FAST NUCES

DATA MINING

FORECASTING SYSTEM

Submitted by:

14TH MAY,2024

SHARJEEL NADIR 21I-1707
HAMAD SIKANDER 21I-2699

Contribution

HAMAD SIKANDER

- Dataset Selection
- Preprocessing
- SARIMA
- ARIMA
- ETS
- SVR
- PROPHET
- Pickle file creation
- Front-end development
- Report

SHARJEEL NADIR

- ANN
- LSTM
- HYBRID
- Pickle file creation
- Integration of models to front end
- Back end development (SQL)
- Report
- Deployment on GIT

ANN

Artificial neural network (ANN) is a computational model that mimics the structure and function of the human brain's neural networks. It comprises interconnected nodes arranged in layers, with each node processing and transmitting information. Through training on labeled data, ANNs can learn complex patterns and relationships, enabling tasks like classification, regression, and pattern recognition in various domains.

Test MAE: 0.8204792737960815
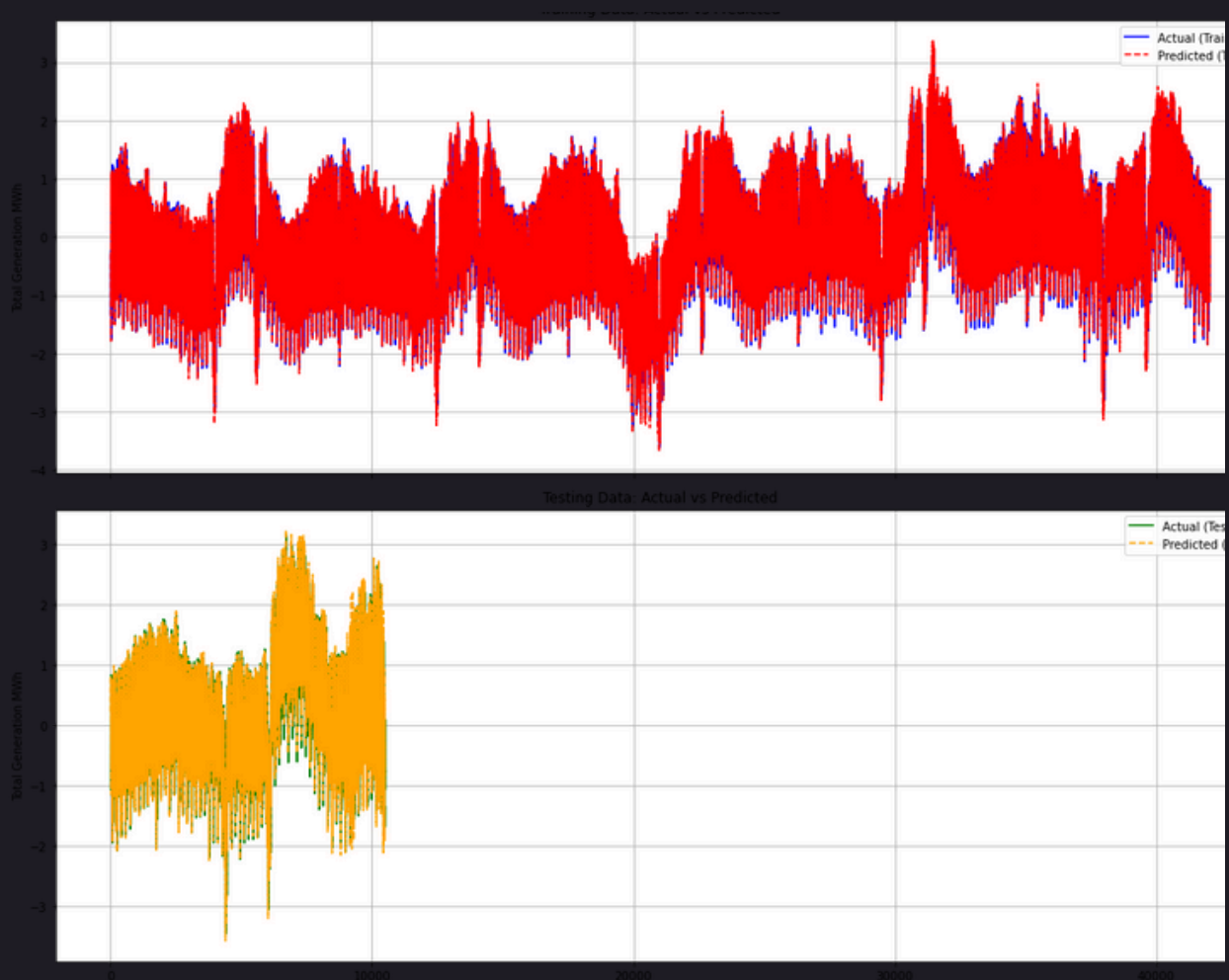
ARIMA

ARIMA (Autoregressive Integrated Moving Average) is a statistical method used for time series forecasting. It models the relationship between a variable and its past values, incorporating elements of autoregression, differencing, and moving averages.

ADF Statistic (Differenced Data): -37.11792891774274
p-value (Differenced Data): 0.0
ADF Statistic (Log-Transformed Data): -14.8385333145051
p-value (Log-Transformed Data): 1.852332501697893e-27
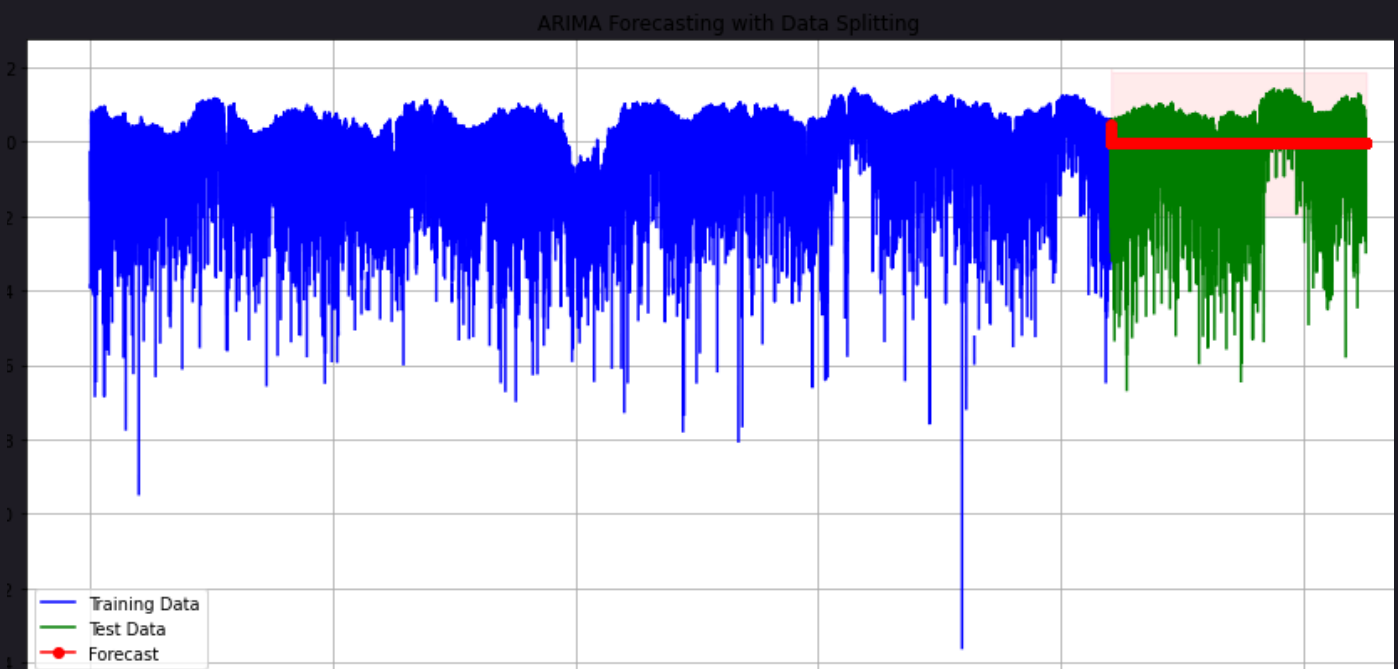
ETS

