tersebut adalah skenario optimal untuk model ARIMA pada kasus konsumsi energi listrik dengan dataset [9].

Setiap hasil prediksi di setiap iterasi akan divisualisasikan menggunakan plot grafik untuk melihat perbedaan data uji dan hasil prediksi. Hal itu dilakukan untuk memudahkan analisis dan penarikan kesimpulan mengenai alasan suatu skenario optimal ataupun tidak pada penelitian ini.

BAB IV

ANALISIS, PERANCANGAN, IMPLEMENTASI DAN PENGUJIAN

IV.1 Analisis

IV.1.1 Pembuatan Model Prediksi

Berikut merupakan ilustrasi dari penggunaan ARIMA untuk melakukan pemilihan dan pembuatan model prediksi, menggunakan data pada dataset [9]. Untuk simulasi ini digunakan *tools* bahasa pemrograman python dengan *library* statsmodel untuk menentukan order dan membuat model ARIMA.

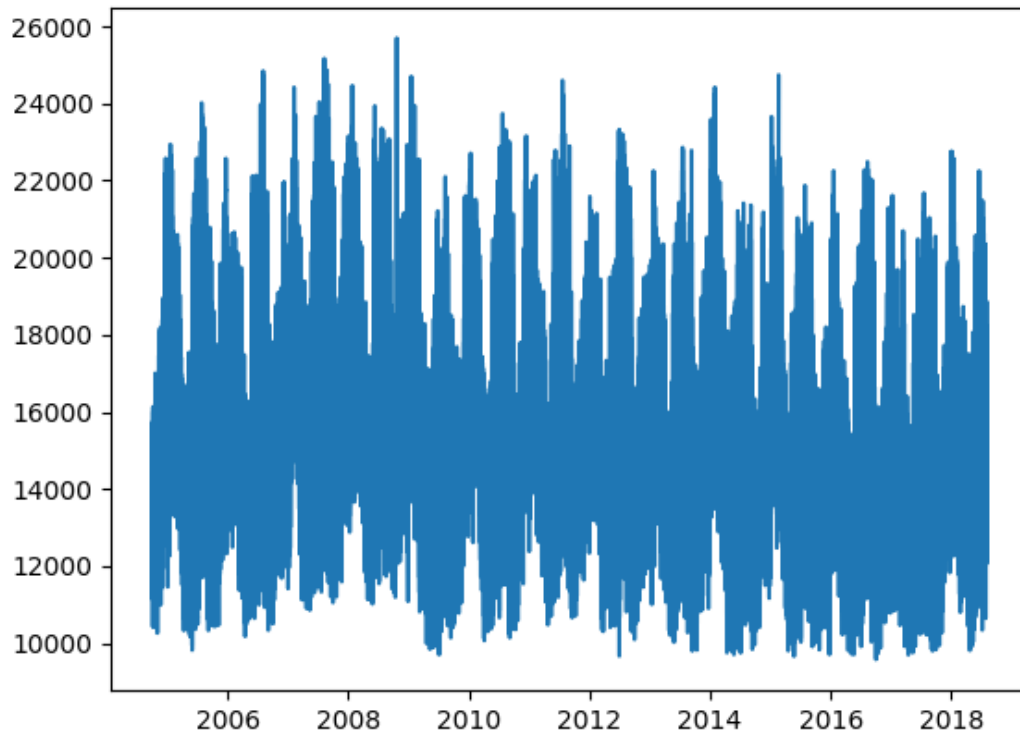
IV.1.1.1 Data Latih

Data yang akan digunakan pada tahapan eksperimen untuk bagian *training* adalah data pada dataset [9] dari 1 Agustus 2006 hingga tanggal 31 Juli 2014 dengan jumlah 70109 data. Contoh isi dari data latih terdapat pada tabel IV.1.

Tabel IV.1 Contoh isi dari data latih

Datetime	AEP_MW
8/1/2006 0:00	19926
8/1/2006 1:00	18397
8/1/2006 2:00	17347
8/1/2006 3:00	16510
...	...
7/31/2014 21:00	16491
7/31/2014 22:00	16418
7/31/2014 23:00	15323

Plot grafik dari data latih disajikan pada gambar IV.1.



Gambar IV.1 Plot grafik pada data latih per jam

Dari plot grafik tidak terlihat adanya kecenderungan tren pada data latih, untuk memastikannya akan dilihat *summary statistic* yaitu rata-rata dan varian dari data latih.

