

## Peramalan Tren Penjualan pada Sistem Informasi Inventory Barang Diecast Menggunakan *Support Vector Regression*

Immanuel Saragih<sup>1,\*</sup>, Ardy Januanto<sup>2</sup>, Bagus Hardiansyah<sup>3</sup>

<sup>1,3</sup>Teknik Informatika, Universitas 17 Agustus 1945 Surabaya, Indonesia

<sup>2</sup>Sistem dan Teknologi Informasi, Universitas 17 Agustus 1945 Surabaya, Indonesia

### ABSTRACT

*Inventory management of diecast products is crucial for Sada Hobby store. The challenge in managing inventory lies in the importation of goods from overseas, which often lacks a definite delivery time, resulting in stock delays. An effective forecasting model is needed to address this challenge, such as Support Vector Regression (SVR), which can handle non-linear data using  $\epsilon$ -sensitive approach. The SVR model predicts sales for each month in the next year, with accuracy measured using Mean Squared Error (MSE). The research results show that the SVR model estimates sales to be 8,423 units with an MSE accuracy of 0.0018. Sales predictions are also made for the Tomica brand (2,423 units, MSE 0.0019), Hotwheels brand (1,299 units, MSE 0.091), and Majorette brand (360 units, MSE 1.244). Although the MSE for these brands is higher than the overall sales prediction, the results are still good. The brand-based prediction process requires comparisons with all other brands, while the overall sales prediction sums up the sales results from all brands.*

**Keywords:** *inventory, prediction, stock, supply, Support Vector Regression*

### ABSTRAK

Pengelolaan persediaan produk *diecast* sangat penting bagi toko Sada Hobby yang berlokasi di Surabaya, Jawa Timur. Tantangan dalam persediaan barang terletak pada impor dari luar negeri yang sering kali tidak memiliki waktu pengiriman yang pasti, mengakibatkan keterlambatan stok barang. Diperlukan model peramalan efektif untuk mengatasi tantangan tersebut, seperti *Support Vector Regression* (SVR), yang mampu memproses data nonlinier menggunakan  $\epsilon$ -sensitive. Model SVR memprediksi penjualan untuk setiap bulan dalam 1 tahun ke depan dengan akurasi diukur menggunakan *Mean Squared Error* (MSE). Hasil penelitian menunjukkan bahwa model SVR memperkirakan penjualan sebesar 8.423 unit dengan akurasi nilai MSE sebesar 0,0018. Prediksi penjualan juga dilakukan untuk merk Tomica (2.423 unit, MSE 0,0019), merek Hotwheels (1.299 unit, MSE 0,091), dan merk Majorette (360 unit, MSE 1,244). Meskipun MSE untuk merek-merek tersebut lebih tinggi daripada prediksi penjualan secara keseluruhan, hasilnya tetap baik. Proses prediksi berdasarkan merk memerlukan perbandingan dengan semua merk lainnya, sedangkan prediksi penjualan secara keseluruhan hanya menjumlahkan hasil penjualan dari semua merk.

**Kata Kunci:** *inventory, persediaan, prediksi, stok barang, Support Vector Regression*

---

Corresponding Author E-mail: imanuelsaragih07@gmail.com\*

Received June 2023; revised September 2023; accepted December 2023; published January 2024

## 1. Pendahuluan

Tren perkembangan teknologi saat ini sangat mempengaruhi seluruh aktivitas masyarakat di mana saja. Kemudahan-kemudahan yang ditawarkan menarik banyak orang untuk memanfaatkannya sebagai penunjang kegiatan mereka, baik itu untuk aktivitas pribadi maupun pekerjaan atau bisnis yang dijalankan. Salah satu teknologi yang sangat berguna dalam membangun usaha adalah sebuah sistem informasi untuk *inventory* barang berbasis web. *Inventory* barang merupakan kegiatan yang umum bagi pemilik usaha dalam menyimpan dan mengelola barang-barang produk jualan mereka.

Salah satu usaha yang membutuhkan sistem informasi *inventory* barang ini adalah toko Sada Hobby yang berlokasi di daerah Sambikerep, Surabaya, Jawa Timur. Usaha tersebut bergerak pada bidang jualan mainan yang memasarkan produknya secara *online* dengan memanfaatkan media sosial Facebook dan *marketplace* Tokopedia. Mainan yang dijual umumnya berupa mobil-mobilan dengan merek seperti Hot Wheels, Tamiya, Tomica, dan lain-lain. Mainan-mainan ini sering dikenal sebagai *diecast toys*.

Sistem informasi *inventory* ini memerlukan sebuah fitur *forecasting* untuk melakukan peramalan penjualan produk untuk membantu pengambilan keputusan terkait penyediaan produk. Beberapa metode peramalan penjualan telah dikembangkan untuk memecahkan masalah pengelolaan persediaan. Beberapa metode yang lebih umum digunakan termasuk regresi linier, *moving average*, dan *exponential smoothing*. Namun, dalam konteks penjualan produk *diecast* yang kompleks, data penjualan seringkali memiliki pola nonlinier yang sulit diprediksi menggunakan metode tradisional.

