

BAB II

TINJAUAN PUSTAKA

2.1. Landasan Teori

2.1.1 Pariwisata

Pariwisata adalah perjalanan dari suatu tempat ketempat lain, bersifat sementara, dilakukan perorangan atau kelompok, sebagai usaha mencari keseimbangan atau keserasian dan kebahagiaan dengan lingkungan dalam dimensi sosial, budaya, alam, dan ilmu pariwisata didefinisikan sebagai bentuk suatu proses kepergian sementara dari seorang, lebih menuju ketempat lain diluar tempat tinggalnya. Dorongan kepergiannya adalah karena berbagai kepentingan baik karena kepentingan ekonomi, sosial, budaya (I.P Sudana, 2013).

2.1.2 Prediksi

Prediksi (*forecasting*) merupakan suatu kegiatan untuk memperkirakan apa yang akan terjadi pada masa yang akan datang. Pengumpulan data yang relevan berupa informasi dapat menghasilkan peramalan yang akurat disertai pemilihan teknik peramalan yang tepat maka pemanfaatan informasi data akan diperoleh secara optimal (Martiningtyas, N. 2004).

Peramalan dibedakan menjadi dua jenis yang berbeda berdasarkan metode peramalan yang digunakan, peramalan dibedakan menjadi metode kualitatif dan metode kuantitatif. Metode kualitatif merupakan metode peramalan yang tidak menggunakan data historis masa lalu, lebih didasarkan pada intuisi. Metode

kuantitatif merupakan metode peramalan yang menggunakan data historis masa lalu, memanipulasi data historis yang tersedia secara memadai dan tanpa intuisi, metode ini umumnya didasarkan pada analisis statistik (Makridakis, Wheelwright and McGee. 1991).

Menurut Makridakis, Wheelwright, dan McGee (1983, h.8-9), peramalan kuantitatif dapat diterapkan bila tiga kondisi terpenuhi yaitu :

- Tersedianya informasi mengenai keadaan waktu yang lalu
- Informasi itu dapat dikuantitatifkan dalam bentuk data numerik
- Dapat diasumsikan bahwa beberapa aspek dari pola di waktu yang lalu akan berlanjut ke waktu yang akan datang.

Karakteristik dari peramalan kuantitatif adalah sebagai berikut
(Kucharavy, D. De Guio, R. 2005):

- Melibatkan proyeksi dari masa lalu ke masa mendatang
- Lebih *Scientific*
- Relatif bebas dari persepsi seseorang
- Lebih bersifat Objektif
- Memungkinkan adanya analisis kesalahan / *error analysis*
- Dapat dilakukan secara berulang (*Reproducible*) Dalam waktu yang sama , hasil yang sama dapat diperoleh dari berbagai teknik yang berbeda

2.1.3 Algoritma K-Means

K-means clustering merupakan salah satu metode data *clustering* non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih cluster/kelompok. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan cluster/kelompok yang lain. Sehingga data yang berada dalam satu cluster/kelompok memiliki tingkat variasi yang kecil (Agusta, Y. 2007).

Langkah-langkah melakukan clustering dengan metode *K-Means* adalah sebagai berikut (Santosa, B. (2007):

- a. Pilih jumlah cluster k.
- b. Inisialisasi k pusat cluster yang dilakukan dengan cara random. Pusat-pusat cluster diberi nilai awal dengan angka-angka random.
- c. Alokasikan semua data/ objek ke cluster terdekat. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat cluster. Jarak antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam cluster mana.
- d. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak *Euclidean* yang dirumuskan sebagai berikut:

$$D(i,j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2} \quad (1)$$

Dimana :

$D(i,j)$ = Jarak data ke i ke pusat *cluster* j

X_{ki} = Data ke i pada atribut data ke j

X_{kj} = Titik pusat ke j pada atribut ke k

- e. Hitung kembali pusat *cluster* dengan keanggotaan cluster yang sekarang. Pusat *cluster* adalah rata-rata dari semua data/ objek dalam *cluster* tertentu. Dapat juga menggunakan median dari cluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.
- f. Tugaskan lagi setiap objek memakai pusat *cluster* yang baru. Jika pusat *cluster* tidak berubah lagi proses *clustering* selesai. Atau, kembali ke langkah nomor 3 sampai pusat *cluster* tidak berubah lagi.

