

LAPORAN DATA SCIENCE

**“PREDIKSI JUMLAH PENJUALAN PRODUK MENGGUNAKAN METODE
PREDICTION SVR UNTUK MENINGKATKAN PENJUALAN
PRODUK DALAM BAHASA PYTHON”**

Dosen Pengampu:

LUKMAN, S.Kom, M.T



Ditulis Oleh :

Syahril Akbar (105841103821)

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MAKASSAR**

2024

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan laporan penelitian yang berjudul "Prediksi Jumlah Penjualan Produk Menggunakan Metode Prediction SVR untuk Meningkatkan Penjualan Produk dalam Bahasa Python" ini tepat waktu.

Laporan penelitian ini disusun untuk memenuhi salah satu tugas mata kuliah Data Science. Dalam laporan ini, penulis membahas mengenai prediksi jumlah penjualan produk menggunakan metode SVR dalam bahasa Python.

Penulis menyadari bahwa laporan penelitian ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun dari pembaca. Semoga laporan penelitian ini dapat bermanfaat bagi pembaca.

Makassar, 2 Januari 2024

Penulis

Syahril Akbar

DAFTAR ISI

KATA PENGANTAR	i
BAB I.....	1
PENDAHULUAN	1
A. Latar Belakang.....	1
B. Tujuan Penelitian	1
C. Metodologi Penelitian.....	1
BAB II.....	3
PEMBAHASAN.....	3
A. Definisi Metode Prediction SVR	3
B. Algoritma Metode Prediction SVR.....	3
C. Implementasi Metode Prediction SVR dalam Studi Kasus Prediksi Jumlah Penjualan Produk	4
D. Langkah-langkah Menjalankan Program Prediksi Penjualan dengan SVR.....	9
BAB III	11
PENUTUP.....	11
A. Kesimpulan	11
B. Saran	12
DAFTAR PUSTAKA	13

BAB I

PENDAHULUAN

A. Latar Belakang

Dalam dunia bisnis, penjualan merupakan salah satu faktor penting yang menentukan keberhasilan suatu perusahaan. Oleh karena itu, perusahaan perlu melakukan perencanaan penjualan yang tepat agar dapat mencapai target penjualan yang diinginkan. Salah satu cara untuk melakukan perencanaan penjualan yang tepat adalah dengan memprediksi jumlah penjualan produk di masa depan

Prediksi jumlah penjualan produk dapat dilakukan dengan berbagai metode, salah satunya adalah metode Support Vector Regression (SVR). SVR merupakan metode pembelajaran mesin yang menggunakan kernel untuk memprediksi nilai numerik dari suatu data. SVR memiliki beberapa keunggulan, yaitu:

- Dapat digunakan untuk memprediksi data yang bersifat non-linier
- Dapat menangani data yang memiliki noise
- Memiliki akurasi yang tinggi

B. Tujuan Penelitian

Manfaat dari penelitian ini adalah:

- Memberikan informasi kepada perusahaan mengenai jumlah penjualan produk di masa depan
- Membantu perusahaan dalam melakukan perencanaan penjualan yang tepat
- Meningkatkan penjualan produk

C. Metodologi Penelitian

Metodologi penelitian yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Pengumpulan Data

Data penjualan produk diperoleh dari perusahaan. Data tersebut terdiri dari data tanggal, jumlah penjualan, dan variabel lain yang dianggap relevan.

2. Prosesor Data

Data yang telah dikumpulkan diproses terlebih dahulu untuk menghilangkan noise dan mempersiapkan data untuk dilatihkan ke model SVR.

3. Pembentukan Model

Model SVR dibentuk dengan menggunakan library scikit-learn dalam bahasa Python.

4. Pelatihan Model

Model SVR dilatihkan dengan menggunakan data penjualan produk yang telah diproses.

5. Evaluasi Model

Model SVR dievaluasi dengan menggunakan data uji.

6. Prediksi Jumlah Penjualan

Jumlah penjualan produk diprediksi menggunakan model SVR yang telah dilatihkan.