Metode yang digunakan dalam penelitian ini adalah *Support Vector Regression* (SVR). SVR adalah metode yang efektif untuk memproses data nonlinear dan menawarkan fleksibilitas yang tinggi untuk data yang kompleks. Keuntungan utama SVR

adalah kemampuannya untuk menemukan pola tersembunyi dalam data yang tidak dapat dijelaskan oleh metode linier tradisional.

Dibandingkan dengan metode lain seperti regresi linier, *moving average*, dan *exponential smoothing*, SVR dipilih karena dapat menangani data nonlinier dengan lebih baik. Dalam hal penjualan produk *die casting*, data penjualan cenderung memiliki pola nonlinier karena faktor-faktor seperti tren popularitas produk dan preferensi kolektor. Metode SVR dapat mengatasi tantangan ini dengan lebih baik karena dapat menangkap pola kompleks dalam data dan menghasilkan prediksi yang lebih akurat.

Oleh karena itu, penelitian ini memilih menggunakan metode SVR karena dapat menangani data penjualan produk *diecast* secara nonlinier. Metode ini diharapkan dapat memberikan perkiraan penjualan yang lebih akurat dan membantu mengelola inventaris produk dengan lebih efisien.

SVR digunakan untuk memproses data yang lebih rumit yang tidak dapat dilakukan oleh *Support Vector Machine* (SVM). Metode ini akan digunakan untuk membuat menu peramalan terhadap tren penjualan di pasar. Hasil dari peramalan tersebut akan diukur tingkat akurasinya menggunakan *Mean Squared Error* (MSE).

Peramalan dengan SVR dapat digunakan pada jenis data yang bersifat nonlinear yang menggunakan pendekatan regresi, di mana keunggulannya adalah melakukan optimasi peramalan pada jenis data nonlinear dengan *trick kernel*. Data-data dari penjualan toko Sada Hobby memiliki variabel-variabel yang beragam khususnya dari sisi merek, sehingga jenis data yang nonlinear ini perlu diproses menggunakan SVR nonlinear yang menggunakan *trick kernel* tersebut [1].

Tujuan dari penelitian ini adalah menghasilkan sebuah sistem informasi yang mampu mendapatkan hasil peramalan terhadap tren penjualan dengan menggunakan metode SVR dan mengukur hasil kinerjanya menggunakan MSE.

Manfaat penelitian ini adalah memudahkan pemilik untuk memprediksi penjualan dalam kurun waktu setahun ke depan.

## 2. Tinjauan Pustaka

### 2.1. Kajian Pustaka

Terdapat beberapa penelitian terdahulu yang menggunakan metode SVR untuk melakukan prediksi seperti prediksi penurunan jumlah kasus penderita COVID-19 yang menggunakan keempat kernel, yaitu kernel linier, *polynomial*, Sigmoid dan *Radial Basis Function* (RBF). Hasil prediksi tersebut dievaluasi dan divalidasi dengan mencari *Root Mean Squared Error* (RMSE), di mana jika nilainya di bawah 1 berarti hasil prediksi tersebut baik. Perbandingan dari hasil nilai RMSE dari prediksi setiap negara adalah Jerman dengan 0,99, Itali 0,101, Indonesia 0,102, Brazil 0,105, dan Amerika Serikat 0,105 [2]. Kemudian terdapat pula penelitian untuk memprediksi penjualan pertalite untuk bulan berikutnya. Hasil prediksi yang diperoleh adalah pertalite akan terjual sebanyak 11.436 liter pada tanggal 8 Januari 2019 [3]. Lalu terdapat penelitian lainnya yang memprediksi penjualan majalah. Berbeda dengan penelitian sebelumnya, penelitian ini menggunakan metode SVR dengan 1 kernel saja, yaitu kernel RBF karena kernel RBF memberikan hasil prediksi yang lebih baik dibandingkan kernel yang lain [4].

Penulis juga menggunakan 1 kernel saja pada penerapan metode SVR untuk melakukan prediksi penjualan *diecast* dalam 1 tahun ke depan. Namun dalam penelitian ini memiliki perbedaan dengan penelitian sebelumnya, di mana penelitian ini menggunakan sebuah algoritma pencari parameter terbaik untuk metode SVR yang akan dijalankan. Dengan algoritma grid ini juga akan diperoleh kernel yang terbaik untuk diterapkan pada SVR.

Algoritma grid search bekerja dengan menggabungkan setiap parameter satu per satu dan membandingkan nilai *error* dari hasil penggabungan tersebut. Kombinasi

parameter dengan nilai *error* terkecil jelas menjadi parameter yang paling optimal untuk dijalankan menggunakan SVR [5].

Sebelum menjalankan metode SVR perlu menerapkan normalisasi data terhadap data penjualan. Hal ini karena penjualan pada data penelitian memiliki rentang nilai yang besar. Permana dan Salisah [6] membuktikan bahwa *dataset* dengan rentang nilai yang jauh tidak perlu dilakukan normalisasi data. Sedangkan *dataset* yang lainnya memerlukan normalisasi karena rentang nilainya besar. Hal ini dapat dilihat dari penelitian Permana dan Salisah [6] yang mencoba untuk mengetahui pengaruh dari normalisasi data dalam melakukan klasifikasi menggunakan algoritma *backpropagation* yang hasilnya hanya nilai akurasi *dataset* Iris yang mencapai akurasi yang tinggi, dengan 98%, selebihnya adalah 30% dan 60%. Sedangkan pengujian dengan menggunakan ketiga jenis normalisasi, hasil yang didapatkan justru berbanding terbalik di mana seluruh *dataset* memperoleh nilai akurasi yang tinggi, di atas 90%.

