

Implementasi Algoritme *Support Vector Regression* Pada Prediksi Jumlah Pengunjung Pariwisata

Mimin Putri Raharyani¹, Rekyan Regasari Mardi Putri², Budi Darma Setiawan³

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya
Email: ¹miminputri2@gmail.com, ²rekyan.rmp@ub.ac.id, ³s.budidarma@ub.ac.id

Abstrak

Pariwisata mempunyai peranan penting bagi pertumbuhan ekonomi suatu daerah. Salah satu faktor yang memengaruhi sektor pendapatan pariwisata adalah jumlah pengunjung. Semakin banyak jumlah pengunjung dapat meningkatkan pendapatan, jika jumlah pengunjung mengalami penurunan maka akan berdampak pada perkembangan tempat wisata yang dapat merugikan pihak pengelola pariwisata. Sistem prediksi jumlah pengunjung diperlukan sebagai gambaran mengenai tingkat jumlah pengunjung pariwisata untuk periode yang akan datang dan dapat memberikan informasi kepada pihak pengelola pariwisata untuk mempersiapkan sarana dan prasarana yang lebih baik serta mampu mengelola pendapatan dan pengeluaran untuk meminimalkan kerugian. Prediksi jumlah pengunjung pariwisata dapat dilakukan dengan menerapkan algoritme *Support vector regression*. Algoritme *Support vector regression* merupakan metode yang dapat menyelesaikan masalah regresi dan menghasilkan kinerja yang baik dalam pengambilan solusi. Pada penelitian ini data yang digunakan sebanyak 72 data jumlah pengunjung bulanan pada pariwisata dari tahun 2010 hingga 2015. Hasil pengujian menunjukkan bahwa rata-rata nilai MAPE minimum yang dihasilkan adalah 9,16% dan nilai MAPE terbaik yang didapatkan adalah 6,98% yang berarti rata-rata selisih antara hasil prediksi dengan data aktual sebesar 115 jumlah pengunjung dengan parameter $\sigma = 925,8409$ $\lambda = 0,3868$, $cLR = 0,0802$, $\epsilon = 1,27E-10$, $complexity = 3234,539$, jumlah iterasi maksimal 5000.

Kata kunci: prediksi, pariwisata, jumlah pengunjung, support vector regression

Abstract

Tourism has an important role for the economic growth of a region. One of the factors affecting the tourism revenue sector is the number of visitors. The more number of visitors can increase revenue, if the number of visitors decreased it will have an impact on the development of tourist attractions that can harm the manager of tourism. The prediction system of the number of visitors is needed as an illustration of the level of the number of tourism visitors for the period to come and can provide information to the managers of tourism to prepare better facilities and infrastructure and able to manage income and expenses to minimize losses. The prediction of the number of visitors to tourism can be done by applying the Support vector regression algorithm. Support vector regression algorithm is a method that can solve regression problems and produce good performance in the solution. In this study data used 72 data on the number of visitors monthly on tourism from 2010 to 2015. Test results show that the average value of MAPE minimum generated is 9,16% and the best MAPE value obtained is 6,98% which means The average difference between the predicted result and the actual data is 115 visitor number with sigma parameter = 925,8409 lambda = 0,3868, cLR = 0,0802, epsilon = 1,27E-10, complexity = 3234,539, maximal iteration 5000.

Keywords: prediction, tourism, visitor number, support vector regression

1. PENDAHULUAN

Sektor pariwisata memiliki pengaruh yang penting bagi pertumbuhan di bidang ekonomi negara (Suchaina, 2014). Sektor pariwisata merupakan salah satu penyumbang devisa

terbesar bagi pertumbuhan ekonomi Negara. Perkembangan industri pariwisata merupakan salah satu cara untuk meningkatkan pendapatan asli daerah dengan menarik wisatawan baik dalam negeri maupun luar negeri (Rantetadung, 2012). Salah satu daerah yang memiliki potensi besar akan keindahan alamnya terdapat di

Provinsi Jawa Timur yaitu Kabupaten Jember.

Kabupaten Jember merupakan daerah yang kaya akan tempat wisata baik wisata alam maupun wisata buatan (Setiawan & Badriyanto, 2015). Salah satu objek wisata yang terlama namun tetap menarik untuk dikunjungi di Kabupaten Jember adalah Taman Wisata Oleng Sibutong. Sebagai taman wisata yang menarik, Taman wisata Oleng Sibutong selama ini ramai dikunjungi oleh wisatawan (PT. Oleng Sibutong, 2016).

Seiring bertambahnya waktu, jumlah pengunjung yang tidak tetap dan berubah setiap harinya dapat mempengaruhi pendapatan dan perkembangan tempat wisata. Jumlah pengunjung dapat menentukan jumlah pengeluaran untuk produksi tiket yang jumlahnya harus disesuaikan dengan jumlah pengunjung guna mengurangi jumlah tiket yang terbuang ketika jumlah pengunjung tidak mencapai target jumlah tiket yang telah diproduksi begitu pula sebaliknya. Peningkatan jumlah pengunjung yang tidak terduga juga dapat menyebabkan kesulitan bagi pihak pengelola pariwisata dalam memberikan fasilitas terbaik untuk pengunjung yang berlibur (Chen, et al., 2015). Terkait dengan masalah itu sebaiknya pihak pengelola pariwisata harus dapat memprediksi dan menetapkan jumlah pengunjung untuk periode yang akan datang.

