CS454-554 Homework 2: Nonparametric Regression

You are going to use the same test dataset and the ten training datasets that you previously used in the first homework.

Use the kernel smoother

$$\hat{g}(x) = \frac{\sum_{t=1}^{N} K\left(\frac{x - x^t}{h}\right) r^t}{\sum_{t=1}^{N} K\left(\frac{x - x^t}{h}\right)}$$

to find your estimate at x. Use the Gaussian kernel as your kernel function $K()$.

1) Try ***h*** values of 0.05, 0.1, 0.25, 0.5, 1, and 5. For each h value, plot your estimate for all test instances and calculate the mean square error. You need to do this 10 times for the 10 training samples.
2) Plot the average (over the 10 samples) test mean square error vs h in a separate plot.

This homework is due **April 11th (Monday), 23:00**.
Your submission should include a short report of your findings, the plots, and your source code.
Upload your submission as **one pdf file** to LMS.