

**CE 5319 Machine Learning for Civil Engineers**  
**Exercise Set 9**  
**Support Vector Regression (SVR) and Rating Curves**

## Purpose(s)

This exercise, you will:

- Fit and visualize Support Vector Regression (SVR) models using `scikit-learn`
- Compare SVR to linear regression on nonlinear hydrologic datasets
- Analyze the impact of hyperparameter  $\epsilon$  on prediction smoothness and flexibility
- Relate machine learning tools to hydrologic modeling practice

## Provided Starter Code (Do Not Edit)

```
# Simulate a nonlinear rating curve (Stage - Discharge)
import numpy as np
import matplotlib.pyplot as plt
from sklearn.preprocessing import StandardScaler

# Seed and simulate
np.random.seed(0)
stage = np.linspace(1.0, 3.5, 30).reshape(-1, 1)
discharge = 30 * (stage.flatten() - 0.8)**2 +
            np.random.normal(0, 5, size=stage.shape[0])

# Plot the raw data
plt.scatter(stage, discharge, color='gray')
plt.xlabel("Stage (ft)")
plt.ylabel("Discharge (cfs)")
plt.title("Simulated Stage-Discharge Data")
plt.grid(True)
plt.tight_layout()
plt.show()
```

**Exercise(s)**

1. SVR Model Fit and Prediction (40 pts)
  - (a) Fit an SVR model to the scaled data using:
    - `kernel='rbf'`
    - `C=100`
    - `epsilon=0.1`
  - (b) Invert the scaling to retrieve predicted discharge in original units.
  - (c) Plot observed vs predicted discharge values.
  - (d) Label the axes and add a legend to your plot.
2. Comparison with Linear Regression (30 pts)
  - (a) Fit a linear regression model to the same scaled data.
  - (b) Plot both SVR and linear regression predictions against observed values.
  - (c) Compute and report the **RMSE** of both models.
  - (d) Interpret the result: Which model fits better? Why?
3. Exploring  $\epsilon$  Sensitivity (20 pts)
  - (a) Fit SVR models using three different `epsilon` values: 0.01, 0.1, and 0.5.
  - (b) Plot the predictions from all three models on the same graph.
  - (c) Include observed data in the plot for reference.
  - (d) Discuss how changing  $\epsilon$  affects the fit and flexibility of the model.
4. Reflection and Application (10 pts) Write a short paragraph (4–5 sentences) answering the following:
  - How might SVR be useful for estimating rating curves in real hydrologic applications?
  - What limitations would you consider before using SVR operationally?