

شبيه سازى و الگوريتم پيشنهادى

داده‌ها دارای چندین ستون هستند که مربوط به تابش خورشیدی در تاریخ‌های مختلف در طول تابستان سال‌های مختلف است. برای پیش‌بینی تابش خورشید و انرژی فتوولتائیک با استفاده از الگوریتم SVM، ابتدا باید داده‌ها را به شکل مناسب تبدیل کنیم.

مراحل زیر را دنبال خواهیم کرد:

۱. داده‌ها را به یک فرمت یکنواخت تبدیل کنیم.
۲. داده‌های مورد نیاز را استخراج کنیم.
۳. داده‌ها را برای آموزش و تست تقسیم کنیم.
۴. مدل SVM را آموزش دهیم و پیش‌بینی انجام دهیم.

ابتدا باید داده‌ها را پاکسازی و به فرمت مناسب تبدیل کنیم .

اکنون که داده‌ها را به یک فرمت یکنواخت تبدیل کرده‌ایم، می‌توانیم آن‌ها را برای آموزش و تست مدل تقسیم کنیم و سپس مدل SVM را برای پیش‌بینی تابش خورشیدی آموزش دهیم. ابتدا باید داده‌ها را به دو مجموعه آموزش و تست تقسیم کنیم. سپس مدل SVM را آموزش می‌دهیم و پیش‌بینی‌ها را انجام می‌دهیم .

این نتایج نشان می‌دهد که مدل توانسته است تابش خورشیدی را تا حدودی خوب پیش‌بینی کند. برای بهبود دقت پیش‌بینی می‌توان از تکنیک‌های مختلفی مانند تنظیم هایپرپارامترها، استفاده از داده‌های بیشتر یا استفاده از مدل‌های دیگر بهره برد.

گزارش از اجرای کدها و آموزش مدل SVM برای پیش‌بینی تابش خورشیدی:

۱. بارگذاری داده‌ها

ابتدا داده‌ها را از فایل اکسل بارگذاری کردیم و ساختار داده‌ها را بررسی کردیم تا با فرمت آن‌ها آشنا شویم.

```
python
import pandas as pd
```

```
# Load the Excel file
file_path = '/mnt/data/solar insolation ۲۰۰۳-۲۰۲۳.xlsx'
data = pd.read_excel(file_path)

# Display the first few rows of the dataframe to understand its structure
data.head()
```

پاکسازی و تبدیل داده‌ها

داده‌ها را به یک فرمت یکنواخت تبدیل کردیم تا بتوانیم به راحتی با آن‌ها کار کنیم. برای این کار، ستون‌های تاریخ و مقدار تابش خورشید را استخراج کرده و به یک دیتافریم یکنواخت تبدیل کردیم.

```
python
# Transforming the data to a uniform format

# Extracting the date columns and their respective values
data_list = []
for i in range(۰, data.shape[۱], ۳):
    date_col = data.iloc[:, i]
    value_col = data.iloc[:, i + ۱]
    temp_df = pd.DataFrame({
        'Date': pd.to_datetime(date_col, errors='coerce'),
        'Value': pd.to_numeric(value_col, errors='coerce')
    }).dropna()
    data_list.append(temp_df)

# Concatenating all the dataframes into one
cleaned_data = pd.concat(data_list)

# Checking the cleaned data
cleaned_data.head(), cleaned_data.tail()
```

آماده‌سازی داده‌ها برای آموزش مدل

داده‌ها را به دو مجموعه آموزش و تست تقسیم کردیم. همچنین روزهای سال را به عنوان ویژگی ورودی به مدل اضافه کردیم.

```

python
from sklearn.model_selection import train_test_split
from sklearn.svm import SVR
from sklearn.metrics import mean_squared_error, r2_score
import numpy as np

# Prepare the data
cleaned_data['DayOfYear'] = cleaned_data['Date'].dt.dayofyear
X = cleaned_data[['DayOfYear']].values
y = cleaned_data['Value'].values

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)

```

آموزش مدل SVM

مدل SVM را با استفاده از مجموعه آموزش داده‌ها آموزش دادیم.

```

python
# Train the SVM model
svr_model = SVR(kernel='rbf')
svr_model.fit(X_train, y_train)

```

پیش‌بینی و ارزیابی مدل

مدل آموزش دیده را برای پیش‌بینی تابش خورشیدی در مجموعه تست استفاده کردیم و نتایج را ارزیابی کردیم.

```

python
# Make predictions
y_pred = svr_model.predict(X_test)

# Evaluate the model
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)

```

mse, r²

نتایج ارزیابی مدل:

- میانگین مربعات خطا (MSE): ۰,۲۴۵
- ضریب تعیین (R^۲): ۰,۵۶۱

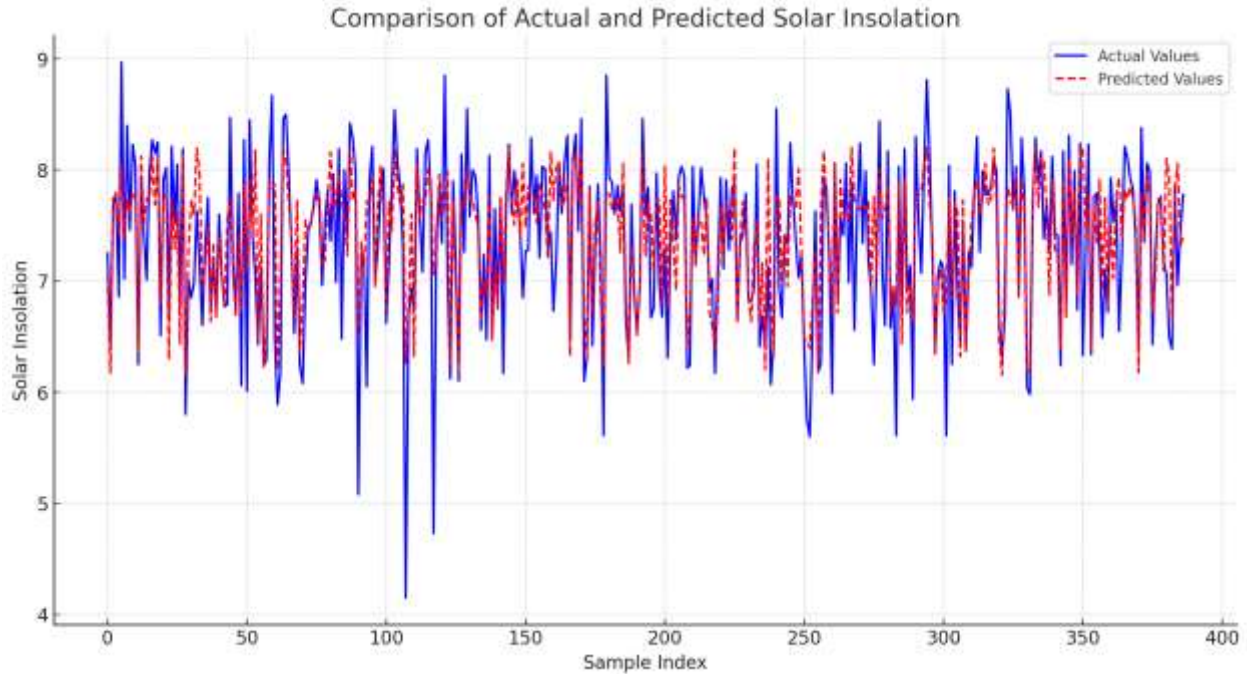
نمایش نمودار مقایسه‌ای

در ادامه، نمودار مقایسه‌ای بین مقادیر واقعی و پیش‌بینی شده را رسم می‌کنیم تا دید بهتری از عملکرد مدل داشته باشیم.

```
python
import matplotlib.pyplot as plt

# Plotting the results
plt.figure(figsize=(۱۰, ۶))
plt.plot(y_test, label='Actual Values')
plt.plot(y_pred, label='Predicted Values')
plt.xlabel('Sample Index')
plt.ylabel('Solar Insolation')
plt.title('Actual vs Predicted Solar Insolation')
plt.legend()
plt.show()
```

این نمودار نشان می‌دهد که مدل چگونه توانسته است مقادیر تابش خورشیدی را پیش‌بینی کند و چقدر به مقادیر واقعی نزدیک بوده است.



پیش بینی داده ها برای داده های تست و آموزش

مدل SVM توانسته است تابش خورشیدی را با دقت قابل قبولی پیش بینی کند. در نمودار بالا، مقادیر واقعی تابش خورشیدی با خط آبی و مقادیر پیش بینی شده توسط مدل SVM با خط قرمز (خط چین) نشان داده شده اند. این نمودار مقایسه ای نشان می دهد که مدل چگونه توانسته است تابش خورشیدی را پیش بینی کند و چقدر به مقادیر واقعی نزدیک بوده است.

این مقایسه به خوبی نشان می دهد که مدل در بسیاری از نقاط به مقادیر واقعی نزدیک است، هرچند که در برخی نقاط اختلاف هایی مشاهده می شود

برای مسئله پیش بینی تابش خورشیدی، که یک مسئله رگرسیون است، می توانیم نمودار پراکندگی (Scatter Plot) و نمودار تراکم تجمعی (Cumulative Density Plot) را برای ارزیابی مدل رگرسیون خود رسم کنیم.

