

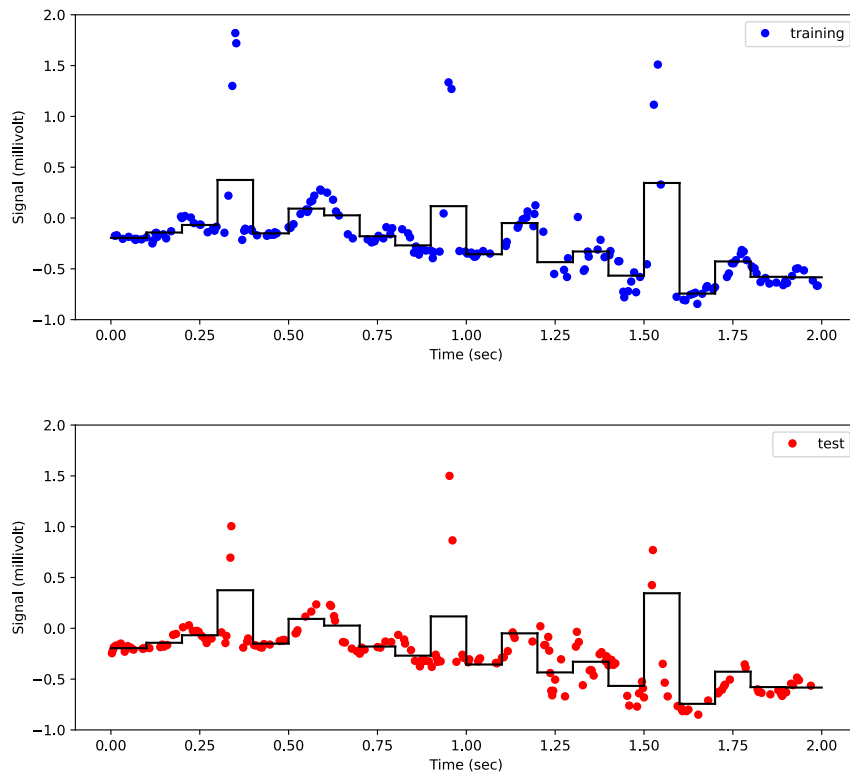
ENGR 421 / DASC 521: Introduction to Machine Learning

Homework 04: Nonparametric Regression

Deadline: April 8, 2022, 11:59 PM

In this homework, you will implement three nonparametric regression algorithms in Python. Here are the steps you need to follow:

1. Read Section 8.8 from the textbook.
2. You are given a univariate regression data set, which contains 180 training data points in the file named `hw04_data_set_train.csv` and 180 test data points in the file names `hw04_data_set_test.csv`.
3. Learn a regressogram by setting the bin width parameter to 0.1 and the origin parameter to 0.0. Draw training data points/test data points and your regressogram in the same figures. Your figures should be like the following figures. (20 points)



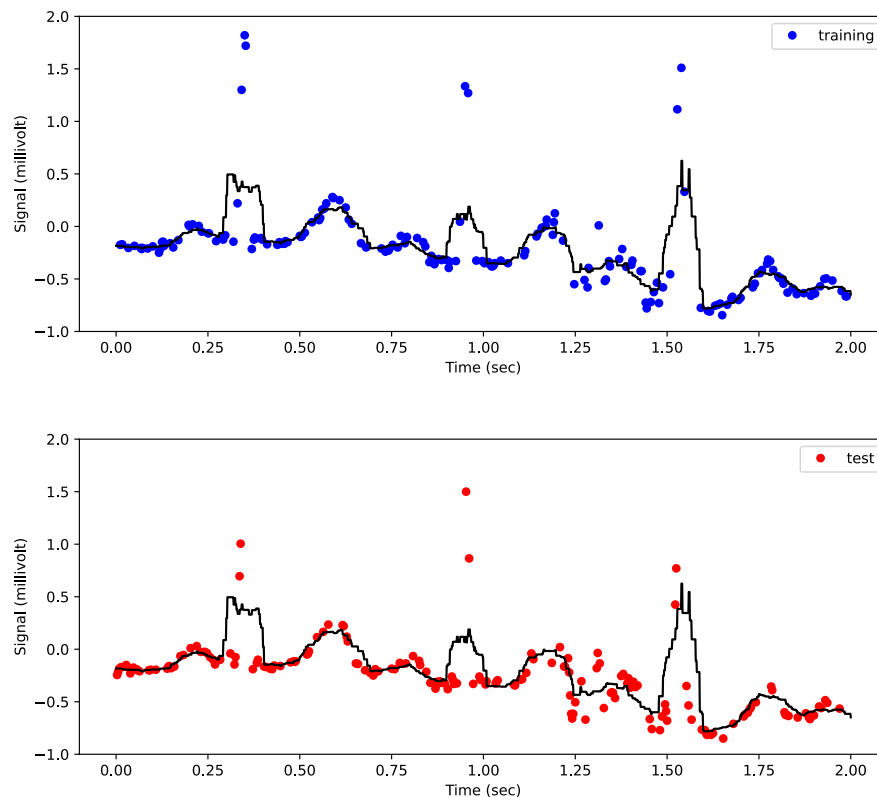
4. Calculate the root mean squared error (RMSE) of your regressogram for test data points. The formula for RMSE can be written as