Algoritme SVR merupakan metode pengembangan dari *Support Vector Machine (SVM)* untuk kasus regresi. SVR memiliki keunggulan dibandingkan dengan model regresi linier berganda (RLB). Keunggulannya adalah dalam hal pemanfaatan data nonlinier secara implisit melalui penerapan fungsi kernel. Fungsi kernel memiliki tujuan untuk memetakan vektor input ke ruang fitur berdimensi tinggi sehingga pada kasus data linier memungkinkan penggunaan model SVR. Pada RLB hal tersebut dapat diatasi dengan model regresi namun biasanya tidak menemukan fungsi transformasi yang sesuai untuk data. Algoritme SVR dapat mengatasi masalah overfitting lebih baik dibandingkan dengan model proses training model regresi biasa dan Jaringan Syaraf Tiruan sehingga menghasilkan performansi yang bagus (Maharesi, 2013). Dibandingkan dengan SVR, ARIMA seringkali belum mampu mengakomodasi data outlier serta tidak mampu memodelkan time series yang nonlinier. Algoritme SVR mampu menangkap nonlinieritas dengan menambahkan fungsi kernel (Suci & Irhamah, 2017). Algoritme SVR

dapat diterapkan pada prediksi jumlah pengunjung pariwisata dikarenakan data jumlah pengunjung yang menghasilkan nilai nonlinier.

Algoritme SVR telah diterapkan secara luas dari permasalahan kehidupan nyata dan berfungsi secara sempurna dalam beberapa kasus dengan menghasilkan kinerja regresi secara akurat (Akande, et al., 2017). Beberapa penelitian telah membuktikan kemampuan SVR dalam beberapa studi kasus, antara lain penelitian tentang peramalan jumlah tamu hotel di Kabupaten Demak menggunakan metode *Support Vector Regression*. Penelitian ini menggunakan data sebanyak 80 jumlah tamu hotel. Data latih yang digunakan sebesar 60% dan 40% untuk data uji. Kernel yang digunakan yaitu kernel polinomial dan kernel linier. Hasil penelitian ini adalah nilai MAPE sebesar 10,2806% untuk data latih dan nilai MAPE sebesar 11,6220% untuk data uji (Adiningtyas, et al., 2015).

Penelitian lain mengimplementasikan algoritme SVR dalam memprediksi kurs rupiah. Pada penelitian ini menggunakan data latih sebanyak 344 data dan data uji sebanyak 147 data dari keseluruhan data dengan jumlah 491 data harian kurs jual Rupiah terhadap Dollar Amerika Serikat. Data tersebut diperoleh melalui situs resmi Bank Indonesia pada bulan Februari 2014. Kernel yang digunakan yaitu kernel linier dan kernel polinomial. Hasil prediksi menggunakan algoritme SVR menggunakan kernel linier dengan akurasi sebesar 100% dengan MAPE 0,3757% sedangkan jika menggunakan kernel polinomial akurasi yang dihasilkan 100% dengan MAPE sebesar 0,4477% (Amanda, et al., 2014).

Berdasarkan uraian latar belakang permasalahan dan solusi penelitian sebelumnya, peneliti mengangkat judul Implementasi Algoritme *Support Vector Regression* pada prediksi jumlah pengunjung pariwisata. Diharapkan dengan adanya sistem ini dapat memberikan hasil prediksi yang baik sehingga dapat membantu pihak pengelola pariwisata dalam menyelesaikan permasalahan untuk meminimalkan biaya pengeluaran dan mengelola pendapatan tempat wisata serta membantu pengelola pariwisata untuk mempersiapkan sarana dan prasarana yang lebih baik dengan memprediksi jumlah pengunjung untuk masa yang akan datang.

2. PARIWISATA

Menurut arti bahasa, pariwisata terdiri dari dua kata yaitu *pari* dan *wisata* berasal dari Bahasa sansekerta. Kata *pari* artinya seluruh, sedangkan *wisata* artinya perjalanan. Pariwisata dapat diartikan suatu perjalanan secara menyeluruh mulai dari awal keberangkatan dari suatu tempat ke beberapa tempat lain hingga kembali ke tempat semula (Nandi, 2008). Terdapat beberapa faktor yang mempengaruhi sektor pariwisata diantaranya adalah sebagai berikut (Rahma & Handayani, 2013):

1. Jumlah Pengunjung Wisata

Jumlah pengunjung merupakan salah satu faktor yang mempengaruhi perkembangan pariwisata. Semakin banyak pengunjung maka semakin besar pendapatan suatu objek wisata tersebut (Rantetadung, 2012).

2. Jumlah Objek Wisata

Objek wisata adalah suatu destinasi yang memiliki fasilitas untuk menarik perhatian wisatawan. Semakin banyak objek wisata maka semakin banyak pula wisatawan yang berkunjung otomatis pendapatan pada suatu daerah tersebut meningkat.

3. Tingkat Hunian Hotel

Penghuni hotel adalah seseorang atau berkelompok yang singgah di hotel untuk kepentingan acara atau liburan. Tingkat hunian hotel sama halnya dengan jumlah pengunjung pada objek wisata. Semakin banyak pengunjung yang menginap atau singgah pada hotel tersebut maka semakin meningkat pula jumlah penghuni hotel.

4. Pendapatan Perkapita

Meningkatnya sektor pariwisata sangat bergantung kepada jumlah kedatangan wisatawan di suatu daerah. Kedatangan wisatawan dapat meningkatkan pendapatan asli daerah dimana pendapatan tersebut diperoleh dari pajak daerah, laba dari objek wisata daerah, biaya retribusi, dan pendapatan lainnya.