Permana dan Salisah [6] menggunakan normalisasi *z-score* untuk normalisasi data, sama seperti penelitian oleh Whendasmoro dan Joseph [7] yang menganalisis penerapan normalisasi data dengan *Z-Score* pada algoritma *K-Nearest Neighbor* (KNN) bahwa hasil dengan akurasi lebih tinggi pada data yang telah dinormalisasi terlebih dahulu.

### 2.2. Forecasting

*Forecasting* adalah suatu upaya untuk meramalkan masa depan seakurat mungkin dengan mempertimbangkan segala informasi yang tersedia, termasuk catatan sejarah atau kejadian lampau dan pemahaman tentang peristiwa yang potensial memengaruhi prediksi [8]. Dalam dunia bisnis, tujuan utama peramalan adalah untuk mengukur aliran permintaan dari bulan-bulan sebelumnya dan memproyeksikan ke bulan-bulan mendatang dengan kesalahan perkiraan minimum. Cara untuk meningkatkan tujuan ini adalah dengan menyaring permintaan

riwayat untuk mencari permintaan *outlier* dan menyesuaikannya [9].

### 2.3. Normalisasi Data

Normalisasi data yang digunakan adalah normalisasi *Z-Score*. *Z-Score* merupakan metode yang menggunakan nilai rata-rata atau *mean* dan standar deviasi untuk memperoleh hasil normalisasi, di mana hasilnya dapat mengurangi efek *outlier* atau data yang mengalami penyimpangan secara ekstrim dari nilai rata-rata [7]. Persamaan dari metode ini disajikan pada Persamaan (1),

$$X_{baru} = \frac{X_{lama} - \bar{x}}{\sigma} \quad (1)$$

di mana  $X_{baru}$  adalah nilai hasil normalisasi,  $X_{lama}$  adalah nilai yang akan dinormalisasi,  $\bar{x}$  adalah nilai rata-rata, dan  $\sigma$  adalah standar deviasi.

### 2.4. Support Vector Regression

*Support Vector Regression* (SVR) merupakan salah satu penerapan dari SVM yang menggunakan regresi. SVM diterapkan untuk pengklasifikasian yang keluarannya berupa bilangan bulat, sedangkan *output* dari SVR adalah berupa bilangan riil atau *continue* [10]. SVR adalah sebuah sistem pembelajaran yang bekerja pada sebuah fitur ruang hipotesis berdimensi tinggi dengan mengimplementasikan fungsi-fungsi tertentu. SVR sangat baik dalam mengatasi masalah *overfitting* sehingga menghasilkan performa yang baik [1]. *Overfitting* adalah suatu performa baik yang dihasilkan oleh model dari data yang telah dilatih dengan prediksi hampir sempurna, tetapi saat diuji ternyata hasil prediksi tidak sesuai. Metode SVR ini bertujuan untuk menemukan sebuah garis pemisah terbaik antara dua kelas data, yang disebut *hyperplane* terbaik. *Hyperplane* terbaik akan dicari dengan menentukan nilai margin dari *hyperplane* tersebut dan mencari titik maksimal dari margin tersebut. Margin merupakan jarak dari *hyperplane* dengan data-data yang paling dekat di dalam dua kelas tersebut. Data-data yang paling dekat itulah yang nantinya akan dijadikan *support*

*vector* [10]. Persamaan umum dari SVR disajikan pada Persamaan (2),

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha \times i) K(x_i, x_j) + b \quad (2)$$

Kebanyakan masalah dalam dunia nyata umumnya kebanyakan bersifat nonlinier, sehingga dalam persoalan ini dibutuhkan fungsi kernel untuk menyelesaikannya. Fungsi kernel akan mengganti *inner product* ( $x_i$  dan  $x_j$ ) untuk menyelesaikan persoalan linier dengan ruang dimensi yang tinggi [10].

Fungsi kernel pada SVR yang umum digunakan adalah kernel linier, kernel polinomial, kernel RBF/gaussian, dan kernel *tangent hyperbolic* (*sigmoid*) [11].

Kernel linier merupakan fungsi kernel paling sederhana di mana hanya menggunakan garis lurus sebagai *hyperplane* ketika hendak pemisahan data serta menganalisisnya [12]. Persamaan kernel linier disajikan pada Persamaan (3),

$$K(x, y) = x \times y \quad (3)$$

di mana pada Persamaan (3),  $K(x, y)$  adalah kernel linier,  $x$  dan  $y$  adalah vektor *input*.

Kernel *polynomial* merupakan fungsi kernel yang biasa dijalankan pada metode SVM dan metode kernel lainnya. Persamaan untuk kernel *polynomial* disajikan pada Persamaan (4),

$$K(x, y) = (x \times y + c)d \quad (4)$$

di mana pada Persamaan (4),  $K(x, y)$  adalah kernel *polynomial*,  $x$  dan  $y$  adalah vektor *input*,  $c$  adalah konstanta, dan  $d$  adalah parameter yang mengontrol kompleksitas model.