ETS (Error, Trend, Seasonality) is a time series forecasting method that decomposes a series into three components: error, trend, and seasonality. It uses exponential smoothing techniques to model each component separately, allowing for more accurate predictions by capturing the underlying patterns in the data.
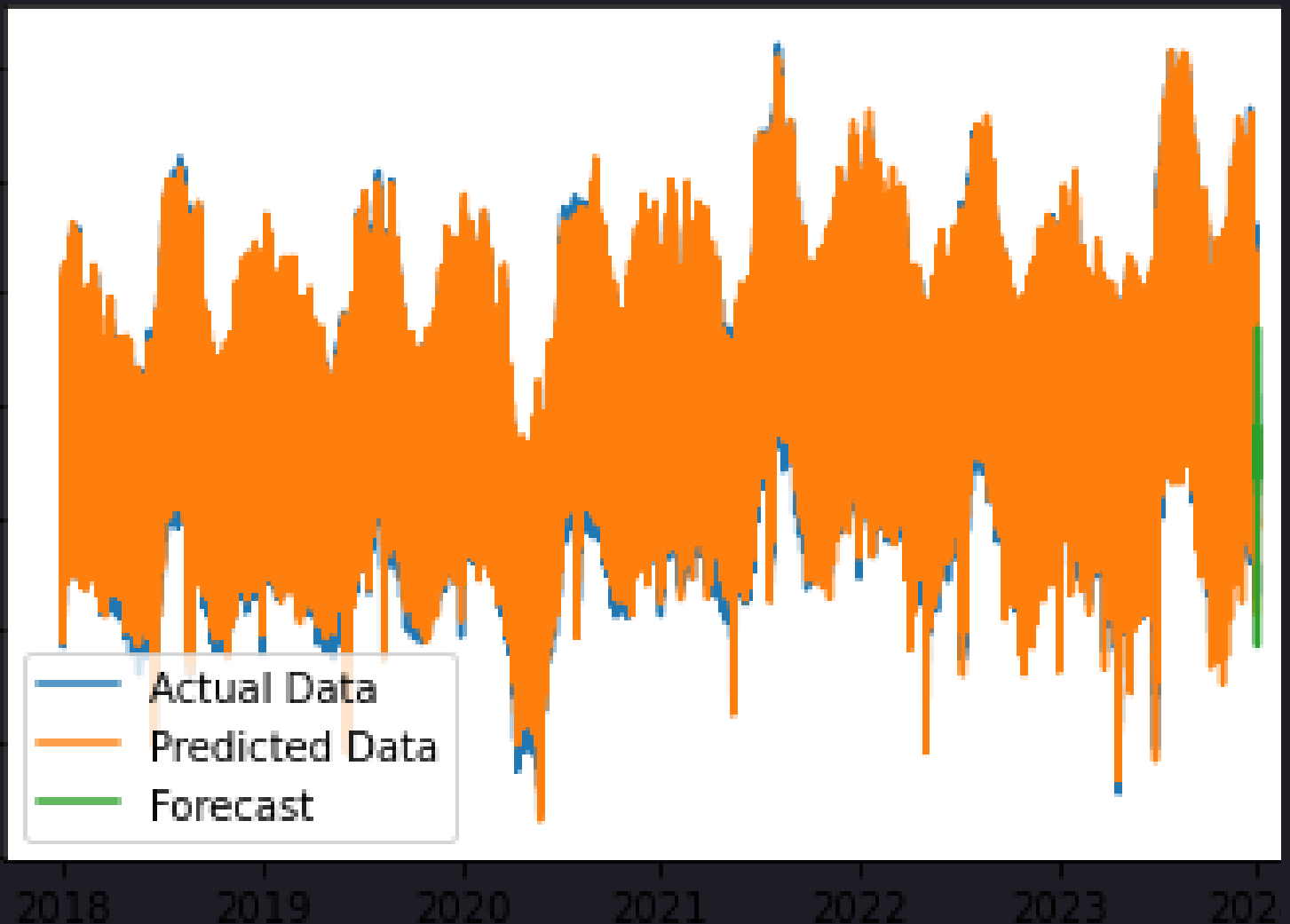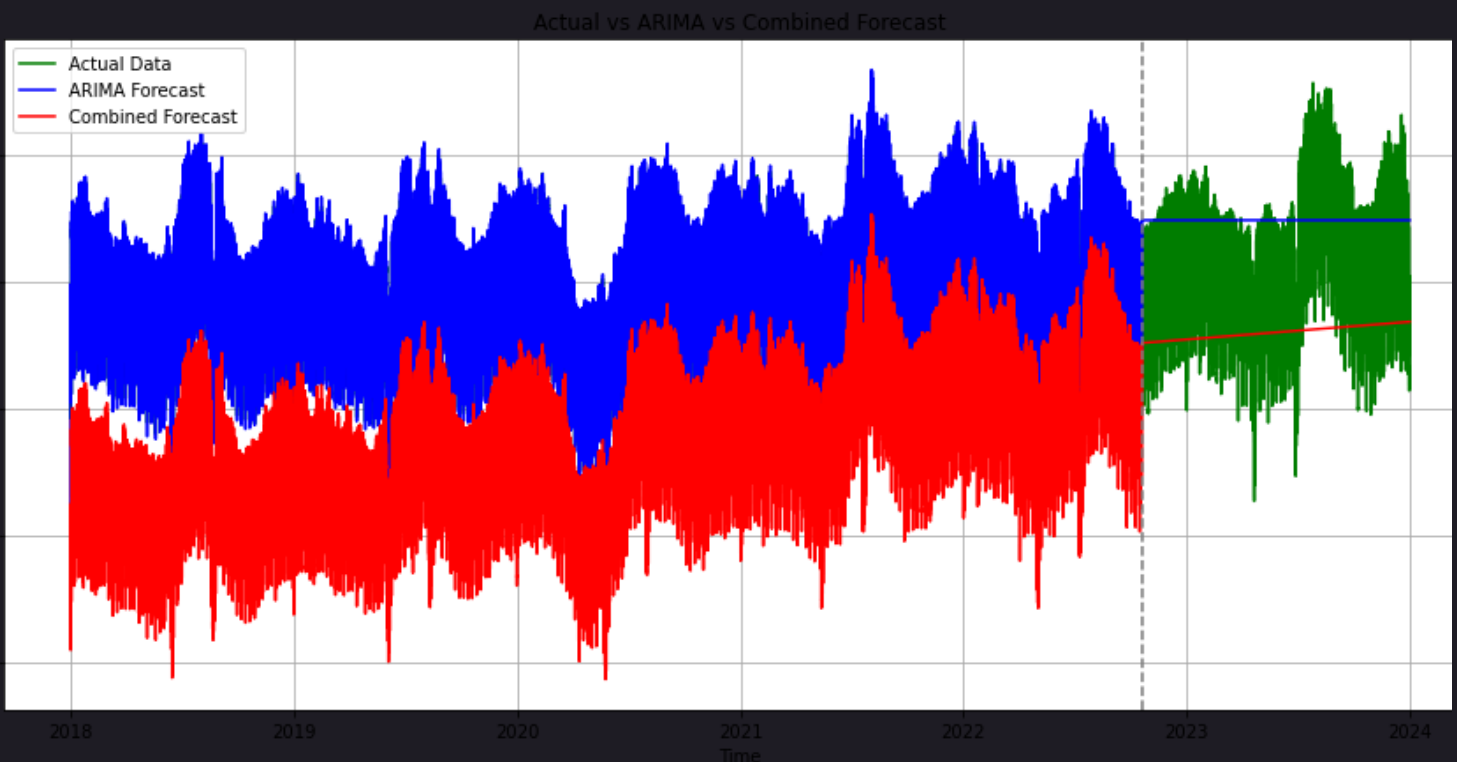
HYBRID