3. DATA DAN METODE

3.1 Data

Data yang digunakan penelitian ini adalah data jumlah pengunjung yang didapatkan dari salah satu pariwisata yang ada di Kabupaten Jember yaitu Taman Wisata Oleng Sibuttong. Data yang diperoleh berupa data bulanan dari Bulan Januari 2010 hingga Bulan Desember 2015 sebanyak 72 data jumlah pengunjung.

3.2 Support Vector Regression (SVR)

Algoritme *Support Vector Regression* (SVR) merupakan pengembangan dari algoritme *Support Vector Machine* (SVM) untuk regresi. Pada kasus regresi, *output* yang dihasilkan dalam bentuk bilangan nyata (*riil*) atau data sekuensial (*kontinue*) (Mustakim, et al., 2016). Terdapat penyempurnaan persamaan fungsi regresi nonlinier berdasarkan penelitian sebelumnya yaitu terdapat pada nilai bias. Nilai b dapat digantikan dengan vektor skalar (λ) sehingga menghasilkan nilai regresi yang lebih baik ditunjukkan pada Persamaan (1). Berikut langkah-langkah proses *sequential learning Support Vector Regression* (Vijayakumar & Wu, 1999):

1. Inisialisasi parameter SVR. Parameter SVR diantaranya yaitu parameter λ (variabel skalar atau *lambda*), cLR (*constant learning rate*), ε (nilai *epsilon*), C (*complexity*), inisialisasi α_i^* dan $\alpha_i = 0$ dan iterasi maksimum.
2. Menghitung Matriks *hessian* dengan persamaan sebagai berikut:

$$R_{ij} = (K(x_i, x_j) + \lambda^2) \quad (1)$$

Keterangan:

R_{ij} = Matriks *hessian* baris ke- i kolom ke- j

$K(x_i, x_j)$ = Fungsi Kernel

λ^2 = Variabel skalar atau *lambda*

i, j = indeks

l = Banyaknya data

Keluaran dari matriks *hessian* adalah nilai parameter *gamma* (γ) yang digunakan pada tahap selanjutnya yaitu proses *sequential learning*. Nilai parameter *gamma* (γ) dapat dihitung dengan Persamaan (2):

$$\gamma = \frac{cLR}{\max(\text{matriks Hessian})} \quad (2)$$

3. Pada data latih lakukan perhitungan Tahap a, b, dan c untuk tiap data *training point* yaitu:

- a. Hitung nilai *error* dengan persamaan:

$$E_i = y_i - \sum_{j=1}^n (\alpha_i^* - \alpha_i) R_{ij} \quad (3)$$

Keterangan:

E_i = Nilai *error* ke- i

y_i = Nilai aktual

α_i^*, α_i = Nilai *Lagrange multiplier*

R_{ij} = Matriks *hessian* baris ke- i kolom ke- j

i, j = indeks

- b. Hitung perubahan nilai *Lagrange multiplier* yang merupakan batas-batas untuk menentukan nilai *Lagrange multiplier* yang terbaru.

$$\delta_{\alpha_i^*} = \min\{\max(\gamma(E_i - \varepsilon), -\alpha_i^*), C - \alpha_i^*\} \quad (4)$$

$$\delta_{\alpha_i} = \min\{\max(\gamma(-E_i - \varepsilon), -\alpha_i), C - \alpha_i\} \quad (5)$$

Keterangan:

$\delta_{\alpha_i^*}, \delta_{\alpha_i}$ = perubahan nilai α_i^* dan α_i

γ = *Learning rate (gamma)*

ε = *Epsilon*

C = Nilai kompleksitas

- c. Hitung nilai *Lagrange multiplier* yang baru yaitu update nilai α_i dan α_i^* .

$$\alpha_i^*(baru) = \delta_{\alpha_i^*} + \alpha_i^* \quad (6)$$

$$\alpha_i(baru) = \delta_{\alpha_i} + \alpha_i \quad (7)$$

4. Langkah ke-3 diulangi sampai iterasi maksimum yang telah diinisialisasi di awal, atau telah mencapai konvergensi dengan syarat $\max(|\delta_{\alpha_i^*}|) < \varepsilon$ dan $\max(|\delta_{\alpha_i}|) < \varepsilon$.
5. Hitung fungsi regresi dengan persamaan berikut:

$$f(x) = \sum_{i=1}^l (\alpha_i^* - \alpha_i) (K(x_i, x_j) + \lambda^2) \quad (8)$$

Keterangan:

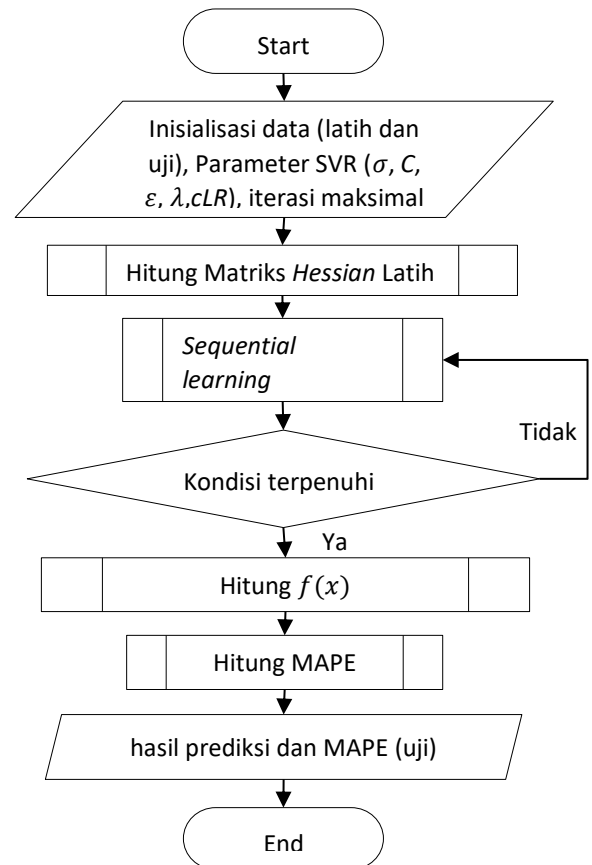
$f(x)$ = Nilai hasil regresi

l = Banyaknya Data

$K(x_i, x_j)$ = Fungsi kernel

λ^2 = Variabel scalar

Diagram alir proses algoritme SVR yang diimplementasikan pada penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Diagram Alir SVR

4. HASIL PENGUJIAN DAN ANALISIS

Pengujian yang dilakukan meliputi pengujian nilai batas bawah dan batas atas parameter algoritme SVR (σ (*sigma*), λ (*lambda*), *cLR* (*constant learning rate*), ε (*epsilon*), C (*complexity*)), pengujian jumlah iterasi dan waktu komputasi, pengujian jumlah fitur, pengujian variasi jumlah data latih dan jumlah data uji. Metode yang biasa digunakan untuk mengukur tingkat akurasi pada prediksi terutama model *Support Vector Regression* yaitu *Mean Absolute Percentage Error* (MAPE). Persamaan MAPE ditunjukkan pada persamaan 11 (Nugroho & Purqon, 2015):

$$MAPE = \frac{1}{n} \sum_{i=1}^l \left| \frac{y_i' - y_i}{y_i} \right| \times 100 \quad (9)$$

Keterangan:

n = Jumlah data

y_i' = Hasil prediksi atau hasil peramalan pada indeks ke- $i=1,2,\dots,l$

y_i = Nilai aktual pada indeks ke- $i=1,2,\dots,l$

l = Banyaknya dimensi data

Penggunaan MAPE pada evaluasi hasil prediksi mengukur rata-rata *error* mutlak terhadap selisih nilai actual dan nilai prediksi. Kriteria MAPE ditunjukkan pada Tabel 1 (Setiyoutami et al., 2012 dikutip dalam Rahmadiani & Anggraeni, 2012).

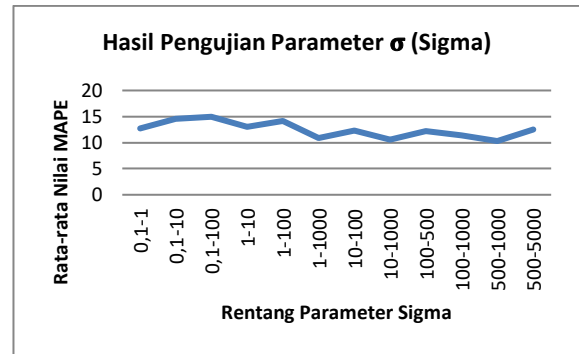
Tabel 1. Kriteria Nilai MAPE

Nilai MAPE	Definisi
< 10%	Sangat Baik
10%-20%	Baik
20%-50%	Cukup
>50%	Buruk

Pengujian parameter menggunakan batas atas dan batas bawah nilai parameter dikarenakan dengan menggunakan nilai rentang maka pencarian nilai parameter yang didapatkan lebih luas jangkauannya sehingga memudahkan dalam pencarian nilai parameter yang terbaik. Semua pengujian dilakukan sebanyak 10 kali percobaan.

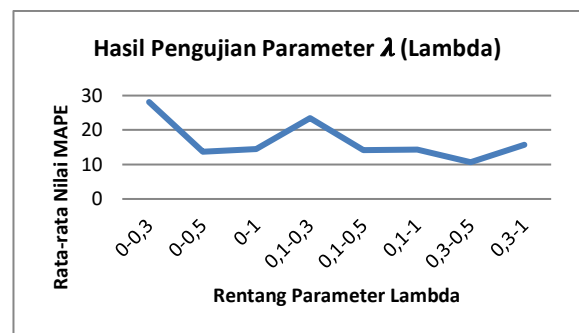
4.1 Hasil Pengujian Batas Parameter σ (*sigma*)

Skenario pengujian dilakukan dengan rincian parameter lain yaitu $\lambda = 0,3-0,4$, $cLR = 0,05-0,2$, $\varepsilon = 0,00001-0,009$, $C = 1000-10000$, jumlah data latih 24 (bulan), jumlah data uji 12 (bulan) dan jumlah iterasi maksimal adalah 10000. Parameter σ (*sigma*) merupakan nilai konstanta dari fungsi kernel *Gaussian Radial Basis Function* yang memiliki pengaruh besar terhadap kinerja pada mesin pembelajaran dengan tujuan untuk mengatur persebaran data ke dalam dimensi fitur yang lebih tinggi (Furi et al., 2015; Li et al., 2005). Pada Gambar 2, diperoleh rata-rata nilai MAPE yang minimum yaitu 10,30 dinyatakan dalam persentase dengan rentang parameter $\sigma = 500-1000$. Nilai MAPE terbaik adalah 8,83. Nilai MAPE tersebut menyatakan bahwa rata-rata selisih antara data aktual dan hasil prediksi sejumlah 136 jumlah pengunjung. Hal tersebut menunjukkan bahwa nilai σ (*sigma*) dengan rentang nilai yang lebih kecil menyebabkan persebaran data yang tidak sesuai sehingga hasil prediksi yang dihasilkan jauh dari data aktual dan nilai *error rate* meningkat.

