

**ANALISIS PERBANDINGAN IMPLEMENTASI SARIMA DAN
SUPPORT VECTOR MACHINE (SVM) DALAM PREDIKSI
JUMLAH WISATAWAN MANCANEgara**

(Studi Kasus : Jumlah Wisatawan Mancanegara yang Datang ke Indonesia
Berdasarkan 19 Pintu Masuk Periode Tahun 2010-2017)

TUGAS AKHIR



**MOLYDAH S
14 611 166**

**JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA**

2018

**ANALISIS PERBANDINGAN IMPLEMENTASI SARIMA DAN
SUPPORT VECTOR MACHINE (SVM) DALAM PREDIKSI
JUMLAH WISATAWAN MANCANEgara**

(Studi Kasus : Jumlah Wisatawan yang Datang ke Indonesia Berdasarkan 19 Pintu
Masuk Periode Tahun 2010-2017)

TUGAS AKHIR

Diajukan sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana
Jurusan Statistika



MOLYDAH S

14 611 166

**JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA**

2018

HALAMAN PERSETUJUAN PEMBIMBING

TUGAS AKHIR

Judul : Analisis Perbandingan Implementasi SARIMA dan *Support Vector Machine* (SVM) dalam Prediksi Jumlah Wisatawan Mancanegara

Nama Mahasiswa : Molydah S

Nomor Mahasiswa : 14 611 166

**TUGAS AKHIR INI TELAH DIPERIKSA DAN DISETUJUI UNTUK
DIUJIKAN**

Yogyakarta, 17 April 2018

الإمامة الإسلامية
Dosen pembimbing

(Ayundyah kesumawati, S.Si., M.Si)

HALAMAN PENGESAHAN

TUGAS AKHIR

ANALISIS PERBANDINGAN IMPLEMENTASI SARIMA DAN SUPPORT VECTOR MACHINE (SVM) DALAM PREDIKSI JUMLAH WISATAWAN MANCANEGARA

Nama Mahasiswa : Molydah S
Nomor Mahasiswa : 14611166

TUGAS AKHIR INI TELAH DIUJIKAN
PADA TANGGAL, 11 MEI 2018

Nama Penguji

1. Andrie Pasca Hendradewa, S.T., M.T
2. Tuti Purwaningsih, S.Stat., M.Si.
3. Ayundyah Kesumawati, S.Si., M.Si.

Tanda Tangan



Mengetahui,
الإسلام جامعة

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



(Drs. Allwar, M.Sc., Ph.D)

HALAMAN PERSEMBAHAN

Tugas akhir ini penulis persembahkan untuk:

Bapak Sudirman dan Ibu Emmah selaku orang tua saya yang selama ini telah berjuang membiayai segala keperluan saya mencurahkan segala kasih sayang dan perhatiannya, dan selalu mendo'akan yang terbaik untuk anak-anaknya.

Terimah kasih atas semua yang telah Bapak Ibu berikan kepada saya. Mohon maaf saya belum bisa membalas semua budi jasa Bapak dan Ibu. Semoga karya ini bisa menjadi salah satu hadiah terindah meskipun masih terdapat banyak kekurangan didalamnya.

I LOVE YOU FOREVER

KATA PENGANTAR



Assalamu'alaikum Wr.Wb.

Alhamdulillahirobbil alamin, Puji dan syukur kehadiran Allah SWT atas berkat rahmat serta kasih-Nya yang selalu dilimpahkan kepada penulis sehingga penulis dapat menyelesaikan Tugas Akhir ini dengan judul “Analisis Perbandingan Implementasi SARIMA Dan *Support Vector Machine* (SVM) dalam Prediksi Jumlah Wisatawan Mancanegara”. Shalawat beserta salam semoga senantiasa terlimpah curahkan kepada Nabi Muhammad SAW, kepada keluarganya, para sahabatnya, hingga kepada umatnya hingga akhir zaman.

Penulisan skripsi ini diajukan untuk memenuhi salah satu syarat untuk memperoleh gelar Sarjana Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia. terselesaikannya skripsi ini tidak terlepas dari bantuan banyak pihak, sehingga pada kesempatan ini dengan segala kerendahan hati dan penuh rasa hormat penulis menghaturkan terima kasih yang sebesar-besarnya bagi semua pihak yang telah memberikan bantuan moril maupun materil, baik langsung maupun tidak langsung dalam penyusunan Tugas Akhir ini hingga selesai terutama kepada :

1. Bapak Allwar, M.Sc selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia Yogyakarta.
2. Bapak Dr. RB. Fajriya Hakim, M.Si selaku Kepala Jurusan Statistika beserta seluruh jajarannya.
3. Ibu Ayundyah Kesumawati, S.Si., M.Si selaku Dosen Pembimbing penulis yang telah memberikan kritik dan saran bimbingan maupun arahan yang sangat berguna dalam penyusunan Tugas Akhir ini.
4. Ibu Kariyam, M.Si selaku DPA penulis yang selalu memberikan saran maupun arahan yang sangat berguna.

5. Dosen-dosen Statistika Universitas Islam Indonesia yang telah mendidik dan berbagi ilmunya kepada penulis.
6. Kedua orang tua tercinta Bapak Sudirman dan Ibu Emmah yang selalu memberikan semangat dan dukungan tiada henti serta selalu mengiringi lewat untaian doa disetiap langkah penulis.
7. Adik penulis Mardatillah yang selalu memberikan semangat yang membara bagi penulis.
8. Sahabat seperjuangan dari umur 12 tahun sampai seterusnya Canda Hartina yang selalu ada.
9. Sahabat seperjuangan Vhina, Rahmi, Muni, dan sekar yang selalu memberi semangat dan dukungan yang luar biasa.
10. Sahabat seperjuangan Lira, Welly, Aisyah, Yumu, Novi, Mita, dan Ayu yang selalu memberi semangat dalam menyelesaikan Tugas Akhir ini.
11. Teman-teman IKA Jogja yang selalu memberi semangat dukungan dan senantiasa berbagi cerita.
12. Teman-teman AKSEL yang selalu berbagi cerita dan pengalaman.
13. Teman- Teman satu bimbingan tugas akhir Siti Rahmi Kurniasari, Maulida Jannati Wulansari, Suci Yulianawati, Herlina Catur Sulistya Ningrum, Riza Indriani Rakhmalia, Afifah Mukhtaroh, Reny Roswita Nazar, Zia Ayu Nuansa Gumilang, Molydah S, Gustiara Dayu Amrinda, Elsa Murni Nasution, Dian Purnama Sari, Purwina Kowala, Galih Rahmatdona Sugito, Yayan Dwi Septian, Achmad Kurniansyah Thalib, dan Syauqi Amri Yahya yang senantiasa selalu berbagi ilmu, cerita dan pengalaman.
14. Sahabat Statistika 2014 (XIX) yang sudah banyak memberikan semangat kepada penulis dalam penyusunan Tugas Akhir ini.
15. Teman-teman KKN unit 86 Desa Bulus Kabupaten Purworejo Wintang Nanas, Yuni, Novi, Hari, Hendy, Teddy, dan Alam.
16. Teman-teman Statistika 2014 yang sudah banyak memberikan semangat dan bantuan dalam memulai dan mengakhiri tugas kerja praktek ini.
17. Semua pihak yang tidak dapat penyusun sebutkan satu per satu terima kasih.

Penyusun menyadari sepenuhnya bahwa laporan Tugas Akhir ini masih jauh dari sempurna, oleh karena itu segala kritik dan saran yang sifatnya membangun selalu penyusun harapkan. Semoga laporan ini dapat bermanfaat bagi penyusun khususnya dan bagi semua yang membutuhkan umumnya. Akhir kata semoga Allah SWT selalu melimpahkan rahmat serta hidayah-Nya kepada kita semua, amin ya robbal 'alamiin

Wassalamu'alaikum, Wr.Wb .

Yogyakarta, 17 April 2017

Penyusun

DAFTAR ISI

KATA PENGANTAR	vi
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiv
DAFTAR ISTILAH	xv
DAFTAR LAMPIRAN	xvi
PERNYATAAN	Error! Bookmark not defined.
INTISARI	xviii
ABSTRAK	xix
<u>BAB I</u> PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.3. Batasan Masalah.....	4
1.4. Tujuan Penelitian.....	5
1.5. Manfaat Penelitian.....	5
<u>BAB II</u> KAJIAN PUSTAKA	6
2.1 Kajian Pustaka	6
<u>BAB III</u> LANDASAN TEORI	10
3.1. Wisatawan Mancanegara.....	10
3.2. Analisis Deskriptif.....	11
3.3. Pengkategorian Variabel Penelitian	12
3.4. <i>Forecasting</i> / Peramalan	13
3.5. Uji Normalitas	14
3.6. Metode SARIMA	15
3.7. Metode <i>Support Vector Machine</i> (SVM)	15
3.8. <i>Mean Absolute Percentage Error</i> (MAPE).....	17
<u>BAB IV</u> METODOLOGI PENELITIAN	18
4.1. Populasi dan Sampel Penelitian.....	18
4.2. Jenis dan Sumber Data	18
4.3. Cara Pengambilan Data	18

4.4.	Variabel dan Definisi Operasional Variabel.....	19
4.5.	Metode Analisis Data	19
4.6.	Proses Analisis Data.....	20
BAB V HASIL DAN PEMBAHASAN		22
5.1.	Analisis deskriptif.....	22
5.2.	Kategorisasi Data.....	23
5.3.	<i>Preprocessing</i> Data	24
5.3.1.	Uji Normalitas.....	24
5.3.2.	Pembuatan Data <i>Training</i> dan Data <i>Testing</i>	25
5.4.	Identifikasi Data dengan Plot <i>Time Series</i>	26
5.4.1.	<i>Time Series</i> Plot Pintu Masuk Kategori Tinggi	26
5.4.2.	<i>Time Series</i> Plot Pintu Masuk Kategori Sedang	26
5.4.3.	<i>Time Series</i> Plot Pintu Masuk Kategori Rendah.....	27
5.5.	Penerapan Metode SARIMA.....	27
5.5.1	Pintu Masuk dengan Kategori Tinggi.....	27
a.	Uji Stasioner	27
b.	Analisis Model SARIMA	30
c.	<i>Mean Absolut Percentage Error</i> (MAPE)	31
d.	Hasil Prediksi Jumlah Wisatawan	32
5.5.2	Pintu Masuk dengan Kategori Sedang.....	32
a.	Uji Stasioner	32
b.	Analisis Model SARIMA	35
c.	<i>Mean Absolut Percentage Error</i> (MAPE)	36
d.	Hasil Prediksi Jumlah Wisatawan	37
5.5.3	Pintu Masuk dengan Kategori Rendah.....	37
a.	Uji Stasioner	37
b.	Analisis Model SARIMA	40
5.6.	Metode <i>Support Vector Machine</i> (SVM)	43
5.6.1.	Pencarian Model SVM terbaik.....	43
a.	Pintu Masuk dengan Kategori Tinggi.....	44
b.	Pintu Masuk dengan Kategori Sedang	46

c. Pintu Masuk dengan Kategori Rendah	48
BAB VI KESIMPULAN DAN SARAN.....	53
6.1. Kesimpulan.....	53
6.2. Saran.....	53
DAFTAR PUSTAKA.....	55
LAMPIRAN	

DAFTAR TABEL

Nomor	Judul	Halaman
1.	Penelitian Terdahulu	6
2.	Aturan Pengkategorisasian	11
3.	Fungsi Kernel dalam SVM	13
4.	Tingkat Akurasi Hasil Peramalan Berdasarkan Nilai MAPE	15
5.	Definisi operasional variabel	19
6.	Hasil Kategorisasi 19 Pintu Masuk di Indonesia	23
7.	Hasil Uji Normalisasi dari 3 Kategori	25
9.	Pembagian Data <i>Training</i> dan Data <i>Testing</i>	25
10.	Output Uji Kestasioneran Data Musiman	28
11.	Output Uji Kestasioneran Data dalam Rata-rata	28
12.	Output Uji Kestasioneran Data dalam Rata-rata	29
13.	Hasil Pengujian Model	31
14.	Hasil Uji Diagnostik	31
15.	Hasil Prediksi	32
16.	Output Uji Kestasioneran Data Musiman	33
17.	Output Uji Kestasioneran Data dalam Rata-rata	33
18.	Output Uji Kestasioneran Data dalam Rata-rata	34
19.	Hasil Pengujian Model	36
20.	Hasil Uji Diagnostik	36
21.	Hasil Prediksi	37
22.	Output Uji Kestasioneran Data Musiman	38
23.	Output Uji Kestasioneran Data dalam Rata-rata	38
24.	Output Uji Kestasioneran Data dalam Rata-rata	39
25.	Hasil Pengujian Model	41
26.	Hasil Uji Diagnostik	42
27.	Hasil Prediksi	43
28.	Nilai Parameter Variasi Fungsi Kernel pada Data <i>Training</i>	44

	Kategori Tinggi	
29.	Nilai MAPE dari Variasi Fungsi Kernel untuk Kategori Tinggi	44
30.	Hasil Prediksi Pintu Masuk dengan Kategori Tinggi	45
31.	Nilai Parameter Variasi Fungsi Kernel pada Data <i>Training</i> Kategori Sedang	46
32.	Nilai MAPE dari Variasi Fungsi Kernel untuk Kategori Sedang	47
33.	Hasil Prediksi Pintu Masuk untuk Kategori Sedang	48
34.	Nilai Parameter Variasi Fungsi Kernel pada Data <i>Training</i> Kategori Sedang	49
35.	Nilai MAPE dari Variasi Fungsi Kernel untuk Kategori Rendah	49
36.	Hasil Prediksi Pintu Masuk untuk Kategori Rendah	51

DAFTAR GAMBAR

Nomor	Judul	Halaman
1.	Perbandingan Pariwisata dengan Komoditi Ekspor lainnya Tahun 2013 – 2016.	2
2.	<i>Flow Chart Analisis Data</i>	20
3.	Jumlah Wisatawan Mancanegara Periode 2010-2017	22
4.	<i>Plot Pintu Masuk dengan Kategori Tinggi</i>	26
5.	<i>Plot Pintu Masuk dengan Kategori Sedang</i>	26
6.	<i>Plot Pintu Masuk dengan Kategori Rendah</i>	27
7.	Plot ACF dan PACF	30
8.	Plot ACF dan PACF	35
9.	Plot ACF dan PACF	40
10.	<i>Plotting Grafik Data Testinging Kategori Tinggi</i>	45
11.	<i>Plotting Grafik Data Testinging Kategori Sedang</i>	48
12.	<i>Plotting Grafik Data Testinging Kategori Rendah</i>	50

DAFTAR ISTILAH

SVM	: <i>Support Vector Machine</i>
MSE	: <i>Mean Square Error</i>
MAPE	: <i>Mean Absolut Percentage Error</i>

DAFTAR LAMPIRAN

No	Judul	Hal
1.	Data Jumlah Wisatawan Setiap Pintu Tahun 2010	58
2.	Data Jumlah Wisatawan Setiap Pintu Tahun 2011	59
3.	Data Jumlah Wisatawan Setiap Pintu Tahun 2012	60
4.	Data Jumlah Wisatawan Setiap Pintu Tahun 2013	61
5.	Data Jumlah Wisatawan Setiap Pintu Tahun 2014	62
6.	Data Jumlah Wisatawan Setiap Pintu Tahun 2015	63
7.	Data Jumlah Wisatawan Setiap Pintu Tahun 2016	64
8.	Data Jumlah Wisatawan Setiap Pintu Tahun 2017	65
9.	Hasil Uji Normalitas Untuk Pintu dengan Kategori Tinggi	66
10.	Hasil Uji Normalitas Untuk Pintu dengan Kategori Sedang	66
11.	Hasil Uji Normalitas Untuk Pintu dengan Kategori Rendah	66
12.	Hasil Kategorisasi dengan SPSS	67
13.	Hasil Defferensiasi Kategori Tinggi	68
14.	Hasil Defferensiasi Kategori Sedang	68
15.	Hasil Defferensiasi Kategori Rendah	69

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang sebelumnya pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan dalam daftar pustaka.