2.1.4 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, Y. D. (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 :

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

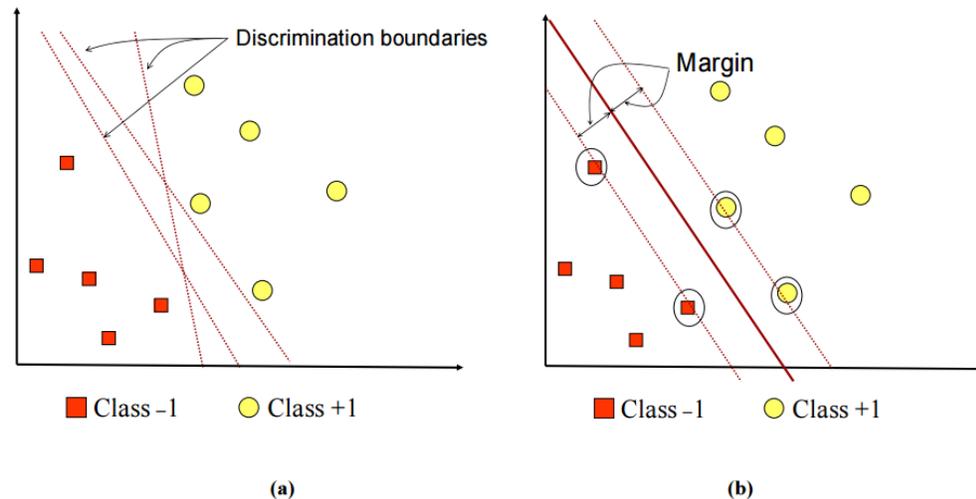
- b. Jika data menyebar jauh dari garis diagonal dan tidak mengikuti arah garis diagonal, maka dapat disimpulkan bahwa model tidak memenuhi asumsi normalitas”.

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

2.1.5 Support Vector Machine (SVM)

Support Vector Machine (SVM) dikembangkan oleh Boser, Guyon, Vapnik, dan pertama kali dipresentasikan pada tahun 1992 di *Annual Workshop on Computational Learning Theory* (Vapnik V.N). *Support vector machine* (SVM) merupakan suatu teknik yang relatif baru dan saat ini telah banyak digunakan untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi, yang sangat populer belakangan ini. *Support vector machine* berada dalam satu kelas dengan *Neural Network* dalam hal fungsi dan kondisi permasalahan yang bisa diselesaikan, keduanya masuk kedalam kelas *supervised learning* (Santoso, B., 2007).

Secara teoritik *Support Vector Machine* dikembangkan untuk menyelesaikan permasalahan klasifikasi pada dua kelas dengan mencari *hyperplane* terbaik. *Hyperplane* merupakan fungsi untuk memisahkan antara dua kelas pada *input space* , sehingga dari data yang tersebar dapat dilakukan klasifikasi dan analisa regresi. Vapnik menerangkan bahwa setiap permasalahan dapat dimodelkan dengan menggunakan SVM (Vapnik V.N) . Ilustrasi SVM dapat dilihat pada Gambar 2.1.



Gambar 2.1 Ilustrasi SVM

Sumber : “*Support Vector Machine Teori dan Aplikasinya dalam BioInformatika*”, 2003

Pada Gambar 2.2(a) menunjukkan alternatif garis pemisah (*discrimination boundaries*) dimana *pattern* yang tergabung pada *class* -1 disimbolkan dengan warna merah (kotak), sedangkan *pattern* pada *class* +1, disimbolkan dengan warna kuning(lingkaran)., sedangkan pada Gambar 2.2(b) diperlihatkan bahwa terdapat garis *hyperplane* yang tepat berada diantara dua buah kelas. Prinsip dasar dari analisis ini adalah menemukan *hyperplane* terbaik yakni dengan meminimalkan kesalahan klasifikasi dan memaksimalkan margin geometriknya seperti pada Gambar 2.2(b). Usaha untuk mencari lokasi *hyperplane* ini merupakan inti dari proses pembelajaran pada SVM (Nugroho, A. S., Witarto, A. B., & Handoko, D. (2003)).

Pengerjaan dengan menggunakan SVM dibagi menjadi 3 (Chiu, D. Y., & Chen, P. J. (2009)),yaitu:

2.1.5.1. Support Vector Machine Linier

Prinsip dasar SVM adalah *linier classifier*, yaitu kasus klasifikasi yang dapat dipisahkan secara linier. Data yang tersedia dinotasikan sebagai $x_i \in R^d$ sedangkan label masing-masing dinotasikan $y_i \in \{-1, +1\}$ untuk $i = 1, 2, \dots, l$, yang mana l adalah jumlah data. Diasumsikan kedua class -1 dan $+1$ dapat terpisah secara sempurna oleh hyperplane berdimensi d , yang didefinisikan :

$$\vec{W} \cdot \vec{X} + b = 0 \quad (2)$$

Pattern x_i yang termasuk *class* -1 (sampel negatif) dapat dirumuskan sebagai pattern yang memenuhi pertidaksamaan dimana nilai menunjukkan adanya penurunan