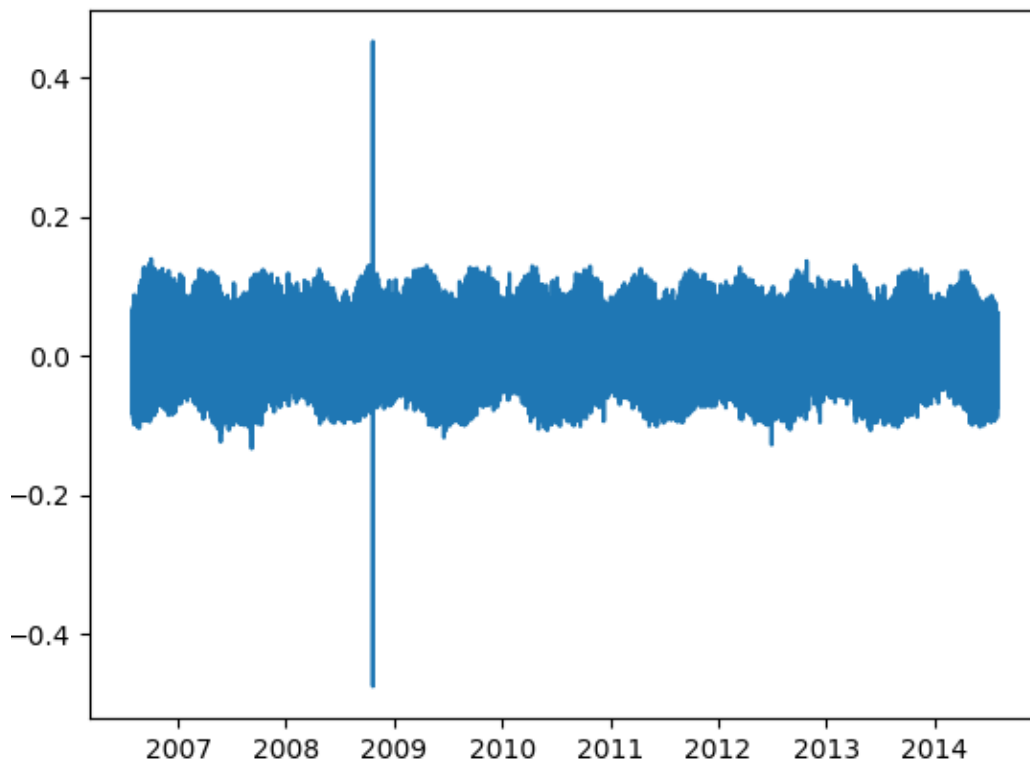
IV.1.1.2 Uji Stasioner

Uji stasioner dilakukan dengan cara melihat *summary statistic* yaitu rata-rata dan varian dari data latih, apakah terjadi perubahan nilai yang signifikan atau tidak dari nilai rata-rata dan varian pada periode waktu tertentu. Data latih dibagi menjadi dua bagian untuk melihat perubahan nilai rata-rata dan varian dari bagian satu dengan bagian kedua. Hasil perhitungan rata-rata dan varian dari bagian satu dan dua dapat dilihat pada tabel IV.2.

Tabel IV.2 Tabel perhitungan rata-rata dan variasi awal

<i>Summary statistic</i>	Bagian 1	Bagian 2	Selisih
Rata-rata	16093.320762	15501.658537	591.662225
Variasi	6993113.701006	6419876.891729	573236.809277

Dari hasil uji stasioner pada Tabel IV.2, hasil uji pertama menunjukkan bahwa data belum dalam bentuk stasioner, sehingga perlu dilakukan proses transformasi ke nilai logaritmanya dan proses *differencing* menggunakan persamaan (4). Plot grafik dari data latih setelah proses transformasi ke nilai logaritmanya dan proses *differencing* disajikan pada gambar IV.2.



Gambar IV.2 Plot grafik pada data latih setelah proses transformasi ke nilai logaritmanya dan proses *differencing*

Setelah dilakukan proses transformasi ke nilai logaritmanya dan proses *differencing*, dilakukan kembali uji stasioner. Hasil perhitungan rata-rata dan varian setelah proses *differencing* dapat dilihat pada tabel IV.3.