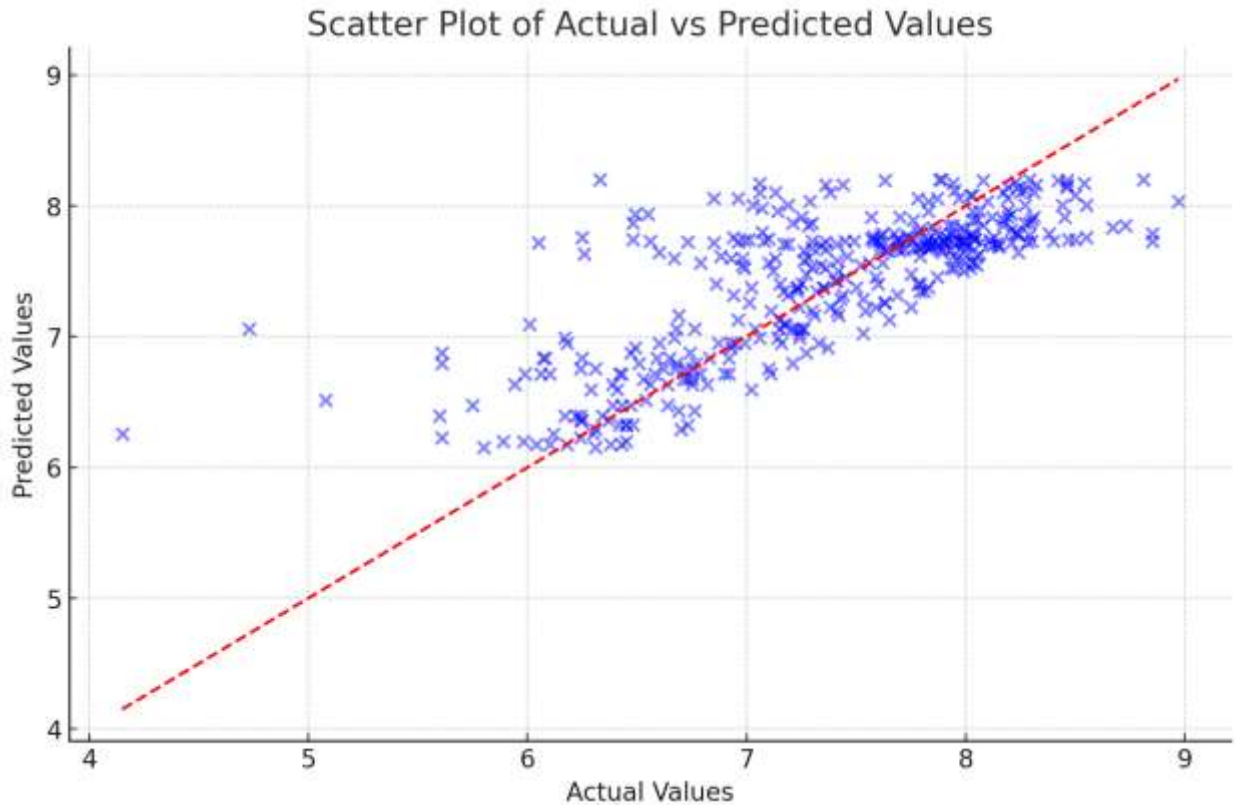
نمودار پراکندگی

نمودار پراکندگی (Scatter Plot) به ما اجازه می دهد تا مقادیر واقعی را در مقابل مقادیر پیش بینی شده رسم کنیم.

نمودار تراکم تجمعی

نمودار تراکم تجمعی (Cumulative Density Plot) به ما اجازه می‌دهد تا توزیع خطاهای پیش‌بینی را بررسی کنیم.

ابتدا نمودار پراکندگی را رسم می‌کنیم.