Gambar 2. Grafik Hasil Pengujian Batas *Sigma*

4.2 Hasil Pengujian Batas Parameter λ (*lambda*)

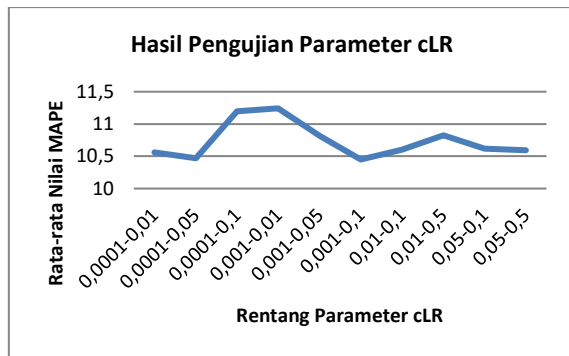
Parameter λ (*lambda*) merupakan vektor skalar yang mempengaruhi ruang pemetaan kernel pada algoritme SVR. Nilai parameter λ (*lambda*) yang tinggi akan memberikan hasil prediksi yang baik dan juga menyesuaikan nilai parameter *sigma* yang merupakan konstanta dari kernel RBF. Hasil pengujian parameter λ (*lambda*) yang telah ditunjukkan pada Gambar 3, diperoleh rata-rata nilai MAPE yang minimum adalah 10,67 dengan rentang parameter $\lambda = 0,3-0,5$. Hal tersebut menunjukkan bahwa nilai λ (*lambda*) dengan rentang nilai yang lebih kecil menyebabkan penskalaan ruang pemetaan kernel tidak sesuai sehingga nilai *error rate* meningkat. Sebaliknya semakin tinggi rentang parameter λ (*lambda*) yang digunakan maka semakin mempengaruhi proses *learning* sehingga menyebabkan nilai *error rate* yang meningkat.

Gambar 3. Grafik Hasil Pengujian Batas *Lambda*

4.3 Hasil Pengujian Batas Parameter *cLR*

Parameter *cLR* merupakan nilai konstanta *learning rate* untuk laju proses pembelajaran. Nilai parameter *cLR* akan mempengaruhi proses pembelajaran pada pelatihan SVR. Hasil pengujian parameter *cLR* yang telah ditunjukkan pada Gambar 4, diperoleh rata-rata nilai MAPE

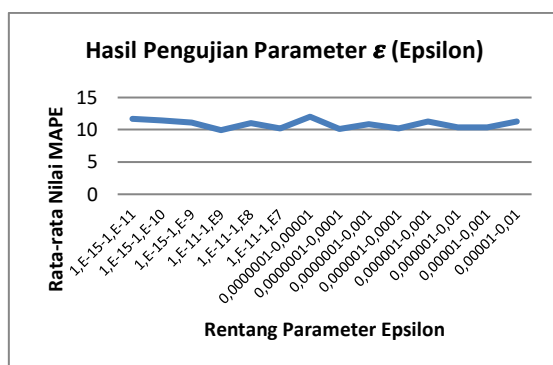
minimum adalah 10,4 dengan rentang parameter $cLR = 0,001-0,1$. Hal tersebut menunjukkan bahwa nilai cLR dengan rentang nilai yang besar dapat mempengaruhi proses pembelajaran sehingga menghasilkan nilai prediksi yang cukup baik namun apabila terlalu besar rawan memberikan hasil prediksi yang buruk.



Gambar 4. Grafik Hasil Pengujian Batas cLR

4.4 Hasil Pengujian Batas Parameter Epsilon

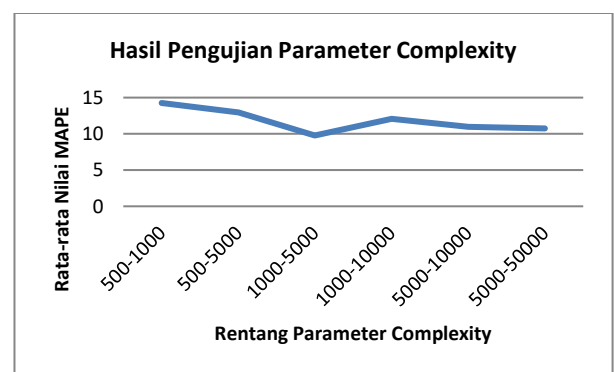
Parameter ϵ (*epsilon*) merupakan batas kesalahan nilai $f(x)$ diasumsikan bahwa nilai $f(x)$ masih berada dalam zona $\epsilon - insensitive$ (Karim, et al., 2013). Jika nilai $f(x)$ melebihi batas toleransi kesalahan ϵ akan dikenakan penalti sebesar konstanta C yang telah ditentukan (Nugroho & Purqon, 2015). Hasil pengujian parameter ϵ yang telah ditunjukkan pada Gambar 5, diperoleh rata-rata nilai MAPE yang minimum adalah 9,94 dengan rentang parameter $\epsilon = 1,E-11-1,E-9$. Hal tersebut menunjukkan bahwa nilai ϵ dengan rentang nilai yang kecil memberikan hasil prediksi yang cukup baik. Semakin kecil batas toleransi kesalahan mengakibatkan semakin banyak melakukan proses pembelajaran *training* dan tidak mentolerir kesalahan.



