

Algoritma SVM untuk Memprediksi Pengunjung Wisata Musium di Jakarta

Fikri Abdul Najib¹, Nur Nafi'iyah²

^{1,2}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Islam Lamongan

ABSTRACT

Various kinds of tourist attractions in Jakarta become vacation or play destinations, ranging from nature tourism, malls, cinemas, amusement parks, or museums. Each individual has a variety of activities and routines, so it requires entertainment and time to let go of boredom. From the data.jakarta.go.id website there is a dataset of museum tour visits from both Indonesian and foreign tourists. From this dataset, it can be used to process and extract information. Digging and processing datasets is a data mining activity, which is implementing an algorithm to explore knowledge. The SVM algorithm is used to predict museum tourism visits in Jakarta, where there are variables of destination, month, type of visitors and number of visitors. There are 10 types of tourist destinations, and 2 types of visitors, namely domestic and foreign tourists. Where the prediction results from SVM on 222 data rows of museum tour visitors in Jakarta is poor. Proven from the difference in the value of real data with very high prediction results, and the error value is very high 2838303.5.

Article History

Received 2020-09-14
Revised 2020-10-01
Accepted 2020-11-04

Key words

*Predictions, Museum Tour
Visitors, SVM*

ABSTRAK

Berbagai macam tempat wisata yang ada di Jakarta menjadi tujuan berlibur atau bermain, mulai dari wisata alam, mall, bioskop, taman hiburan, atau museum. Setiap individu mempunyai aktivitas dan rutinitas bermacam-macam, sehingga membutuhkan hiburan dan waktu untuk melepaskan kejenuhan. Dari website data.jakarta.go.id didapatkan dataset kunjungan wisata museum baik dari wisatawan Indonesia maupun luar Indonesia. Dari dataset tersebut dapat dimanfaatkan untuk diolah dan digali informasinya. Menggali dan mengolah dataset adalah suatu kegiatan data mining, yaitu menerapkan suatu algoritma untuk menggali pengetahuan. Algoritma SVM digunakan untuk memprediksi kunjungan wisata museum di Jakarta, di mana terdapat variabel tempat destinasi, bulan, jenis pengunjung dan jumlah pengunjung. Tempat destinasi ada 10 jenis wisata, dan jenis pengunjung ada 2, yaitu wisatawan dalam negeri dan luar negeri. Di mana hasil prediksi dari SVM pada data 222 baris pengunjung wisata museum di Jakarta jelek. Dibuktikan dari nilai selisih data nyata dengan hasil prediksi sangat tinggi, dan nilai errornya sangat tinggi 2838303,5.

PENDAHULUAN

Beberapa penelitian terdahulu telah melakukan data mining dalam peramalan kunjungan wisatawan luar negeri ke Indonesia. Misalnya penelitian Naufal 2017, melakukan prediksi kunjungan wisatawan dari luar negeri ke Indonesia. Tempat yang digunakan pengambilan data wisatawan di enam tempat pintu masuk, yaitu: di Ngurah Rai, Batam, Tanjung Uban, Husein Sastranegara, Tanjung Pinang, Sepinggan. Hasil prediksi menggunakan algoritma SVM antara 62%-64% [1]. Penelitian Molydah, 2018 melakukan data mining untuk peramalan kunjungan wisatawan dari luar negeri ke Indonesia. Yang membedakan antara penelitian Naufal dengan Molydah adalah metode. Metode penelitian Molydah menggunakan SARIMA dan SVM. Hasil SARIMA dalam memprediksi kunjungan wisatawan luar negeri ke Indonesia nilai MAPE 7,07% dan SVM nilai MAPE 9,59% [2].

Penelitian terkait peramalan lainnya, yaitu penelitian Barry Render 2012. Manajer perusahaan tidak mengetahui kondisi yang terjadi di masa depan. Menambah persediaan bahan produksi dan menambah produksi selalu dilakukan walaupun masa berikutnya tidak ada permintaan. Dan jumlah persediaan produksi semakin meningkat banyak dan hasil keuntungan tidak terlihat signifikan. Tujuan peramalan adalah mengurangi ketidakpastian serta membuat pemikiran mengenai perkiraan barang yang harus distok atau diproduksi. Berbagai cara dilakukan untuk meminimalisir ketidakpastian, yaitu dengan peramalan atau prediksi. Di mana dibutuhkan dataset bertahun-tahun yang lampau untuk dapat mengenali pola penjualan atau produksi di perusahaan. Beberapa teknik

untuk peramalan, yaitu: regresi linear sederhana atau berganda, neural network, dan berbasis rata-rata (moving average)[3].

