

BAB 2

LANDASAN TEORI

2.1 *Tinjauan Pustaka*

Penelitian sebelumnya dengan metode yang sama tetapi dengan topik yang berbeda dibahas dalam tinjauan literatur ini. Topik yang dibahas seperti masukan input, metode serta hasil akan dibahas pada tinjauan literatur ini.

Tabel 2. 1 Tinjauan Pustaka

Penulis; Tahun	Judul	Deskripsi Umum Penelitian
Maryana, T., Kusrini, K., & Al Fatta, H; 2019	Analisis perbandingan prediksi obat dengan menggunakan metode abc analisis dan <i>SVR</i> pada aplikasi “morbis”	Penelitian ini memfokuskan pada aplikasi analisis data untuk pengelolaan persediaan obat di rumah sakit. Ada 2 metode yang digunakan, Metode ABC Analysis digunakan untuk mengklasifikasikan obat ke dalam tiga kategori berdasarkan tingkat kepentingannya dan <i>SVR</i> yang akan digunakan dalam memprediksi penjualan obat dengan mengkombinasikan preprocessing data, yaitu linier scaling dan normalisasi z. Data yang diambil adalah data penjualan obat dari bulan

Penulis; Tahun	Judul	Deskripsi Umum Penelitian
		<p>Januari hingga November 2019 dengan proses klasifikasi obat menggunakan metode ABC menghasilkan 1.202 item obat. Metode <i>SVR</i> dengan preprocessing linier scaling memberikan hasil akurasi terbaik dengan <i>MAPE</i> (Mean Absolute Percentage Error) sebesar 17,51%., Sedangkan Sebagai perbandingan, <i>MAPE</i> untuk metode menggunakan normalisasi z adalah 18,09%, dan untuk klasifikasi ABC Analysis adalah 33,15%. Menunjukkan bahwa <i>SVR</i> dengan penambahan preprocessing linier scaling memberikan performa prediksi yang lebih baik dibandingkan dengan metode lainnya yang diuji dalam penelitian tersebut (Maryana & Al Fatta, 2019).</p>

Penulis; Tahun	Judul	Deskripsi Umum Penelitian
Firmanda,A.B.E.F.,AS,A.H ., Tholib, A., & Guterres, J. X; 2024	Implementasi <i>Grid Search</i> dalam Meningkatkan Kinerja Model Support Vector Regression(<i>SVR</i>) untuk Prediksi Penjualan	Produk pada Meuble Rohman Jaya Penelitian ini membahas penggunaan GridSearch apakah optimal untuk Prediksi Penjualan Produk pada Meuble Rohman Jaya. Dengan metode yang dipakai <i>Support Vector Regression (SVR)</i> , dan dioptimalkan menggunakan teknik GridSearch untuk menemukan kombinasi parameter yang optimal. Data yang digunakan adalah data penjualan bulanan Meubel Rohman Jaya dari Januari 2021 hingga Desember 2022. Hasil menunjukkan prediksi menggunakan optimasi parameter <i>Grid Search</i> lebih optimal dengan akurasi <i>MAPE</i> 0,45%, sedangkan jika tidak menggunakan optimasi parameter <i>Grid Search</i> mendapatkan <i>MAPE</i> 40,39%(Baidowi et al., 2024).

Penulis; Tahun	Judul	Deskripsi Umum Penelitian
Ishlah, A. W.,Sudarno, S., & Kartikasari, P; 2023	Implementasi GridSearchcv Pada <i>SUPPORT VECTOR REGRESSION (SVR)</i> Untuk Peramalan Harga Saham.	<p>Penelitian ini membahas penerapan metode <i>SVR</i> dalam memprediksi Harga saham PT Anabatic Technologies Tbk(ATIC) dan mengoptimalkan performa model dengan menggunakan GridSearchCV untuk tuning hyperparameter. Metode yang digunakan Support Vector regression dengan menggunakan GridSearchCV untuk tuning hyperparameter. Data yang digunakan adalah data historis harian harga penutupan saham PT Anabatic Technologies Tbk(ATIC) mulai tanggal 1 Januari 2017 sampai 31 Desember 2021, yang terdiri dari 1262 data. Analisis dilakukan dengan mengukur kinerja model menggunakan beberapa ukuran error, yaitu Mean Square Error(MSE), Symmetric</p>