Yogyakarta, 17 April 2018



Molydah s

**ANALISIS PERBANDINGAN IMPLEMENTASI SARIMA DAN
SUPPORT VECTOR MACHINE (SVM) DALAM PREDIKSI
JUMLAH WISATAWAN MANCANEGERA**

(Studi Kasus : Jumlah Wisatawan yang Datang ke Indonesia Berdasarkan 19
Pintu Masuk Periode Tahun 2010-2017)

Oleh : Molydah S
Program Studi Statistika Fakultas MIPA
Universitas Islam Indonesia

INTISARI

Pariwisata merupakan salah satu sektor ekonomi terbesar dan merupakan sumber pendapatan utama bagi banyak negara di dunia. Negara Indonesia merupakan salah satu negara yang menjadikan sektor pariwisata sebagai penunjang pertumbuhan ekonomi di mana pada tahun 2017 penerimaan devisa dari sektor pariwisata sudah mengalahkan minyak bumi dan gas. Kekayaan alam dan budaya yang dimiliki Indonesia tersebar secara berlimpah sehingga berpotensi untuk menjadikan setiap wilayah Indonesia memiliki objek wisata yang menarik wisatawan. Jumlah wisatawan yang datang ke Indonesia sewaktu-waktu dapat berubah dan berdampak bagi suatu kondisi tertentu. Upaya untuk meminimalisir jumlah wisatawan yang tidak tentu tersebut, maka dalam penelitian ini akan diramalkan jumlah wisatawan mancanegara yang datang ke Indonesia berdasarkan 3 kategori pintu masuk tinggi, sedang, dan rendah. Pada penelitian ini digunakan metode SARIMA dan SVM dengan variasi fungsi kernel untuk meramalkan jumlah wisatawan mancanegara yang datang ke Indonesia dari 3 kategori pintu masuk yaitu tinggi, sedang, dan rendah. Hasil yang diperoleh pada penelitian ini menunjukkan bahwa pada kategori tinggi metode SVM memiliki nilai MAPE yang lebih kecil sebesar 3,45%, sedangkan pada kategori sedang dan rendah metode SARIMA memiliki nilai MAPE yang lebih kecil sebesar 7,07% dan 9,59% dibanding metode SVM.

Kata kunci: Peramalan, Wisatawan Mancanegara, SARIMA, SVM, MAPE.

**COMPARATIVE ANALYSIS OF SARIMA AND IMPLEMENTATION
SUPPORT VECTOR MACHINE (SVM) IN PREDICTION
TOTAL FOREIGNER TOURIST**

(Case Study: Number of Foreign Tourists Arrival To Indonesia Based on 19
Entrance Year 2010-2017)

By: Molydah S
Departement of Statistics Faculty Mathematics and Natural Sciences
Islamic University of Indonesia

ABSTRAK

Tourism is one of the largest economic sectors and is a major source of income for many countries in the world. The country of Indonesia is one of the countries that make the tourism sector as a supporter of economic growth in which in 2017 foreign exchange earnings from the tourism sector has beaten the oil and gas. Indonesia's natural wealth and culture are spread abundantly so it has the potential to make every region of Indonesia has a tourist attraction that attracts tourists. The number of tourists who come to Indonesia at any time may change and affect for a certain condition. Efforts to minimize the number of tourists that are not necessarily. then in this study will be predicted the number of foreign tourists who come to Indonesia based on 3 categories of high, medium, and low entrances. In this research used SARIMA and SVM method with kernel function variation to predict number of foreign tourists coming to Indonesia from 3 categories of entrance that is high, medium, and low. The results obtained in this study indicate that in the high category of SVM method has a smaller MAPE value of 3.45%, while in the medium and low category SARIMA method has a smaller MAPE value of 7.07% and 9.59% SVM method.

Keywords: Forecasting, foreigner tourist, SARIMA, SVM, MAPE.

BAB I

PENDAHULUAN

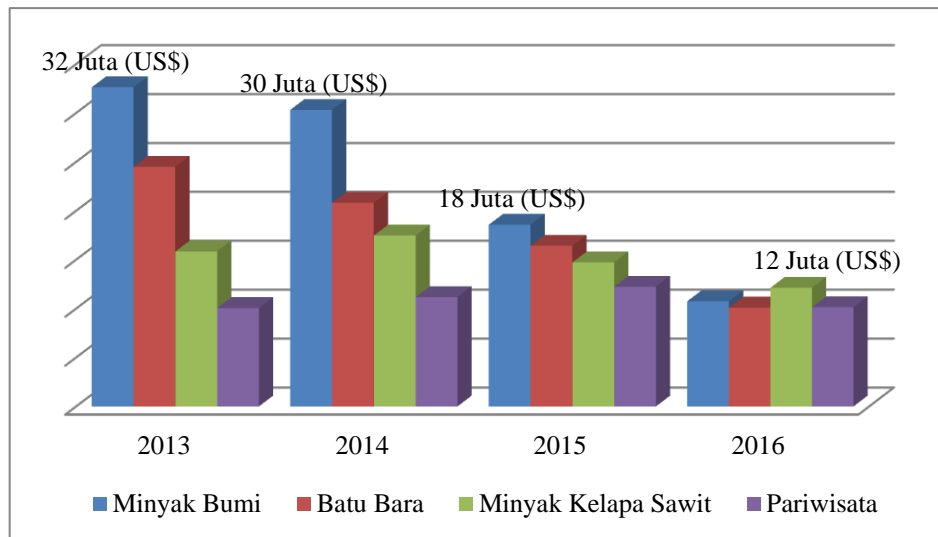
1.1. Latar Belakang

Pada era globalisasi saat ini, sektor pariwisata merupakan salah satu sektor ekonomi terbesar dan mempunyai tingkat pertumbuhan paling pesat dan menjadi salah satu sumber utama pendapatan bagi banyak negara di dunia. Berdasarkan laporan *World Travel and Tourism Council (WTTC)* pada tahun 2016 sumbangan pariwisata terhadap produk domestik bruto (PDB) dunia mencapai 10% sementara penerimaan dari kunjungan wisatawan internasional menyumbang 7% dari total ekspor barang dan jasa dunia. Dilihat dari penciptaan lapangan kerja, satu dari 10 tenaga kerja diciptakan oleh sektor pariwisata (Lestari dkk, 2017).

Indonesia merupakan salah satu negara di ASEAN yang memiliki banyak destinasi wisata yang patut diperhitungkan baik di tingkat regional maupun dalam kancah internasional (Indrasetianingsih dkk, 2017). Berdasarkan laporan akuntabilitas kinerja kementerian pariwisata tahun 2016 bahwa kunjungan wisatawan mancanegara ke Indonesia mengalami pertumbuhan yang cukup drastis sebesar 15,54% jika dibandingkan dengan beberapa negara tetangga ASEAN seperti Negara Thailand yaitu sebesar 9,7% periode Januari-November 2016, Singapura 7,9% periode Januari-November 2016 dan Malaysia 4,4% periode Januari-Oktober 2016.

Pertumbuhan kunjungan wisatawan mancanegara ke Indonesia memberikan dampak positif. mengingat pariwisata merupakan salah satu sektor yang menjadi motor penggerak dalam pertumbuhan ekonomi Negara Indonesia. Dilansir dalam *travel detik.com* deputy Bidang Pengembangan Kelembagaan Kepariwisataan RI mengungkapkan bahwa di tahun 2017 devisa dari sektor pariwisata sudah mengalahkan minyak bumi dan gas, sedangkan untuk urutan

pertama sebagai penyumbang terbesar masih dikuasai oleh kelapa sawit. Berikut perbandingan antara penerimaan devisa pariwisata dengan komoditi ekspor lainnya dari tahun 2013-2016 adalah sebagai berikut :



Sumber : Laporan Kementerian Pariwisata RI 2016

Gambar 1.1 Grafik Perbandingan Pariwisata dengan Komoditi Ekspor lainnya Tahun 2013–2016

Perkembangan sektor pariwisata ditargetkan menjadi mesin pertumbuhan ekonomi baru di Indonesia. Pada tahun 2019 kementerian Pariwisata menargetkan jumlah wisatawan mancanegara yang berkunjung ke Indonesia sebesar 20 juta. Pencapaian target tersebut melalui Kementerian Pariwisata dan Ekonomi Kreatif Indonesia gencar mempromosikan diri sebagai tujuan wisata untuk turis-turis asing dengan kampanye "*Wonderful Indonesia*". Berinvestasi dalam kampanye-kampanye promosi seperti ini tidak lain untuk menyebarkan citra positif Indonesia ke negara luar guna mendorong minat wisatawan untuk berkunjung.

Potensi besar yang mampu dihasilkan oleh sektor pariwisata mengakibatkan beberapa negara di dunia melakukan kompetisi untuk mengembangkan pariwisatanya. Perkembangan pariwisata ditentukan oleh banyaknya kunjungan wisatawan baik wisatawan nusantara maupun mancanegara. Perkembangan pariwisata juga menyangkut dengan keputusan sektor publik dan swasta seperti infrastruktur, transportasi, akomodasi, promosi, dan layanan lainnya. Apabila sektor pariwisata dapat tumbuh dan berkembang dengan baik,

tentu saja akan menjadi penunjang lajunya pembangunan suatu wilayah. Mengingat potensi besar yang mampu dihasilkan oleh sektor pariwisata, maka sangat perlu dilakukan studi yang mendalam sehubungan dengan peramalan kunjungan wisatawan mancanegara dan domestik pada kurun waktu tertentu dalam rangka memprediksi besaran kontribusi sektor pariwisata pada pendapatan suatu daerah. “Menurut Herawati (2016) peramalan juga dibutuhkan untuk perencanaan kebijakan yang efektif”.

Pada penelitian sebelumnya beberapa metode telah dikembangkan dalam meramalkan jumlah wisatawan di masa mendatang mulai dari metode yang sederhana sampai metode yang kompleks. Seringkali dalam peramalan data *time series* menunjukkan perilaku yang bersifat musiman. Musiman didefinisikan sebagai kecenderungan data *time series* yang berulang setiap periode. Pada umumnya data jumlah wisatawan mancanegara mengandung pola musiman yang bergerak pada beberapa periode setiap tahun, dimana pada musim–musim tertentu seperti liburan sekolah, tahun baru, liburan musim panas jumlah kunjungan wisatawan di berbagai tempat pariwisata lebih meningkat daripada hari-hari biasa. Hal tersebut mengindikasikan bahwa jumlah kunjungan wisatawan mengandung pola musiman. Metode peramalan yang dapat menangani pola musiman yaitu metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Hal tersebut diperkuat dengan penelitian yang pernah dilakukan oleh Ainy (2015) yaitu meramalkan jumlah kunjungan wisatawan mancanegara di kabupaten lombok tengah pada tahun 2010-2015 menggunakan metode SARIMA. Penelitian tersebut menyimpulkan bahwa metode SARIMA merupakan analisis terbaik untuk data *time series* terutama *seasonal*.

Pada peramalan jumlah wisatawan dimasa mendatang juga dapat diterapkan metode kecerdasan buatan yang relatif baru dalam prediksi yaitu metode *Support Vector Machine* (SVM). Metode SVM baik digunakan dalam kasus klasifikasi maupun regresi. Metode SVM adalah salah satu dari sekian banyak metode yang dapat digunakan untuk menyelesaikan berbagai jenis permasalahan termasuk peramalan. Penyelesaian sebuah permasalahan, metode SVM mampu menangani permasalahan nonlinier dengan adanya fungsi kernel

yang membuat metode ini dapat digunakan untuk peramalan *time series* (Naufal, 2017). Penerapan metode SVM yang relatif baru dalam prediksi dilakukan peneliti untuk dijadikan sebagai pembanding dari metode SARIMA. Selain itu, metode SVM ini juga sudah banyak digunakan dalam menyelesaikan persoalan *time series* seperti pada penelitian yang dilakukan oleh Naufal (2017) yaitu meramalkan jumlah wisatawan mancanegara berdasarkan 6 pintu masuk dengan menggunakan metode SVM, didapatkan rata-rata akurasi dari 6 model memiliki MAPE 10% yang artinya model peramalannya tergolong baik.

Berdasarkan latar belakang diatas, maka peneliti tertarik mengambil penelitian dengan tema yang berjudul “**Analisis Perbandingan Implementasi SARIMA dan *Support Vector Machine* (SVM) dalam Prediksi Jumlah Wisatawan Mancanegara**”.

1.2. Rumusan Masalah

Berdasarkan latar belakang permasalahan di atas maka rumusan masalah yang dapat diidentifikasi dalam penelitian ini adalah :

1. Bagaimana penerapan metode SARIMA dalam meramalkan jumlah wisatawan mancanegara yang datang ke Indonesia melalui 3 kategori pintu masuk ?
2. Bagaimana penerapan metode SVM dalam meramalkan jumlah wisatawan mancanegara yang datang ke Indonesia melalui 3 kategori pintu masuk ?
3. Bagaimanakah hasil perbandingan ramalan jumlah wisatawan mancanegara yang datang ke Indonesia melalui 3 kategori pintu masuk untuk beberapa periode kedepan dengan menggunakan metode SARIMA dan SVM ?

1.3. Batasan Masalah

Batasan masalah dilakukan agar penelitian yang dilakukan tidak melebar. Penelitian ini terfokus pada jumlah wisatawan mancanegara yang masuk ke Indonesia melalui 19 pintu periode tahun Januari 2010 – November 2017.

1.4. Tujuan Penelitian

Berdasarkan permasalahan di atas maka tujuan penelitian ini adalah:

1. Mengetahui penerapan metode SARIMA dalam meramalkan jumlah wisatawan mancanegara yang datang ke Indonesia melalui 3 kategori pintu masuk.
2. Mengetahui penerapan metode SVM dalam meramalkan jumlah wisatawan mancanegara yang datang ke Indonesia melalui 3 kategori pintu masuk.
3. Mengetahui hasil perbandingan ramalan jumlah wisatawan mancanegara yang masuk ke Indonesia melalui 3 kategori pintu masuk untuk beberapa periode kedepan menggunakan metode SARIMA dan SVM.

1.5. Manfaat Penelitian

Adapun manfaat yang dapat diberikan dari tugas akhir ini adalah:

1. Memberikan informasi yang akurat kepada pihak Badan Pusat Statistik dan Dinas Pariwisata serta industri pariwisata yang ada pada setiap daerah mengenai nilai ramalan jumlah wisatawan mancanegara yang datang ke Indonesia pada periode ke depan untuk setiap pintu masuk.
2. Pihak-pihak terkait dapat melakukan perencanaan dan pengambilan keputusan terkait upaya peningkatan jumlah wisatawan yang berdampak pada keuntungan.
3. Selain itu, membantu industri pariwisata dari tiap daerah untuk membangun strategi dalam meningkatkan keuntungan seperti menambah dan mengembangkan tempat-tempat wisata dari setiap daerah yang dapat menarik wisatawan mancanegara untuk berkunjung dimasa mendatang.

BAB II

KAJIAN PUSTAKA

2.1. Kajian Pustaka

Dalam mencapai tujuan penelitian yang di paparkan pada bab sebelumnya. hal yang harus dilakukan pertama kali oleh penulis adalah melakukan kajian pustaka atau studi literatur dari beberapa jurnal terdahulu. Hasil dari bab kajian pustaka ini akan menjadi dasar bagi peneliti dalam menyelesaikan laporan penelitian. Berikut jurnal yang mendasari penelitian ini:

Tabel 2.1 Penelitian Terdahulu

Judul :	Peramalan Jumlah Kunjungan Wisatawan Mancanegara Di Kabupaten Lombok Tengah Pada Tahun 2010-2015 Menggunakan Metode Sarima
Penulis; Tahun:	Sofyani Ramdhatul Ainy ; 2015
Deskripsi Umum Penelitian :	Penelitian ini bertujuan untuk menentukan model SARIMA terbaik dalam meramalkan jumlah kunjungan wisatawan asing di Kabupaten Lombok Tengah menggunakan metode SARIMA. Pada penelitian ini dikatakan metode SARIMA merupakan analisis terbaik untuk data <i>time series</i> terutama <i>seasonal</i> . Hasil penelitian diperoleh bahwa data wisatawan asing yang berkunjung ke Kabupaten Lombok Tengah mengandung pola tren yang terus meningkat setiap tahun. Model terbaik adalah SARIMA (2,1,3) (1,1,0) ¹²
Judul:	Perbandingan Beberapa Metode <i>Time Series</i> Pada Peramalan Jumlah Kunjungan Wisatawan Mancanegara: Studi Kasus Di Kabupaten Karimun. Provinsi Kepulauan Riau.