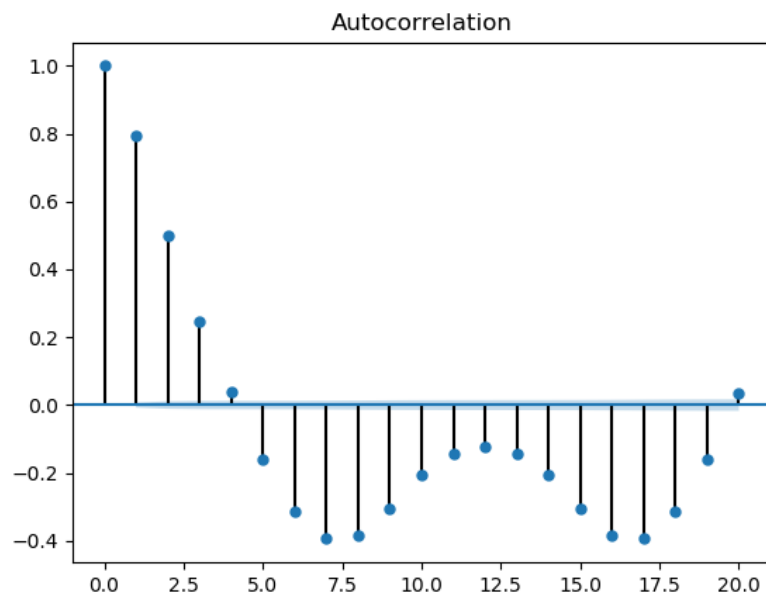
Tabel IV.3 Tabel perhitungan rata-rata dan variasi setelah proses transformasi ke nilai logaritmanya dan proses *differencing*

<i>Summary statistic</i>	Bagian 1	Bagian 2	Selisih
Rata-rata	-0.000007	-0.000001	0.000006
Variasi	0.001419	0.001345	0.000074

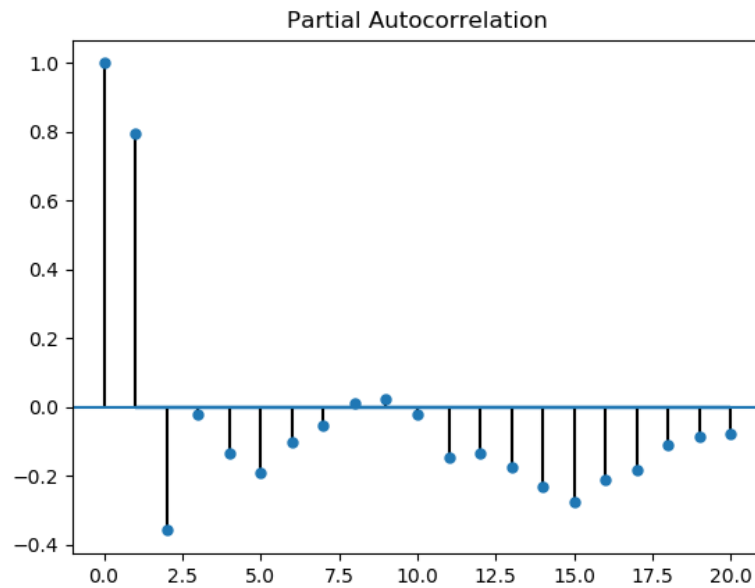
Dari hasil uji stasioner pada Tabel IV.2, menunjukkan perubahan yang tidak signifikan yang berarti data telah stasioner dengan satu kali proses *differencing*. Sehingga nilai d yang didapat adalah 1.

IV.1.1.3 Penentuan Dugaan Awal AR(p) dan MA(q)

Untuk menentukan dugaan awal nilai p dan q dilakukan dengan memerhatikan plot grafik ACF dan PACF, seperti pada Gambar IV.3 untuk ACF dan Gambar IV.4 untuk PACF.



Gambar IV.3 Grafik ACF data latih



Gambar IV.4 Grafik PACF data latih

Dari grafik ACF pada gambar IV.3 terlihat bahwa *cutoff* dimulai pada lag ke-1 sehingga dugaan awal nilai q sebesar 1 untuk MA. Dari grafik PACF pada gambar IV.4 terlihat bahwa *cutoff* terakhir terjadi pada *lag* ke-20, sehingga dugaan awal nilai p untuk AR sebesar 20.