Penulis; Tahun	Judul	Deskripsi Umum Penelitian
		Mean Absolute Percentage Error(<i>SMAPE</i>), dan koefisien determinasi R^2 . Model <i>SVR</i> Tanpa GridSearchCV, $MSE:0,02993634$, $SMAPE:0,2319903$ (atau 23,19903%), $R^2:0,1545886$ (atau 15,45886%). Model <i>SVR</i> Dengan GridSearchCV, $MSE:0,001237173$, $SMAPE:0,1167301$, $R^2:0,9206643$. Hasil akurasi yang didapatkan menunjukkan performa model <i>SVR</i> dengan parameter yang dioptimalkan menggunakan GridSearchCV lebih baik dibandingkan dengan model <i>SVR</i> tanpa optimasi hyperparameter (Isilah et al., 2023).
Maulana, N. D., Setiawan, B. D., & Dewi, C; 2019	Implementasi metode <i>Support Vector Regression</i> (<i>SVR</i>) dalam	Penelitian ini membahas penggunaan metode <i>Support Vector Regression</i> yang digunakan berfokus

Penulis; Tahun	Judul	Deskripsi Umum Penelitian
	peramalan penjualan roti. (Studi Kasus : Harum Bakery)	dalam prediksi penjualan roti. Menggunakan metode <i>SVR</i> dengan data time series yang terdiri dari 4 fitur. Data yang dipakai adalah histori data permintaan darah mulai bulan September 2013 sampai Agustus 2014. Dengan kinerja <i>SVR</i> mendapat hasil evaluasi <i>MAPE</i> (Kesalahan Rata-rata Persentase) yang terbaik sebesar 3,899% dengan parameter sebagai berikut: 12 data diuji menggunakan lambda 10, sigma 0,5, coefisien rasio pembelajaran 0,01, epsilon 0,01, dan jumlah iterasi 5000. Dengan demikian, penelitian ini menunjukkan hasil yang baik dalam hal akurasi penggunaan metode <i>SVR</i> untuk peramalan(Dini Maulana et al., 2019).
Ruliana, R., Rais, Z., Marni, M., & Ahmar, A.S;	Implementation of the SUPPORT	Penelitian ini membahas tentang prediksi inflasi di

Penulis; Tahun	Judul	Deskripsi Umum Penelitian
2024.	VECTOR REGRESSION(<i>SVR</i>) Method in Inflation Prediction in Makassar City.	Kota Makassar menggunakan metode SUPPORT VECTOR REGRESSION(<i>SVR</i>). Dengan menggunakan metode <i>Support Vector Regression (SVR)</i> yang dioptimalkan menggunakan Teknik <i>Grid Search Optimization</i> . Data yang dipakai data dari Indeks Harga Konsumen di Makassar dari mulai bulan Januari 2014 sampai Desember 2022. Hasil akurasi yang diperoleh menunjukkan nilai Root Mean Square Error(<i>RMSE</i>) sebesar 0,029, yang menunjukkan bahwa model memiliki kemampuan yang baik dalam mengikuti pola data inflasi di Makassar untuk periode prediksi dari Januari hingga Desember 2023 menentukan parameter terbaik(Ruliana et al., 2024).
Drajat Indra Purnama; 2020	peramalan jumlah penumpang dating	Penelitian ini membahas tentang peramalan jumlah

Penulis; Tahun	Judul	Deskripsi Umum Penelitian
	melalui transportasi udara di Sulawesi tengah menggunakan SUPPORT VECTOR REGRESSION(SVR)	penumpang melalui transportasi udara di Sulawesi tengah. dengan menggunakan SVR dan kernel RBF serta bantuan <i>Grid Search</i> Untuk mendapatkan parameter optimal dari model. Data yang digunakan adalah data sekunder dari jumlah penumpang yang datang melalui transportasi udara di Sulawesi Tengah pada periode Januari 2011 sampai November 2019. Nilai Mean Absolute Percentage Error (<i>MAPE</i>) digunakan untuk mengetahui akurasi model. Pada data latih, nilai <i>MAPE</i> yang diperoleh adalah sebesar 5,01 persen, yang menunjukkan kemampuan peramalan sangat baik. Sedangkan pada data uji, nilai <i>MAPE</i> yang diperoleh adalah sebesar 16,65 persen, yang menunjukkan kemampuan

Penulis; Tahun	Judul	Deskripsi Umum Penelitian
		peramalan baik (Indra Purnama & Setianingsih, 2020).