Kernel RBF merupakan kernel yang sering digunakan pada metode SVR untuk menganalisis data, sebab melalui pengambilan nilai yang sesuai, tingkat akurasi yang diperoleh sama dengan linier dan *polynomial* [13]. Parameter *gamma* berfungsi dalam mengontrol kecepatan *learning* dan parameter *cost* digunakan untuk membatasi nilai *alfa* selama proses *training*.

$$K(x, y) = \exp\left(-\frac{1}{2\sigma^2} \|x - x_i\|^2\right) \quad (5)$$

Di mana pada Persamaan (5),  $K(x, y)$  adalah kernel RBF,  $x$  dan  $y$  adalah vektor *input*,  $\exp$  adalah fungsi eksponensial,  $\sigma$  adalah sigma

untuk mengontrol lebar dari fungsi kernel, dan  $\|x - x_i\|^2$  adalah jarak Euclidean antara vektor input  $x$  dan  $x_i$ .

Kernel *Tangent Hyperbolic (Sigmoid)* merupakan fungsi kernel trigonometrik hiperbolik yang berbentuk “S”. Kernel ini adalah kombinasi dari berbagai fungsi eksponen [13].

$$K(x, y) = \tan(\sigma(x \times y) + c) \quad (6)$$

di mana  $K(x, y)$  pada Persamaan (6) adalah kernel Sigmoid,  $\tan$  adalah tangen,  $x, y$  adalah vektor *input*, dan  $c$  adalah konstanta.

Dalam menentukan parameter yang optimal untuk digunakan pada SVR, diperlukan sebuah algoritma, di mana salah satunya yang sering digunakan adalah algoritma *grid search*. Model yang ingin dilatih akan diambil oleh algoritma *grid search*, dengan melatih setiap nilai parameter yang tidak sinkron dan menghitung nilai kesalahannya. Nilai *error* yang terendah akan dijadikan sebagai parameter terbaik. Selama proses pengoptimalan algoritma ini, rentang parameter dibagi ke dalam *grid* dan di semua titik untuk memperoleh parameter terbaik. Penerapan algoritma *grid search* harus disertai dengan beberapa metrik kinerja, yang nilainya ditentukan menggunakan *cross-validation* pada data *training* [14].

Kemudian dalam mengukur akurasi dari hasil prediksi yang dilakukan SVR, alat ukur yang digunakan adalah MSE. MSE merupakan salah satu alat ukur untuk menguji hasil peramalan dengan menghitung selisih data aktual dengan data hasil prediksi, kemudian dikuadratkan dan dibagi dengan jumlah data [15]. Persamaan (7) digunakan untuk menghitung MSE,

$$MSE = \sum \frac{(y' - y)^2}{n} \quad (7)$$

di mana  $n$  adalah jumlah data,  $y$  adalah data aktual, dan  $y'$  adalah data hasil prediksi.

Hasil prediksi yang sangat baik diperoleh bila nilai MSE yang dihasilkan sangat kecil atau di bawah 1. Jika hasilnya di atas 10, maka hasil peramalannya dapat dikategorikan buruk. Hal ini dapat ditoleransi apabila variasi data yang digunakan untuk melakukan prediksi sangat besar.

### 3. Metode

#### 3.1. Pengumpulan Data

Pada tahap ini, data-data penelitian dikumpulkan sesuai dengan perumusan masalah yang telah disusun agar masalah tersebut dapat dipecahkan. Data-data tersebut diperoleh dari *file* yang berisi data terkait yang tersimpan pada komputer pemilik toko. Data-data yang dikumpulkan berupa total pesanan masuk, barang terjual, semua merek dari barang yang terjual, serta *marketplace* tempat terjualnya barang tersebut.

**Tabel 1.** Deskripsi sampel data penjualan

Tanggal	Pesanan Masuk	Barang Terjual	Merk		
			Matchbox	Tomica	Kyosho
1-Jan-22	1	1	0	0	0
2-Jan-22	5	5	1	2	2
3-Jan-22	10	10	9	1	0
4-Jan-22	11	11	10	0	1
5-Jan-22	2	2	0	2	0
6-Jan-22	1	1	0	0	0

#### 3.2. Jumlah Data

Dalam penelitian ini, jumlah data yang digunakan adalah data penjualan yang terkumpul selama periode 1 tahun di tahun 2022. Data penjualan tersebut mencakup penjualan produk *diecast* pada dua *marketplace*, yaitu Facebook dan Tokopedia. Setiap transaksi penjualan mencakup informasi seperti total pesanan masuk, jumlah barang terjual, merek barang yang terjual, dan *marketplace* tempat terjualnya barang tersebut (Facebook atau Tokopedia).

Dengan menggunakan jumlah data yang cukup besar dan meliputi periode waktu yang representatif, penelitian ini diharapkan dapat memberikan pemahaman yang lebih baik tentang pola penjualan produk *diecast* dan membantu dalam pengembangan model prediksi yang lebih akurat. Ilustrasi data penjualan disajikan pada Tabel 1.