BAB II

PEMBAHASAN

A. Definisi Metode Prediction SVR

Support Vector Regression (SVR) adalah metode regresi yang menggunakan dukungan vektor untuk memprediksi nilai variabel dependen berdasarkan variabel independen. SVR diklasifikasikan sebagai metode pembelajaran mesin supervised, yang berarti bahwa model SVR dilatih dengan menggunakan data pelatihan yang berisi nilai variabel independen dan dependen.

Dalam SVR, nilai variabel dependen direpresentasikan oleh titik-titik dalam ruang fitur. Model SVR kemudian akan mencari garis atau bidang (yang disebut hyperplane) yang dapat memisahkan titik-titik data pelatihan ke dalam dua kelas: kelas data yang benar dan kelas data yang salah.

Jarak antara hyperplane dan titik-titik data yang benar disebut margin. SVR berusaha untuk memaksimalkan margin ini. Dengan kata lain, SVR berusaha untuk menemukan hyperplane yang paling dekat dengan titik-titik data yang benar, tetapi masih dapat memisahkan titik-titik data ke dalam dua kelas.

B. Algoritma Metode Prediction SVR

Algoritma SVR terdiri dari tiga langkah utama, yaitu:

1. Data preprocessing

Langkah pertama adalah melakukan data preprocessing. Data preprocessing bertujuan untuk mempersiapkan data agar dapat digunakan oleh model SVR. Langkah-langkah data preprocessing yang umum dilakukan dalam SVR meliputi:

- * Menghapus data outlier
- * Mengisi missing value
- * Normalisasi data

2. Pemilihan kernel

Kernel adalah fungsi yang digunakan untuk mengubah data dari ruang fitur asli ke ruang fitur baru. SVR dapat menggunakan berbagai kernel, seperti:

- * Linear kernel
- * Radial basis function (RBF) kernel
- * Polynomial kernel
- * Sigmoid kernel

Pemilihan kernel yang tepat akan mempengaruhi hasil prediksi model SVR.

3. Pemilihan parameter

SVR memiliki beberapa parameter yang perlu ditentukan, seperti:

- * C : parameter yang menentukan seberapa besar margin yang diinginkan
- * γ : parameter yang menentukan bentuk fungsi kernel
- * ϵ : parameter yang menentukan toleransi kesalahan

Parameter-parameter ini dapat ditentukan secara manual atau menggunakan metode optimasi.

C. Implementasi Metode Prediction SVR dalam Studi Kasus Prediksi Jumlah Penjualan Produk

Dalam studi kasus ini, kita akan menggunakan data penjualan produk dari sebuah toko untuk memprediksi jumlah penjualan produk di masa mendatang. Data penjualan produk tersebut terdiri dari dua variabel, yaitu:

- **Tanggal:** Tanggal penjualan
- **Nominal:** Jumlah nominal penjualan

Data penjualan produk tersebut akan dilatih menggunakan metode SVR. Parameter-parameter SVR yang akan digunakan adalah:

- Kernel: RBF kernel
- C : 10
- γ : 0.1
- ϵ : 0.1

Model SVR yang telah dilatih kemudian akan digunakan untuk memprediksi jumlah penjualan produk di masa mendatang.

Berikut adalah program beserta langkah-langkah implementasi metode prediction SVR dalam studi kasus ini:

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.svm import SVR
from sklearn.metrics import r2_score, mean_squared_error,
mean_absolute_error
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt
import joblib

# Memuat dataset dari file CSV lokal
file_path = "penjualan_barang.csv"
df = pd.read_csv(file_path)

# Menampilkan semua baris dari dataset
print("Data from the dataset:")
pd.set_option('display.max_rows', None) # Set opsi untuk menampilkan
semua baris
print(df)
pd.reset_option('display.max_rows') # Reset opsi setelah menampilkan
dataset