$$\vec{w} \cdot \vec{x}_i + b \leq -1 \quad (3)$$

sedangkan pattern x_i yang termasuk *class* $+1$ (sampel positif) menunjukkan adanya kenaikan pada nilai

$$\vec{w} \cdot \vec{x}_i + b \geq 1 \quad (4)$$

Margin terbesar dapat ditemukan dengan memaksimalkan nilai jarak antara hyperplane dan titik terdekatnya, yaitu $1/\|\vec{w}\|$. Hal ini dapat dirumuskan sebagai *Quadratic Programming* (QP) problem, yaitu mencari titik minimal persamaan (5), dengan memperhatikan constraint persamaan (6).

$$\min_{\vec{w}} \tau(w) = \frac{1}{2} \|\vec{w}\|^2 \quad (5)$$

$$y_i(\vec{x}_i \cdot \vec{w} + b) - 1 \geq 0, \forall_i \quad (6)$$

Problem ini dapat dipecahkan dengan berbagai teknik komputasi, di antaranya *Lagrange Multiplier*.

$$L(\vec{w}, b, \alpha) = \frac{1}{2} \|\vec{w}\|^2 - \sum_{i=1}^l \alpha_i (y_i (\vec{x}_i \cdot \vec{w} + b) - 1) \quad (7)$$

α_i adalah Lagrange multipliers, yang bernilai nol atau positif ($\alpha_i > 0$). Nilai optimal dari persamaan (7) dapat dihitung dengan meminimalkan L terhadap \vec{w} dan b , dan memaksimalkan L terhadap α_i . Dengan memperhatikan sifat bahwa pada titik optimal *gradient* $L=0$, persamaan (7) dapat dimodifikasi sebagai maksimalisasi problem yang hanya mengandung α_i saja, sebagaimana seperti pada persamaan (8) di bawah.

Maximize :

$$\sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j \vec{x}_i \cdot \vec{x}_j \quad (8)$$

Subject to :

$$\alpha_i \geq 0 (i = 1, 2, \dots, l) \quad \sum_{i=1}^l \alpha_i y_i = 0 \quad (9)$$

Dari hasil perhitungan ini diperoleh α_i yang kebanyakan bernilai positif. Data yang berkorelasi dengan α_i yang positif inilah yang disebut dengan *support vector*.

Penjelasan di atas berdasarkan asumsi bahwa kedua belah *class* dapat terpisah secara sempurna oleh *hyperplane*. Akan tetapi, umumnya dua buah *class* pada input space tidak dapat terpisah secara sempurna. Hal ini menyebabkan *constraint* pada persamaan (6) tidak dapat terpenuhi, sehingga

optimisasi tidak dapat dilakukan. Untuk mengatasi masalah ini, SVM dirumuskan ulang dengan memperkenalkan teknik *softmargin*. Dalam *softmargin*, persamaan (6) dimodifikasi dengan memasukkan *slack variable* ξ_i ($\xi_i > 0$) sebagai berikut .

$$y_i(\vec{x}_i \cdot \vec{w} + b) \geq 1 - \varepsilon_i, \forall_i \quad (10)$$

Dengan demikian persamaan (5) diubah menjadi :

$$\min_{\vec{w}} \tau(\vec{w}, \varepsilon) = \frac{1}{2} \|\vec{w}\|^2 + C \sum_{i=1}^l \varepsilon_i \quad (11)$$

Paramater C dipilih untuk mengontrol *tradeoff* antara margin dan *error* klasifikasi . Nilai C yang besar berarti akan memberikan penalti yang lebih besar terhadap error klasifikasi tersebut.

2.1.5.2 Support Vector Machine Linier Separation for nonseparable data

Dalam banyak kasus terkadang data tidak dapat dipisahkan dengan menggunakan pemisah linear. Namun, *hyperplane* dengan tingkat *error* yang paling minimum dapat dicari. Sehingga nantinya akan ada variabel *slack* non-negatif ξ_i , $i = 1, \dots, m$. Sehingga diperoleh persamaan dibawah ini:

$$(w \cdot x) + b \geq +1 - \varepsilon_i \rightarrow y_i = +1 \quad (12)$$

$$(w \cdot x) + b \leq +1 - \varepsilon_i \rightarrow y_i = -1 \quad (13)$$

Untuk menentukan *hyperplane* dan meminimasi *error* yang dihasilkan, fungsi tujuan dari permasalahannya adalah sebagai berikut:

$$\min_{w,b,\varepsilon} \frac{1}{2} \|W\|^2 + C \sum_{i=1}^p \varepsilon_i \quad (14)$$