2.2 Peramalan

Peramalan adalah proses menghitung perkiraan jumlah sesuatu di masa depan berdasarkan data sebelumnya, terutama statistika. Peramalan adalah perkiraan permintaan yang akan datang yang didasarkan pada beberapa variabel, biasanya data deret dan waktu, yang digunakan untuk mengestimasi kebutuhan data yang akan datang (Ita Bela & Bhakti, 2022).

2.3 Machine Learning

Machine Learning yaitu salah satu kecerdasan buatan dengan fokus pada algoritma yang memungkinkan komputer untuk belajar dari data dan menggunakannya dalam membuat prediksi atau pengambilan keputusan berdasarkan pola dan informasi yang ada pada data (Siswoyo, 2020). Ciri khas Machine learning adalah adanya proses pelatihan, pembelajaran, atau pelatihan. Dalam peramalan, ada tiga tahapan yang harus dilewati saat membangun metode peramalan: 1) Menganalisis data sebelumnya. Langkah ini bertujuan untuk mengumpulkan pola dari data yang relevan. 2) Memilih teknik atau metode yang akan digunakan. Banyak sekali metode algoritma yang bisa digunakan, pemilihan algoritma dapat mempengaruhi hasil peramalan. Hasil ramalan dihitung dengan menghitung error atau kesalahan terkecil. Akibatnya, tidak ada teknik peramalan yang benar-benar cocok untuk semua jenis data. 3) Proses transformasi data sebelumnya dengan metode yang dipilih. Perubahan dilakukan sesuai kebutuhan (Hasibuan & Karim, 2022).

2.4 Normalisasi Data

Normalisasi yaitu proses mengubah nilai dari suatu data dalam skala lebih kecil. Normalisasi harus dilakukan untuk memastikan bahwa data dalam rentang yang sama. Metode pengukuran min-max digunakan untuk normalisasi dataset.

Nilai asli dapat diubah menjadi nilai interval menggunakan metode ini, seperti yang ditunjukkan dalam persamaan (Mukhtar et al., 2022). Rumus normalisasi data ada pada persamaan 2.1 dibawah :

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (2.1)$$

keterangan :

X : nilai asli data

Xmin : nilai minimum dari dataset.

Xmax : nilai maximum dari dataset.

Xnorm : nilai yang telah dinormalisasi.

2.5 *Support Vector Regression (SVR)*

Sistem *SVR* menggunakan fitur ruang hipotesis berdimensi besar untuk melakukan fungsi yang lebih spesifik. Performa *SVR* sangalaht baik dalam mengatasi overfitting (Indra Purnama & Setianingsih, 2020).

Overfitting sendiri merupakan suatu performa baik yang dihasilkan oleh model data yang dilatih dengan prediksi kurang lebih sempurna, tetapi ketika diuji hasilnya tidak sesuai. Metode ini sendiri bertujuan untuk menemukan hyperplane terbaik antara dua kelas data. Hyperplane terbaik akan ditemukan dengan menentukan nilai margin hyperplane dan menemukan titik maksimal margin.

Margin adalah jarak data-data yang paling dekat dari hyperline di dalam dua kelas tersebut, dan data-data yang paling dekat itu yang nantinya akan dijadikan support vector (Saragih et al., 2023).

Dasar Algoritma support vector untuk regresi adalah menghitung fungsi linier, dimana α adalah pengali langrange non-negative. Rumus umum *Support Vector Regression (SVR)* ada pada persamaan 2.2 di bawah :

$$F(X) = \sum(\alpha_i - \alpha_i^*)K(X_i, X^*) + b \quad (2.2)$$

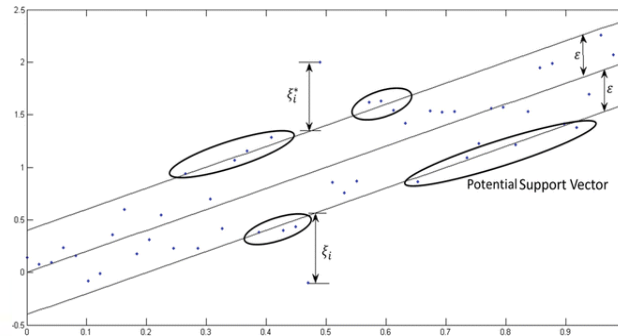
keterangan:

α_i, α_i^* = Lagrange multipliers

K = Fungsi Kernel

b = Bias

Konsep SVM konsepnya membagi data menjadi dua kelas, sementara *SVR* membatasi semua data ke dalam satu zona dengan meminimalkan nilai epsilon(ϵ).