Penulis;Tahun:	Erie Sadewo; 2013
Deskripsi Umum Penelitian :	Penelitian ini mengenai peramalan jumlah kunjungan wisatawan di Kabupaten Karimun menggunakan beberapa metode <i>time series</i> . Berdasarkan pengujian performa ketepatan hasil peramalan didapati bahwa model DMA dan ARIMA (4,1,0) merupakan yang paling baik digunakan dalam peramalan.
Judul :	Peramalan Jumlah Wisatawan Mancanegara Yang Datang Ke Indonesia Berdasarkan Pintu Masuk Menggunakan Metode Support Vector Machine (SVM)
Penulis; Tahun:	Muhammad Farhan Naufal ; 2017
Deskripsi Umum Penelitian :	Pada penelitian ini, digunakan metode SVM untuk meramalkan jumlah wisatawan mancanegara yang datang ke Indonesia dari 6 pintu masuk. Berdasarkan penelitian tersebut didapatkan rata-rata akurasi dari 6 model memiliki MAPE sekitar 10%, sehingga dapat disimpulkan bahwa model peramalan secara keseluruhan tergolong baik.
Judul :	<i>Hybridizing GMDH And Least Squares SVM Support Vector Machine For Forecasting Tourism Demand</i>
Penulis; Tahun:	Ruhaidah Samsudin. Puteh Saad & Ani Shabri ; 2010
Deskripsi Umum Penelitian	Penelitian ini adalah untuk menguji kelayakan model <i>hybrid</i> dalam peramalan jumlah wisatwan ke Johor Malaysia periode tahun 1970 sampai 2008 dengan membandingkannya dengan model GMDH dan LSSVM. Penelitian ini menggunakan metode <i>Group of Method Data Handling</i> yang dikombinasikan dengan <i>Least Squares Support Vector Machine</i> (LSSVM). Pengkombinasian kedua metode tersebut dinamakan GLSSVM

	dalam melakukan peramalan. Berdasarkan penelitian tersebut didapatkan bahwa metode GLLSVM memiliki hasil yang cukup akurat atau hanya terjadi kesalahan ramalan sebesar 0,05 % pada model yang diperoleh.
Judul :	Peramalan Beban Listrik Bulanan Sektor Industri Menggunakan Support Vector Machine dengan Variasi Fungsi Kernel
Penulis ;Tahun:	Luqman Assaffat ; 2015
Deskripsi Umum Penelitian :	Penelitian ini menyajikan peramalan beban listrik bulanan di sektor industri menggunakan SVM dengan berbagai fungsi Kernel. Ada 3 jenis data pelatihan yang aplikasikan pada SVM, diantaranya tipe bulan dalam setahun, data produksi dan data deret waktu setelah beban listrik. Adapun variasi fungsi Kernel yang digunakan pada penelitian ini adalah fungsi kernel Linear, kuadratik, Gaussian RBF, polinomial dan Multilayer Perceptron. Pada penelitian ini dihasilkan MAPE peramalan terkecil sebesar 5,33% dengan Fungsi kernel Gaussian RBF.
Judul :	Perbaikan Peramalan Produksi Padi di Kabupaten Kendal dengan Menggunakan Metode SVM
Penulis;Tahun :	Ayu Andita dan Wellie Sulistijanti ; 2018
Deskripsi Umum Penelitian :	Penelitian ini menggunakan data produksi padi dari bulan Januari 2013. Peramalan dinyatakan baik jika memiliki nilai Mean Square Error (MSE) yang kecil. Peramalan pada produksi padi dengan metode Seasonal Autoregressive Integrated Moving Average (SARIMA) diperoleh MSE training sebesar 0,043918 dan MSE testing 77,118 terlihat bahwa nilai MSE antara keduanya sangat jauh berbeda. Adapun nilai MSE yang diperoleh dengan metode (SVM) untuk data <i>training</i> sebesar

	<p>0,14 dan MSE untuk data testing sebesar 0,57. Terlihat bahwa nilai MSE yang dihasilkan sangat dekat tidak jauh berbeda dengan keduanya.</p>
Penelitian Terbaru:	<p>Pada penelitian ini. akan diramalkan jumlah wisatawan mancanegara yang datang ke Indonesia berdasarkan 3 kategori pintu masuk dengan membandingkan metode SARIMA dan metode <i>Support Vector Machine</i> (SVM) dengan beberapa variasi fungsi kernel diantaranya linear. polynomial. dan RBF. Penentuan model terbaik dilihat dari nilai akurasi terbaik berdasarkan nilai rata-rata MAPE terkecil.</p>

BAB III

LANDASAN TEORI

3.1. Wisatawan Mancanegara

Menurut *United Nation World Tourism Organization* (UNWTO), wisatawan mancanegara adalah setiap orang yang berkunjung ke suatu negara di luar tempat tinggalnya untuk beberapa keperluan tanpa bermaksud memperoleh penghasilan di tempat yang dikunjungi dan lamanya kunjungan tersebut tidak lebih dari 12 (dua belas) bulan. Definisi ini mencakup dua kategori tamu mancanegara yaitu :

- a. Wisatawan (*tourist*) adalah setiap pengunjung yang datang ke suatu tempat dan tinggal di tempat tersebut tidak lebih dari dua belas 12 bulan dengan maksud kunjungan antara lain : berlibur, rekreasi dan olahraga bisnis, mengunjungi teman dan keluarga, misi, menghadiri pertemuan, konferensi, kunjungan dengan alasan kesehatan, belajar, dan keagamaan.
- b. Pelancong (*excursionist*) adalah setiap pengunjung seperti definisi di atas yang tinggal ditempat tersebut kurang dari dua puluh empat jam di tempat yang dikunjungi (termasuk *cruise passenger* yaitu setiap pengunjung yang tiba di suatu tempat dengan menggunakan kapal atau kereta api, dan mereka tidak menginap di akomodasi yang tersedia di negara tersebut) (BPS: Statistik Kunjungan Wisatawan Mancanegara, 2016).

Definisi wisatawan menurut Norval adalah setiap orang yang datang dari suatu negara yang alasannya bukan untuk menetap atau bekerja di situ secara teratur dan yang di negara di mana ia tinggal untuk sementara itu membelanjakan uang yang didapatkannya di lain tempat, sedangkan menurut Soekadijo wisatawan adalah pengunjung di negara yang dikunjunginya setidaknya-tidaknya tinggal dalam kurun waktu 24 jam dan datang berdasarkan motivasi sebagai berikut (Astuti, 2016).

- a. Mengisi waktu senggang atau untuk bersenang-senang, berlibur, untuk alasan kesehatan, studi, keluarga, dan sebagainya.
- b. Melakukan perjalanan untuk keperluan bisnis.
- c. Melakukan perjalanan untuk mengunjungi pertemuan-pertemuan atau sebagai utusan (ilmiah, administrative, diplomatic, keagamaan, olahraga dan sebagainya).
- d. Dalam rangka pelayaran pesiar, jika kalau ia tinggal kurang dari 24 jam.

3.2. Analisis Deskriptif

Analisis deskriptif adalah suatu cara untuk menggambarkan atau mendeskripsikan persoalan berdasarkan data yang dimiliki yakni dengan cara menata data tersebut sedemikian rupa, sehingga dengan mudah dapat dipahami tentang karakteristik data, dijelaskan dan berguna untuk keperluan selanjutnya. Jadi dalam hal ini terdapat aktivitas atau proses pengumpulan data dan pengolahan data berdasarkan tujuannya.

Menurut Nurjannah (2016) statistik deskriptif merupakan bidang ilmu statistika yang mempelajari cara-cara pengumpulan, penyusunan, dan penyajian data suatu penelitian. Kegiatan yang termasuk dalam kategori tersebut adalah kegiatan *collecting* atau pengumpulan data, *grouping* atau pengelompokan data, penentuan nilai dan fungsi statistik, serta yang terakhir termasuk pembuatan grafik dan gambar.

Pada umumnya terdapat dua metode yang biasa digunakan untuk menjelaskan suatu data, yaitu:

1. Tabel

Penyajian data dalam bentuk tabel bertujuan untuk mengelompokkan nilai-nilai pengamatan ke dalam beberapa kelompok yang masing-masing mempunyai karakteristik yang sama. Bentuk tabel yang sering digunakan adalah tabel distribusi frekuensi, tabel distribusi frekuensi relatif dan tabel kontingensi untuk data kualitatif dengan banyak kategori dalam baris maupun kolom.

2. Grafik atau diagram

Penyajian data dalam bentuk grafik atau diagram bertujuan untuk memvisualisasikan data secara keseluruhan dengan menonjolkan karakteristik-

karakteristik tertentu dari data tersebut. Jenis grafik atau diagram yang akan digunakan pada penelitian ini adalah diagram batang.

- a. Diagram Batang
- b. Grafik Garis

3.3. Pengkategorian Variabel Penelitian

Penelitian ini menggunakan 2 variabel yaitu variabel periode waktu dalam bulan dan jumlah wisatawan mancanegara yang datang ke Indonesia melalui 19 pintu masuk periode tahun 2010 - 2017. Dari 19 pintu masuk tersebut akan dikategorisasikan menjadi 3 yaitu tinggi, sedang, dan rendah. Pengkategorian tersebut dilakukan berdasarkan jumlah wisatawan yang kedatangan ke Indonesia dari setiap pintu. Sebelum masuk ke tahap pengkategorisasian terlebih dahulu peneliti melakukan uji normalitas. Data yang tidak berdistribusi normal, pengelompokannya didasarkan pada nilai kuartil (Q_1 , Q_2 , Q_3) (Fatmawati, 2013). Berikut rumus yang digunakan dalam mencari nilai Q_1 , Q_2 , dan Q_3 untuk data tunggal.

Untuk n ganjil

$$Q_i = \frac{i(n+1)}{4} \dots \dots \dots (1)$$

Untuk n genap

$$Q_1 = x_{\frac{n+2}{4}} \dots \dots \dots (2)$$

$$Q_2 = 1/2(x_{n/2} + x_{\frac{n}{2}+1}) \dots \dots \dots (3)$$

$$Q_3 = x_{(3n+2)/4} \dots \dots \dots (4)$$

Dimana :

Q_i = Kuartil ke- i

n = Banyak Data

Data jumlah wisatawan yang akan dikategorikan merupakan data tunggal dan berjumlah 19, sehingga dalam mencari nilai kuartil peneliti menggunakan rumus kuartil untuk n ganjil. Berikut aturan yang berlaku pada pengkategorisasian pintu masuk.

Tabel 3.1 Aturan Pengkategorisasian

Total Wisatawan	Kategori
Total Wisatawan < Q_1	Rendah
$Q_1 < \text{Total Wisatawan} < Q_3$	Sedang
$Q_3 > \text{Total Wisatawan}$	Tinggi

Tabel diatas merupakan aturan pengkategorisasian pintu masuk. Jika total wisatawan lebih kecil dari nilai Q_1 maka akan masuk ke kategori rendah, dan jika nilai Q_1 lebih kecil dari total wisatawan namun lebih besar dari nilai Q_3 maka akan masuk kedalam kategori sedang, dan apabila total wisatawan lebih besar dari nilai Q_3 maka akan masuk kedalam kategori tinggi. Pada penelitian ini, untuk mengetahui kategorisasi dari setiap pintu dilihat berdasarkan hasil SPSS.

3.4. Forecasting / Peramalan

Peramalan merupakan suatu dugaan terhadap permintaan yang akan datang berdasarkan pada beberapa variabel peramal. sering berdasarkan data deret waktu historis” (Febrina, Arina, & Ekawati , 2013). Menurut Gaspersz (2002) peramalan adalah metode untuk memperkirakan suatu nilai dimasa depan dengan menggunakan data masa lalu. Peramalan juga dapat diartikan sebagai seni dan ilmu untuk memperkirakan kejadian di masa yang akan datang, sedangkan aktivitas peramalan merupakan suatu fungsi bisnis yang berusaha memperkirakan penjualan dan penggunaan suatu produk sehingga produk-produk itu dapat dibuat dalam kuantitas yang tepat.

Menurut Assauri (1984) peramalan yang baik adalah peramalan yang dilakukan dengan mengikuti langkah-langkah atau prosedur penyusunan yang baik yang akan menentukan kualitas atau mutu dari hasil peramalan yang disusun. Pada dasarnya ada 3 langkah peramalan yang penting yaitu:

1. Menganalisa data yang lalu, tahap ini berguna untuk pola yang terjadi pada masa lalu.
2. Menentukan data yang dipergunakan, Metode yang baik adalah metode yang memberikan hasil ramalan yang tidak jauh berbeda dengan kenyataan yang terjadi.
3. Memproyeksikan data yang lalu dengan menggunakan metode yang dipergunakan dan mempertimbangkan adanya beberapa faktor perubahan perubahan kebijakan-kebijakan yang mungkin terjadi, termasuk perubahan kebijakan pemerintah, perkembangan potensi masyarakat perkembangan teknologi dan penemuan-penemuan baru.

3.5. Uji Normalitas

Uji normalitas adalah uji untuk mengukur apakah data kita memiliki distribusi normal ataukah tidak. Model data yang baik adalah model yang memiliki distribusi normal atau mendekati normal, sehingga layak dilakukan pengujian secara statistik (Pakar, 2003). Ketentuan dalam uji normalitas adalah sebagai berikut :

- a. Jika probabilitas $> 0,05$ maka distribusi dari populasi adalah normal.
- b. Jika probabilitas $< 0,05$ maka populasi tidak berdistribusi secara normal.

Dasar pengambilan keputusan uji normalitas juga dapat dilihat dari plotnya (Santosa, 2007):

- a. Jika data menyebar disekitar garis diagonal dan mengikuti arah garis diagonal, maka dapat disimpulkan bahwa model memenuhi asumsi normalitas.
- b. Jika data menyebar jauh dari garis diagonal dan tidak mengikuti arah garis diagonal, maka dapat disimpulkan bahwa model tidak memenuhi asumsi normalitas”.

Ada beberapa cara uji statistik non parametrik yang digunakan untuk menguji normalitas data. Pada penelitian ini akan digunakan uji statistik *Kolmogorov-Smirnov* menggunakan *software* SPSS.

3.6. Metode SARIMA

Seasonal Autoregressive Integrated Moving Average (SARIMA) adalah pengembangan dari model ARIMA pada data runtun waktu yang memiliki pola musiman. Adapun Notasi SARIMA adalah sebagai berikut.

$$\text{SARIMA}(p,d,q)(P, D, Q)^S$$

Dimana

(p,d,q) : bagian yang tidak musiman dari model

(P,D,Q) : bagian musiman dari model

S : jumlah periode per musim

Adapun rumus umum dari SARIMA(p,d,q)(P, D, Q)^S sebagai berikut:

$$\Phi_{p^{(B)} : p^{(B)}(1-B)^d(1-B^S)^D} Z_t = \gamma_{q^{(B)} \ominus q^{(B^S)} a_t}$$

Dimana :

$p^{(B)}$: *AR non seasonal*

$\Phi_{p^{(B^S)}}$: *AR seasonal*

$(1 - B)^d$: *differencing non seasonal*

$(1 - B^S)^D$: *differencing seasonal*

$\gamma_{q^{(B)} \ominus}$: *MA non seasonal*

$\ominus q^{(B^S)} a_t$: *MA seasonal*

Tingkat keakuratan ramalan yang dihasilkan SARIMA dapat diketahui berdasarkan nilai *Mean Percentage Error* (MAPE). Semakin kecil nilai MAPE yang diperoleh maka semakin kecil pula peluang *erornya*.

3.7. Metode *Support Vector Machine* (SVM)

Dewasa ini SVM telah berhasil diaplikasikan dalam problema dunia nyata. Metode SVM memiliki kelebihan dibandingkan metode *artificial neural network* (ANN). Pada metode ANN semua data latih akan dipelajari selama proses pelatihan, sedangkan pada metode SVM berbeda karena hanya sejumlah data terpilih saja yang berkontribusi untuk membentuk model yang akan digunakan. Hal tersebut menjadi kelebihan SVM karena tidak semua data latih akan

dipandang untuk setiap iterasi pelatihannya. Metode SVM dapat mengatasi masalah klasifikasi dan regresi dengan linier ataupun nonlinier. Selain itu, SVM juga memiliki kemampuan untuk mengatasi *overfitting* dan tidak membutuhkan data yang terlalu besar dan dapat digunakan untuk melakukan prediksi. Teknik pencarian grid menggunakan validasi silang 10 kali lipat digunakan untuk menentukan nilai terbaik parameter SVM dalam proses peramalan.