Batasan :

$$y_i(w \cdot x + b) \geq 1 - \varepsilon, \varepsilon \geq 0, i = 1, \dots, p \quad (15)$$

Dalam kasus optimasi untuk data yang *nonseparable*, rumus *Langrange* juga dapat digunakan dalam menemukan solusi yang paling optimal.

$$\max W(\alpha) = \sum_{i=1}^P \alpha_i - \sum_{i=1}^P \sum_{j=1}^P \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (16)$$

Dimana α merupakan pengali *Langrange*

Batasan:

$$\sum_{i=1}^P \alpha_i y_i = 0, 0 \leq \alpha_i \leq C, i = 1, 2, \dots, p \quad (17)$$

2.1.5.3 Support Vector Machine Non-Linear Separation

Pada kasus pemisahan non-linear digunakan fungsi *mapping*, yang biasa disebut dengan fungsi *kernel*. Fungsi ini dapat melakukan *mapping* ruang input dari data *training* ke *feature space* dimensi yang lebih tinggi. Fungsi kernel dapat dilihat pada persamaan berikut :

$$\phi(x_i)\phi(x_j) := k(x_i, x_j) \quad (18)$$

Pada pemisahan nonlinear rumus *Langrange* juga dapat digunakan untuk menemukan solusi.

$$\max W(\alpha) = \sum_{i=1}^P \alpha_i - \sum_{i=1}^P \sum_{j=1}^P \alpha_i \alpha_j y_i y_j k(x_i, x_j) \quad (19)$$

Fungsi apapun yang dapat memenuhi *Mercer's condition* dapat digunakan sebagai fungsi kernel. Pada fungsi kernel, nantinya akan ditentukan nilai parameter seperti nilai C (*cost*) dan (γ) (*gamma*). Jika diadopsi dengan fungsi kernel Radial (20) , maka fungsi kernel untuk SVM adalah sebagai berikut :

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (20)$$

Fungsi kernel lainnya dapat dilihat pada tabel 2.1

Table 2.1 Jenis kernel SVM

Jenis Kernel	Definisi
Linier K	$K = (x, y) = x \cdot y$
Polynomial $K(x, y)$	$K(x, y) = (x \cdot y + c)^d$
Gaussian RBF	$K(x_i, x_j) = \exp(-\gamma \ x_i - x_j\ ^2)$
Sigmon (tangen hiperbolik)	$K(x, y) = \tanh(\sigma(x \cdot y) + c)$
Invers multiquadric $K(x, y)$	$K(x, y) = \frac{1}{\sqrt{\ x - y\ ^2 + c^2}}$

Fungsi umum yang digunakan untuk non linear SVM dapat dituliskan pada persamaan :

$$f(\phi(\vec{x})) = \langle \vec{w} \cdot \phi(\vec{w}) \rangle + b \quad (21)$$

Dimana :

x = vektor input

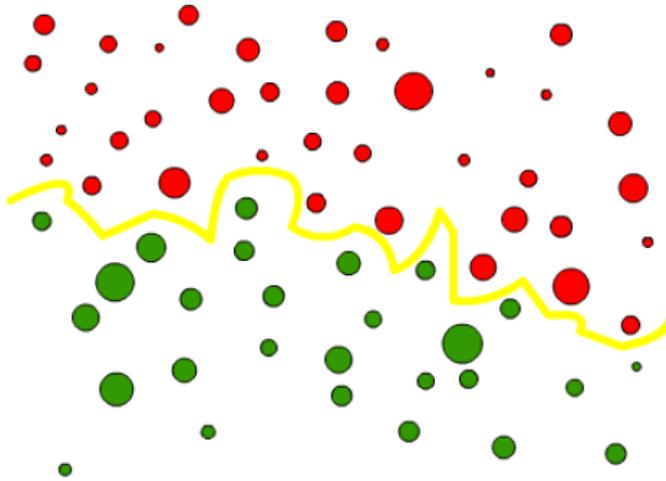
w = parameter bobot

b = bias

Dengan memasukkan *Lagrange multipliers* dan memanfaatkan konstrain pengoptimalan , maka rumus (21) dapat menghasilkan fungsi sebagai berikut:

$$f(x, a_i, a_i^*) = \sum_{i=1}^n (a_i - a_i^*) K(x, x_i) + b \quad (22)$$

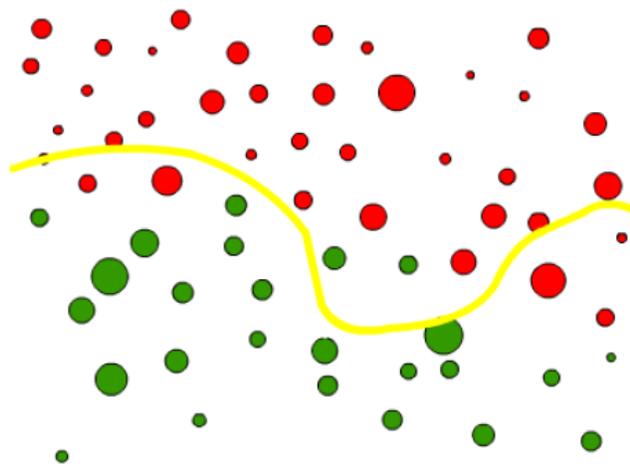
Ilustrasi pemisahan nonlinear dapat dilihat pada Gambar 2.2 dan Gambar 2.3.



Gambar 2.2 Pemisahan dengan kernel

Sumber : (<http://onlinesvr.altervista.org/Theory/Theory%20-%2001%20-%20Introduction.htm>)

Gambar 2.3 menunjukkan contoh pemisahan nonlinear pada dua kelas data.



Gambar 2.3 Pemisahan dengan kernel 2

Sumber : (<http://onlinesvr.altervista.org/Theory/Theory%20-%2001%20-%20Introduction.htm>)

Namun, jika pemilihan parameternya salah, maka akan dihasilkan pemisahan yang kurang optimal seperti terlihat pada Gambar 2.4.

SVM untuk melakukan peramalan kedepan dapat menggunakan persamaan (23) (Chiu, D. Y., & Chen, P. J. (2009))

$$f(x) = \text{Sign} \left(\sum_{i=1}^p y_i a_i K(x, x_i) + y_i - W \cdot x_i \right) \quad (23)$$

2.1.5.4 Grid Search Optimization.

Terdapat beberapa algoritma untuk menentukan parameter optimal pada model SVM, salah satunya adalah menggunakan algoritma grid search. Algoritma ini membagi jangkauan parameter yang akan dioptimalkan kedalam grid dan melintasi semua titik untuk mendapatkan parameter yang optimal. Dalam aplikasinya, algoritma grid search harus dipandu oleh beberapa metrik kinerja, biasanya diukur dengan cross-validation pada data training. Oleh karena itu disarankan untuk mencoba beberapa variasi pasangan parameter pada hyperplane SVM. Pasangan parameter yang menghasilkan akurasi terbaik yang didapatkan dari uji *crossvalidation* merupakan parameter yang optimal. Parameter optimal tersebut yang selanjutnya digunakan untuk model SVM terbaik. Setelah itu, model SVM tersebut digunakan untuk memprediksi data testing untuk mendapatkan generalisasi tingkat akurasi model. Menurut Leidiyana, *cross-validation* adalah pengujian standar yang dilakukan

untuk memprediksi *error rate*. Data training dibagi secara random ke dalam beberapa bagian dengan perbandingan yang sama kemudian *error rate* dihitung bagian demi bagian, selanjutnya hitung rata-rata seluruh *error rate* untuk mendapatkan *error rate* secara keseluruhan (Yasin, H., Prahutama, A., & Utami, T. W), (2014).

2.1.5.5 Evaluasi Hasil Peramalan

Hasil peramalan tidak ada yang dapat dipastikan benar seluruhnya, selalu ada penyimpangan nilai atau perbedaan nilai dengan kenyataan. Berikut adalah evaluasi kinerja yang dilakukan untuk mengetahui seberapa baik peramalan yang dihasilkan atau seberapa kecil penyimpangan yang dialami (Prasetyo, E.,2014) . Adapun evaluasi yang dilakukan berupa (Makridakis, Wheelwright and Mcgee. 1991):

2.1.5.5.1 Mean Square Error (MSE)

Mean Square Error (MSE) menghasilkan error yang menunjukkan perbedaan dari hasil estimasi dengan hasil yang diestimasi. Perbedaan yang muncul diakibatkan oleh keacakan pada data atau kurang akuratnya estimasi yang diperoleh.

Rumus MSE secara umum dapat dituliskan :

$$MSE = \frac{\sum_{t=1}^n (x_t - f_t)^2}{n} \times 100\% \quad (24)$$

Dimana :

n = Jumlah Sampel

x_t = Nilai Aktual Indeks pada periode ke-t

f_t = Nilai Prediksi Indeks pada periode ke-t

2.1.5.5.2 Mean Absolute Percentage Error (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.

Rumus MAPE secara umum dapat dituliskan :

$$MAPE = \frac{\sum_{t=1}^n \left| \frac{x_t - f_t}{x_t} \right|}{n} \times 100\% \quad (25)$$

Dimana:

n = Jumlah Sampel

x_t = Nilai Aktual Indeks pada periode ke-t

f_t = Nilai Prediksi Indeks pada periode ke-t

Tabel 2.6 menunjukkan perbandingan tingkat akurasi hasil peramalan berdasarkan nilai MAPE (Makridakis, S. W. 1999).