$$\sqrt{\frac{\sum_{i=1}^{N_{test}} (y_i - \hat{y}_i)^2}{N_{test}}}.$$

Your output should be like the following sentence. (10 points)

```
Regressogram => RMSE is 0.25643022943721944 when h is 0.1
```

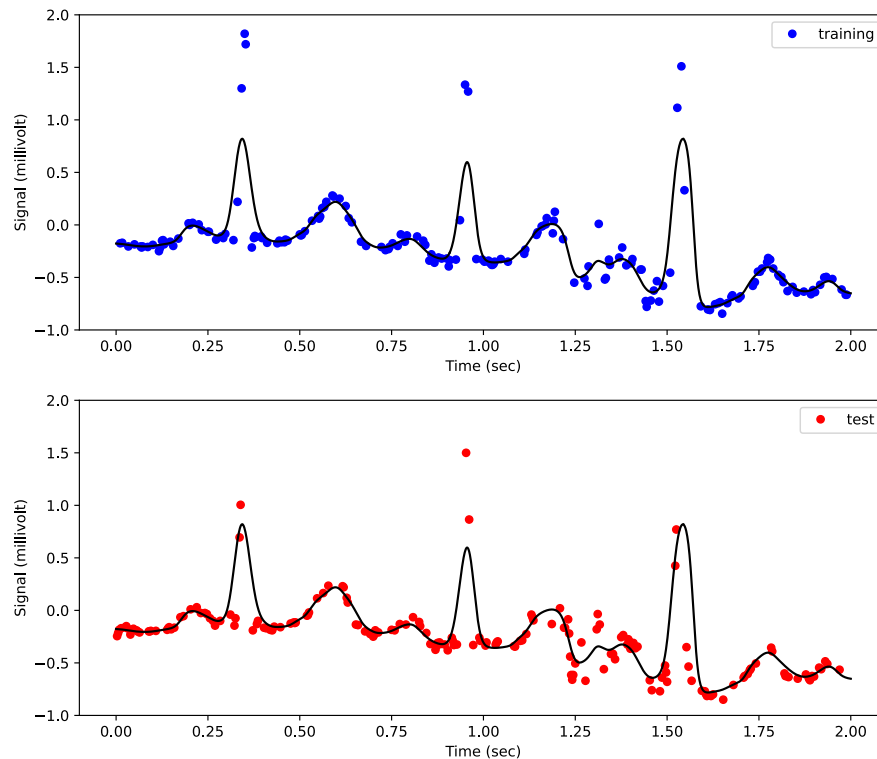
5. Learn a running mean smoother by setting the bin width parameter to 0.1. Draw training data points/test data points and your running mean smoother in the same figures. Your figures should be like the following figures. (20 points)



6. Calculate the RMSE of your running mean smoother for test data points. Your output should be like the following sentence. (10 points)

```
Running Mean Smoother => RMSE is 0.24178102764267204 when h is 0.1
```

7. Learn a kernel smoother by setting the bin width parameter to 0.02. Draw training data points/test data points and your kernel smoother in the same figures. Your figures should be like the following figure. (30 points)



8. Calculate the RMSE of your kernel smoother for test data points. Your output should be like the following sentence. (10 points)

```
Kernel Smoother => RMSE is 0.1986411340867904 when h is 0.02
```

What to submit: You need to submit your source code in a single file (.py file) named as *STUDENTID.py*, where *STUDENTID* should be replaced with your 7-digit student number.

How to submit: Submit the file you created to Blackboard. Please follow the exact style mentioned and do not send a file named as *STUDENTID.py*. Submissions that do not follow these guidelines will not be graded.

Late submission policy: Late submissions will not be graded.

Cheating policy: Very similar submissions will not be graded.