Metode SVM di perkenalkan pertama kali oleh Vapnik pada tahun 1992 sebagai konsep unggulan dalam bidang pattern recognition. Algoritma ini dapat memilih model otomatis dan tidak memiliki masalah *overfitting*". Penelitian lain dilakukan oleh Kyoung-jae, dimana metode SVM sangat baik untuk prediksi karena metode ini dapat meminimalkan kesalahan klasifikasi dan penyimpangan data pada data *traininging*. Menurut Cristianini (2000) prinsip dasar SVM yaitu sebagai *linear classifier*, dan selanjutnya dikembangkan agar dapat bekerja pada problem non-linear dengan memasukkan konsep *Kernel trick* pada ruang kerja berdimensi tinggi".

Secara sederhana konsep SVM adalah sebagai usaha mencari *hyperlane* terbaik yang berfungsi sebagai pemisah dua buah *class* pada *input space*. Banyak teknik *machine learning* yang dikembangkan dengan asumsi kelinearan, sehingga algoritma yang di hasilkan terbatas untuk kasus-kasus yang linear. maka untuk mengatasinya kita bisa menggunakan metode kernel (Enri, 2018). Fungsi kernel yang biasa digunakan dalam SVM sebagai berikut:

Tabel 3.2 Fungsi Kernel dalam SVM

Jenis kernel	Definisi
Linear K	$K = (x, y) = x, y$
Polynomial	$K(x, y) = (x, y + c)^d$
Gaussian RBF	$K(x_i, x_j) = \exp(-\gamma x_i - x_j)^2$
Sigmoid	$K(x, y) = \tan(\alpha(x, y) + c)$

Menurut Fei dan Sun (2008) SVM juga menerapkan prinsip minimalisasi risiko structural yang membuatnya memiliki kemampuan generalisasi yang sangat baik dalam situasi sampel kecil (Andita & Sulistijanti, 2018).

3.8. Mean Absolute Percentage Error (MAPE)

Hasil dari suatu peramalan tidak dapat dipastikan benar seluruhnya. selalu ada penyimpangan nilai atau perbedaan nilai dengan kenyataan. Salah satu cara untuk mengetahui seberapa baik peramalan yang dihasilkan atau seberapa kecil penyimpangan yang dialami yaitu dengan menghitung nilai MAPE.

MAPE merupakan perhitungan yang menunjukkan nilai absolut rata-rata perbedaan antara nilai aktual dan nilai prediksi. Hasil peramalan dikatakan semakin akurat jika nilai MAPE semakin kecil (Naufal, 2017).

Rumus MAPE secara umum dapat dituliskan

$$MAPE = \frac{\sum_{t=1}^n \left| \frac{x_t - f_t}{x_t} \right|}{n} \times 100\% \dots\dots\dots(5)$$

Dimana :

n = Jumlah Sampel

x_t = Nilai Aktual Indeks pada periode ke-t

f_t = Nilai Prediksi Indeks pada periode ke-t

Tabel 3.3 menunjukkan perbandingan tingkat akurasi hasil peramalan berdasarkan nilai MAPE.

Tabel 3.3 Tingkat Akurasi Hasil Peramalan Berdasarkan Nilai MAPE

MAPE	Hasil Peramalan
<10%	Sangat Baik
10-20%	Baik
20-50%	Layak/Cukup
>50%	Buruk

Tabel 3.3 merupakan acuan peneliti dalam menentukan keakurasian suatu peramalan. Jika nilai MAPE yang didapatkan dalam suatu peramalan itu kurang dari 10% maka tingkat keakurasian dari hasil peramalan tersebut bisa dikatakan sangat baik. sedangkan jika nilai MAPEnya itu berada diantara 10 - 20% maka tingkat keakurasian dari hasil peramalan tersebut baik. begitupun seterusnya.

BAB IV

METODOLOGI PENELITIAN

4.1. Populasi dan Sampel Penelitian

Populasi yang digunakan pada penelitian ini adalah jumlah wisatawan mancanegara yang datang ke Indonesia melalui 19 pintu masuk. Sampel penelitian yang digunakan adalah 3 kategori pintu masuk wisatawan mancanegara ke Indonesia. Kategorisasi tersebut berdasarkan jumlah wisatawan dari setiap pintu masuk. Data yang digunakan pada penelitian ini diperoleh dari situs resmi Badan Pusat Statistik terkait jumlah wisatawan mancanegara yang datang ke Indonesia berdasarkan 19 pintu masuk periode tahun 2010-2017 untuk setiap bulannya.

4.2. Jenis dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder dalam bentuk bulanan periode tahun 2010-2017. Sumber data sekunder dalam penelitian ini di ambil dari situs resmi Badan Pusat Statistik.

4.3. Cara Pengambilan Data

Penelitian ini menggunakan dua metode pengumpulan data. yaitu:

a. Studi Pustaka

Penelitian ini mengumpulkan data dan teori yang relevan terhadap permasalahan yang akan di teliti dengan melakukan studi pustaka terhadap literature dan bahan pustaka lainnya seperti artikel, jurnal, buku dan penelitian terdahulu.

b. Studi Dokumenter

Jenis data yang digunakan pada penelitian ini adalah data sekunder diperoleh dari situs resmi Badan Pusat Statistik terkait jumlah wisatawan mancanegara ke Indonesia periode tahun 2010-2017 untuk setiap bulannya.

4.4. Variabel dan Definisi Operasional Variabel

Pada **Tabel 4.1** berisi tentang penjelasan dan definisi operasional dari masing-masing variabel penelitian:

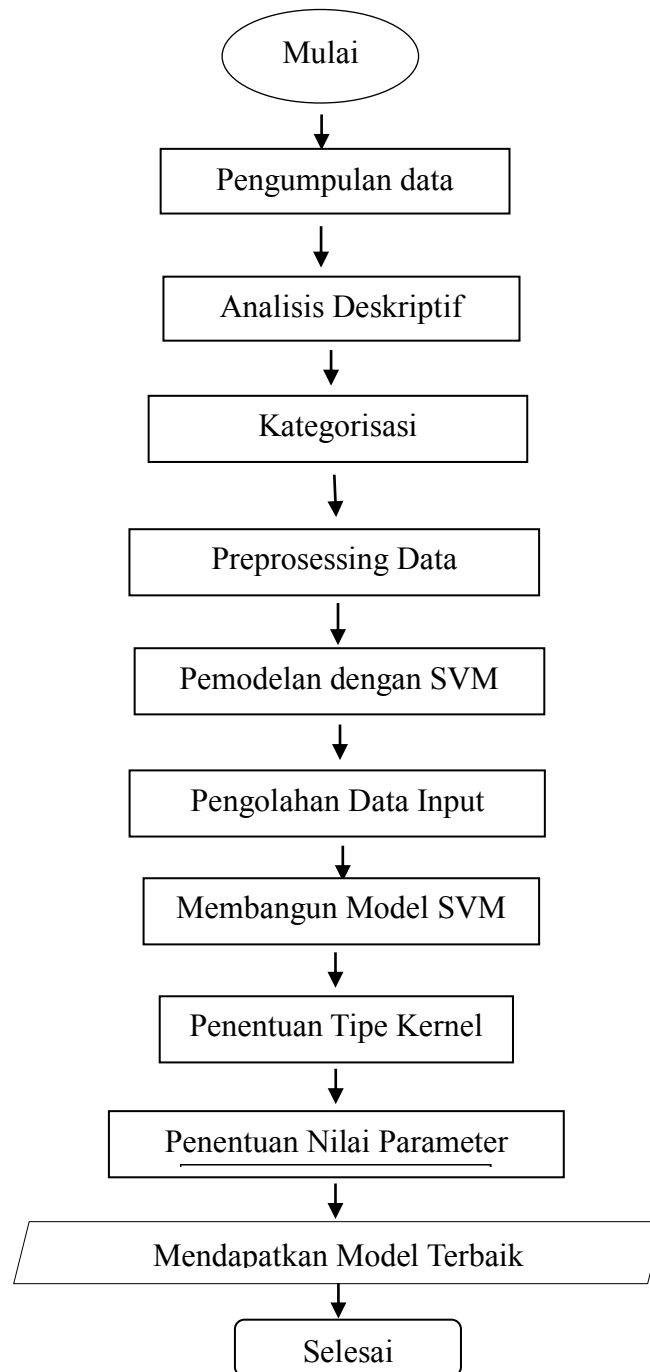
Tabel 4.1 Definisi operasional variabel

Variabel	Definisi Operasional Variabel
Periode Waktu (X)	Menunjukkan periode waktu dalam bulan jumlah wisatawan mancanegara yang masuk ke Indonesia.
Jumlah Wisatawan Mancanegara (Y)	Jumlah Wisatawan Mancanegara yang masuk ke Indonesia per tiap bulan dari periode tahun 2010-2017.

4.5. Metode Analisis Data

Analisis data dalam penelitian ini menggunakan beberapa *software* diantaranya *software* R 3.4.2, SPSS 22.0, dan *Microsoft Excel* yang didalamnya memuat analisis deskriptif, uji normalitas, kategorisasi data, dan analisis peramalan menggunakan SARIMA dan SVM. Hasil *output* dari analisis tersebut akan di analisis sehingga diketahui penerapan metode SARIMA dan SVM dalam meramalkan jumlah wisatawan mancanegara yang datang ke Indonesia melalui 3 kategori pintu masuk, yaitu tinggi, sedang dan rendah. Selain itu, juga akan diketahui prediksi jumlah wisatawan mancanegara yang datang ke Indonesia untuk beberapa periode kedepan.

4.6. Proses Analisis Data



Gambar 4.1 *Flow Chart Analisis Data*

Gambar 4.1 merupakan *flowchart* dari analisis data. Langkah awal dari analisis ini adalah mengumpulkan data yang diperlukan dalam analisis. Pada penelitian ini digunakan data kunjungan wisatawan mancanegara dengan dua variabel yaitu periode dalam bulan (X) dan jumlah wisatawan mancanegara (Y).

Data tersebut kemudian di analisis menggunakan deskriptif. Setelah itu data jumlah wisatawan mancanegara dari 19 pintu masuk dikategorikan menjadi 3 kategori. Tahap berikutnya adalah peramalan menggunakan metode SVM dengan tahapan sebagai berikut:

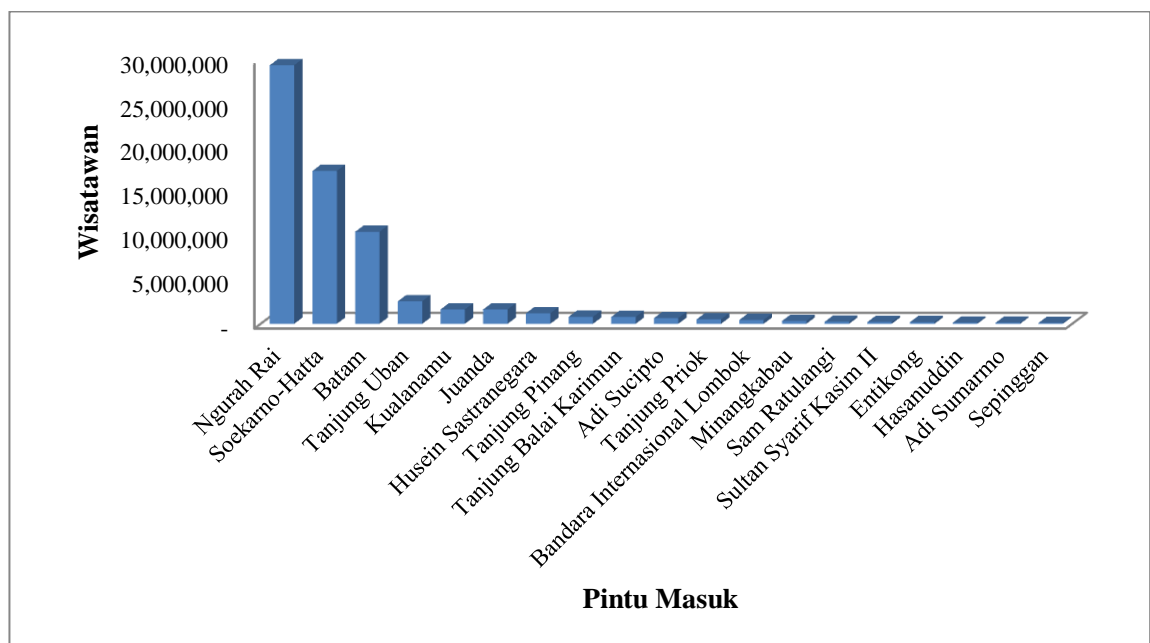
1. Pra-Proses Data yaitu sebelum proses pengolahan data hendaknya yang perlu diperhatikan adalah pemeriksaan dan penanganan data, apakah ada data yang *missing* atau tidak, jika tidak maka dilanjutkan dengan uji normalitas. Data yang tidak berdistribusi normal akan dilakukan transformasi dan untuk data yang berdistribusi normal lanjut ketahap berikutnya. Tahap berikutnya adalah pembagian data yang meliputi data pelatihan (*training*) dan data pengujian(*testing*).
2. Pemodelan dengan SVM, dimana pada tahap ini mulai dilakukan *training* SVM, sebelum masuk ke tahap ini hal yang terlebih dahulu dilakukan yaitu dengan menentukan tipe kernel dan nilai parameter. Setelah mendapatkan model SVM terbaik, selanjutnya digunakan untuk memprediksi data *testing* guna mendapatkan generalisasi tingkat akurasi model.
3. Menentukan tipe kernel. Pada fungsi kernel, akan ditentukan nilai parameter seperti nilai *C (cost)*, *Epsilon* dan *gamma*.
4. Penentuan nilai parameter seperti *epsilon*, *C (cost)* dan (*gamma*) dilakukan dengan implementasi algoritma *grid-search*. *Grid-search* akan melatih banyak pasangan model dari *range* nilai yang telah ditentukan. Setelah itu, peneliti mendapatkan tipe kernel dan parameter terbaik dalam meramalkan. Evaluasi sebuah hasil peramalan harus memiliki nilai akurasi yang baik. Pada penelitian ini untuk menentukan hasil peramalan yang baik digunakan nilai dari rata-rata MAPE terkecil.

BAB V

HASIL DAN PEMBAHASAN

5.1. Analisis deskriptif

Hal pertama yang dilakukan dalam penelitian yaitu melakukan analisis deskriptif dari variabel yang digunakan. Manfaat dari analisis deskriptif sendiri adalah untuk memberikan gambaran mengenai suatu data. Analisis deskriptif dalam penelitian ini digunakan untuk melihat gambaran umum mengenai karakteristik data jumlah wisatawan mancanegara yang datang melalui 19 pintu masuk.



Gambar 5.1 Total Wisatawan Mancanegara Periode 2010-2017

Gambar 5.1 menggambarkan total wisatawan mancanegara yang datang ke Indonesia melalui 19 pintu masuk periode tahun 2010-2017. Berdasarkan grafik diatas dapat diketahui bahwa wisatawan mancanegara yang datang ke Indonesia paling banyak melalui pintu Ngurah Rai yaitu sebanyak 29,336,709 wisatawan. Seperti yang yang dilansir dalam kompas.com lonjakan wisatawan yang masuk melalui pintu Ngurah Rai tak lepas dari rasa penasaran wisatawan

terhadap keindahan budaya dan alam yang ditawarkan oleh Bali. Tercatat 2 tahun terakhir wisatawan mancanegara yang paling banyak mengunjungi Bali adalah wisatawan dari Negara China. Adapun pintu yang paling sedikit dilalui oleh wisatawan mancanegara untuk datang ke Indonesia adalah pintu Sepinggan yaitu sebanyak 96.465 wisatawan.

5.2. Kategorisasi Data

Total jumlah wisatawan yang masuk ke Indonesia melalui 19 pintu akan dikategorisasikan menjadi 3 yaitu tinggi, sedang, dan rendah. Setiap pintu akan masuk ke salah satu kategori baik itu tinggi, sedang, maupun rendah. Adapun hasil kategorisasinya dapat dilihat pada **Tabel 5.1**.