Table 2.2 Hasil Parameter Berdasarkan nilai MAPE

MAPE	HASIL PERAMALAN
<10%	Sangat Baik
10-20%	Baik
20-50%	Layak/ Cukup
>50%	Buruk

2.1.5.6 Evaluasi Turning Point

Evaluasi *turning point* digunakan untuk memprediksi perubahan arah data. Evaluasi *turning point* yang digunakan pada penelitian tugas akhir ini adalah:

2.1.5.6.1 Directional Change Accuration (DCA)

DCA merupakan tes nonparametric yang digunakan untuk mengukur keakuratan arah peramalan. Fokus dari DCA adalah untuk membenarkan prediksi perubahan arah pada variabel yang dipertimbangkan. DCA tidak membutuhkan data kuantitatif, hanya memperhatikan tanda $Z_t^{\rightarrow\rightarrow}$ dan Z_t dimana $Z_t^{\rightarrow\rightarrow}$ adalah nilai prediksi pada periode t

sedangkan Z_t adalah nilai aktual pada periode t . Sehingga hanya perlu melihat apakah arah data berubah dari rendah ke tinggi yang ditandai dengan (+) ataupun sebaliknya (-) pada nilai prediksi. Kemudian menentukan berapa jumlah tanda yang sama dengan data aktual lalu dibagi dengan total jumlah tanda tersebut (Pesaran, M. H., & Timmermann, A.1992).

2.2 Penelitian Terkait

Penelitian terkait ini dilaksanakan dengan survey literatur di internet dan perpustakaan, dan diskusi dengan dosen atau orang-orang yang memiliki banyak ide, pengetahuan, dan pengalaman. Adapun beberapa penelitian terkait yang digunakan dalam penelitian ini, bisa dilihat pada Tabel 2.3.

Tabel 2.3 Penelitian Terkait

NO	Penulis	Judul	Metode	Hasil Penelitian
1	Ni Putu Nanik Hendayanti, Maulida Nurhidayati (2020)	Perbandingan Metode Seasonal Autoregressive Integrated Moving Average (SARIMA) dengan Support Vector Regression (SVR) dalam Memprediksi Jumlah Kunjungan Wisatawan Mancanegara ke Bali	Seasonal ARIMA dan Support Vector Regression (SVR)	Metode SARIMA lebih kecil dari metode SVR, dari Penelitian yang dilakukan dapat disimpulkan bahwa model SARIMA merupakan metode yang lebih baik untuk memprediksi jumlah kunjungan wisatawan mancanegara ke Bali.

Tabel 2.3 Penelitian Terkait (Lanjutan 1)

NO	Penulis	Judul	Metode	Hasil Penelitian
2	<i>Fikri Abdul Najib, Nur Nafi'iyah</i> (2020)	Algoritma SVM untuk Memprediksi Pengunjung Wisata Musium di Jakarta	Algoritma SVM (Support Vector Machine)	Hasil prediksi dari SVM pada data 222 baris pengunjung wisata musium di Jakarta jelek. Dibuktikan dari nilai selisih data nyata dengan hasil prediksi sangat tinggi, dan nilai errornya sangat tinggi 2838303,5.
3	Amin Nur Rais, Rousyati, Indra Jiwana Thira, Desiana Nur Kholifah, Nani Purwati, Yustina Meisella Kristania (2020)	Evaluasi Metode <i>Forecasting</i> Pada Data Kunjungan Wisatawan Mancanegara ke Indonesia	Metode <i>Forecasting</i>	Hasil penelitian menunjukan peramalan menggunakan SMA yang menghasilkan nilai MAPE sebesar 5.96% lebih unggul dibandingkan dengan metode SES yang menghasilkan nilai MAPE sebesar 6.06%. Berdasarkan kriteria MAPE, baik hasil peramalan menggunakan metode SMA maupun metode SES termasuk pada kriteria MAPE dengan nilai baik sekali.