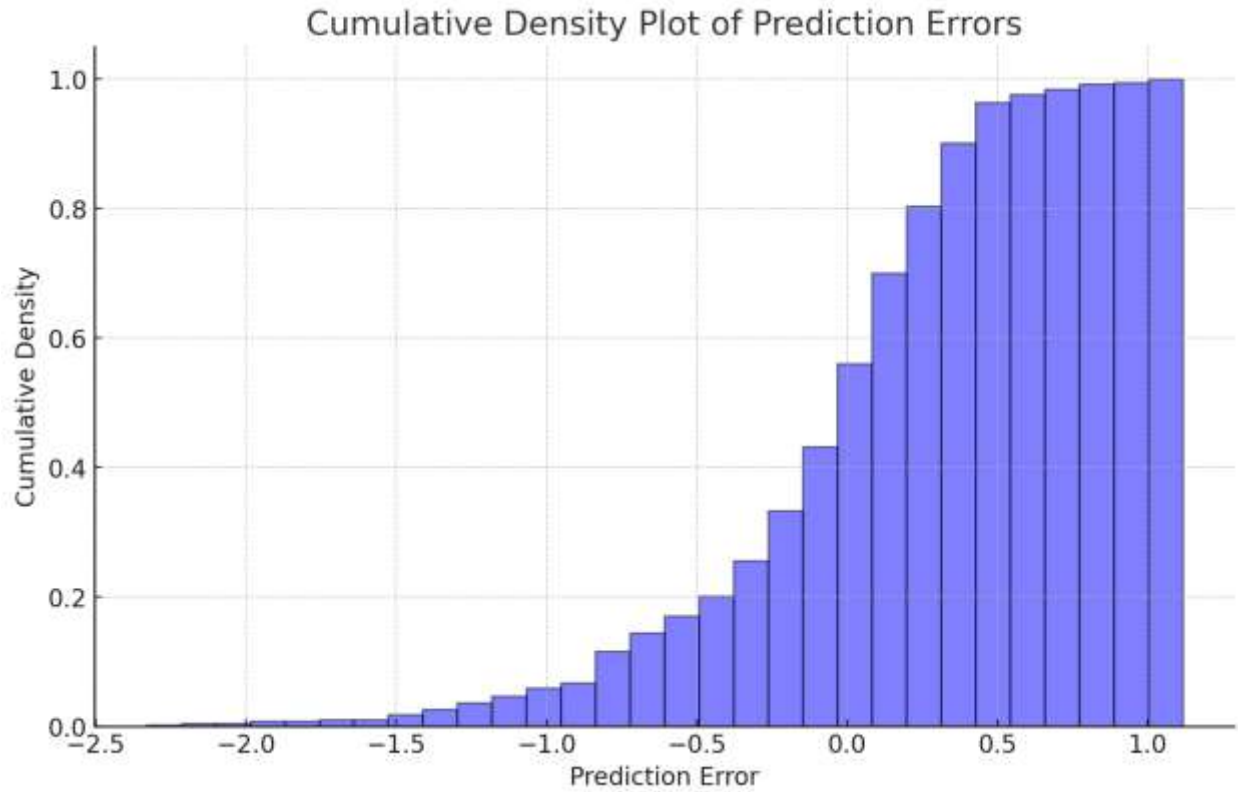
نمودار پراکندگی

نمودار پراکندگی بالا مقادیر واقعی را در مقابل مقادیر پیش‌بینی شده نشان می‌دهد. خط قرمز خط ایده‌آل است که نشان می‌دهد پیش‌بینی‌ها دقیقاً با مقادیر واقعی منطبق هستند. هرچه نقاط به این خط نزدیک‌تر باشند، مدل دقت بیشتری دارد.

نمودار تراکم تجمعی (Cumulative Density Plot)

برای بررسی توزیع خطاهای پیش‌بینی، می‌توانیم نمودار تراکم تجمعی (Cumulative Density Plot) را رسم کنیم. این نمودار نشان می‌دهد که چه درصدی از خطاها در محدوده‌های مختلف قرار دارند.

ابتدا خطاها را محاسبه می‌کنیم و سپس نمودار تراکم تجمعی را رسم می‌کنیم.



نمودار تراکم تجمعی

نمودار تراکم تجمعی بالا توزیع خطاهای پیش‌بینی را نشان می‌دهد. این نمودار به ما کمک می‌کند تا بفهمیم چه درصدی از خطاها در محدوده‌های مختلف قرار دارند. به عنوان مثال، اگر بیشتر نقاط نزدیک به ابتدای نمودار تراکم تجمعی قرار داشته باشند، این نشان می‌دهد که اکثر خطاها کوچک هستند و مدل به خوبی پیش‌بینی کرده است.

نتیجه‌گیری

- نمودار پراکندگی نشان داد که مدل در بسیاری از نقاط به مقادیر واقعی نزدیک است، هرچند که در برخی نقاط اختلاف‌هایی مشاهده می‌شود.
- نمودار تراکم تجمعی نشان داد که توزیع خطاهای پیش‌بینی بیشتر در محدوده‌های کوچک قرار دارد، که نشان‌دهنده دقت نسبی خوب مدل است.

این تحلیل‌ها نشان می‌دهند که مدل SVM توانسته است با دقت قابل قبولی تابش خورشیدی را پیش‌بینی کند. برای بهبود دقت پیش‌بینی‌ها می‌توان از تکنیک‌های مختلفی مانند تنظیم پارامترها و استفاده از داده‌های بیشتر بهره

برد