#### 3.3. Pembagian Data

Kemudian selanjutnya adalah pembagian data. Pertama data akan diproses menjadi data yang layak untuk dimodelkan melalui normalisasi. Dalam normalisasi data, teknik yang umum digunakan adalah normalisasi *Min-Max*. Normalisasi *Min-Max* dilakukan dengan mengubah rentang nilai data ke rentang antara 0 dan 1. Dengan

demikian, semua variabel pada data memiliki skala yang seragam, memudahkan proses pembelajaran dan mengurangi efek dominasi variabel dengan skala yang lebih besar. Ilustrasi dari data normalisasi disajikan pada Tabel 2.

Setelah data dinormalisasi, data akan dibagi menjadi data *training* dan data *testing*. Data *training* digunakan untuk melatih model prediksi, sedangkan data *testing* digunakan untuk menguji performa model yang sudah dilatih. Pembagian ini penting untuk menguji sejauh mana model yang dibangun dapat melakukan prediksi yang akurat pada data yang belum pernah dilihat sebelumnya.

**Tabel 2.** Sampel data normalisasi

Tanggal	Pesanan Masuk	Barang Terjual	Merk		
			Matchbox	Tomica	Kyosho
1-Jan-22	-0.9966	-0.9966	-1,	-1,	-1,
2-Jan-22	-0.9828	-0.9828	-0.9773	-0.9874	-0.9412
3-Jan-22	-0.9656	-0.9656	-0.7955	-0.9937	-1,
4-Jan-22	-0.9622	-0.9622	-0.7727	-1.0000	-0.9706
5-Jan-22	-0.9931	-0.9931	-1,	-0.9874	-1,
6-Jan-22	-0.9966	-0.9966	-1,	-1,	-1,

### 3.4. Pembuatan Model SVR

Setelah data *training* dan data *testing* sudah dibagi, langkah selanjutnya adalah membangun model SVR. Pada tahap ini, data *training* digunakan untuk melatih model SVR. Data *training* terdiri dari fitur-fitur (seperti total pesanan masuk, jumlah barang terjual, dll.) dan label yang merupakan target prediksi (misalnya, total penjualan).

Dalam proses pelatihan model SVR, terdapat satu tahap penting yang disebut dengan algoritma *Grid Search*. Algoritma *Grid Search* digunakan untuk mencari kombinasi parameter yang optimal untuk model SVR. Parameter-parameter ini, seperti C (parameter toleransi kesalahan), gamma (parameter kernel), dan epsilon (parameter margin), mempengaruhi kinerja dan kemampuan prediktif model SVR.

Setelah model dilatih, model tersebut dapat digunakan untuk melakukan prediksi penjualan produk *diecast* di masa depan berdasarkan fitur-fitur yang diberikan.

### 3.5. Skenario Pengujian

Dalam skenario pengujian, dilakukan evaluasi terhadap model SVR yang digunakan untuk *forecasting*. Pengujian

dilakukan menggunakan metode *K-Fold Cross Validation*. Metode ini umumnya menggunakan *10 fold cross-validation* karena penelitian sebelumnya menunjukkan bahwa prediksi hasil dengan *10 fold* tersebut memiliki bias yang rendah dibandingkan dengan metode lainnya, sehingga hasil prediksinya mendekati kenyataan.

Selanjutnya, dilakukan validasi data baik secara kuantitatif maupun kualitatif. Validasi data kuantitatif menghasilkan ringkasan penjualan berdasarkan merek, disertai dengan nilai MSE. MSE memberikan indikasi seberapa baik model SVR dalam memprediksi penjualan berdasarkan perbedaan kuadrat antara nilai aktual dan nilai prediksi. Semakin rendah nilai MSE, semakin akurat prediksi model.

Validasi data kualitatif dilakukan dengan membandingkan penjualan aktual dengan hasil prediksi berdasarkan merek yang dihasilkan oleh model SVR. Hasil perbandingan ini dapat disajikan dalam bentuk grafik untuk memvisualisasikan kesesuaian antara data aktual dan prediksi. Dengan demikian, dapat dilihat secara langsung sejauh mana model SVR mampu menggambarkan pola penjualan yang terjadi.

Melalui skenario pengujian ini, penelitian ini dapat mengevaluasi performa model SVR secara holistik. Hasil validasi kuantitatif dan kualitatif memberikan pemahaman yang lebih mendalam tentang keakuratan prediksi model dan sejauh mana model tersebut dapat menggambarkan tren penjualan berdasarkan merek.

## 4. Hasil dan Pembahasan

### 4.1. Penerapan Model *Support Vector Regression*

Dalam mengimplementasikan model SVR, *software* yang digunakan adalah Google Colaboratory dan mengimpor SVR dari *library* *sklearn.svm*, karena SVR merupakan bagian dari SVM. Kemudian dari data penelitian selama setahun dibagi menjadi 2 bagian, yaitu data *training* dan data *testing*. Bobot data *training* lebih besar dengan ukuran 80% dari total data dan bobot

data *testing* adalah 20%. Setelah itu dicari kernel yang cocok serta parameter-parameternya yang optimal menggunakan algoritma *grid search*. Hasil terbaik yang didapatkan dengan menggunakan 10 *cross fold validation* adalah kernel terbaik yaitu kernel *tangen hyperbolic (sigmoid)* dengan  $\epsilon = 0,001$  dan parameter  $C = 100$ . Kemudian dilakukan prediksi barang terjual pada data *training*, dengan hasilnya disajikan pada Gambar 1.

```

Hasil prediksi pada saat training model SVR
Barang Terjual Aktual  Barang Terjual Prediksi
0      2      4
1      0      3
2      1      3
3      4      7
4      1      3
..      ..      ..
287    3      5
288    5      6
289    2      6
290    4      6
291    6      9
    
```

Gambar 1. Hasil prediksi SVR barang terjual pada data training

Prediksi yang dilakukan berdasarkan penjualan harian dengan total 292 hari untuk data *training*. MSE yang diperoleh pada data training sebesar 0,00182, yang berarti hasil prediksinya sangat akurat. Gambar 2 adalah prediksi pada data *testing*.