Gambar 5. Grafik Hasil Pengujian Batas *Epsilon*

4.5 Hasil Pengujian Batas Parameter C

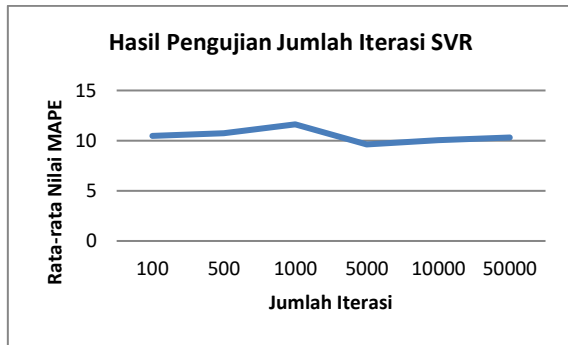
Parameter C (*complexity*) merupakan nilai penalti terhadap $f(x)$ jika melebihi batas kesalahan $\epsilon-insensitive$ (Karim et al., 2013). Semakin besar nilai parameter C (*complexity*) memberikan hasil prediksi yang bagus dikarenakan semakin besar nilai C (*complexity*) maka fungsi tersebut semakin tidak memberikan toleransi terhadap kesalahan nilai $f(x)$ (Furi, et al., 2015). Hasil pengujian parameter C (*complexity*) yang telah ditunjukkan pada Gambar 6, diperoleh rata-rata nilai MAPE yang minimum adalah 9,78 dengan rentang parameter $C = 1000-5000$. Hal tersebut menunjukkan bahwa nilai C (*complexity*) dengan rentang nilai yang besar memberikan hasil prediksi yang cukup baik. Namun apabila terlalu besar dapat menyebabkan fungsi regresi menjadi tidak stabil sehingga nilai MAPE meningkat.



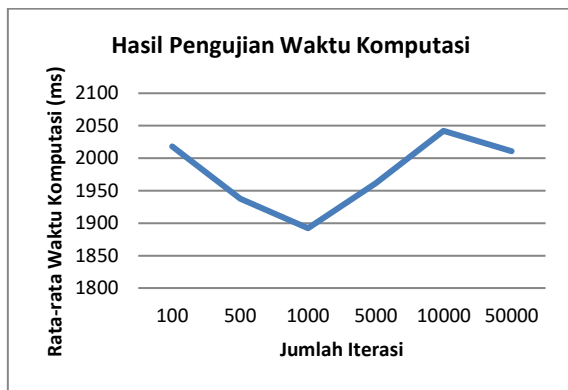
Gambar 6. Grafik Hasil Pengujian Batas C

4.6 Hasil Pengujian Jumlah Iterasi

Berdasarkan Gambar 7 dan 8, jumlah iterasi sangat mempengaruhi hasil dari *sequential training* α_i^* dan α^* yang diperoleh, semakin besar jumlah iterasi maka kemampuan SVR dalam melakukan observasi terhadap pola data pun meningkat. Jika jumlah iterasi terlalu besar maka observasi terhadap pola data pun menjadi tidak stabil sehingga menyebabkan nilai *error rate* yang meningkat. Hasil pengujian waktu komputasi berdasarkan jumlah iterasi memperlihatkan bahwa semakin banyak jumlah iterasi, waktu komputasi yang dibutuhkan berbeda-beda. Hal ini disebabkan pada setiap percobaan nilai parameter berbeda dan jumlah iterasi yang berhenti karena kondisi konvergensi juga berbeda sehingga hasil dapat berubah-ubah.



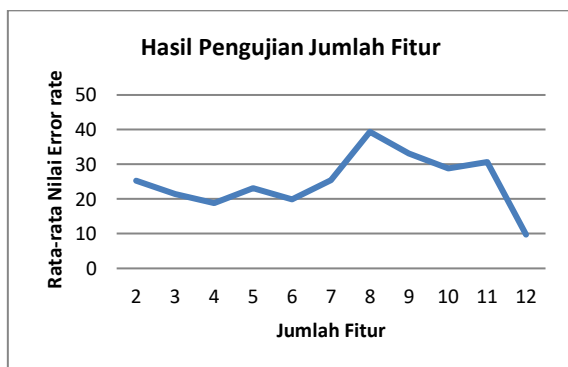
Gambar 7. Hasil Pengujian Jumlah Iterasi



Gambar 8. Hasil Pengujian Waktu Komputasi

4.7 Hasil Pengujian Jumlah Fitur

Berdasarkan Gambar 9, hasil pengujian rata-rata nilai MAPE semakin banyak jumlah fitur yang digunakan tidak menjamin semakin kecil nilai MAPE yang dihasilkan. Hal ini terjadi dikarenakan terdapat pengaruh nilai parameter algoritme SVR dalam mempelajari fitur yang digunakan. Nilai parameter yang digunakan lebih sesuai dengan fitur berjumlah 12 sehingga menghasilkan nilai MAPE yang lebih kecil dibandingkan dengan jumlah fitur lainnya.