Tabel 2.3 Penelitian Terkait (Lanjutan 2)

NO	Penulis	Judul	Metode	Hasil Penelitian
4	M. Sanzabi Libianto , Tibyani , Yuita Arum Sari (2019)	Prediksi Jumlah Kunjungan Wisatawan Mancanegara Pada Negara Singapura Menggunakan Algoritme <i>Extreme Learning Machine</i>	Extreme Learning Machine	Berdasarkan hasil yang telah didapatkan sebelumnya maka sistem yang telah dibuat digolongkan “sangat bagus”, hal tersebut didasarkan nilai dari MAPE yang dihasilkan dibawah 10%. Sehingga dapat disimpulkan bahwa, algoritme ELM mampu menyelesaikan permasalahan dari prediksi jumlah kunjungan wisatawan mancanegara pada negara Singapura.
5	<i>Gustientiedinaa,</i> <i>M.Hasmil Adiyaa,</i> <i>Yenny Desnelitab</i> (2019)	Penerapan Algoritma K- Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru	Algoritma K-Means	Dari hasil clusterisasi pada data obat – obatan dapat ditarik kesimpulan bahwa kelompok obat yang termasuk pemakaian sedikit rata rata permintaan obat setiap tahunnya kurang dari 18000 buah, dan obat yang termasuk pemakaian

				<p>sedang rata rata permintaan obat setiap tahunnya diantara 18000– 70000 buah, sedangkan obat yang masuk kedalam kelompok obat yang pemakaian tinggi rata – rata permintaan obat setiap tahunnya diatas 70000 buah.</p>
--	--	--	--	--

Tabel 2.3 Penelitian Terkait (Lanjutan 3)

NO	Penulis	Judul	Metode	Hasil Penelitian
6	Ayu Mutmainnah, Edy Widodo (2018)	APPLICATION OF SUPPORT VECTOR MACHINE (SVM) METHODS ON STOCK PRICE FORECASTING OF PT TELEKOMUNIKASI INDONESIA TBK.	SUPPORT VECTOR MACHINE (SVM)	Gambaran umum harga saham PT Telkom Tbk. periode Januari-Juli 2018 yaitu harga saham menurun hingga bulan Mei dan kembali meningkat pada bulan Juni-Juli. Harga saham terendah berada pada bulan Mei dan harga saham tertinggi pada bulan Januari.
7	Agus Darmawan, Nunu Kustian, Wanti Rahayu (2018)	IMPLEMENTASI DATA MINING MENGUNAKAN MODEL SVM UNTUK PREDIKSI KEPUASAN PENGUNJUNG TAMAN TABEBUYA	SUPPORT VECTOR MACHINE (SVM)	Hasil dari penelitian melakukan evaluasi dan validasi dapat disimpulkan bahwa algoritma SVM memiliki akurasi dan performa secara rata-rata yaitu sebesar 86,00% dan nilai AUC sebesar 0.947.
8	MUHAMMAD FARHAN NAUFAL (2017)	PERAMALAN JUMLAH WISATAWAN MANCANEGERA YANG DATANG KE INDONESIA BERDASARKAN PINTU MASUK MENGUNAKAN	SUPPORT VECTOR MACHINE (SVM)	Hasil yang diperoleh dari uji coba penelitian ini menunjukkan bahwa model peramalan secara keseluruhan tergolong baik. Rata-rata akurasi dari 6 model memiliki MAPE sekitar 10% dengan nilai terkecil yakni 4.50% pada pintu masuk Ngurah Rai. Selain itu, hasil dari SVM juga memiliki akurasi perubahan arah data atau Directional

		METODE SUPPORT VECTOR MACHINE (SVM)		Change Accuracy (DCA) yang cukup baik. Hal ini dibuktikan dengan hasil rata-rata DCA secara keseluruhan sebesar 62.64% dengan nilai tertinggi yakni 64.01% pada pintu masuk Husein Sastranegara.
--	--	-------------------------------------	--	--

Tabel 2.3 Penelitian Terkait (Lanjutan 4)

NO	Penulis	Judul	Metode	Hasil Penelitian
9	Ivo Colanus Rally Drajana (2017)	METODE <i>SUPPORT VECTOR MACHINE</i> DAN <i>FORWARD SELECTION</i> PREDIKSI PEMBAYARAN PEMBELIAN BAHAN BAKU KOPRA	SUPPORT VECTOR MACHINE DAN FORWARD SELECTION	Hasil eksperimen penelitian ini menunjukkan dimana <i>series</i> pembayaran algoritma SVM dan <i>Forward Selection</i> memberikan kinerja performa yang terbaik dibandingkan SVM, SVM dan <i>Backward Elimination</i> serta BPNN dan <i>Feature Selection</i> .
10	Windha Mega Pradnya Duhita (2016)	CLUSTERING MENGGUNAKAN METODE K-MEANS UNTUK MENENTUKAN STATUS GIZI BALITA	METODE K-MEANS	Analisa hasil data output, pengelompokan status gizi balita menggunakan tabel Growth Chart, dan yang terakhir melakukan pengujian dengan membandingkan hasil pengelompokan algoritma K-means dan tabel Growth Chart. Dengan membandingkan hasil pengelompokan menggunakan tabel Growth Chart dan algoritma K-Means didapat 17 data yang memiliki kelompok yang sama. Dari angka ini algoritma K-Means hanya memiliki nilai akurasi 34% benar.