Gambar 2. 1 Ilustrasi *SVR*

Pada Gambar 2.1 *Support Vector Regression (SVR)* menemukan hyperplane terbaik dan membentuk margin dengan batas toleransi error ϵ . Data di dalam margin dianggap sudah terprediksi dengan baik dan tidak berkontribusi dalam pembentukan model, sedangkan data di luar margin memiliki nilai error (ξ_i) dan berfungsi sebagai support vector. Support vector ini digunakan untuk mendefinisikan.

Garis tengah menunjukkan hyperplane regresi, dan dua garis lainnya menunjukkan batas toleransi kesalahan ϵ . Nilai ξ_i menunjukkan penyimpangan data dari batas toleransi, yaitu jarak data aktual yang di luar margin terhadap margin. *SVR* dapat menangkap pola non-linear dengan memetakan data ke dimensi yang lebih besar melalui penggunaan kernel.

2.6 *Kernel Radial Basis Function (RBF)*

Banyak metode pengolahan data dengan pendekatan linier, sehingga banyak algoritma yang terbatas dalam kasus tidak linier, di dalam kasus dunia nyata umumnya banyak menggunakan kasus tidak linier, maka dari itu untuk mengatasinya dapat menggunakan fungsi kernel (Caraka et al., 2017). Kernel RBF sangat cocok digunakan dalam hal tersebut karena dapat menangani data-data non linear dengan sangat baik. *Kernel Radial Basis Function (RBF)* dapat dilihat pada persamaan 2.3 berikut.

$$K(x, x_i) = \exp \left(-\frac{\|x - x_i\|^2}{2\sigma^2} \right) \quad (2.3)$$

Keterangan :

x = nilai fitur dari normalisasi data yang digunakan untuk peramalan

x_i = nilai fitur dari normalisasi data sebelumnya yang digunakan untuk peramalan

σ = nilai dari Radial Basis

2.7 *Grid Search Optimization*

Dalam menggunakan metode *SVR* sering sekali permasalahan saat akan menentukan parameter model apa yang akan digunakan. Ada salah satu cara yang digunakan untuk mencari parameter model yang optimal adalah dengan menggunakan *Grid Search Optimization*. Dalam aplikasinya optimasi *Grid Search* biasanya sering diukur menggunakan cross-validation (*Grid SearchCV*). *Grid SearchCV* adalah implementasi *Grid Search* yang menggunakan cross-validation dari library scikit-learn. Pemilihan grid adalah pemilihan kombinasi model dan hyperparameter dengan menguji kombinasi satu per satu dan memvalidasi setiap kombinasi. Cross-validation membagi data pelatihan menjadi beberapa bagian (fold) dengan tujuan adalah untuk menemukan kombinasi kinerja model terbaik yang dapat digunakan sebagai model prediksi. Untuk menghasilkan indeks evaluasi untuk model klasifikasi (Toha et al., 2022).

Metode *Grid Search* prosesnya hampir sama dengan metode trial and error, dengan memeriksa setiap kombinasi hyperparameter dan membandingkan nilai kesalahan terkecil yang dihasilkan oleh hyperparameter. Hyperparameter yang biasanya disesuaikan, seperti C , γ , dan ϵ , bertanggung jawab atas pengaturan perilaku model dalam *Support Vector Regression (SVR)*:

- a) Parameter regularisasi C : mengatur jumlah kesalahan yang diizinkan model. Nilai C yang kecil membuat model lebih sederhana (cenderung underfitting), sedangkan nilai C yang besar membuat model lebih kompleks dan berkonsentrasi pada data pelatihan (cenderung overfitting).
- b) γ (coefficient kernel untuk RBF): mengatur bagaimana kurva

fungsi keputusan dibentuk. Kurva yang dihasilkan oleh gamma kecil lebih halus, yang dikenal sebagai underfitting, sementara kurva yang dihasilkan oleh gamma besar lebih tajam dan sensitif terhadap data pelatihan.