IV.1.1.4 Penentuan Kombinasi Order ARIMA

Dari dugaan awal nilai p dan q pada tahapan sebelumnya, dapat dibuat kombinasi beberapa model yang mungkin seperti pada tabel IV.4.

Tabel IV.4 Kombinasi model ARIMA yang diduga

No.	Order			Model ARIMA	Model AR/MA
	p	d	q		
1	0	1	1	ARIMA(0,1,1)	MA(1)
2	1	1	0	ARIMA(1,1,0)	AR(1)
3	1	1	1	ARIMA(1,1,1)	-
4	2	1	0	ARIMA(2,1,0)	AR(2)
5	2	1	1	ARIMA(2,1,1)	-
6	3	1	0	ARIMA(3,1,0)	AR(3)
7	3	1	1	ARIMA(3,1,1)	-
8	4	1	0	ARIMA(4,1,0)	AR(4)

9	4	1	1	ARIMA(4,1,1)	-
10	5	1	0	ARIMA(5,1,0)	AR(5)
...
36	18	1	0	ARIMA(18,1,0)	AR(18)
37	18	1	1	ARIMA(18,1,1)	-
38	19	1	0	ARIMA(19,1,0)	AR(19)
39	19	1	1	ARIMA(19,1,1)	-
40	20	1	0	ARIMA(20,1,0)	AR(20)
41	20	1	1	ARIMA(20,1,1)	-

IV.1.1.5 Penentuan Order Model Terbaik

Dari beberapa model tersebut ditentukan model terbaik menggunakan nilai AIC. Model terbaik yang digunakan adalah model yang memiliki nilai AIC terkecil. Tabel IV.5 menunjukkan nilai AIC dari masing-masing kombinasi model.

Tabel IV.5 Hasil AIC dari masing-masing kombinasi model ARIMA yang diduga

Model	AIC
ARIMA(0,2,1)	
ARIMA(1,2,0)	
ARIMA(1,2,1)	
ARIMA(2,2,0)	
ARIMA(2,2,1)	
ARIMA(3,2,0)	
ARIMA(3,2,1)	
ARIMA(4,2,0)	
ARIMA(4,2,1)	
ARIMA(5,2,0)	
...	
ARIMA(18,2,0)	
ARIMA(18,2,1)	
ARIMA(19,2,0)	
ARIMA(19,2,1)	
ARIMA(20,2,0)	
ARIMA(20,2,1)	

Dari beberapa kombinasi model ARIMA pada tabel IV.5 model terbaik yang memiliki nilai AIC terkecil akan dipilih untuk digunakan sebagai model untuk prediksi konsumsi energi listrik.

IV.1.2 Analisis Kebutuhan

IV.1.2.1 Analisis Kebutuhan Data

Untuk mengembangkan aplikasi pada penelitian ini maka dibutuhkan data konsumsi energi listrik sebagai *input* untuk pembuatan model prediksi dengan menggunakan metode ARIMA. Berikut merupakan rincian data yang dibutuhkan dalam aplikasi yang dibuat:

1. Data konsumsi energi listrik

Data konsumsi energi listrik merupakan data primer yang dibutuhkan untuk pembuatan model prediksi. Data konsumsi energi listrik dalam penelitian ini didapat dari dataset [9] yang sudah dijelaskan pada subbab pengumpulan dan penyiapan data.

2. Order model

Order model merupakan nilai p , d , dan q yang menunjukkan order model terbaik yang dihasilkan dari tahapan penentuan model terbaik. Order model ini dibutuhkan pada tahapan pembuatan model prediksi menggunakan metode ARIMA.

IV.1.2.2 Analisis Kebutuhan Aplikasi

Aplikasi yang akan dibangun merupakan aplikasi untuk melakukan penentuan model prediksi terbaik, dan pembuatan model prediksi dari setiap parameter cuaca suatu lokasi. Aplikasi yang dibangun digunakan untuk keperluan eksperimen ada penelitian ini. Penentuan model terbaik dan pembuatan model terbaik pada aplikasi ini menggunakan metode ARIMA. Kebutuhan yang harus dipenuhi aplikasi dapat dilihat pada Tabel IV.6.