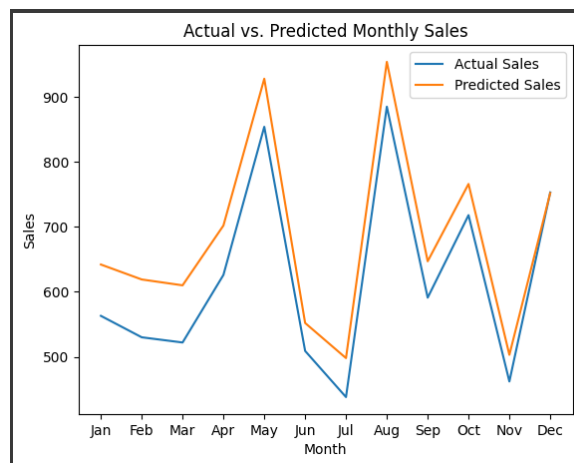
```

Hasil prediksi pada saat testing SVR
Barang Terjual Aktual  Barang Terjual Prediksi
0      1      4
1      1      3
2      1      5
3      520    497
4      7      11
..      ..      ..
68     2      5
69     0      3
70     3      5
71     2      4
72     3      6
    
```

Gambar 2. Hasil Prediksi SVR barang terjual pada data testing

Data *testing* dilakukan dalam kisaran 73 hari karena bobotnya lebih rendah dari data *training*. MSE yang diperoleh pada data *testing* sedikit berbeda dengan data *training*, yaitu sebesar 0,00235. Hasilnya tetap sangat akurat meski sedikit lebih besar dari nilai MSE data *training*.

Gambar 3 adalah grafik representasi hasil dari prediksi beserta perbandingan data aktual barang terjual dapat terlihat lebih jelas.



Gambar 3. Grafik Perbandingan Data Penjualan Aktual dan Prediksi

Gambar 3 diperoleh dengan menjumlahkan data *training* dan data *testing* beserta hasil prediksinya. Lalu data aktual dan hasil prediksi yang masih berupa harian dijumlahkan berdasarkan bulanan. Dapat dilihat bahwa garis dari hasil prediksi memiliki kemiripan pola dan jarak yang berdekatan, sehingga dapat dipahami bahwa hasil prediksinya baik. Tabel 3 ilustrasi dari perbandingan antara data penjualan aktual dan penjualan hasil prediksi secara bulanan dalam kurun waktu 1 tahun.

Tabel 3. Perbandingan data penjualan aktual dan data penjualan prediktif

Bulan	Data Penjualan Aktual	Data Penjualan Prediktif
1	563	642
2	530	619
3	522	610
4	626	702
5	854	928
6	509	552
7	438	498
8	885	954
9	591	647
10	718	766
11	462	503
12	753	752

#### 4.2. Pengujian Cross Validation

Pada pengujian hasil prediksi menggunakan SVR, *cross validation* digunakan sebagai alat evaluasi model yang sudah dibangun untuk mengukur kinerjanya

agar dapat diterapkan ke data yang baru. Pengujian dilakukan terhadap parameter  $C$ ,  $\gamma$ , dan kernel untuk memperoleh parameter terbaik serta MSE sebagai metrik pengukuran yang digunakan dengan pengujian mulai dari 3 lipatan/*fold* hingga 10-*fold cross validation*. Gambar 4 adalah ilustrasi dari hasil pengujian ini.

**Tabel 4.** Hasil pengujian *cross validation*

Fold	C	Gamma	kernel	MSE
3	100	0,001	Sigmoid	0,0454
5	100	0,001	Sigmoid	0,0185
7	100	0,001	Sigmoid	0,0161
10	100	0,001	Sigmoid	0,0148

Berdasarkan hasil pengujian keseluruhan, ditemukan bahwa parameter terbaik untuk model SVR adalah  $C$  dengan nilai 100,  $\gamma$  sebesar 0,001, dan menggunakan kernel *sigmoid*. Selain itu, hasil prediksi model SVR menunjukkan bahwa nilai MSE cenderung semakin kecil ketika jumlah lipatan bertambah. Pada lipatan ke-10, diperoleh nilai MSE sebesar 0,0148, yang merupakan nilai MSE terendah dibandingkan dengan lipatan lainnya. Jumlah lipatan yang digunakan hanya sampai 10 *fold* karena umum digunakan dan dianggap pilihan yang baik karena akurasi serta efisiensi komputasionalnya baik.