- c) Epsilon (jalur epsilon dalam *SVR*): berfungsi untuk menentukan margin toleransi kesalahan. Nilai epsilon yang besar membuat model lebih toleran terhadap kesalahan kecil, sedangkan nilai epsilon yang kecil membuat model lebih sensitif terhadap perbedaan kecil antara prediksi dan data asli.

2.8 UML

Unified Modelling Language Adalah pemodelan dengan visual bahasa yang memungkinkan pemodelan sistem dan komunikasi dengan menggunakan diagram dan teks-teks pendukung. Use case diagram, class diagram, sequence diagram dan activity diagram adalah beberapa contoh pemodelan UML (Syarif & Nugraha, 2020).

2.8.1 Use Case Diagram







Use Case Diagram adalah model untuk kelakuan sistem informasi yang akan dibangun yang biasanya terdiri dari actor dan relationshipnya (Syarif & Nugraha, 2020). Simbol dan fungsinya dapat dilihat pada Gambar 2.2 di bawah.

NO	GAMBAR	NAMA	KETERANGAN
1		<i>Actor</i>	Menspesifikasikan himpunan peran yang pengguna mainkan ketika berinteraksi dengan <i>use case</i> .
2		<i>Include</i>	Menspesifikasikan bahwa <i>use case</i> sumber secara <i>eksplisit</i> .
3		<i>Association</i>	Apa yang menghubungkan antara objek satu dengan objek lainnya.
4		<i>System</i>	Menspesifikasikan paket yang menampilkan sistem secara terbatas.
5		<i>Use Case</i>	Deskripsi dari urutan aksi-aksi yang ditampilkan sistem yang menghasilkan suatu hasil yang terukur bagi suatu aktor

Gambar 2. 2 Simbol Use Case Diagram

2.8.2 Activity Diagram

Activity Diagram merupakan aliran kerja dari sebuah sistem, proses bisnis atau menu yang ada pada perangkat lunak. Activity Diagram sangat berguna untuk menunjukkan logika bisnis, skenario penggunaan (use case) dan alur proses yang kompleks (Syarif & Nugraha, 2020). Simbol dan fungsinya dapat dilihat pada Gambar 2.3 di bawah.

NO	GAMBAR	NAMA	KETERANGAN
1		<i>Activity</i>	Memperlihatkan bagaimana masing-masing kelas antarmuka saling berinteraksi satu sama lain.
2		<i>Action</i>	State dari sistem yang mencerminkan eksekusi dari suatu aksi.
3		<i>Initial Node</i>	Bahamian object dibentuk atau diawali.
4		<i>Activity Final Node</i>	Bahamian object dibentuk dan dihancurkan.
5		<i>Fork Node</i>	Satu aliran yang pada tahap tertentu berubah menjadi beberapa aliran.
6		<i>Decision Node</i>	Menunjukkan aktivitas yang harus dipilih.

Gambar 2. 3 Simbol Activity Diagram

2.8.3 Class Diagram

Class Diagram menggambarkan struktur dari sistem, serta mendefinisikan kelas-kelas yang digunakan untuk membangun sistem antar kelasnya. (Syarif & Nugraha, 2020).

1. Asosiasi

Asosiasi adalah salah satu kelas dapat memiliki atribut yang mengacu pada kelas lain, atau kelas tersebut dapat berinteraksi dengan kelas lain dalam situasi tertentu.

2. Agregasi

Agregasi adalah hubungan di antara dua kelas di mana salah satu kelas adalah bagian dari kelas lain, tetapi keduanya dapat berdiri sendiri.








3. Pewarisan

Pewarisan adalah kemampuan untuk mewariskan metode dan atribut dari kelas induk (superclass) ke kelas turunan (subclass). Dalam pewarisan,

subclass dapat mewarisi sifat-sifat superclass dan dapat menambahkan atau mengganti perilaku yang ada jika diperlukan.