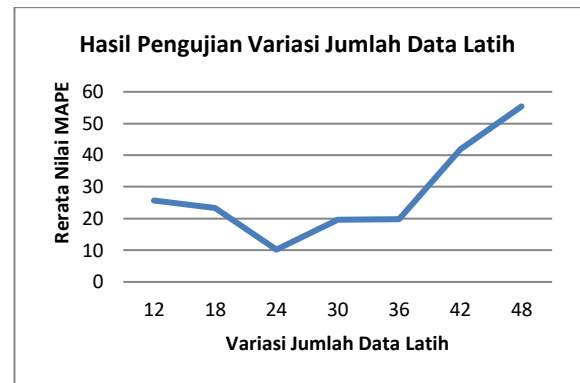
Gambar 9. Grafik Hasil Pengujian Jumlah Fitur

Pada grafik hasil pengujian jumlah fitur dapat dilihat bahwa jumlah fitur terbaik sejumlah 12 dikarenakan pada setiap bulan

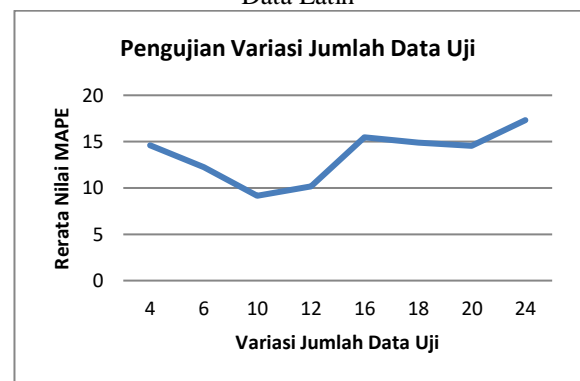
Januari pada data jumlah pengunjung selalu mengalami peningkatan sehingga dengan 12 fitur dapat memberikan hasil prediksi yang lebih baik dengan nilai MAPE minimum.

4.8 Hasil Pengujian Variasi Jumlah data latih dan Jumlah data uji

Berdasarkan Gambar 10 dan Gambar 11 dapat disimpulkan bahwa semakin banyak jumlah data latih tidak menjamin nilai MAPE yang dihasilkan akan menurun dikarenakan pola data yang berubah. Terlihat bahwa jumlah data latih yang memberikan hasil dengan rata-rata nilai MAPE yang minimum adalah 24. Sedangkan semakin sedikit jumlah data uji menghasilkan nilai MAPE yang semakin kecil. Hal tersebut dikarenakan pola data yang cenderung stabil.

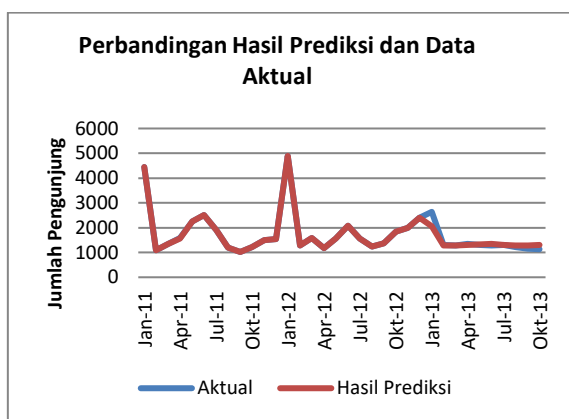


Gambar 10. Grafik Hasil Pengujian Variasi Jumlah Data Latih



Gambar 11. Grafik Hasil Pengujian Variasi Jumlah Data Uji

Sehingga berdasarkan pengujian yang telah dilakukan, diperoleh hasil prediksi dengan MAPE terbaik. Visualisasi grafik berdasarkan hasil pengujian batas atas dan batas bawah nilai parameter algoritme SVR, pengujian jumlah iterasi, dan pengujian variasi jumlah data latih dan data uji terbaik sehingga menghasilkan nilai MAPE terbaik ditunjukkan pada Gambar 12.



Gambar 12 Grafik Perbandingan Hasil Prediksi dan Data Aktual

5. KESIMPULAN

Kesimpulan yang dapat diambil berdasarkan penelitian ini adalah sebagai berikut:

1. Algoritme *Support vector regression* (SVR) dapat diterapkan dalam kasus prediksi jumlah pengunjung dengan cara sebagai berikut:
 - a. Proses pertama pada algoritme SVR ini adalah pemilihan data latih dan data uji dan juga pemilihan fitur. Fitur yang digunakan pada proses prediksi ini adalah bulan sebelumnya dan bersifat sekuensial. Proses selanjutnya adalah perhitungan *matriks hessian* yang digunakan untuk penambahan penskalaan ruang ke dimensi yang lebih tinggi. Setelah mendapatkan nilai *hessian*, proses selanjutnya adalah proses *sequential learning* untuk memperoleh fungsi regresi $f(x)$ yang sesuai sehingga memberikan hasil prediksi yang baik. Prediksi yang dilakukan adalah prediksi secara sekuensial.
 - b. Evaluasi yang digunakan pada penelitian ini adalah *Mean Absolute Percentage Error* (MAPE). Semakin kecil nilai MAPE maka semakin baik akurasi sistem yang dihasilkan.
2. Pengaruh masing-masing parameter terhadap prediksi pengunjung pariwisata menggunakan algoritme *Support Vector Regression* bervariasi. Rentang nilai σ , λ , dan cLR yang besar cenderung memberikan hasil prediksi yang baik namun terlalu besar rawan menghasilkan nilai MAPE yang tinggi. Rentang nilai ϵ yang kecil menghasilkan nilai MAPE yang kecil namun hasil cenderung fluktuasi. Rentang nilai *Complexity* yang besar memberikan hasil yang baik namun jika

terlalu besar tidak menjamin menghasilkan nilai MAPE yang lebih kecil.