### 4.3. Validasi Kuantitatif

Validasi data kuantitatif melibatkan penilaian keseluruhan penjualan berdasarkan merek dan disertai dengan nilai MSE. Tabel 5 adalah validasi kuantitatif untuk hasil aktual dan prediksi penjualan setiap merek *diecast* dalam setahun. Nilai MSE juga disajikan sebagai ukuran akurasi prediksi. Contohnya, untuk merek Hotwheels, prediksi model SVR sebesar 1.359 unit dengan total penjualan aktual 1.545 unit, menghasilkan nilai MSE sebesar 0,091, menunjukkan prediksi yang relatif akurat. Namun, ada merek-merek seperti Efsi, Konami, Dragon Cando, Takara, Majorette, dan lainnya dengan nilai MSE di atas 1, menandakan

perbedaan yang signifikan dengan penjualan aktual. Tabel 5 memberikan pemahaman tentang performa prediksi model SVR untuk setiap merek *diecast*, memungkinkan identifikasi merek dengan prediksi yang baik dan yang perlu diperbaiki.

### 4.4. Validasi Kualitatif

Dalam validasi kualitatif, informasi mengenai hasil prediksi penjualan untuk setiap merek akan disajikan dalam bentuk grafik. Empat grafik akan dipilih berdasarkan hasil prediksi penjualan dari merek populer.

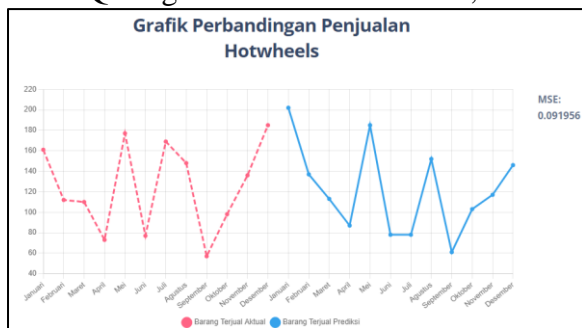
**Tabel 5.** Hasil prediksi penjualan dan nilai Gambar 4 adalah perbandingan hasil

Merek	Total Aktual Penjualan 1 Tahun	Total Prediksi Penjualan 1 Tahun	Nilai MSE
Hotwheels	1545	1359	0.091
Tamiya	84	82	0.014
Efsi	3	0	1.238
Konami	31	5	1.022
Dragon Cando	9	0	1.195
Takara	98	9	1.259
Majorette	528	365	1.244
Matchbox	449	431	1.049
Tomica	2628	2920	0.001
Hachette	121	76	0.212
Kyosho	369	433	0.460
Mercury Speedy	1	0	0
Aoshima	3	0	1.223
Herpa	238	6	1.228
CMS	253	51	1.232
Epoch	150	2	0.951
Bandai	55	3	1.262
Maisto	118	0	1.241
Realtoy	36	0	1.245
ChoroQ	510	513	0.005
Brekina	12	49	13.684
Kintoy	5	0	1.237
Kinsmart	45	2	1.058
Siku	161	20	1.241
Shinsei	39	2	0.314
Asahi	1	0	1.250

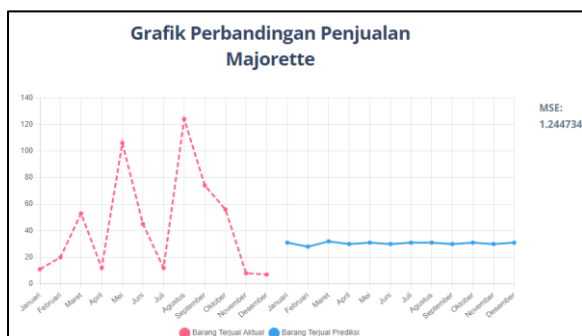
penjualan aktual (label warna merah) dengan hasil penjualan prediksi (label warna biru) merek mainan Hotwheels dengan nilai MSE sebesar 0,091956. Gambar 5 adalah perbandingan hasil penjualan aktual (label warna merah) dengan hasil penjualan prediksi (label warna biru) merek mainan Majorette dengan nilai MSE sebesar



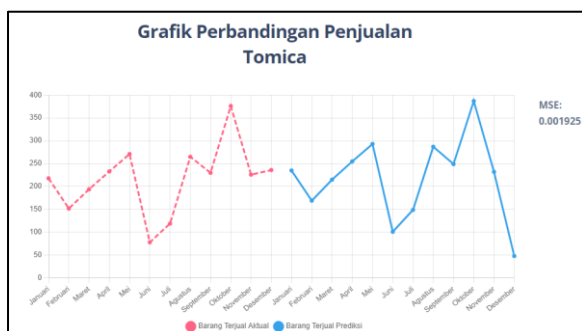
1,244734. Gambar 6 adalah perbandingan hasil penjualan aktual (label warna merah) dengan hasil penjualan prediksi (label warna biru) merek mainan Tomica dengan nilai MSE sebesar 0,001925. Gambar 7 adalah perbandingan hasil penjualan aktual (label warna merah) dengan hasil penjualan prediksi (label warna biru) merek mainan ChoroQ dengan nilai MSE sebesar 0,005145.