4. Hubungan Dinamis

Tahapan pesan diberikan dari kelas satu ke kelas lainnya dapat dilihat pada Gambar 2.4 di bawah.

No	Gambar	Nama	Keterangan
1		<i>Generalization</i>	Hubungan dimana objek anak (<i>descendent</i>) berbagi perilaku dan struktur data dari objek yang ada diatasnya objek induk (<i>ancestor</i>).
2		<i>N-ary Association</i>	Upaya untuk menghindari asosiasi dengan lebih dari dua objek
3		<i>Class</i>	Himpunan dari objek-objek yang berbagi atribut serta operasi yang sama
4		<i>Collaboration</i>	Deskripsi dari urutan aksi-aksi yang ditampilkan sistem yang menghasilkan suatu hasil yang terukur bagi suatu aktor.
5		<i>Realization</i>	Operasi yang benar-benar dilakukan oleh suatu objek.
6		<i>Dependency</i>	Hubungan dimana perubahan yang terjadi pada suatu elemen mandiri (<i>independent</i>) akan mempengaruhi elemen yang bergantung padanya elemen yang tidak mandiri.
7		<i>Association</i>	Apa yang menghubungkan antara objek satu dengan objek lainnya.

Gambar 2. 4 Simbol Class Diagram

2.9 Evaluasi

Evaluasi adalah suatu proses atau kegiatan untuk memilih, mengumpulkan, menganalisis, dan menyampaikan informasi. Informasi ini dapat digunakan sebagai dasar untuk pengambilan keputusan dan pembuatan program selanjutnya (Purnama, 2020). Dalam konteks analisis data ada evaluasi model untuk menghitung matriks seperti, *MAE*, *MSE*, *RMSE*, dan *MAPE* untuk mengukur akurasi prediksi.

2.9.1 Mean Absolute Error (MAE)

MAE menghitung rata-rata kesalahan absolut dalam prediksi dari selisih absolut antara nilai prediksi dan nilai asli. *MAE* sering digunakan dalam evaluasi. Kualitas model meningkat dengan nilai *MAE* yang lebih rendah.

Rumus *MAE* dapat dilihat pada Persamaan 2.4 berikut.

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (2.4)$$

keterangan :

n = jumlah data

y_t = nilai data asli

\hat{y}_t = nilai data prediksi

2.9.2 Mean Squared Error (MSE)

MSE menemukan rata-rata dari selisih kuadrat antara nilai prediksi dan nilai asli. *MSE* menemukan jumlah kesalahan kuadrat rata-rata pada prediksi. Kualitas model meningkat dengan nilai *MSE* yang lebih rendah.

Rumus *MSE* dapat dilihat pada persamaan 2.5 berikut :

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2 \quad (2.5)$$

keterangan:

n = jumlah sampel data

y_t = nilai data asli

\hat{y}_t = nilai data prediksi

2.9.3 Root Mean Squared Error (RMSE)

RMSE sendiri merupakan akar kuadrat *MSE*, selisih kuadrat antara nilai prediksi dan nilai asli dihitung rata-ratanya. Setelah itu, diambil akar kuadratnya. Dengan nilai *RMSE* yang kecil, kualitas modelnya juga sangat baik.

Rumus *RMSE* dapat dilihat pada persamaan 2.6 berikut :

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (2.6)$$

keterangan:

n = jumlah sampel dari data

y_t = nilai dari data asli

\hat{y}_t = nilai dari data prediksi

2.9.4 Mean Absolute Percentage Error (MAPE)

MAPE metode yang paling populer dalam evaluasi, *MAPE* menghitung rata-rata dari selisih persentase antara nilai prediksi dan nilai asli. Kualitas model berkorelasi positif dengan nilai *MAPE* yang lebih rendah.

Rumus *MAPE* dapat dilihat pada persamaan 2.7 berikut :

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100\% \quad (2.7)$$

keterangan:

n = jumlah dari sampel data yang

digunakan y_t = nilai dari data asli

yang digunakan

\hat{y}_t = nilai dari data prediksi yang digunakan

Tabel 2. 2 Hasil Evaluasi *MAPE*

RANGE <i>MAPE</i>	EVALUASI
<10%	Model peramalan sangat baik
10-20%	Model peramalan baik
20-50%	Model peramalan layak
>50%	Model peramalan buruk