Pengertian prediksi adalah mengestimasi data yang akan datang. Estimasi adalah memperkirakan data berdasarkan kumpulan data masa lalu. Agar dapat melakukan prediksi diperlukan suatu pemodelan atau membuat model dari data masa lalu. Cara memodelkan data dengan menggali informasi data masa lalu atau mengolah data. Membuat model bisa menggunakan persamaan matematis atau rumus regresi [4].

Kemudian dari beberapa penelitian di atas maka dilakukan prediksi kunjungan wisata museum di Jakarta dengan algoritma SVM. Support vector machine (SVM) adalah algoritma pelatihan untuk mempelajari klasifikasi dan aturan regresi dari data, misalnya SVM dapat digunakan untuk mempelajari polinomial, fungsi basis radial (RBF) dan multi-layer perceptron (MLP). SVM pertama kali disarankan oleh Vapnik pada 1960-an untuk klasifikasi dan baru-baru ini menjadi bidang penelitian intensif karena perkembangan teknik dan teori ditambah dengan ekstensi untuk regresi dan estimasi kepadatan [5].

Di mana tujuan penelitian ini, adalah mengetahui hasil prediksi kunjungan wisata museum di Jakarta menggunakan algoritma SVM. Algoritma SVM dipilih untuk memprediksi karena hasil akurasi prediksi di penelitian “Peramalan Jumlah Wisatawan Mancanegara ke Indonesia” baik 64% [1].

TINJAUAN PUSTAKA

SVM melakukan training agar menghasilkan bobot, dan bias. Di mana dalam melakukan training akan melakukan update nilai bobot dan bias seperti Persamaan 1. Hasil dari training adalah suatu persamaan garis seperti Persamaan 2 [5].

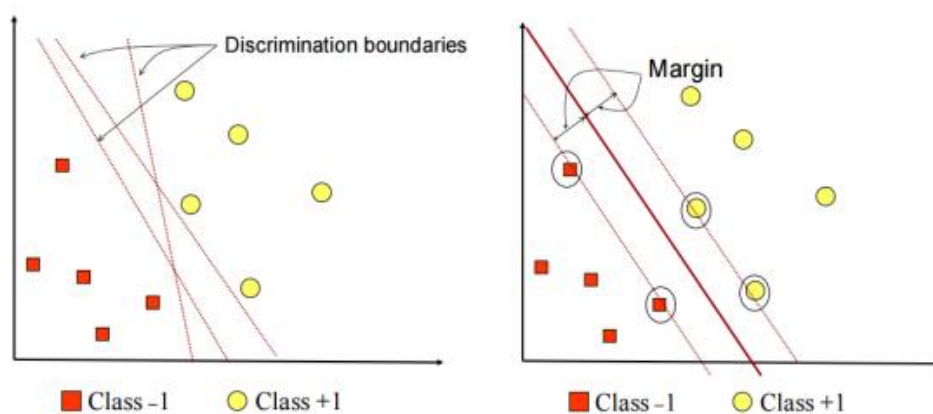
$$w^T x_i + b \geq 1 \quad (1)$$

$$w^T x_i + b \leq -1$$

Di mana x_i adalah variabel input. Hasil training berupa Persamaan 2.

$$f_{w,b}(x) = \text{sgn}(w^T x + b) \quad (2)$$

SVM akan menghasilkan 2 kelas dari suatu garis persamaan. Gambar 1 merupakan ilustrasi hyperplane dari proses training SVM. SVM (Support Vector Machine) adalah algoritma yang masuk di jaringan saraf tiruan. Cara kerja SVM akan melatih data di masa lalu agar menghasilkan suatu persamaan atau hyperplane seperti Persamaan 2. Dari Persamaan hasil belajar atau training digunakan untuk mengujicoba data baru [5].



Gambar 1. Ilustrasi Hasil Hyperplane SVM

Setiap data latih dinyatakan dalam (x,y) , x adalah variabel atau atribut *feature set*, y adalah label kelas. Klasifikasi linier SVM, dinyatakan dalam Persamaan 3 [6].

$$\begin{aligned}w.x_i + b &= 0 \\w.x_i + b &\leq -1 \\w.x_i + b &\geq +1\end{aligned}\tag{3}$$

Di mana w dan b adalah vektor bobot atau bobot, dan x_i adalah variabel input.