3. Evaluasi hasil untuk mengukur tingkat error pada solusi permasalahan prediksi jumlah pengunjung pariwisata menggunakan nilai MAPE. Hasil pengujian menunjukkan bahwa rata-rata nilai MAPE yang minimum yang diperoleh yaitu 9,16% dengan rentang parameter σ (σ) = 500-1000, λ (λ) = 0,3-0,5, parameter cLR (*constant learning rate*) = 0,001-0,1, parameter ϵ (ϵ) = 1,E-11-1,E9, parameter *complexity* = 1000-5000. MAPE terbaik yang dihasilkan adalah 6,98% dengan parameter yang didapatkan yaitu σ = 925,8409, λ = 0,3868, cLR = 0,0802, ϵ = 1,27E-10, *complexity* = 3234,539, jumlah iterasi maksimal 5000. Nilai MAPE yang dihasilkan < 10% dan dapat dikategorikan baik untuk memprediksi jumlah pengunjung pariwisata. Berdasarkan Nilai MAPE tersebut rata-rata selisih antara hasil prediksi dengan data aktual adalah sebesar 115 jumlah pengunjung.

6. DAFTAR PUSTAKA

- Adiningtyas, D. T., Safitri, D. & Mukid, M. A., 2015. Peramalan Jumlah Tamu Hotel di Kabupaten Demak Menggunakan Metode Support Vector Regression. *Jurnal Gaussian*, 4(4), pp. 785-794.
- Akande, K. O., Owolabi, T. O., Olatunji, S. O. & AbdulRaheem, A., 2017. A Hybrid Particle Swarm Optimization and Support Vector Regression Model for modelling permeability prediction of hydrocarbon reservoir. *Journal of Petroleum Science and Engineering*, Volume 150, pp. 43-53.
- Amanda, R., Yasin, H. & Prahutama, A., 2014. Analisis support vector regression (SVR) dalam Memprediksi Kurs Rupiah Terhadap Dollar Amerika Serikat. *jurnal gaussian*, Volume 3, pp. 849-857.
- Chen, R., Liang, C. Y., Hong, W. C. & Gu, D. X., 2015. Forecasting Holiday Daily Tourist Flow Based On Seasonal Support Vector Regression With Adaptive Genetic Algorithm. *Applied Soft Computing*, Volume 26, pp. 435-443.
- Furi, R. P., J. & Saepudin, D., 2015. *Prediksi Financial Time Series Menggunakan Independent Component Analysis dan*

- Support Vector Regression Studi Kasus : IHSB dan JII*, Bandung: s.n.
- Karim, M. S., Ginardi, H. & Fatichah, C., 2013. Prediksi Nilai Mata kuliah berdasarkan nilai Prasyarat Menggunakan Metode Support Vector Regression. *jurnal teknik pomits*, Volume 2.
- Maharesi, R., 2013. Penggunaan Support Vector Regression (SVR) Pada Prediksi Return Saham Syariah BEI. *Proceeding PESAT*, Volume 5.
- Mustakim, Buono, A. & Hermadi, I., 2016. Performance comparison between support vector regression and artificial neural network for prediction of oil palm production. *jurnal ilmu komputer dan informasi*, Volume 9, pp. 1-8.
- Nandi, 2008. Pariwisata dan Pengembangan Sumberdaya Manusia. *GEA*, 8(1).
- Nugroho, N. A. & Purqon, A., 2015. *Analisis 9 saham Sektor Industri di Indonesia Menggunakan Metode SVR*. s.l., SKF.
- PT. Oleng Sibuttong, 2016. *Arjuno Oleng Sibuttong*. [Online] Available at: <https://www.olengsibuttong.com/> [Accessed 5 Januari 2017].
- Rahmadiani, A. & Anggraeni, W., 2012. Implementasi Fuzzy Neural Network untuk Memperkirakan Jumlah Kunjungan Pasien Poli Bedah di Rumah Sakit Onkologi Surabaya. *Jurnal Tekni POMITS*, I(1), pp. 1-5.
- Rahma, F. N. & Handayani, H. R., 2013. Pengaruh Jumlah Kunjungan Wisatawan, Jumlah Onyek wisata dan pendapatan perkapita terhadap Penerimaan Sektor Pariwisata D Kabupaten Kudus. *Diponegoro Journal of Economics*, 2(2), pp. 1-9.
- Rantetadung, M., 2012. Analisis Pengaruh dukungan pemerintah dan kunjungan wisatawan terhadap pendapatan asli daerah di Kabupaten Nabire. *Jurnal Agroforensi*, Volume 1, pp. 25-32.
- Setiawan, A. Y. & Badriyanto, B. S., 2015. Perkembangan Industri Pariwisata di Kabupaten Jember Tahun 2003-2014. *Artikel Ilmiah Mahasiswa 2015*.
- Setiyoutami, A., Anggraeni, W. & Kusumawardani, R. P., 2012. *Prediksi Jumlah Kunjungan Pasien Poli Bedah di Rumah Sakit Onkologi Surabaya Menggunakan Fuzzy Time Series*, Surabaya: s.n.
- Suchaina, 2014. Pengaruh Kualitas Fasilitas Sarana dan Prasarana Terhadap Peningkatan Jumlah Pengunjung Wisata Danau Ranu Grati. *jurnal psikologi*, Volume II, pp. 89-109.
- Suci, K. W. & Irhamah, 2017. Peramalan Curah Hujan Sebagai Pendukung Kalender Tanam Padi di Pos Kadungadem Bojonegoro Menggunakan ARIMA, Support Vector Regression dan Genetic Algorithm-SVR. *Jurnal Sains dan Seni*, 6(1), pp. 55-61.
- Vijayakumar, S. & Wu, S., 1999. *Sequential Support Vector Classifiers and Regression*. Genoa, Italy, Saitama: RIKEN Brain Science Institute, The Institute for Physical and Chemical Research.