Tabel IV.6 Kebutuhan aplikasi eksperimen

No	Requirement	ID
1.	Aplikasi dapat membaca data konsumsi energi listrik dari <i>file</i> csv	REQ-01
2.	Aplikasi dapat melakukan uji stasioner pada data latih	REQ-02

3.	Aplikasi dapat melakukan proses transformasi ke nilai logaritma dari data latih	REQ-03
4.	Aplikasi dapat melakukan proses <i>differencing</i> pada data latih	REQ-04
5.	Aplikasi dapat menampilkan plot ACF dan PACF	REQ-05
6.	Aplikasi dapat menentukan setiap order ARIMA yang mungkin dari nilai awal p, d, dan q	REQ-06
7.	Aplikasi dapat menentukan model terbaik dari setiap order ARIMA yang mungkin	REQ-07
8.	Aplikasi dapat melakukan pembuatan model prediksi dari data latih dan order ARIMA yang sudah ditentukan	REQ-08

IV.2 Perancangan

IV.3 Implementasi

Pada tahap ini dijelaskan mengenai lingkungan implementasi, pengecekan kebutuhan aplikasi yang berhasil diimplementasi, dan model *deployment* yang digunakan.

Hardware	
RAM	8 GB
Hardisk	500 GB
Processor	AMD Ryzen 3 2200G 3.5GHz
CPU Cores	4
Software	
Operating System	Windows 10 Pro 64-bit
Bahasa Pemrograman	Python
Library	statsmodel
	pandas
	matplotlib
	sklearn
	math
	numpy

Hasil implementasi berupa aplikasi untuk melakukan pembuatan model prediksi. Aplikasi digunakan untuk mendukung kegiatan eksperimen, bukan untuk digunakan oleh *user* secara umum.

IV.4 Pengujian

BAB V
HASIL DAN PEMBAHASAN

BAB VI

KESIMPULAN DAN SARAN

VI.1 Kesimpulan

VI.2 Saran

DAFTAR PUSTAKA

- [1] K. A. Al-zahra, K. Moosa and B. H. Jasim, "A comparative Study of Forecasting the Electrical Demand in Basra City using Box-Jenkins and Modern Intelligent Techniques," *Iraq J. Electrical and Electronic Engineering*, vol. 11, 2015.
- [2] F. Kaytez, M. C. Taplamacioglu, E. Cam and F. Hardalac, "Forecasting Electricity Consumption: A Comparison of Regression Analysis, Neural Networks And Least Squares Support Vector Machines," *Electrical Power and Energy Systems*, vol. 67, pp. 431-438, 2015.
- [3] Y. Lu, "Time Series Forecasts of Renewable Energy Consumption in the United States," 2018.
- [4] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, p. 159 – 175, 2003.
- [5] V. Sineglazov, E. Chumachenko and V. Gorbatyuk, "An algorithm for solving the problem of forecasting," *Aviation*, vol. 17, pp. 9-13, 2013.
- [6] V. Yakovyna and O. Bachkai, "The Comparison of Holt-Winters and Box-Jenkins Methods for Software Failures Prediction," 2017.
- [7] K. Sakulkitbanjong and C. Pongchavalit, "Time Series Analysis and Forecasting of Forest Fire Weather," 2017.
- [8] M. I. Fauzan, "Analisis Pemrosesan Paralel dalam Mendukung Layanan Prediksi Cuaca," 2018.

- [9] R. Mulla, "Hourly Energy Consumption," 30 August 2018. [Online]. Available: <https://www.kaggle.com/robikscube/hourly-energy-consumption>. [Accessed 30 January 2019].
- [10] B. N. Adha, "Analisis Prakiraan Kebutuhan Energi Listrik Pada PT. PLN(PERSERO) WS2JB Area Palembang dengan Menggunakan Metode Explanatory," 2018.
- [11] F. S. Purnomo, "Penggunaan Metode ARIMA (Autoregressive Integrated Moving Average) untuk Prakiraan Beban Konsumsi Listrik Jangka Pendek (Short Term Forecasting)," 2015.
- [12] M. Priyono, *Metode Penelitian Kuantitatif*, Surabaya: Zifatama Publishing, 2016.
- [13] Y. Yohny, R. Goejantoro and S. Wahyuningsih, "Metode Trend Non Linear Untuk Forecasting Jumlah Keberangkatan Tenaga Kerja," *Ekspansional*, vol. 4, Mei 2013.
- [14] C. Deb, F. Zhang, J. Yang, S. E. Lee and K. W. Shah, "A review on time series forecasting techniques for building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 902-924, 2017.