Tabel 5.1 Hasil Kategorisasi 19 Pintu Masuk di Indonesia

No	Pintu Masuk	Total Wisatawan	Kategori
1	Ngurah Rai	29.336.709	Tinggi
2	Soekarno-Hatta	17.388.359	Tinggi
3	Batam	10.470.776	Tinggi
4	Tanjung Uban	2.560.336	Tinggi
5	Kualanamu	1.635.039	Sedang
6	Juanda	1.632.673	Sedang
7	Husein Sastranegara	1.197.379	Sedang
8	Tanjung Pinang	792.618	Sedang
9	Tanjung Balai Karimun	778.715	Sedang
10	Adi Sucipto	652.536	Sedang
11	Tanjung Priok	506.413	Sedang
12	Bandara Internasional Lombok	434.841	Sedang
13	Sam Ratulangi	324.912	Sedang
14	Minangkabau	227.401	Sedang
15	Sultan Syarif Kasim II	201.529	Sedang
16	Entikong	189.846	Rendah

No	Pintu Masuk	Total Wisatawan	Kategori
17	Hasanuddin	123.918	Rendah
18	Adi Sumarmo	118.201	Rendah
19	Sepinggan	96.465	Rendah

Tabel 5.1 menunjukkan hasil kategorisasi dari 19 pintu. Setiap pintu masuk ke salah satu kategori, baik itu kategori tinggi, sedang maupun rendah. Kategorisasi tinggi menunjukkan bahwa jumlah wisatawan yang masuk ke Indonesia melalui pintu tersebut sangat banyak, begitupun dengan kategori lainnya. Berdasarkan hasil kategorisasi diatas peneliti akan meramalkan ke 3 kategori tersebut berdasarkan nilai rata-rata untuk setiap kategorinya.

5.3. *Preprocessing Data*

Tahapan *Preprocessing* data dipersiapkan dengan melakukan *cleaning* data dengan tujuan agar menghilangkan data-data yang hilang ataupun data-data yang dianggap tidak normal. Data penelitian yang digunakan bersumber dari BPS dan tidak terdapat *missing values* maka tahapan tersebut akan dilewati. Tahapan selanjutnya yaitu melakukan uji normalisasi data. Data yang tidak berdistribusi normal akan dilakukan transformasi, dan untuk data yang berdistribusi normal akan lanjut ke tahap berikutnya. Data dibagi ke dalam dua bagian. yaitu data *training* dan data *testing*.

5.3.1. Uji Normalitas

Uji normalitas dilakukan untuk melihat apakah data berdistribusi normal atau tidak. Jika data tidak berdistribusi normal maka dilakukan proses transformasi. Jika data telah berdisrtibusi normal, tahapan selanjutnya adalah data siap untuk diolah lebih lanjut pada tahapan pemodelan SVM. Hasil uji normalitas dari 3 kategori pintu masuk wisatawan mancanegara dapat dilihat pada **Tabel 5.2**.

Tabel 5.2 Hasil Uji Normalisasi dari 3 Kategori

Kategori	P_Value	Keputusan	Keterangan
Tinggi	0,200	$0,200 > 0,05$	Normal
Sedang	0,138	$0,138 > 0,05$	Normal
Rendah	0,200	$0,200 > 0,05$	Normal

Tabel 5.2 merupakan hasil uji normalitas dengan *Kolmogorov Smirnov*. Berdasarkan hasil tersebut didapatkan bahwa ke 3 kategori pintu masuk memiliki nilai *p_value* yang lebih besar dari nilai *alpha* sehingga dapat disimpulkan bahwa ke 3 pintu masuk tersebut berdistribusi normal dan lanjut ke tahapan berikutnya.

5.3.2. Pembuatan Data *Training* dan Data *Testing*

Pembuatan data *training* dilakukan dalam rangka meningkatkan kinerja dari SVM terhadap data *testing* dalam memilih parameter dan fungsi kernel terbaik untuk model. Pembagian data *training* dan data *testing* dengan presentase 70:30 dari total data yang digunakan tidak lain dari hasil uji coba yang dilakukan dari beberapa macam pembagian data salah satunya dengan presentase 80:20. Berdasarkan hasil uji coba tersebut didapatkan bahwa pembagian data *training* dan data *testing* dengan presentase 70:30 dibanding 80:20 karena memiliki nilai akurasi yang lebih baik, sehingga diterapkan pada penelitian ini.

Model SVM yang didapatkan dari data *training* yang nantinya akan di terapkan pada data *testing*. Adapun pembagian data *training* dan data *testing* terlihat pada **Tabel 5.3**.

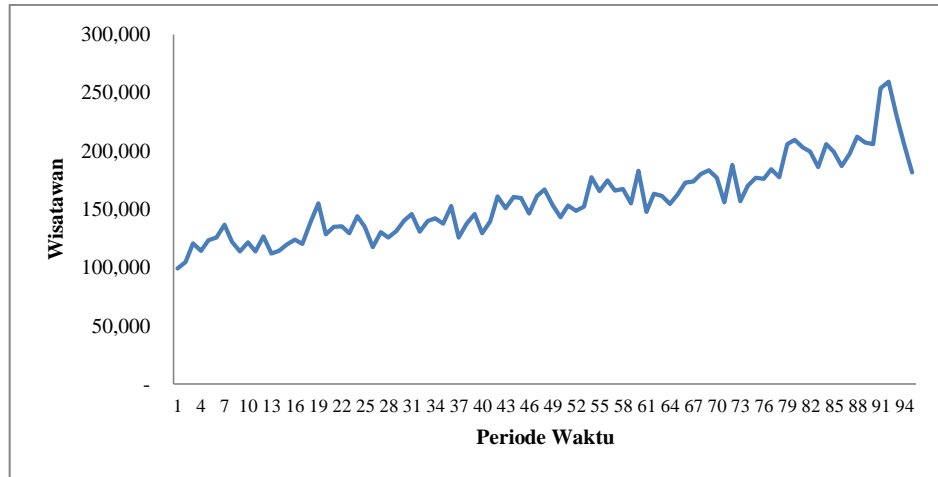
Tabel 5.3 Pembagian Data *Training* dan Data *Testing*

Pintu Masuk	Periode	Total Data	Data <i>Training</i>	Data <i>Testing</i>
Kategori Tinggi	2010-2017	95	67	28
Kategori Sedang	2010-2017	95	67	28
Kategori Rendah	2010-2017	95	67	28

Setelah didapatkan pembagian data untuk *training* dan *testing*, proses selanjutnya adalah dengan menganalisis trend dari data *time series*.

5.4. Identifikasi Data dengan Plot *Time Series*

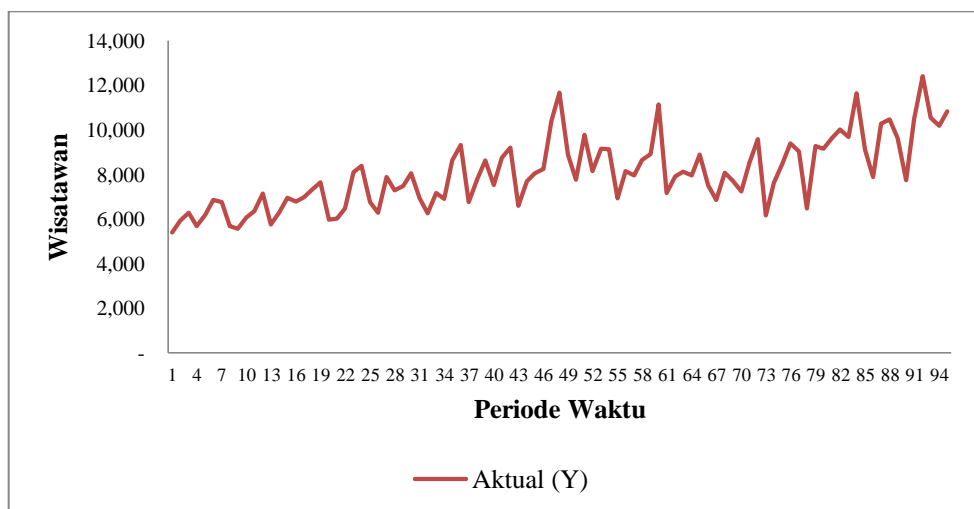
5.4.1. *Time Series* Plot Pintu Masuk Kategori Tinggi



Gambar 5.2 *Plot* Pintu Masuk Kategori Tinggi Periode 2010-2017

Gambar 5.2 menunjukkan grafik jumlah wisatawan mancanegara yang masuk ke Indonesia melalui pintu dengan kategori tinggi. Grafik tersebut memiliki dua sumbu yaitu X dan Y, dimana sumbu X menyatakan periode waktu dalam bulan sedangkan sumbu Y menyatakan jumlah wisatawan mancanegara. Selain itu, grafik tersebut menunjukkan bahwa data pada pintu masuk dengan kategori tinggi cenderung memiliki pola *trend* naik dan disertai adanya pola musiman atau dapat disebut *seasonal trend*.

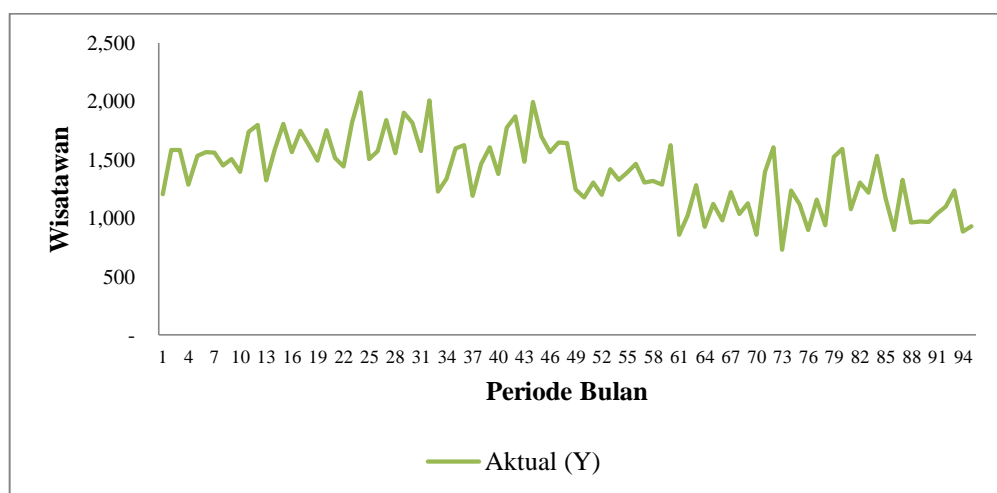
5.4.2. *Time Series* Plot Pintu Masuk Kategori Sedang



Gambar 5.3 *Plot* Pintu Masuk Kategori Sedang Periode 2010-2017

Gambar 5.3 menunjukkan grafik jumlah wisatawan mancanegara yang masuk ke Indonesia melalui pintu dengan kategori sedang. Grafik tersebut memiliki dua sumbu yaitu X dan Y, dimana sumbu X menyatakan periode waktu dalam bulan sedangkan sumbu Y menyatakan jumlah wisatawan mancanegara. Selain itu, grafik tersebut memperlihatkan pintu dengan kategori sedang cenderung memiliki pola data yang *irregular/* acak.

5.4.3. Time Series Plot Pintu Masuk Kategori Rendah



Gambar 5.4 Plot Pintu Masuk Kategori Rendah Periode 2010-2017

Gambar 5.4 menunjukkan grafik jumlah wisatawan mancanegara yang datang ke Indonesia melalui pintu dengan kategori rendah. Grafik tersebut memiliki dua sumbu yaitu X dan Y, dimana sumbu X menyatakan periode waktu dalam bulan sedangkan sumbu Y menyatakan jumlah wisatawan mancanegara. Selain itu, grafik tersebut memperlihatkan bahwa pintu dengan kategori rendah cenderung memiliki pola yang *irregular/* acak.

5.5. Penerapan Metode SARIMA

5.5.1 Pintu Masuk dengan Kategori Tinggi

a. Uji Stasioner

Syarat utama dalam analisis SARIMA adalah adanya kestasioneran data. Kestasioneran data musiman dapat dihitung dengan uji KPSS. Adapun *output* uji KPSS dapat dilihat pada tabel berikut.

Tabel 5.4 *Output Uji Kestasioneran Data Musiman*

KPSS Testing for Level Stationarity	
Nilai <i>P_Value</i> :	0,01

Berdasarkan hasil uji kestasioneran diatas dapat dibuat hipotesis sebagai berikut.

a. Hipotesis

H_0 : Data stasioner

H_1 : Data tidak stasioner

b. Tingkat signifikansi

$\alpha = 5\%$ atau 0,05

c. Daerah kritis

Tolak H_0 jika $p\text{-value} < \alpha$

d. Keputusan

$p\text{-value} (0,01) < \alpha (0,05)$ maka tolak H_0

e. Kesimpulan

Dengan tingkat kepercayaan 95% maka didapatkan hasil tolak H_0 , artinya bahwa data tidak stasioner.

Berdasarkan uji KPSS dapat disimpulkan bahwa data tersebut tidak stasioner, sehingga perlu dilakukan differensiasi untuk data musiman. Adapun hasil differensiasinya dapat dilihat pada **lampiran 13**.

Langkah selanjutnya adalah mengidentifikasi kestasioneran data dalam rata-rata. Uji kestasioneran data dapat dihitung dengan uji Adf. Adapun *output* dari uji Adf dapat dilihat pada **Tabel 5.5**.

Tabel 5.5 *Output Uji Kestasioneran Data dalam Rata-rata*

<i>Augmented Dickey-Fuller Testing</i>	
Nilai <i>P_Value</i>	0,3837

Berdasarkan hasil uji kestasioneran diatas dapat dibuat hipotesis sebagai berikut.

a. Hipotesis

H_0 : Data mengandung unit root (tidak stasioner)

H_1 : Data tidak mengandung unit root (stasioner)

b. Tingkat signifikansi

$$\alpha = 5\% \text{ atau } 0,05$$

c. Daerah kritis

$$\text{Tolak } H_0 \text{ jika } |ADF| > |\text{critical value}|$$

$$\text{Tolak } H_0 \text{ jika } p\text{-value} < \alpha$$

d. Keputusan

$$p\text{-value } (0,5083) > \alpha (0,05) \text{ maka gagal tolak } H_0$$

e. Kesimpulan

Dengan tingkat kepercayaan 95% maka didapatkan hasil gagal tolak H_0 , A artinya bahwa data mengandung unit root (tidak stasioner dalam rata-rata).

Berdasarkan uji Adf didapatkan data tidak stasioner dalam rata-rata maka dilakukan differensiasi. Selanjutnya dilakukan kembali pengecekan apakah data sudah stasioner dalam rata-rata. Kestasioneran data dapat dihitung dengan uji Adf. Adapun *output* dari uji Adf sebagai berikut.

Tabel 5.6 *Output Uji Kestasioneran Data dalam Rata-rata*

<i>Augmented Dickey-Fuller Testing</i>	
Nilai <i>P_Value</i>	0,01

Berdasarkan uji kestasioneran diatas dapat dibuat hipotesis sebagai berikut.

a. Hipotesis

H_0 : Data mengandung unit root (tidak stasioner)

H_1 : Data tidak mengandung unit root (stasioner)

b. Tingkat signifikansi

$$\alpha = 5\% \text{ atau } 0,05$$

c. Daerah kritis

$$\text{Tolak } H_0 \text{ jika } |ADF| > |\text{critical value}|$$

$$\text{Tolak } H_0 \text{ jika } p\text{-value} < \alpha$$

d. Keputusan

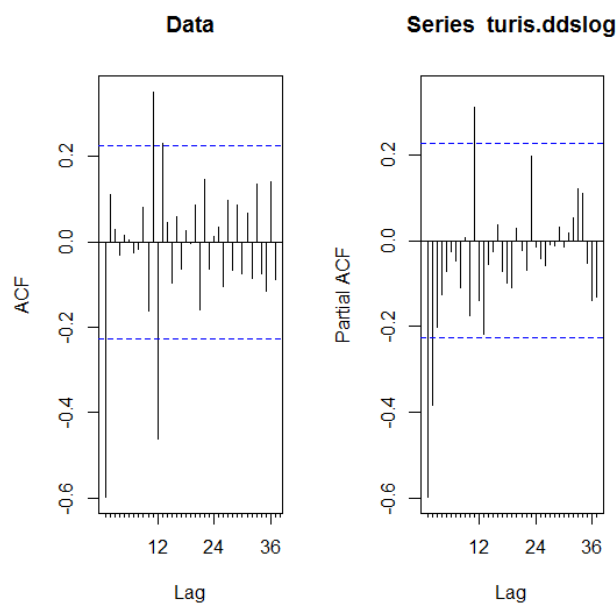
$$p\text{-value } (0,01) < \alpha (0,05) \text{ maka tolak } H_0$$

e. Kesimpulan

Dengan tingkat kepercayaan 95% maka didapatkan hasil tolak H_0 . Artinya dapat disimpulkan bahwa data tidak mengandung unit root (stasioner dalam rata-rata).

b. Analisis Model SARIMA

Penentuan model dilakukan ketika syarat untuk metode SARIMA sudah terpenuhi yaitu data sudah stasioner. Penentuan model SARIMA awal dilihat berdasarkan plot ACF dan PACF dari data yang sudah stasioner. Adapun hasil plotnya dapat dilihat pada gambar berikut.