Hybrid forecasting models combine the strengths of different forecasting methods, such as statistical approaches like ARIMA or ETS with machine learning techniques like neural networks or decision trees.

Train RMSE: 2.6411
Test RMSE: 1.3624

Actual vs ARIMA vs Combined Forecast

LTSM

LSTM (Long Short-Term Memory) is a type of recurrent neural network (RNN) architecture designed to overcome the vanishing gradient problem in traditional RNNs, allowing for the capture of long-term dependencies in sequential data.

Test Loss: 0.8831592798233032
Test MAE: 0.7474406957626343

Actual vs Predicted vs Forecasted
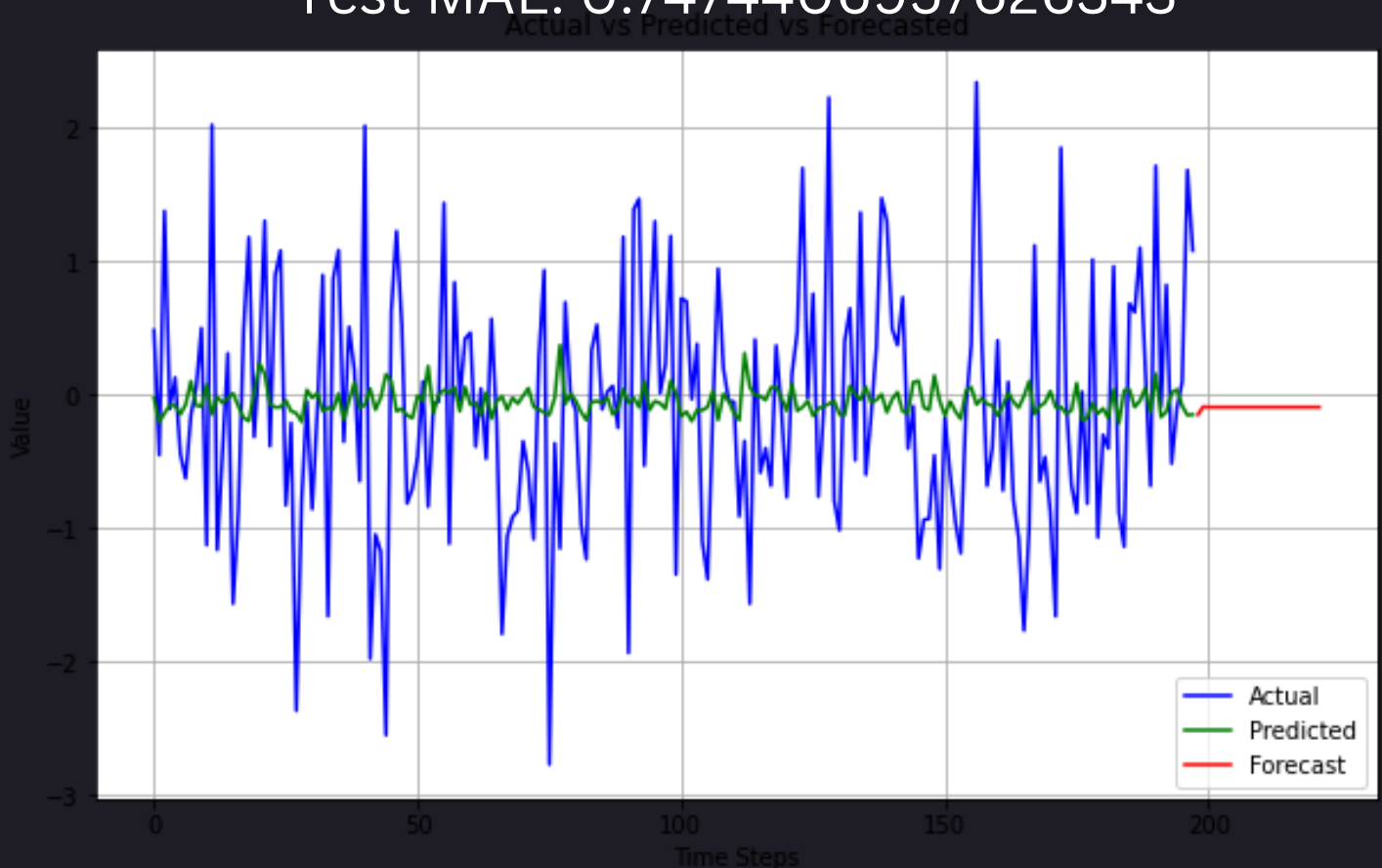
SARIMA

SARIMA (Seasonal Autoregressive Integrated Moving Average) is an extension of the ARIMA model that incorporates seasonality. It's designed to handle time series data with both trend and seasonal components by integrating seasonal differencing into the modeling process.