SVM mempunyai karakteristik yang sama seperti jaringan saraf tiruan. Jaringan saraf tiruan melakukan training atau pembelajaran agar dapat mengenali pola. Hasil training berupa bobot dan bias atau suatu vektor. Sehingga SVM membutuhkan proses training atau pembelajar agar menghasilkan suatu model [6].

Data mining merupakan disiplin ilmu yang mempelajari data atau dataset agar mengetahui pola data di masa lalu. Data mining erat kaitannya dengan dataset, teknik pengenalan pola dari data serta pemodelan atau mengvisualisasikan. Teknik pengenalan pola dari data di masa lalu disebut dengan *machine learning*. Proses machine learning akan menggali data di masa lalu yang hasilnya berupa model atau suatu persamaan matematika. Sehingga machine learning sangat erat kaitannya dengan statistik [7, 11]. Salah satu teknik menggali data atau machine learning adalah jaringan saraf tiruan. Jaringan saraf tiruan akan bekerja seperti cara kerja otak manusia. Jaringan saraf tiruan sering disebut dengan artificial neural network (ANN). ANN akan belajar dari data masa lalu yang selanjutnya akan dimodelkan dalam bentuk persamaan, atau bobot atau vektor. SVM masuk dalam teknik ANN [7-8].

ANN ataupun SVM merupakan teknik mendapatkan model atau pengenalan pola dari data di masa lalu. ANN ataupun SVM adalah algoritma data mining yang sering digunakan untuk melakukan prediksi, klasifikasi. Data yang akan diolah dipelajari atau ditraining sampai batas iterasi yang telah ditentukan. Semakin besar dataset waktu training juga akan lama, dan semakin bervariasi data maka iterasi yang dibutuhkan untuk mengenali pola juga harus banyak [7-8].

Algoritma SVM sering digunakan untuk memprediksi data yang akan datang. Algoritma SVM membutuhkan data beberapa tahun di masa lalu, kemudian dipelajari agar menghasilkan persamaan atau model atau hyperplane. Algoritma SVM digunakan untuk memprediksi kunjungan wisata Taman Tabebuaya dengan hasil akurasi 86% [9]. Bukti lainnya bahwa algoritma SVM mempunyai performa yang baik adalah nilai error atau nilai selisih dari data nyata sangat kecil dalam memprediksi kunjungan wisata ke Bali [10].

METODE

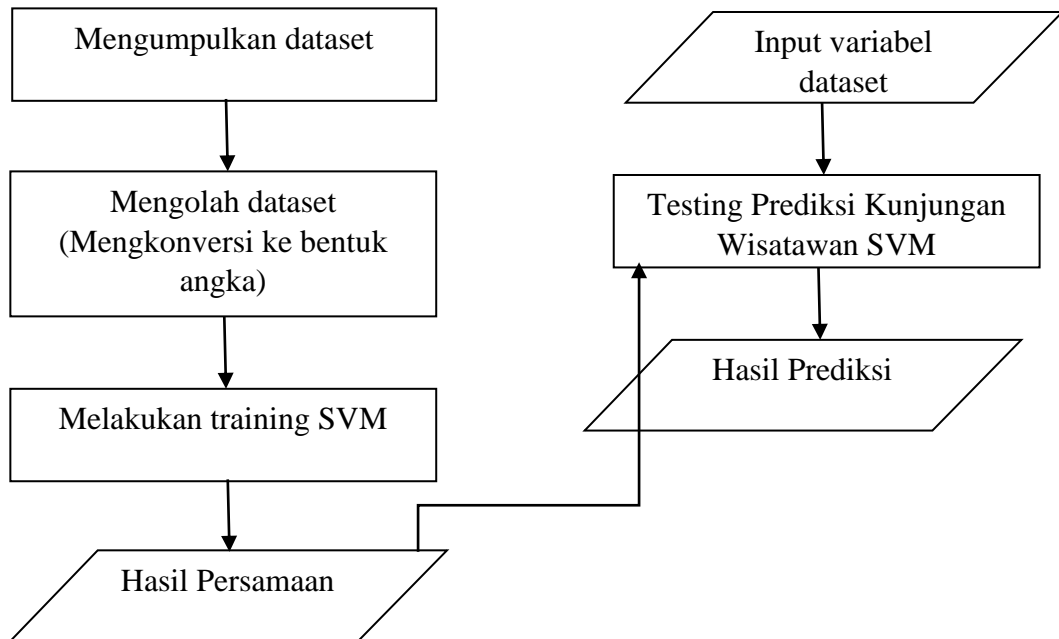
Penelitian ini melakukan prediksi pengunjung wisata museum di Jakarta. Agar dapat menyelesaikan penelitian ini, maka harus menyiapkan suatu aplikasi untuk mengolah data tersebut. Aplikasi yang dibutuhkan untuk mengolah adalah Matlab 2017b. Aplikasi Matlab termasuk dalam *machine learning*, di mana terdapat beberapa algoritma untuk mengolah data. Data yang diolah disimpan dalam file Excel agar dapat memudahkan dalam menganalisa. Alur dari sistem dalam bentuk flowchart dan diagram DFD secara berturut-turut seperti Gambar 2 dan Gambar 3.