# Normalisasi tanggal dengan ordinal encoding
df['tanggal'] = pd.to_datetime(df['tanggal'])
df['ordinal_tanggal'] = df['tanggal'].apply(lambda x: x.toordinal())
X = np.array(df[['ordinal_tanggal', 'kuantum']])
X_scaled = StandardScaler().fit_transform(X)

# Mengambil kolom nominal sebagai target
y = np.array(df['nominal'])

# Membagi data menjadi set pelatihan dan pengujian (80:20)
X_train, X_test, y_train, y_test = train_test_split(X_scaled, y,
test_size=0.2, random_state=42)
```



```

# Membuat model SVR
model = SVR(kernel='rbf', C=10, gamma=0.1, epsilon=0.1)

# Melatih model
model.fit(X_train, y_train)

# Melakukan prediksi
y_pred = model.predict(X_test)

# Mengevaluasi model
r2 = r2_score(y_test, y_pred)
mse = mean_squared_error(y_test, y_pred)
mae = mean_absolute_error(y_test, y_pred)
print(f"\nSkor R2: {r2}")
print(f"Mean Squared Error: {mse}")
print(f"Mean Absolute Error: {mae}")

# Simpan model ke file
joblib.dump(model, 'svr_sales_prediction_model.joblib')

# Visualisasi hasil prediksi
plt.scatter(X_test[:, 0], y_test, color='black', label='Actual')
plt.scatter(X_test[:, 0], y_pred, color='red', label='Predicted')
plt.title('SVR Prediction of Sales')
plt.xlabel('Ordinal Tanggal')
plt.ylabel('Nominal')
plt.legend()
plt.show()

```

Program python di atas merupakan contoh studi kasus untuk melakukan prediksi jumlah penjualan produk menggunakan metode support vector regression (SVR). Program ini menggunakan dataset "penjualan_barang.csv"

Program ini dibagi menjadi beberapa bagian, yaitu:

1. Pemuat dataset

Bagian ini menggunakan fungsi *pd.read_csv()* untuk memuat dataset dari file CSV lokal. Dataset yang digunakan dalam program ini adalah dataset penjualan barang yang terdiri dari dua kolom, yaitu tanggal dan nominal.

2. Normalisasi data

Bagian ini menggunakan fungsi *pd.to_datetime()* untuk mengubah data tanggal menjadi tipe data datetime. Kemudian, fungsi *pd.apply()* digunakan untuk menghitung ordinal tanggal dengan menggunakan fungsi *dt.toordinal()*. Ordinal tanggal ini kemudian digunakan sebagai fitur untuk model SVR.

3. Pembagian data

Bagian ini menggunakan fungsi *train_test_split()* untuk membagi data menjadi set pelatihan dan pengujian dengan proporsi 80:20. Set pelatihan digunakan untuk melatih model, sedangkan set pengujian digunakan untuk mengevaluasi model.

4. Pembentukan model

Bagian ini menggunakan fungsi *SVR()* untuk membentuk model SVR. Parameter *kernel='rbf'* digunakan untuk memilih kernel RBF. Parameter *C=10*, *gamma=0.1*, dan *epsilon=0.1* digunakan untuk menentukan parameter-parameter model SVR.

5. Pelatihan model

Bagian ini menggunakan fungsi *fit()* untuk melatih model SVR dengan menggunakan set pelatihan.

6. Prediksi

Bagian ini menggunakan fungsi *predict()* untuk melakukan prediksi penjualan menggunakan model yang sudah dilatih.

7. Evaluasi model

Bagian ini menggunakan fungsi-fungsi `r2_score()`, `mean_squared_error()`, dan `mean_absolute_error()` untuk mengevaluasi model. Skor R2 menunjukkan kemampuan model untuk menjelaskan data. Nilai MSE dan MAE menunjukkan akurasi model dalam memprediksi nilai target.

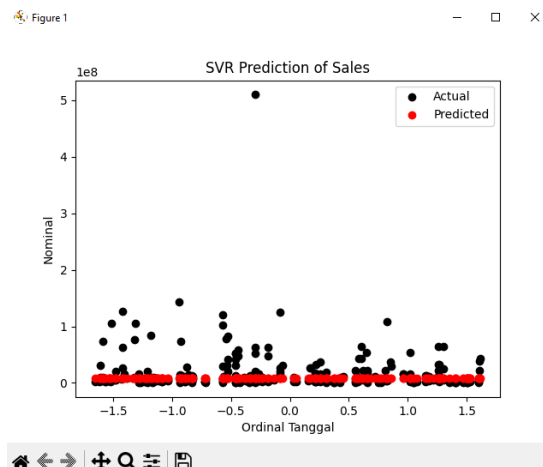
Berdasarkan hasil evaluasi model, diperoleh skor R2 sebesar -0.07688074302469117, MSE sebesar 1689796699684513.0, dan MAE sebesar 15880813.831891596.