Gambar 5.5 Plot ACF dan PACF

Berdasarkan plot ACF dan PACF dilakukan identifikasi model awal. Pada plot ACF terdapat 1 garis yang melewati garis batas yaitu tepat pada lag 12 sehingga $Q=1$ dan pada plot PACF tidak terdapat garis pada lag 12, 24 dan 36 yang melewati garis batas sehingga $P=0$, sehingga dugaan modelnya adalah $SARIMA(2,1,1)(0,1,1)^{12}$.

Model SARIMA awal yang dihasilkan dapat dijabarkan atau di kombinasikan menjadi 4 model lagi sehingga total model yang dihasilkan adalah 5 model. Tahap selanjutnya yaitu melakukan pengujian untuk mengetahui model manakah yang

estimasi yang cocok dengan data. Adapun hasil pengujian dapat dilihat pada tabel berikut.

Tabel 5.7 Hasil Pengujian Model

	Model 1 SARIMA (2,1,1)(0,1,1)	Model 2 SARIMA (2,1,0)(0,1,1)	Model 3 SARIMA (2,1,1)(0,1,1)	Model 4 SARIMA (1,1,1)(0,1,1)	Model 5 SARIMA (1,1,1)(0,1,1)
AIC	X	-206,04	X	-119,07	-209,58
p-value	Terdapat yang tidak signifikan	signifikan	Terdapat yang tidak signifikan	signifikan	Signifikan

Pada model 1 dan model 3 terdapat nilai yang tidak signifikan maka tidak dapat digunakan untuk peramalan. sehingga pilihan model yang memungkinkan adalah SARIMA model 2. SARIMA model 4 dan SARIMA model 5. Penentuan model terbaik dilihat dari hasil diagnostik yang terdapat pada tabel berikut.

Tabel 5.8 Hasil Uji Diagnostik

	Model 2 SARIMA (0.1.1)(2.1.0)	Model 4 SARIMA (0.1.1)(1.1.1)	Model 5 SARIMA (0.1.1)(1.1.1)
AIC	-206.04	-119.07	-209.58
p-value	signifikan	Signifikan	Signifikan
Diagnostic	terpenuhi	Tidak terpenuhi	terpenuhi

Pemilihan metode terbaik untuk digunakan sebagai peramalan dapat dilihat dari diagnostiknya terpenuhi. p-value yang signifikan serta nilai AIC yang terkecil. Dari hasil tabel diatas maka didapatkan bahwa model terbaik untuk dijadikan model peramalan adalah model 2 yaitu ARIMA(0.1.1)(2.1.0)¹².

c. *Mean Absolut Percentage Error (MAPE)*

Hasil dari suatu peramalan tidak dapat dipastikan benar seluruhnya. selalu ada penyimpangan nilai atau perbedaan nilai dengan kenyataan. Salah satu cara untuk mengetahui seberapa baik peramalan yang dihasilkan atau seberapa kecil penyimpangan yang dialami yaitu dengan menghitung nilai MAPE. Berdasarkan

hasil perhitungan didapatkan nilai MAPE dari hasil peramalan menggunakan metode SARIMA untuk data pintu masuk dengan kategori tinggi yaitu sebesar 4,5% dimana nilainya kurang dari 10% yang berarti tingkat keakuratan dari hasil peramalannya sangat baik.

d. Hasil Prediksi Jumlah Wisatawan

Tabel 5.9 Hasil Prediksi Jumlah Wisatawan

Periode	Hasil Prediksi
Dec-17	226.7448
Jan-18	219.7228
Feb-18	206.0411
Mar-18	217.7275
Apr-18	233.4739
May-18	228.1928
Jun-18	226.3076
Jul-18	279.5159
Aug-18	285.6217
Sep-18	253.0764
Oct-18	226.4937
Nov-18	199.8307

Berdasarkan hasil prediksi tersebut didapatkan jumlah wisatawan mancanegara yang masuk melalui pintu dengan kategori tinggi pada bulan Desember 2017 sampai April 2018 mengalami fase naik turun namun tidak drastis.

5.5.2 Pintu Masuk dengan Kategori Sedang

a. Uji Stasioner

Syarat utama dalam analisis SARIMA adalah adanya kestasioneran data. Kestasioneran data musiman dapat dihitung dengan uji KPSS. Adapun *output* uji KPSS dapat dilihat pada tabel berikut.

Tabel 5.10 *Output Uji Kestasioneran Data Musiman*

<i>KPSS Testing for Level Stationarity</i>	
Nilai <i>P_Value</i> :	0,01

Berdasarkan hasil uji kestasioneran diatas dapat dibuat hipotesis sebagai berikut.

- a. Hipotesis
 - H_0 : Data stasioner
 - H_1 : Data tidak stasioner
- b. Tingkat signifikansi
 - $\alpha = 5\%$ atau 0,05
- c. Daerah kritis
 - Tolak H_0 jika $p\text{-value} < \alpha$
- d. Keputusan
 - $p\text{-value} (0,01) < \alpha (0,05)$ maka tolak H_0
- e. Kesimpulan
 - Dengan tingkat kepercayaan 95% maka didapatkan hasil tolak H_0 , artinya data tidak stasioner.

Berdasarkan uji KPSS dapat disimpulkan bahwa data tersebut tidak stasioner, sehingga perlu dilakukan differensiasi untuk data musiman. Adapun hasil differensiasinya dapat dilihat pada **lampiran 13**.

Langkah selanjutnya adalah mengidentifikasi kestasioneran data dalam rata-rata. Uji kestasioneran data dapat dihitung dengan uji Adf. Adapun *output* dari uji Adf dapat dilihat pada tabel berikut.

Tabel 5.11 *Output Uji Kestasioneran Data dalam Rata-rata*

<i>Augmented Dickey-Fuller Testing</i>	
Nilai <i>P_Value</i>	0,6169

Berdasarkan hasil uji kestasioneran diatas dapat dibuat hipotesis sebagai berikut.

- a. Hipotesis
 - H_0 : Data mengandung unit root (tidak stasioner)
 - H_1 : Data tidak mengandung unit root (stasioner)

b. Tingkat signifikansi

$$\alpha = 5\% \text{ atau } 0,05$$

c. Daerah kritis

$$\text{Tolak } H_0 \text{ jika } |ADF| > |\text{critical value}|$$

$$\text{Tolak } H_0 \text{ jika } p\text{-value} < \alpha$$

d. Keputusan

$$p\text{-value } (0,5083) > \alpha (0,05) \text{ maka gagal tolak } H_0$$

e. Kesimpulan

Dengan tingkat kepercayaan 95% maka didapatkan hasil gagal tolak H_0 .

Artinya bahwa data mengandung unit root (tidak stasioner dalam rata-rata).

Berdasarkan uji Adf didapatkan data tidak stasioner dalam rata-rata maka dilakukan differensiasi. Selanjutnya dilakukan pengecekan apakah data sudah stasioner dalam rata-rata. Kestasioneran data dapat dihitung dengan uji Adf. Adapun *output* dari uji Adf sebagai berikut.

Tabel 5.12 *Output* Uji Kestasioneran Data dalam Rata-rata

<i>Augmented Dickey-Fuller Testing</i>	
Nilai <i>P_Value</i>	0,01

Berdasarkan uji kestasioneran diatas dapat dibuat hipotesis sebagai berikut.

a. Hipotesis

H_0 : Data mengandung unit root (tidak stasioner)

H_1 : Data tidak mengandung unit root (stasioner)

b. Tingkat signifikansi

$$\alpha = 5\% \text{ atau } 0,05$$

c. Daerah kritis

$$\text{Tolak } H_0 \text{ jika } |ADF| > |\text{critical value}|$$

$$\text{Tolak } H_0 \text{ jika } p\text{-value} < \alpha$$

d. Keputusan

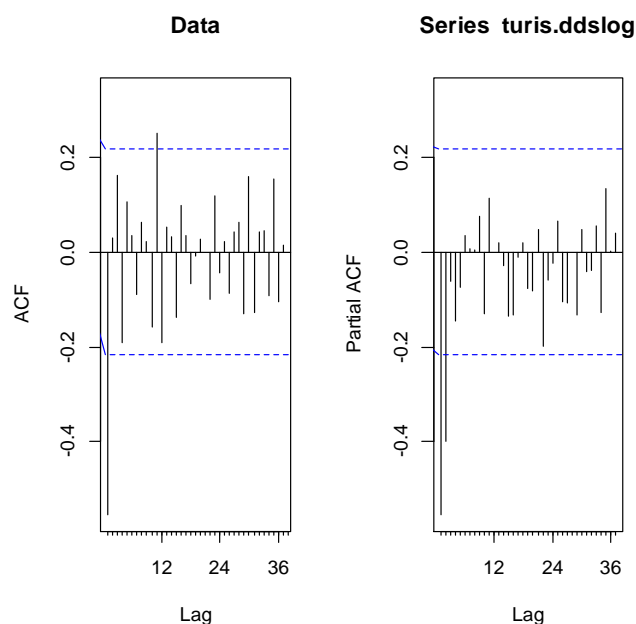
$$p\text{-value } (0,01) < \alpha (0,05) \text{ maka tolak } H_0$$

e. Kesimpulan

Dengan tingkat kepercayaan 95% maka didapatkan hasil tolak H_0 . Artinya dapat disimpulkan bahwa data tidak mengandung unit root (stasioner dalam rata-rata).

b. Analisis Model SARIMA

Penentuan model dilakukan ketika syarat untuk metode SARIMA sudah terpenuhi yaitu data sudah stasioner. Penentuan model SARIMA awal dilihat berdasarkan plot ACF dan PACF dari data yang sudah stasioner. Adapun hasil plotnya dapat dilihat pada gambar berikut.



Gambar 5.6 Plot ACF dan PACF

Berdasarkan plot ACF dan PACF dilakukan identifikasi model awal. Adapun dugaan model awalnya adalah $SARIMA(2,1,1)(0,1,0)^{12}$. Model SARIMA awal yang dihasilkan dapat dijabarkan atau di kombinasikan menjadi 4 model lagi sehingga total model yang dihasilkan adalah 5 model. Tahap selanjutnya yaitu melakukan pengujian untuk mengetahui model manakah yang estimasinya cocok dengan data. Adapun hasil pengujiannya dapat dilihat pada tabel berikut.

Tabel 5.13 Hasil Pengujian Model

	Model 1 SARIMA (2,1,1)(0,1,0)	Model 2 SARIMA (2,1,0)(1,0,1)	Model 3 SARIMA (1,1,1)(1,0,1)	Model 4 SARIMA (1,1,0)(1,0,1)	Model 5 SARIMA (0,1,1)(1,0,1)
AIC	x	-131,16	X	-119,12	-130,35
p-value	Terdapat yang tidak signifikan	Signifikan	Terdapat yang tidak signifikan	signifikan	signifikan

Pada model 1 dan model 3 terdapat nilai yang tidak signifikan maka tidak dapat digunakan untuk peramalan, sehingga pilihan model yang memungkinkan adalah SARIMA model 2, model 4 dan SARIMA model 5. Penentuan model yang terbaik dilakukan dengan uji diagnostik untuk masing-masing model yang signifikan. Adapun hasil diagnostik dapat dilihat pada tabel berikut.

Tabel 5.14 Hasil Uji Diagnostik

	Model 2 SARIMA (2,1,0)(1,0,1)	Model 4 SARIMA (1,1,0)(1,0,1)	Model 5 SARIMA (0,1,1)(1,0,1)
AIC	-131,16	-119,12	-130,35
p-value	Signifikan	Signifikan	signifikan
diagnostik	terpenuhi	Tidak terpenuhi	terpenuhi

Pemilihan metode terbaik untuk digunakan sebagai peramalan dapat dilihat dari diagnostiknya terpenuhi, *p-value* yang signifikan serta nilai AIC yang terkecil. Berdasarkan hasil tabel diatas maka didapatkan bahwa model terbaik untuk dijadikan model peramalan adalah model 5 yaitu ARIMA(0,1,1)(1,0,1)¹².

c. *Mean Absolut Percentage Error (MAPE)*

Hasil dari suatu peramalan tidak dapat dipastikan benar seluruhnya. selalu ada penyimpangan nilai atau perbedaan nilai dengan kenyataan. Salah satu cara untuk mengetahui seberapa baik peramalan yang dihasilkan atau seberapa kecil penyimpangan yang dialami yaitu dengan menghitung nilai MAPE. Berdasarkan hasil perhitungan didapatkan nilai MAPE dari hasil peramalan menggunakan

metode SARIMA untuk data pintu masuk dengan kategori tinggi yaitu sebesar 7,07% dimana nilainya kurang dari 10% yang berarti tingkat keakurasian dari hasil peramalannya sangat baik.

d. Hasil Prediksi Jumlah Wisatawan

Tabel 5.15 Hasil Prediksi Jumlah Wisatawan

Periode	Hasil Prediksi
Dec-17	13049
Jan-18	10205
Feb-18	8831
Mar-18	11514
Apr-18	11718
May-18	10786
Jun-18	8661
Jul-18	11755
Aug-18	13910
Sep-18	11796
Oct-18	11411
Nov-18	12127

Berdasarkan hasil prediksi tersebut didapatkan jumlah wisatawan mancanegara yang masuk melalui pintu dengan kategori tinggi pada bulan Desember 2017 sampai April 2018 mengalami juga fase naik turun namun tidak drastis.

5.5.3 Pintu Masuk dengan Kategori Rendah

a. Uji Stasioner

Syarat utama dalam analisis SARIMA adalah adanya kestasioneran data. Kestasioneran data musiman dapat dihitung dengan uji KPSS. Adapun *output* uji KPSS dapat dilihat pada tabel berikut.

Tabel 5.16 *Output Uji Kestasioneran Data Musiman*

KPSS Testing for Level Stationarity	
Nilai <i>P_Value</i> :	0,01

Berdasarkan hasil uji kestasioneran diatas dapat dibuat hipotesis sebagai berikut.

a. Hipotesis

H_0 : Data stasioner

H_1 : Data tidak stasioner

b. Tingkat signifikansi

$\alpha = 5\%$ atau 0,05

c. Daerah kritis

Tolak H_0 jika $p\text{-value} < \alpha$

d. Keputusan

$p\text{-value} (0,01) < \alpha (0,05)$ maka tolak H_0

e. Kesimpulan

Dengan tingkat kepercayaan 95% maka didapatkan hasil tolak H_0 artinya bahwa data tidak stasioner.

Berdasarkan uji KPSS dapat disimpulkan bahwa data tersebut tidak stasioner. sehingga perlu dilakukan differensiasi untuk data musiman. Adapun hasil differensiasinya dapat dilihat pada **lampiran 14**.

Langkah selanjutnya adalah mengidentifikasi kestasioneran data dalam rata-rata. Uji kestasioneran data dapat dihitung dengan uji Adf. Adapun *output* dari uji Adf dapat dilihat pada **Tabel 5.17**.

Tabel 5.17 *Output Uji Kestasioneran Data dalam Rata-rata*

<i>Augmented Dickey-Fuller Testing</i>	
Nilai <i>P_Value</i>	0.4335

Berdasarkan hasil uji kestasioneran diatas dapat dibuat hipotesis sebagai berikut.

a. Hipotesis

H_0 : Data mengandung unit root (tidak stasioner)

H_1 : Data tidak mengandung unit root (stasioner)

b. Tingkat signifikansi

$$\alpha = 5\% \text{ atau } 0,05$$

c. Daerah kritis

Tolak H_0 jika $|ADF| > |\text{critical value}|$

Tolak H_0 jika $p\text{-value} < \alpha$

d. Keputusan

$p\text{-value} (0,5083) > \alpha (0,05)$ maka gagal tolak H_0

e. Kesimpulan

Dengan tingkat kepercayaan 95% maka didapatkan hasil gagal tolak H_0 .

Artinya bahwa data mengandung unit root (tidak stasioner dalam rata-rata).

Berdasarkan uji Adf didapatkan data tidak stasioner dalam rata-rata maka dilakukan differensiasi. Selanjutnya dilakukan kembali pengecekan apakah data sudah stasioner dalam rata-rata. Kestasioneran data dapat dihitung dengan uji Adf. Adapun *output* dari uji Adf sebagai berikut.