Sistem yang dibangun adalah memprediksi kunjungan wisatawan ke museum Jakarta dengan SVM. Sistem menggunakan tool Matlab untuk melakukan training. Dataset dikumpulkan terlebih dahulu dan disimpan dalam bentuk excel, serta dilakukan konversi. Tabel yang dibuat terdapat variabel atau kolom nama destinasi wisata, bulan, jenis pengunjung, serta jumlah pengunjung. Tabel awal sebelum dilakukan konversi ke bentuk angka seperti Tabel 1. Fitur-fitur yang digunakan dalam proses training adalah nama destinasi wisata, bulan, jenis pengunjung (seperti dalam tabel 1). Selanjutnya Tabel 1 dikonversi ke bentuk angka dengan pemodelan seperti Tabel 2.

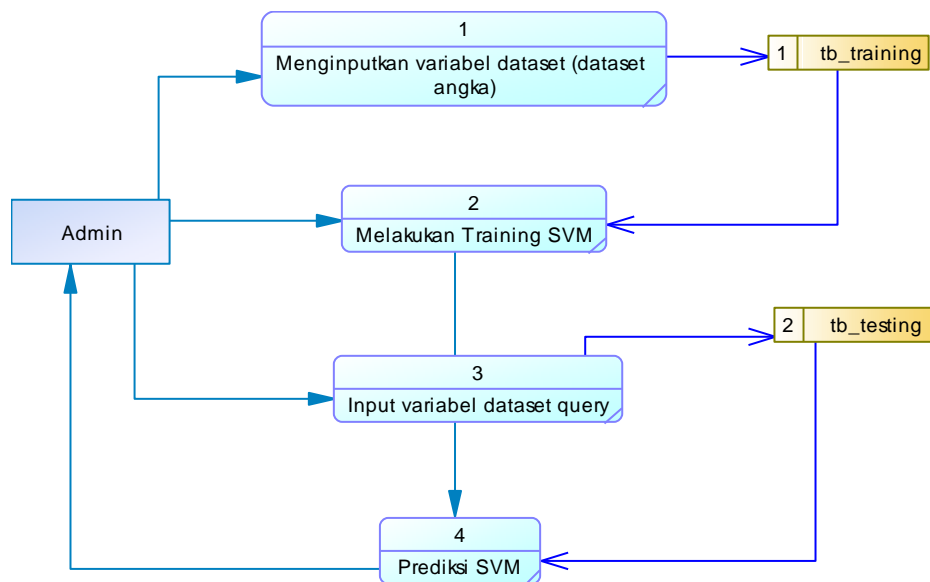
Tabel 1. Contoh Dataset Training

nama destinasi wisata	bulan	jenis pengunjung	jumlah pengunjung
M. Seni Rupa dan Keramik	2015-01	Wisatawan Mancanegara	14
M. Seni Rupa dan Keramik	2015-01	Wisatawan Nusantara	21141
M. Seni Rupa dan Keramik	2015-02	Wisatawan Mancanegara	100
M. Seni Rupa dan Keramik	2015-02	Wisatawan Nusantara	9137
M. Seni Rupa dan Keramik	2015-03	Wisatawan Mancanegara	502

M. Seni Rupa dan Keramik	2015-03	Wisatawan Nusantara	16300
M. Seni Rupa dan Keramik	2015-04	Wisatawan Mancanegara	870
M. Seni Rupa dan Keramik	2015-04	Wisatawan Nusantara	12588
M. Seni Rupa dan Keramik	2015-05	Wisatawan Mancanegara	331
M. Seni Rupa dan Keramik	2015-05	Wisatawan Nusantara	2902
M. Seni Rupa dan Keramik	2015-06	Wisatawan Mancanegara	146