2.3 Matriks Penelitian / State Of The Art

Martiks penelitian menjelaskan tentang perbedaan antara penelitian yang dilakukan dengan penelitian terkait.

Terdapat beberapa indikator yang menunjukkan perbedaan dan persamaan antara penelitian terdahulu dengan penelitian yang dilakukan .

Berikut matriks penelitian terlihat pada tabel 2.4;

Tabel 2.4 matriks Penelitian

no	Penulis	Judul	Perbedaan antara penelitian terkait		
			Menggunakan Forecasting	Cangkupan wilayah Kabupaten/Kota	Metode K-Means
1	Mohammad Faiz Aulia (2021)	PREDIKSI JUMLAH WISATAWAN YANG DATANG KE KABUPATEN TASIKMALAYA MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM)	V	V	V

Tabel 2.4 matriks Penelitian (Lanjutan 1)

no	Penulis	Judul	Perbedaan antara penelitian terkait		
			Menggunakan Forecasting	Cangkupan wilayah Kabupaten/Kota	Metode K-Means
2	Ni Putu Nanik Hendayanti, Maulida Nurhidayati (2020)	Perbandingan Metode Seasonal Autoregressive Integrated Moving Average (SARIMA) dengan Support Vector Regression (SVR) dalam Memprediksi Jumlah Kunjungan Wisatawan Mancanegara ke Bali	V	V	
2	<i>Fikri Abdul Najib, Nur Nafi'iyah</i> (2020)	Algoritma SVM untuk Memprediksi Pengunjung Wisata Musium di Jakarta	V	V	

Tabel 2.4 matriks Penelitian (Lanjutan 2)

no	Penulis	Judul	Perbedaan antara penelitian terkait		
			Menggunakan Forecasting	Cangkupan wilayah Kabupaten/Kota	Metode K- Means
3	Amin Nur Rais, Rousyati, Indra Jiwana Thira, Desiana Nur Kholifah, Nani Purwati, Yustina Meisella Kristania (2020)	Evaluasi Metode <i>Forecasting</i> Pada Data Kunjungan Wisatawan Mancanegara ke Indonesia	V		
4	M. Sanzabi Libianto , Tibyani , Yuita Arum Sari (2019)	Prediksi Jumlah Kunjungan Wisatawan Mancanegara Pada Negara Singapura Menggunakan Algoritme <i>Extreme Learning Machine</i>	V		

Tabel 2.4 matriks Penelitian (Lanjutan 3)

no	Penulis	Judul	Perbedaan antara penelitian terkait		
			Menggunakan Forecasting	Cangkupan wilayah Kabupaten/Kota	Metode K-Means
5	<i>Gustientiedinaa, M.Hasmil Adiyaa, Yenny Desnelitab</i> (2019)	Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru		V	V
6	Ayu Mutmainnah, Edy Widodo (2018)	APPLICATION OF SUPPORT VECTOR MACHINE (SVM) METHODS ON STOCK PRICE FORECASTING OF PT TELEKOMUNIKASI INDONESIA TBK.	V		
7	Agus Darmawan, Nunu Kustian, Wanti Rahayu (2018)	IMPLEMENTASI DATA MINING MENGGUNAKAN MODEL SVM UNTUK PREDIKSI KEPUASAN PENGUNJUNG TAMAN TABEBUYA	V	V	

Tabel 2.4 matriks Penelitian (Lanjutan 4)

no	Penulis	Judul	Perbedaan antara penelitian terkait		
			Menggunakan Forecasting	Cangkupan wilayah Kabupaten/Kota	Metode K- Means
8	MUHAMMAD FARHAN NAUFAL (2017)	PERAMALAN JUMLAH WISATAWAN MANCANEGERA YANG DATANG KE INDONESIA BERDASARKAN PINTU MASUK MENGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM)	V		V
9	Ivo Colanus Rally Drajana (2017)	METODE <i>SUPPORT VECTOR MACHINE</i> DAN <i>FORWARD SELECTION</i> PREDIKSI PEMBAYARAN PEMBELIAN BAHAN BAKU KOPRA	V		

Tabel 2.4 matriks Penelitian (Lanjutan 5)

no	Penulis	Judul	Perbedaan antara penelitian terkait		
			Menggunakan Forecasting	Cangkupan wilayah Kabupaten/Kota	Metode K- Means
10	Windha Mega Pradnya Duhita (2016)	CLUSTERING MENGUNAKAN METODE K-MEANS UNTUK MENENTUKAN STATUS GIZI BALITA			V