Enhanced SARIMA Forecasting Visualization
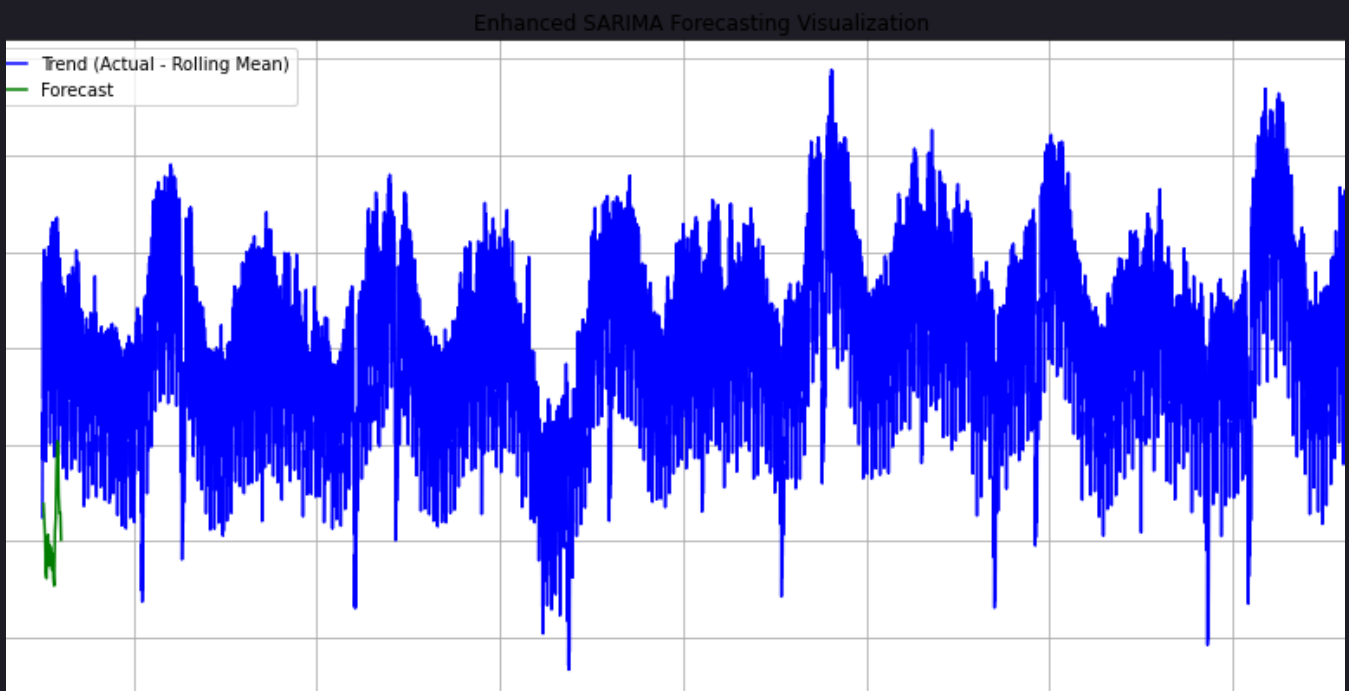
SVR

SVR is particularly effective in high-dimensional spaces and for cases where the data doesn't necessarily follow a linear pattern. It utilizes a kernel trick to transform the input space into a higher-dimensional feature space, allowing for more flexible modeling of nonlinear relationships.

SVR Test MSE: 0.003422161661193253
SVR Test MAE: 0.0488071974973757