Gambar 2. Alur Prediksi Algoritma SVM



Gambar 3. Diagram Flow Sistem

Tabel 2. Hasil Konversi Data

nama destinasi wisata	bulan	jenis pengunjung	jumlah pengunjung
1	1	1	14
1	1	2	21141
1	2	1	100
1	2	2	9137
1	3	1	502
1	3	2	16300
1	4	1	870
1	4	2	12588
1	5	1	331
1	5	2	2902
1	6	1	146
1	6	2	7558
1	7	1	176
1	7	2	17712

Di mana ada 10 nama destinasi wisata dalam dataset dan ada 2 jenis pengunjung, sehingga dilakukan konversi ke bentuk angka seperti Tabel 3 dan Tabel 4.

Tabel 3. Konversi Variabel Nama Destinasi Wisata

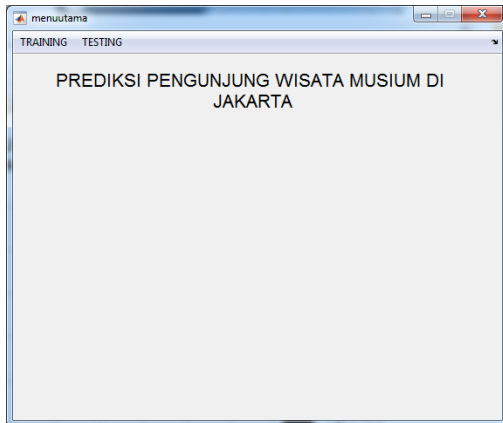
Nama Destinasi Wisata	Hasil Konversi
M. Seni Rupa dan Keramik	1
Monumen Nasional	2
Museum Bahari	3
Museum Joang '45	4
Museum Nasional	5
Museum Sejarah Jakarta	6
Museum Tekstil	7
Museum Wayang	8
Pel. Sunda	9
KelapaTaman Arkeologi P. Onrust	10

Tabel 4. Konversi Variabel Jenis Pengunjung

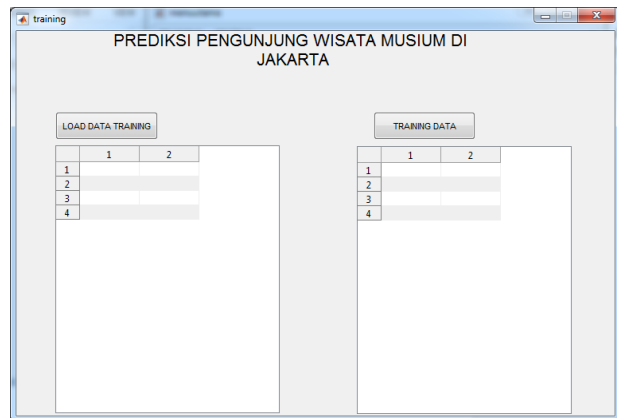
Jenis Pengunjung	Hasil Konversi
Wisatawan Mancanegara	1
Wisatawan Nusantara	2

HASIL DAN PEMBAHASAN

Penelitian ini terkait memprediksi pengunjung wisata di musium Jakarta. Di mana data yang digunakan adalah para pengunjung wisata serta jenis wisatawan di tahun 2015. Adapun tampilan dari program pertama adalah Gambar 4. Di mana dalam Gambar 4 terdapat dua menu, menu Training digunakan untuk mentraining dataset serta menampilkan hasil prediksi. Adapun tampilan menu Training dalam Gambar 5. Dalam Gambar 5 ada button untuk mengambil file excel di dalam folder dan ditampilkan di tabel GUI, selanjutnya akan dilakukan training untuk menghasilkan bobot serta hasil prediksi. Hasil prediksi tersebut akan ditampilkan di tabel GUI, seperti Gambar 6.



Gambar 4. Tampilan Program



Gambar 5. Menu Training



Gambar 6. Tampilan Saat Dilakukan Training

Berdasarkan Gambar 4 terdapat menu untuk testing, di mana user cukup menginputkan 3 variabel, yaitu nama destinasi, bulan, dan jenis pengunjung maka akan menampilkan prediksi pengunjung, seperti Gambar 7.