Tabel 5.18 *Output Uji Kestasioneran Data dalam Rata-rata*

<i>Augmented Dickey-Fuller Testing</i>	
Nilai <i>P_Value</i>	0.01

Berdasarkan uji kestasioneran diatas dapat dibuat hipotesis sebagai berikut.

a. Hipotesis

H_0 : Data mengandung unit root (tidak stasioner)

H_1 : Data tidak mengandung unit root (stasioner)

b. Tingkat signifikansi

$$\alpha = 5\% \text{ atau } 0,05$$

c. Daerah kritis

Tolak H_0 jika $|ADF| > |\text{critical value}|$

Tolak H_0 jika $p\text{-value} < \alpha$

d. Keputusan

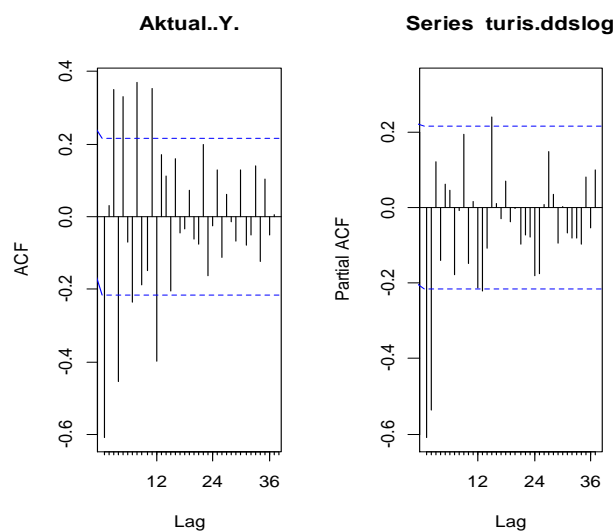
$p\text{-value} (0,01) < \alpha (0,05)$ maka tolak H_0

e. Kesimpulan

Dengan tingkat kepercayaan 95% maka didapatkan hasil tolak H_0 . Artinya dapat disimpulkan bahwa data tidak mengandung unit root (stasioner dalam rata-rata).

b. Analisis Model SARIMA

Penentuan model dilakukan ketika syarat untuk metode SARIMA sudah terpenuhi yaitu data sudah stasioner. Penentuan model SARIMA awal dilihat berdasarkan plot ACF dan PACF dari data yang sudah stasioner. Adapun hasil plotnya dapat dilihat pada gambar berikut.



Gambar 5.7 Plot ACF dan PACF

Berdasarkan plot ACF dan PACF dilakukan identifikasi model awal. Berdasarkan plot ACF dan PACF dugaan model awalnya adalah SARIMA(2,1,3)(0,1,1)¹².

Model SARIMA awal yang dihasilkan dapat dijabarkan atau di kombinasikan menjadi 4 model lagi sehingga total model yang dihasilkan adalah 5 model. Tahap selanjutnya yaitu melakukan pengujian untuk mengetahui model manakah yang estimasinya cocok dengan data. Adapun hasil pengujiannya dapat dilihat pada **Tabel 5.19**.

Tabel 5.19 Hasil Pengujian Model

	Model 1 SARIMA (2,1,3)(0,1,1)	Model 2 SARIMA (2,1,2)(0,1,1)	Model 3 SARIMA (2,1,1)(0,1,1)	Model 4 SARIMA (2,1,0)(0,1,1)
AIC	-79,08	X	-79,08	-78,25
p-value	signifikan	Terdapat yang tidak signifikan	signifikan	Signifikan
	Model 5 SARIMA (1,1,3)(0,1,1)	Model 6 SARIMA (1,1,2)(0,1,1)	Model 7 SARIMA (1,1,1)(0,1,1)	Model 8 SARIMA (0,1,3)(0,1,1)
AIC	X	X	-71,99	X
p-value	Terdapat yang tidak signifikan	Terdapat yang tidak signifikan	signifikan	Terdapat yang tidak signifikan
	Model 9 SARIMA (0,1,2)(0,1,1)	Model 10 SARIMA (0,1,1)(0,1,1)	Model 11 SARIMA (0,1,0)(0,1,1)	
AIC	-73,23	-70,24	X	
p-value	signifikan	Signifikan	Terdapat yang tidak signifikan	

Pada model 2, 5, 6, 8, dan 11 terdapat nilai yang tidak signifikan maka tidak dapat digunakan untuk peramalan. sehingga pilihan model yang memungkinkan adalah SARIMA model 1, 3, 4, 7, 9 dan 10. Penentuan model terbaik dilihat dari hasil diagnostik yang terdapat pada tabel berikut.

Tabel 5.20 Hasil Uji Diagnostik

	Model 1 SARIMA (2,1,3)(0,1,1)	Model 3 SARIMA (2,1,1)(0,1,1)	Model 4 SARIMA (2,1,0)(0,1,1)
AIC	-79,08	-79,08	-78,25
p-value	Signifikan	Signifikan	Signifikan
diagnostik	Terpenuhi	Tidak terpenuhi	Terpenuhi
	Model 7 SARIMA (1,1,1)(0,1,1)	Model 9 SARIMA (0,1,2)(0,1,1)	Model 10 SARIMA (0,1,1)(0,1,1)
AIC	-71,99	-73,23	-70,24
p-value	Signifikan	Signifikan	signifikan
diagnostik	Tidak terpenuhi	Tidak Terpenuhi	Tidak terpenuhi

Pemilihan metode terbaik untuk digunakan sebagai peramalan dapat dilihat dari diagnostiknya terpenuhi, *p-value* yang signifikan serta nilai AIC yang terkecil. Dari hasil tabel diatas maka didapatkan bahwa model terbaik untuk di jadikan model peramalan adalah model 2 yaitu ARIMA(2,1,0)(0,1,1)¹².

c. Mean Absolut Percentage Error (MAPE)

Hasil dari suatu peramalan tidak dapat dipastikan benar seluruhnya. selalu ada penyimpangan nilai atau perbedaan nilai dengan kenyataan. Salah satu cara untuk mengetahui seberapa baik peramalan yang dihasilkan atau seberapa kecil penyimpangan yang dialami yaitu dengan menghitung nilai MAPE. Berdasarkan hasil perhitungan didapatkan nilai MAPE dari hasil peramalan menggunakan metode SARIMA untuk data pintu masuk dengan kategori tinggi yaitu sebesar 9,59% dimana nilainya kurang dari 10% yang berarti tingkat keakurasian dari hasil peramalannya baik.

d. Hasil Prediksi Jumlah Wisatawan

Tabel 5.21 Hasil Prediksi Jumlah Wisatawan

Periode	Hasil Prediksi
Dec-17	1353
Jan-18	740
Feb-18	823
Mar-18	1025
Apr-18	760
May-18	898
Jun-18	836
Jul-18	952
Aug-18	1006
Sep-18	921
Oct-18	818
Nov-18	1353

Berdasarkan hasil prediksi tersebut didapatkan jumlah wisatawan mancanegara yang masuk melalui pintu dengan kategori tinggi pada bulan Desember 2017 sampai April 2018 mengalami juga fase naik turun namun tidak drastis.

5.6. Metode *Support Vector Machine* (SVM)

5.6.1. Pencarian Model SVM terbaik

Pengolahan data jumlah wisatawan mancanegara yang datang ke Indonesia dibagi menjadi dua data set yaitu data *training* dan data *testing* untuk setiap pintu masuk. Pencarian model terbaik dilakukan dengan menguji data *set* menggunakan beberapa variasi fungsi kernel yaitu fungsi kernel RBF, Polynomial, dan Linear. Selain itu, dalam meningkatkan kinerja SVM terhadap data *set* maka dilakukan pencarian parameter terbaik yang ditentukan dengan *tuning* model SVM pada program R. Pada sub bab ini akan dijelaskan secara rinci hasil *training* dan pencarian Model SVM terbaik untuk meramalkan jumlah wisatawan dari 3 kategori pintu masuk.

a. Pintu Masuk dengan Kategori Tinggi

Model pertama adalah model pada pintu masuk dengan kategori tinggi. Model terbaik ditentukan berdasarkan hasil optimasi parameter-parameter dengan tipe kernel yang sesuai. Pada **Tabel 5.22** dapat dilihat nilai parameter model SVM dari variasi fungsi kernel yang didapatkan dari proses *training*.

Tabel 5.22 Nilai Parameter Variasi Fungsi Kernel untuk Kategori Tinggi

Parameter	Fungsi Kernel		
	RBF	Linear	Polynomial
<i>Epsilon</i>	0,1	0,3	0,6
<i>Cost</i>	128	0,25	512
<i>Gamma</i>	4	0,0019	0,0625

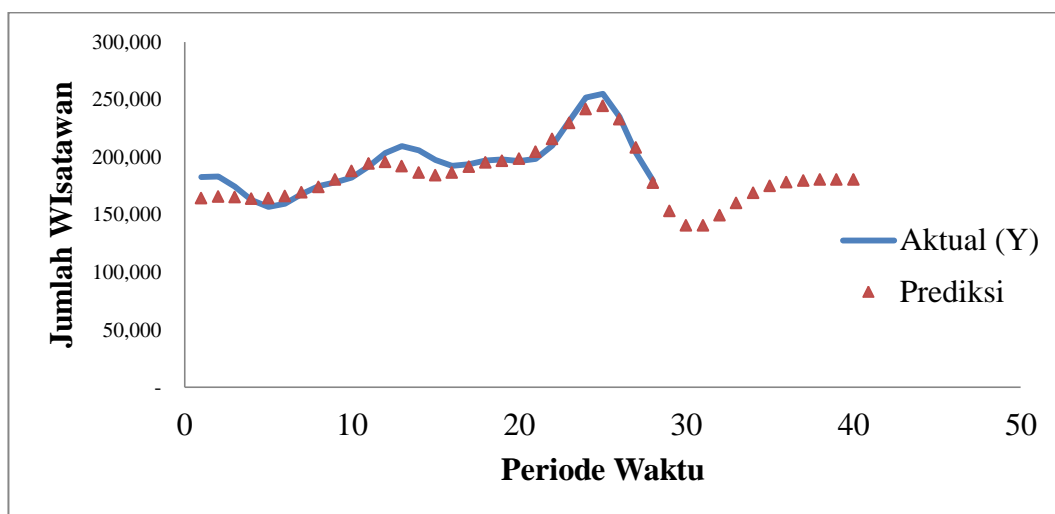
Tabel 5.22 merupakan hasil optimasi parameter dari masing-masing fungsi kernel. Penentuan parameter yang akan digunakan dalam meramalkan jumlah wisatawan mancanegara yang datang melalui pintu masuk dengan kategori tinggi berdasarkan pada nilai akurasi yang dihasilkan dari masing-masing fungsi kernel. Adapun tabel akurasi dari model dapat dilihat pada **Tabel 5.23**.

Tabel 5.23 Nilai MAPE dari Variasi Fungsi Kernel untuk Kategori Tinggi

Data Set	Fungsi Kernel		
	RBF	Linear	Polynomial
<i>Data training</i>	3,99%	5,56%	6,94%
<i>Data testing</i>	3,45%	6,51%	7,97%

Tabel 5.23 merupakan hasil akurasi yang didapatkan dari rata-rata MAPE setiap fungsi kernel. Bila ditinjau dari data *training*, nilai MAPE terkecil adalah nilai MAPE dari kernel RBF sebesar 5,35%, sedangkan pada data *testing* nilai MAPE terkecil juga terdapat pada kernel RBF sebesar 4,59%. Sehingga dapat disimpulkan bahwa model terbaik dalam meramalkan jumlah wisatawan mancanegara yang masuk melalui pintu dengan kategori tinggi adalah fungsi kernel RBF dengan parameter paling optimal yaitu nilai *epsilon* sebesar 0,3, *cost* sebesar 32 dan *gamma* sebesar 4.

Setelah mendapatkan model terbaik dalam meramalkan jumlah wisatawan mancanegara yang masuk melalui pintu dengan kategori tinggi, maka dilakukan *plotting* grafik antara data aktual dengan data prediksi dari data *testing*. Adapun hasil *plotting* grafik antara data aktual dengan data prediksi dari data *testing* dapat dilihat pada **Gambar 5.8**. Pada gambar tersebut juga terlihat peramalan untuk beberapa periode kedepan selama 12 bulan.



Gambar 5.8 *Plotting* Grafik Data *Testing* Kategori Tinggi

Plot pada **Gambar 5.8** menunjukkan bahwa data *testing* cukup dalam mengikuti pola data. Selain itu, plot **Gambar 5.8** juga menunjukkan hasil peramalan pada bulan desember 2017- November 2018. Berikut hasil peramalan kunjungan wisatawan mancanegara untuk beberapa bulan kedepan menggunakan model terbaik.

Tabel 5.24 Hasil Prediksi Pintu Masuk dengan Kategori Tinggi

Periode	Hasil Prediksi
01-des-17	152.698
01-Jan-18	140.216
01-Feb-18	140.535
01-Mar-18	148.961
01-Apr-18	159.665

Periode	Hasil Prediksi
01-May-18	168.684
01-Jun-18	174.648
01-Jul-18	177.918
01-Aug-18	179.443
01-Sep-18	180.057
01-Okt-18	180.272
01-Nov-18	180.338

Berdasarkan hasil prediksi tersebut didapatkan jumlah wisatawan mancanegara yang datang ke Indonesia melalui pintu masuk dengan kategori tinggi dari bulan Desember 2017 – November 2018 mengalami kenaikan secara terus menerus.

b. Pintu Masuk dengan Kategori Sedang

Model kedua adalah model pada pintu masuk dengan kategori sedang. Model terbaik ditentukan berdasarkan hasil optimasi parameter-parameter dengan tipe kernel yang sesuai. Pada **Tabel 5.25** dapat dilihat nilai parameter model SVM dari variasi fungsi kernel yang didapatkan dari proses *training*.

Tabel 5.25 Nilai Parameter dari Variasi Fungsi Kernel untuk Kategori Sedang

Parameter	Fungsi Kernel		
	RBF	Linear	Polynomial
<i>Epsilon</i>	0,5	0,6	0
<i>Cost</i>	2	1	512
<i>Gamma</i>	0,25	0,0019	4

Tabel 5.25 Hasil optimasi parameter dari masing-masing fungsi kernel. Penentuan parameter yang akan digunakan dalam meramalkan jumlah wisatawan mancanegara yang datang melalui pintu masuk dengan kategori sedang berdasarkan pada nilai akurasi yang dihasilkan dari masing-masing fungsi kernel.

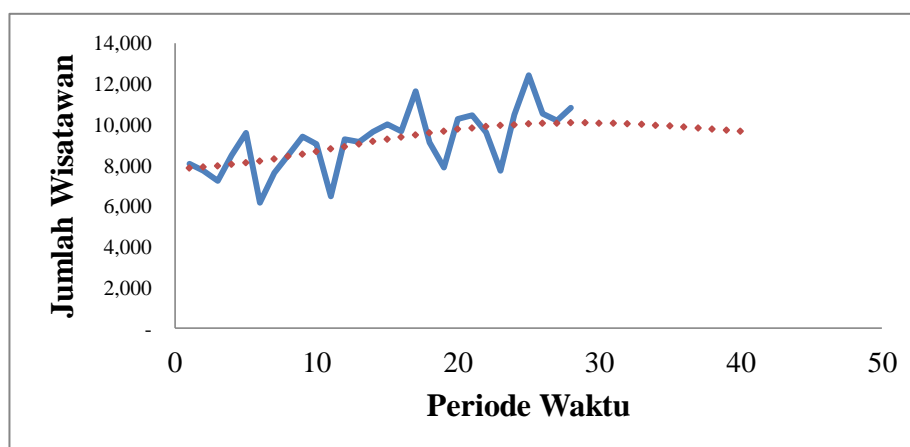
Adapun tabel akurasi dari model dapat dilihat pada **Tabel 5.26**.

Tabel 5.26 Nilai MAPE dari Variasi Fungsi untuk Kategori Sedang

Data Set	Fungsi Kernel		
	RBF	Linear	Polynomial
Data <i>training</i>	9,25%	9,68%	10,76%
Data <i>testing</i>	9,79%	10,09%	10,11%

Tabel 5.26 Hasil akurasi yang didapatkan dari rata-rata MAPE setiap fungsi kernel. Bila ditinjau dari data *training*, nilai MAPE terkecil adalah nilai MAPE dari kernel RBF sebesar 9,25%, sedangkan pada data *testing* nilai MAPE terkecil juga terdapat pada kernel RBF sebesar 9,79%. Sehingga dapat disimpulkan bahwa model terbaik dalam meramalkan jumlah wisatawan mancanegara yang masuk melalui pintu dengan kategori sedang adalah fungsi kernel RBF dengan nilai parameter paling optimal yaitu nilai *epsilon* sebesar 0,5, *cost* sebesar 2 dan *gamma* sebesar 0,25.