**Gambar 4.** Perbandingan hasil penjualan aktual dengan hasil penjualan prediksi merek Hotwheels



**Gambar 5.** Perbandingan hasil penjualan aktual dengan hasil penjualan prediksi merek Majorette



**Gambar 6.** Perbandingan hasil penjualan aktual dengan hasil penjualan prediksi merek Tomica

Kesimpulan yang diperoleh dari penelitian pembuatan fitur peramalan atau *forecasting* penjualan menggunakan SVR dapat memberikan informasi yang sangat berguna bagi untuk mengatur penyediaan barang-barang di masa mendatang di mana untuk hasil akurasi yang baik saat menggunakan alat ukur MSE.



**Gambar 7.** Perbandingan hasil penjualan aktual dengan hasil penjualan prediksi merek ChoroQ

Nilai MSE yang diperoleh dari penelitian ini menunjukkan bahwa pada variabel penjualan dapat memprediksi dengan sangat baik. Kemudian untuk hasil prediksi untuk merk yang banyak terjual juga baik meski nilai MSE lebih rendah dibandingkan nilai MSE variabel penjualan. Merek-merek yang kurang diminati dan jarang terjual cenderung tidak menghasilkan nilai prediksi yang baik di mana merek Brekina memiliki nilai MSE yang paling tinggi, yaitu 13,684 yang berarti hasil prediksinya sangat buruk.

Penelitian di masa mendatang dapat mengukur akurasi dari hasil prediksi menggunakan Root Mean Squarred Error (RMSE) dan Mean Absolute Percent Error (MAPE) sehingga tingkat akurasi benar-benar teruji.

Identifikasi merek-merek yang memiliki nilai MSE yang tinggi menunjukkan prediksi yang kurang akurat. Analisis lebih lanjut bisa berfokus pada karakteristik dan faktor yang mempengaruhi penjualan pada merek-merek tersebut. Perbaikan model prediksi atau pertimbangkan pendekatan yang lebih sesuai

## 5. Penutup

untuk meningkatkan akurasi prediksi pada merek-merek tersebut juga dapat dilakukan.

## 6. Daftar Pustaka

- [1] D. I. Purnama and S. Setianingsih, "Support Vector Regression (SVR) Model for Forecasting Number of Passengers on Domestic Flights at Sultan Hasanudin Airport Makassar," *J. Mat. Stat. dan Komputasi*, vol. 16, no. 3, pp. 391–403, 2020.
- [2] D. Suprayogi and H. F. Pardede, "Support Vector Regression Dalam Prediksi Penurunan Jumlah Kasus Penderita COVID-19," *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 7, no. 2, pp. 63–70, 2022, doi: <https://doi.org/10.31328/jointecs.v7i2.3687>.
- [3] R. S. Laminullah, H. Annur, and I. S. K. Idris, "Prediksi Penjualan Peralite Menggunakan Metode Support Vector Regression," *J. Tek. Elektro CosPhi*, vol. 4, no. 1, pp. 12–14, 2020.
- [4] X. Yu, Z. Qi, and Y. Zhao, "Support Vector Regression for Newspaper/Magazine Sales Forecasting," *Procedia Comput. Sci.*, vol. 17, pp. 1055–1062, 2013, doi: <https://doi.org/10.1016/j.procs.2013.05.134>.
- [5] G. H. Saputra, A. H. Wigena, and B. Sartono, "Penggunaan Support Vector Regression dalam Pemodelan Indeks Saham Syariah Indonesia dengan Algoritme Grid Search," *Indones. J. Stat. Its Appl.*, vol. 3, no. 2, pp. 148–160, 2019.
- [6] I. Permana and F. N. Salisah, "The Effect of Data Normalization on the Performance of the Classification Results of the Backpropagation Algorithm," *Indones. J. Inform. Res. Softw. Eng.*, vol. 2, no. 1, pp. 67–72, 2022, doi: <https://doi.org/10.57152/ijirse.v2i1.311>.
- [7] R. G. Whendasmoro and J. Joseph, "Analisis Penerapan Normalisasi Data dengan Menggunakan Z-Score pada Kinerja Algoritma K-NN," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 4, pp. 872–876, 2022, doi: <http://dx.doi.org/10.30865/jurikom.v9i4.4526>.
- [8] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*. Australia: OTexts: Melbourne, 2018.
- [9] N. T. Thomopoulos, *Demand Forecasting for Inventory Control*. Springer, Cham, 2014. doi: <https://doi.org/10.1007/978-3-319-11976-2>.
- [10] R. P. Furi, J. Jondri, and D. Saepudin, "Prediksi Financial Time Series Menggunakan Independent Component Analysis dan Support Vector Regression Studi Kasus: IHSB dan JII," in *Proceedings of Engineering*, 2015, vol. 2, no. 2, pp. 3608–3618. [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/2745>
- [11] M. Y. Darsyah, "Klasifikasi Tuberkulosis dengan Pendekatan Metode Supports Vector Machine (SVM)," *J. Stat. Univ. Muhammadiyah Semarang*, vol. 2, no. 2, pp. 37–41, 2014.
- [12] R. A. Putri, W. S. Winahju, and M. Mashuri, "Penerapan Metode Ridge Regression dan Support Vector Regression (SVR) untuk Prediksi Indeks Batubara di PT. XYZ," *J. Sains dan Seni ITS*, vol. 9, no. 1, pp. D64–D71, 2020, doi: <http://dx.doi.org/10.12962/j23373520.v9i1.51021>.