Gambar 7. Tampilan Melakukan Ujicoba Prediksi

Penelitian ini menerapkan algoritma SVM untuk memprediksi pengunjung wisata di Jakarta. Di mana data disimpan terlebih dahulu di Excel kemudian dibaca dan diproses training menggunakan fungsi `fitcsvm(input_var,output_var)`. Dan diprediksi dengan fungsi `predict(training,query_test)`. Selanjutnya melakukan proses training menggunakan fungsi `fitcsvm`, dan proses memprediksi menggunakan fungsi `predict`. Hasil prediksi dalam Tabel 5. Dengan nilai perhitungan akurasi menggunakan Persamaan 4:

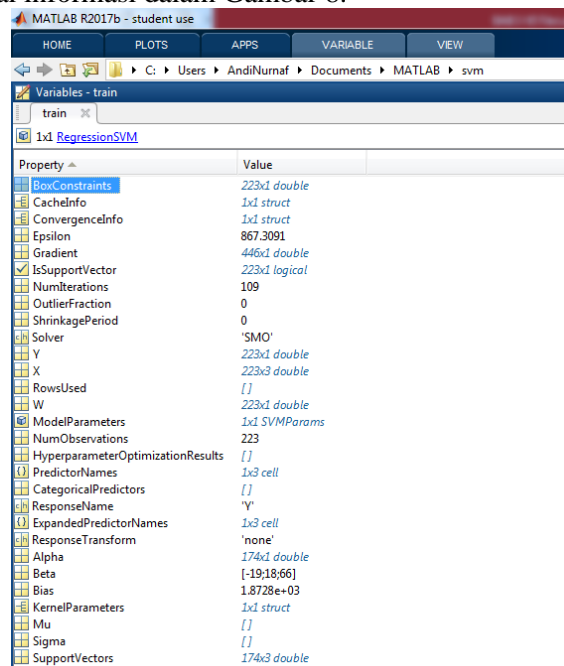
$$RMS\ Error = \sqrt{\frac{\sum_i (y - y')^2}{n}} \quad (4)$$

Dengan y adalah data nyata, dan y' adalah hasil prediksi, dan n total data uji, dengan nilai RMS error adalah 1987,06.

Tabel 5. Hasil Prediksi Ujicoba Pertama

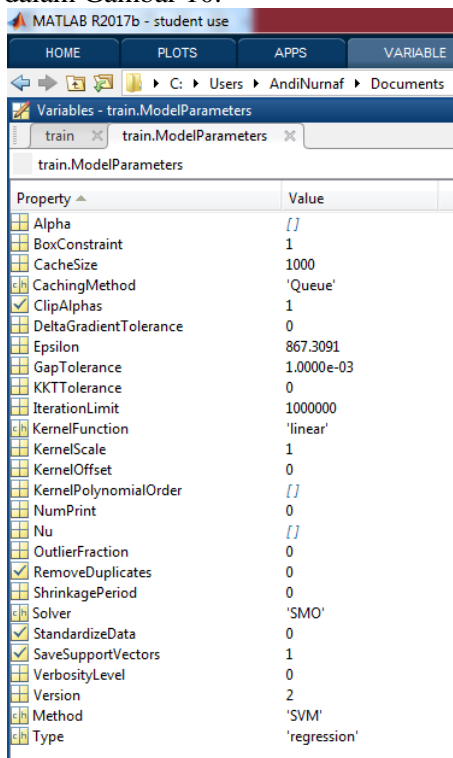
nama destinasi wisata	jenis bulan	jumlah pengunjung	Prediksi	Selisih Error
1	1	14	1937.81	1923.809118
1	1	21141	2003.81	19137.19088
1	2	100	1955.81	1855.809118
1	2	9137	2021.81	7115.190882
1	3	502	1973.81	1471.809118
1	3	16300	2039.81	14260.19088
1	4	870	1991.81	1121.809118
1	4	12588	2057.81	10530.19088
1	5	331	2009.81	1678.809118
1	5	2902	2075.81	826.1908821
1	6	146	2027.81	1881.809118

Di mana terlihat dalam Tabel 5 hasil prediksi yang diperoleh sangat jelek, terbukti dari nilai selisih antara nilai data nyata terhadap hasil prediksi sangat tinggi. Adapun proses training dalam fungsi Matlab mempunyai informasi dalam Gambar 8.