Setelah mendapatkan model terbaik dalam meramalkan jumlah wisatawan mancanegara yang masuk melalui pintu dengan kategori sedang, maka dilakukan *plotting* grafik antara data aktual dengan data prediksi dari data *testing*. Adapun hasil *plotting* grafik antara data aktual dengan data prediksi dan peramalan untuk beberapa periode kedepan dapat dilihat pada **Gambar 5.8**.



Gambar 5.8 *Plotting* Grafik Data *Testing* Kategori Sedang

Plot pada **Gambar 5.8** menunjukkan kemampuan data *testing* dalam

mengikuti pola data. Selain itu, plot diatas juga menunjukkan hasil prediksi pada bulan Desember 2017-November 2018. Berikut hasil peramalan kunjungan wisatawan mancanegara untuk beberapa bulan kedepan menggunakan model terbaik.

Tabel 5.27 Hasil Prediksi Pintu dengan Kategori Sedang

Periode	Hasil Prediksi
01-des-17	10.076
01-Jan-18	10.068
01-Feb-18	10.052
01-Mar-18	10.030
01-Apr-18	10.000
01-May-18	9.965
01-Jun-18	9.925
01-Jul-18	9.880
01-Aug-18	9.831
01-Sep-18	9.780
01-Okt-18	9.727
01-Nov-18	9.672

Berdasarkan hasil prediksi tersebut didapatkan jumlah wisatawan mancanegara yang masuk melalui pintu dengan kategori sedang pada bulan Desember 2017 sampai April 2018 mengalami kenaikan namun tidak drastis, dan untuk bulan Mei 2018-November 2018 mengalami penurunan jumlah wisatawan.

c. Pintu Masuk dengan Kategori Rendah

Model kedua adalah model pada pintu masuk dengan kategori rendah. Model terbaik ditentukan berdasarkan hasil optimasi parameter-parameter dengan tipe kernel yang sesuai. Pada **Tabel 5.28** dapat dilihat nilai parameter model SVM dari variasi fungsi kernel yang didapatkan dari proses *training*.

Tabel 5.28 Nilai Parameter dari Variasi Fungsi Kernel Kategori Rendah

Parameter	Fungsi Kernel		
	RBF	Linear	Polynomial
Epsilon	0,7	0,3	0,5
Cost	16	2	512
Gamma	4	0,0019	4

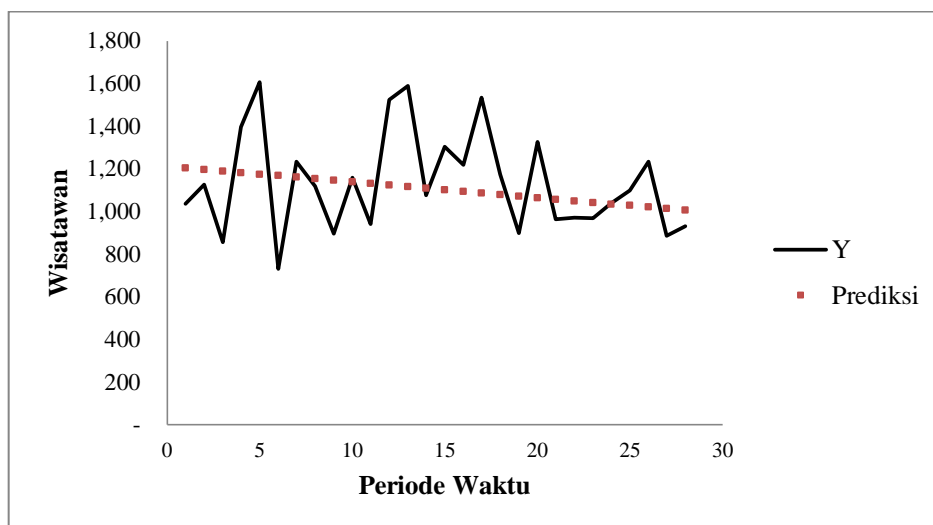
Tabel 5.28 merupakan hasil optimasi parameter dari masing-masing fungsi kernel. Penentuan parameter terbaik dalam meramalkan jumlah wisatawan mancanegara yang datang melalui pintu dengan kategori rendah berdasarkan pada nilai akurasi yang dihasilkan dari masing-masing fungsi kernel. Adapun tabel akurasi dari model dapat dilihat pada **Tabel 5.29**.

Tabel 5.29 Nilai MAPE dari Variasi Fungsi Kernel dengan Kategori Rendah

Data Set	Fungsi Kernel		
	RBF	Linear	Polynomial
Data <i>training</i>	10,13%	12,73%	13,69%
Data <i>testing</i>	16,98%	16,25%	16,33%

Tabel 5.29 merupakan hasil akurasi dari rata-rata MAPE setiap fungsi kernel. Bila ditinjau dari data *training*, nilai MAPE terkecil adalah nilai MAPE dari kernel RBF sebesar 10,13%, sedangkan pada data *testing* nilai MAPE terkecil terdapat pada kernel Linear sebesar 16,25%. Berdasarkan nilai akurasi tersebut dapat disimpulkan bahwa model terbaik dalam meramalkan jumlah wisatawan mancanegara yang masuk melalui pintu dengan kategori rendah adalah fungsi kernel Linear dengan parameter paling optimal yaitu nilai *epsilon* sebesar 0,3 *cost* sebesar 2 dan *gamma* sebesar 0,0019.

Setelah mendapatkan model terbaik maka dilakukan *plotting* grafik antara data aktual dengan data prediksi dari data *testing*. Adapun hasil *plotting* grafik antara data aktual dengan data prediksi dapat dilihat pada **Gambar 5.9**. Pada gambar tersebut juga terlihat peramalan untuk beberapa periode kedepan selama 12 bulan.



Gambar 5.9 *Plotting* Grafik Data *Testing* Pintu dengan Kategori Rendah

Gambar 5.9 menunjukkan bahwa data *testing* cukup mampu dalam mengikuti pola data. Selain itu, plot diatas juga menunjukkan hasil prediksi pada bulan Desember 2017-November 2018. Berikut hasil prediksi wisatawan mancanegara untuk beberapa bulan kedepan menggunakan model terbaik.

Tabel 5.30 Hasil Peramalan Pintu dengan Kategori Rendah

Periode	Hasil Prediksi
01-des-17	981
01-Jan-18	972
01-Feb-18	964
01-Mar-18	955
01-Apr-18	947
01-May-18	938
01-Jun-18	929
01-Jul-18	921
01-Aug-18	912
01-Sep-18	904
01-Okt-18	895
01-Nov-18	887

Berdasarkan hasil prediksi tersebut didapatkan prediksi jumlah wisatawan mancanegara yang masuk ke Indonesia melalui pintu dengan kategori rendah dari bulan Desember 2017 sampai November 2018 terus mengalami fase naik turun, dimana pada bulan Desember 2017 sampai bulan Agustus 2018 jumlah wisatawan yang datang mengalami kenaikan namun, beberapa bulan berikutnya jumlah wisatawan yang datang ke Indonesia melalui pintu dengan kategori rendah mengalami penurunan namun tidak drastis.

5.7 Perbandingan Metode SARIMA dan SVM

Tabel berikut menunjukkan nilai rata-rata MAPE dari setiap metode yang digunakan yaitu SARIMA dan SVM.

Tabel 5.31 Hasil Perbandingan Metode SARIMA dan SVM

Pintu Masuk	Nilai MAPE dari Setiap Metode	
	SARIMA	SVM
Kategori Tinggi	4,5 %	3,45 %
Kategori Sedang	7,07 %	9,79 %
Kategori Rendah	9,59 %	16,25%

Berdasarkan **Tabel 5.31** diatas dapat dilihat untuk data pintu masuk dengan kategori tinggi metode SVM memiliki nilai akurasi yang lebih baik dibandingkan metode SARIMA. Adapun untuk pintu masuk dengan kategori sedang dan rendah metode SARIMA memiliki nilai akurasi yang lebih baik dibanding metode SVM.

BAB VI

KESIMPULAN DAN SARAN

6.1. Kesimpulan

Berdasarkan penelitian yang telah dilakukan. maka penulis dapat menyimpulkan beberapa hal sebagai berikut:

1. Berdasarkan hasil pengujian 3 kategori pada pintu masuk. metode SARIMA dapat digunakan sebagai pilihan dalam meramalkan jumlah wisatawan mancanegara yang datang ke Indonesia dari berbagai pintu masuk karena nilai MAPE yang diperoleh dari ke 3 pengujian tersebut tergolong baik dimana nilai MAPE yang dihasilkan dibawah 10% .
2. Berdasarkan hasil pengujian 3 kategori pada pintu masuk metode SVM dapat digunakan sebagai pilihan dalam meramalkan jumlah wisatawan mancanegara yang datang ke Indonesia dengan menggunakan fungsi kernel RBF karena nilai akurasi yang didapatkan pada pengujian tersebut tergolong baik yaitu nilai MAPE yang diperoleh dibawah 10% .
3. Berdasarkan hasil ramalan dari ke 3 kategori pintu masuk didapatkan nilai peramalan pada bulan desember 2017- November 2018 menggunakan metode SARIMA jumlah wisatawan mancanegara yang masuk baik melalui pintu masuk dengan kategori tinggi, sedang, dan rendah mengalami fase naik turun, sedangkan hasil peramalan dengan menggunakan metode SVM jumlah wisatawan mancanegara yang masuk baik melalui pintu masuk dengan kategori tinggi, sedang, dan rendah secara umum mengalami kenaikan secara terus menerus

6.2. Saran

1. Penelitian ini menggunakan metode SVM untuk peramalan wisatawan mancanegara pada 3 kategori pintu masuk di Indonesia berdasarkan nilai rata-rata. diharapkan pada penelitian selanjutnya dilakukan dengan metode lain dan melakukan pengujian untuk setiap pintu masuk. Hal

tersebut digunakan untuk membandingkan akurasi peramalan dari tiap metode untuk melihat metode apa yang paling cocok digunakan dalam meramalkan wisatawan mancanegara di Indonesia.

2. Pada penelitian berikutnya diharapkan untuk kedepannya dapat menguji coba pada banyak dataset agar lebih meyakinkan bahwa metode SVM paling cocok pada pola data *tren seasonal*.
3. Pada penelitian berikutnya diharapkan untuk kedepannya dapat menggunakan metode regresi dummy agar bisa mendapat hasil peramalan yang lebih akurat.

DAFTAR PUSTAKA

- Andita, A., & Sulistijanti, W. (2018). Perbaikan Peramalan Produksi Padi di Kabupaten Kendal dengan Menggunakan Metode Support Vector Machine (SVM).
- Apriyono, A., & Taman, A. (2013). Analisis Overreaction pada Saham Perusahaan Manufaktur di Bursa Efek Indonesia (BEI) Periode 2005-2009. *Jurnal Nomina*, 2(2).
- Asdhiana, M. (2017, September Jumat). *Jumlah Wisman ke Bali Naik Signifikan*. Retrieved from Kompas.com.
- Assaffat, L. (2017). Peramalan Beban Listrik Industri Menggunakan Support Vector Machine dengan Variasi Fungsi kernel .
- Astuti, U. P. (2016). *Analisis Prediksi Jumlah Wisatawan mancanegara yang Masuk Melalui Pintu Kedatangan Entikong Menggunakan Metode Sarima dengan Software Eviews* . Semarang: Universitas Negeri Semarang.
- BPS. *Konsep dan Definisi Statistik Kunjungan Wisatawan Manca Negara*. Retrieved from <http://www.bps.go.id> (Diakses pada Tanggal 13 April 2018).
- BPS. *Pariwisata Statistik Sektoral* Retrieved from <http://www.bps.go.id> (Diakses pada Tanggal 13 April 2018)
- Caraka, R. E. (2017). Peramalan Crude Palm Oil (CPO) Menggunakan Support Vector Regression Kernel Radial Basis. *Jurnal Matematika*, 7(1).
- Dhuhita, W. M. (2015). Clustering Menggunakan Metode K-Means untuk Menentukan Status Gizi Balita. *Jurnal Informatika*, 15(2).
- Dwi S, N. (2015). Penerapan Algoritma Support Vector Machine untuk Prediksi Harga Emas. *Jurnal Informatika UPGRIS*, 1.
- Enri, U. (2018). Optimasi Parameter Support Vector Machines untuk Prediksi Nilai Tukar Rupiah Terhadap Dollar Amerika Serikat. *Jurnal Gerbang*, 8(1).
- Fatmawati, A. (2013, September). *Cara Mengkategorikan Data Menjadi 3 Kategori*. Retrieved from <https://freyadefunk.wordpress.com>

- Febrina, M., Arina, F., & Ekawati, R. (2013). Peramalan Jumlah Permintaan Produksi Menggunakan Metode Jaringan Syaraf Tiruan (JST) Backpropagation. *Jurnal Teknik Industri*, 1(2).
- Herawati, S. (2016). Peramalan Kunjungan Wisatawan Mancanegara. *Jurnal Infotel*, 8(1).
- Indrasetianingsih, A., Damayanti, I., & Susanto, T. (2017). Analisis Arima Box Jenkins untuk Peramalan Jumlah Kunjungan wisatawan Mancanegara di Indonesia. *Seminar Nasional Matematika dan Aplikasinya*.
- Kaloh, M. (2013). Peramalan Kunjungan Wisatawan Interasional dan Domestik ke Manado.
- Kementerian, B. P. (2016). *Laporan Akuntabilitas Kinerja Kementerian Pariwisata Tahun 2016*.
- Lestari, N., & Wahyuningsih, N. (2012, September). Peramalan Kunjungan Wisata dengan Pendekatan Model SARIMA. *Jurnal SAINS*, 1(1).
- Lestari, T. K., Rubenta, A., & Rufiadi, R. (2017). *Neraca Satelit Pariwisata Nasional 2017*. Jakarta.
- Naufal, M. F. (2017). Peramalan Jumlah Wisatawan Mancanegara yang Datang ke Indonesia Berdasarkan Pintu Masuk Menggunakan Metode Support Vector Machine (SVM).
- Nugroho, A. S., Witarto, A. B., & Handoko, D. (2003). Support Vector Machine Teori dan Aplikasinya dalam Bioinformatika.
- Priscilia, K., Umbara, R. F., & Jondri. (2014). Perbandingan Peramalan Indeks Harga Saham Gabungan Menggunakan Support Vector Machine dan Jaringan Saraf Tiruan .
- Rukini, Arini, P. S., & Nawangsih, E. (2015). Peramalan Jumlah Kunjungan Wisatawan Mancanegara (Wisman) ke Bali Tahun 2019: Metode ARIMA. *Jurnal Ekonomi Kuantitatif Terapan*, VIII(2).
- Sadewo, E. (2013, Januari). Perbandingan Beberapa Metode Time Series pada Peramalan Jumlah Kunjungan Wisatawan Mancanegara.
- Samsudin, R., Shabri, A., & Saad, P. (2010). A Comparison of Time Series Forecasting using Support Vector Machine and Artificial Neural Network Model. *Journal of Applied Sciences*, 10(11).

- Samsudin, R., Saad, P., & Shabri, A. (2010). Hybridizing GMDH And Least Squares SVM Support Vector Machine For Forecasting Tourism Demand. *IJRRAS* 3.
- Statistik, B. P. (n.d.). *Konsep dan Definisi Statistik Kunjungan Wisatawan Manca Negara*. Retrieved from <http://www.bps.go.id>
- Wamad, S. (2017, September Minggu). *Kalahkan Minyak dan Gas, Pariwisata Penyumbang Devisa Nomor 2 RI*. Retrieved from Detik Travel: <https://travel.detik.com/travel-news>
- Wardah, S., & Iskandar. (2016). Analisis Peramalan penjualan Produk Keripik Pisang Kemasan Bungkus. *Jurnal Teknik Industri*, 10(3).
- Xu, X., Law, R., & Tang, L. (2016). Forecasting tourism demand by extracting fuzzy TakagieSugeno rules from trained SVMs. *ScienceDirect*.
- Yasin, H., Prahutama, A., & Utami, T. W. (2014). Prediksi Harga Saham Menggunakan Support Vector Machine Reression dengan Algoritma Grid Serach. *Media Statistika*, 7(1).