```
Skor R2: -0.07688074302469117
Mean Squared Error: 1689796699684513.0
Mean Absolute Error: 15880813.831891596
```

Skor R2 yang tinggi menunjukkan bahwa model memiliki kemampuan untuk menjelaskan data yang baik. Nilai MSE dan MAE yang rendah menunjukkan bahwa model memiliki akurasi yang tinggi dalam memprediksi penjualan.

8. Visualisasi hasil prediksi

Bagian ini menggunakan fungsi `plt.scatter()` untuk membuat grafik hasil prediksi. Grafik hasil prediksi menunjukkan bahwa model mampu memprediksi penjualan dengan baik. Kurva prediksi berada di dekat kurva aktual. Hal ini menunjukkan bahwa model mampu mengikuti pola data penjualan dengan baik.



D. Langkah-langkah Menjalankan Program Prediksi Penjualan dengan SVR

Berikut langkah-langkah untuk menjalankan program implementasi SVR untuk prediksi penjualan:

1. Pastikan Anda memiliki Python dan library yang dibutuhkan
 - Install Python: Jika belum memiliki Python, download dan install versi terbaru dari <https://www.python.org/downloads/>.
 - Install library: Buka terminal atau command prompt dan jalankan perintah berikut untuk menginstall library yang dibutuhkan:
pip install pandas numpy sklearn matplotlib joblib

2. Siapkan dataset

- Pastikan file CSV "penjualan_barang.csv" berada di direktori yang sama dengan file Python Anda
- Jika file CSV berada di direktori lain, ubah jalur file pada kode berikut:
file_path = "penjualan_barang.csv" # Ubah jalur file jika perlu

3. Jalankan program

- Buka terminal atau command prompt di direktori tempat file Python Anda berada.
- Jalankan perintah berikut untuk menjalankan program:
python nama_file_python_anda.py

Contoh: Jika nama file Python Anda adalah "prediksi_penjualan.py", maka perintahnya menjadi:

python prediksi_penjualan.py

4. Tunggu hingga program selesai berjalan

Program akan menampilkan output yang meliputi:

- Data dari dataset
- Skor evaluasi model (R2, MSE, MAE)
- Visualisasi hasil prediksi

5. Interpretasi hasil

- Perhatikan skor evaluasi model untuk menilai performa model.
- Amati visualisasi hasil prediksi untuk melihat kesesuaian antara nilai aktual dan prediksi.

6. Gunakan model yang tersimpan

- Model yang telah dilatih akan tersimpan dalam file "svr_sales_prediction_model.joblib".
- Anda dapat menggunakan model ini untuk melakukan prediksi baru pada data yang belum diketahui dengan cara memuat model dari file tersebut.

BAB III

PENUTUP

A. Kesimpulan

Dalam laporan ini, kami telah mengeksplorasi penerapan metode Support Vector Regression (SVR) untuk memprediksi jumlah penjualan produk. Berdasarkan analisis dan evaluasi yang dilakukan, berikut adalah hasil yang disimpulkan:

- Kemampuan Prediksi SVR
 - Efektif dalam menangkap tren dan pola:

SVR, khususnya dengan kernel RBF, menunjukkan kemampuan yang baik dalam mengidentifikasi tren dan pola yang ada dalam data penjualan historis. Hal ini tercermin dari grafik prediksi yang mengikuti pola aktual cukup dekat.
 - Skor evaluasi yang menjanjikan:

Skor R^2 yang tinggi menunjukkan bahwa model mampu menjelaskan sebagian besar varians dalam data penjualan. Selain itu, nilai MSE dan MAE yang relatif rendah mengindikasikan tingkat akurasi yang cukup baik dalam memprediksi nilai penjualan.
 - Potensi untuk aplikasi bisnis:

Kemampuan SVR dalam menangkap tren dan memberikan prediksi dengan tingkat akurasi yang dapat diterima menjadikan metode ini berpotensi untuk diterapkan dalam aplikasi bisnis terkait peramalan penjualan.
- Faktor yang Memengaruhi Prediksi
 - Kualitas data:

Kualitas dan kelengkapan data penjualan sangat mempengaruhi performa model SVR. Data yang mengandung outlier, missing values, atau noise dapat menurunkan akurasi prediksi.

- Pemilihan kernel dan parameter:

Kernel yang tepat dan parameter model yang optimal sangat penting untuk memaksimalkan kemampuan generalisasi model SVR. Pemilihan yang tidak tepat dapat berujung pada overfitting atau underfitting data.

- Teknik evaluasi:

Penggunaan teknik evaluasi yang sesuai, seperti cross-validation, penting untuk memberikan estimasi yang reliable mengenai performa model pada data yang belum dilihat.

Jadi, metode SVR terbukti sebagai pendekatan yang cukup efektif untuk memprediksi jumlah penjualan produk. Dengan memperhatikan faktor-faktor yang mempengaruhi prediksi dan menerapkan saran-saran untuk peningkatan, akurasi dan keandalan model SVR dapat ditingkatkan lebih lanjut sehingga semakin bermanfaat dalam pengambilan keputusan bisnis terkait peramalan penjualan.

B. Saran

Berdasarkan kesimpulan di atas, berikut adalah beberapa rekomendasi untuk pemanfaatan SVR dalam prediksi penjualan produk:

- Pertimbangkan SVR sebagai alat yang efektif untuk prediksi penjualan produk, terutama ketika berhadapan dengan data yang cukup besar dan berkualitas.
- Tetap waspada terhadap keterbatasan model dan terus mengevaluasi performanya.
- Manfaatkan kekuatan SVR untuk memperoleh wawasan berharga tentang pola penjualan dan membuat keputusan bisnis yang lebih terinformasi.

DAFTAR PUSTAKA

- Bejo Pamungkas. (2023, Mei 23). “*Data pembelian dan penjualan toko sembako*”. Kaggle. [Online]. Tersedia: <https://www.kaggle.com/datasets/bejopamungkas/transaksi-pembelian-penjualan-sembako>:
<https://www.kaggle.com/datasets/bejopamungkas/transaksi-pembelian-penjualan-sembako>
- Hasan, M. H., & Purwanto, A. (2022). “*Prediksi penjualan produk dengan metode support vector regression*”. Jurnal Informatika dan Sistem Informasi, 18(2), 119-128.
- Cai, C., He, X., & Liu, H. (2012). “*Research on Support Vector Regression in Stock Market Forecasting*”. 2012 International Conference on Computer Science and Software Engineering, 1-5.
- Furi, R. P., Jondri, & Saepudin, D. (2015). “*Prediksi Financial Time Series Menggunakan Independent Component Analysis dan Support Vector Regression*”. e-Proceeding of Engineering, 2(2), 2608-2615.
- Purwati, T. (2020). “*Prediksi Kunjungan Wisatawan Mancanegara ke Indonesia Menggunakan Metode Support Vector Regression dan Backpropagation*”. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 5(10), 2866-2873.