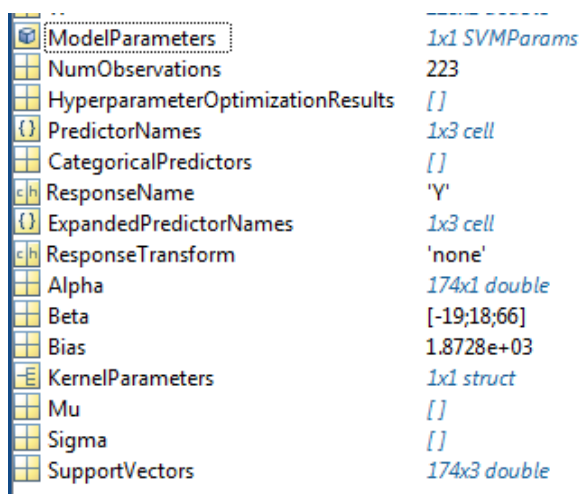


Gambar 8. Informasi Fungsi SVM dalam Matlab

Iterasi dari ujicoba pada Gambar 8 sebanyak 109, sedangkan dalam ujicoba pertama di Gambar 8 menggunakan fungsi SMO, terlihat dalam Gambar 9. Dan informasi nilai bias serta bobot masing-masing variabel input dalam Gambar 10.



Gambar 9. Informasi Model Optimasi



Gambar 10. Informasi Bias dan Bobot

KESIMPULAN

Penelitian ini dapat disimpulkan sebagai berikut: Sistem prediksi kunjungan wisata museum di Jakarta ini dibuat dengan menggunakan aplikasi Matlab. Data terlebih dahulu disimpan di dalam file Excel, selanjutnya dibaca dan dilakukan prediksi. Dalam penelitian ini menerapkan dua cara, yaitu data murni diolah untuk ditraining dan diprediksi, hasilnya kurang bagus. Dan cara kedua data dinormalisasi dengan skaling (0,1) dan hasilnya masih kurang maksimal. Dalam mengimplementasikan algoritma SVM untuk memprediksi kunjungan wisata museum di Jakarta, peneliti menggunakan fungsi fitcsvm. Di mana fungsi tersebut digunakan untuk mentraining serta memprediksi data dengan fungsi predict. Dan hasil prediksi kunjungan wisata menggunakan algoritma SVM kurang baik hasilnya.

DAFTAR PUSTAKA

- [1] M. F. NAUFAL, "PERAMALAN JUMLAH WISATAWAN MANCANEGERA YANG DATANG KE INDONESIA BERDASARKAN PINTU MASUK MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM)," Departemen Sistem Informasi ITS, Surabaya, 2017.
- [2] S. Molydah, "ANALISIS PERBANDINGAN IMPLEMENTASI SARIMA DAN SUPPORT VECTOR MACHINE (SVM) DALAM PREDIKSI JUMLAH WISATAWAN MANCANEGERA," Statistik UII, Yogyakarta, 2018.
- [3] B. Render, R. M. Stair and M. E. Hanna, Quantitative Analysis for Management, New Jersey: Pearson University of Houston, 2012.
- [4] I. M. Tirta, Pengantar Metode Simulasi Statistika dengan Aplikasi R dan S+, Jember: Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember, 2004.
- [5] R. Burbidge and B. Buxton, An Introduction to Support Vector Machines for Data Mining, Computer Science UK.
- [6] E. Prasetyo, Data Mining Konsep dan Aplikasi Menggunakan Matlab, Yogyakarta: Andi, 2012.
- [7] A. Hermawan, Jaringan Saraf Tiruan Teori dan Aplikasi, Yogyakarta: Andi, 2006.
- [8] S. Susanto and D. Suryadi, Pengantar Data Mining Menggali Pengetahuan dari Bongkahan Data, Yogyakarta: Andi, 2010.
- [9] A. Darmawan, N. Kustian and W. Rahayu, "Implementasi Data Mining Menggunakan Model SVM untuk Prediksi Kepuasan Pengunjung Taman Tabebuya," *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, vol. 2, no. 3, pp. 299-307, 2018.
- [10] I. Gunawan, "Penerapan Algoritma Support Vector Machine (SVM) untuk Memprediksi Tingkat Kunjungan Wisatawan Asing," SKRIPSI Jurusan Teknik Elektro-Fakultas Teknik UM, Malang, 2019.
- [11] B. Santosa, Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis Teori dan Aplikasi, Yogyakarta: Graha Ilmu, 2007.