Submit the report as hard-copy before the lecture on Wednesday. Add the source code (with comments) to end of your report.

1. (4pt) Review bias/variance dilemma. Replicate the bias-variance-plot (bias, variance and error vs. model complexity) given in Figure 4.6 in the textbook or in lecture notes [W04.pdf, pg.21]. For this:
   - Generate 20 values $\{x^t\}$ from the range $[0, 5]$ from uniform distribution.
   - Generate 100 different samples X_s. Each sample contains 20 instances $\{x^t, y_s^t\}$ where $y_s^t = f(x^t) + \epsilon$, $f(x) = 2 \times \sin(1.5 \times x)$, and $\epsilon \sim \mathcal{N}(0, 1)$.
   - For each sample, fit polynomial models of order 1,2,3,4, and 5.
   - Plot bias, variance and error of these models.
2. (3pt) Review cross-validation. Replicate the cross-validation error plot (training and validation error vs. model complexity plot) given in Figure 4.7 in the textbook or in lecture notes [W04.pdf, pg22]. For this:
   - Use the same 100 samples generated above.
   - Split each sample to training and validation sets, fit polynomial models of order 1,2,3,4, and 5.
   - Plot mean training and validation error (mean square error) of each model.
3. (1pt) Use of real dataset. Download Iris Data Set from UCI Machine Learning Repository. Randomly split it into training and test sets.
   - Consider only two classes: Iris Setosa and Iris Versicolour.
   - Assuming Gaussian distribution, apply parametric classification, and find which particular feature (sepal length, sepal width, petal length, petal width) is most successful in classifying instances.
   - For this, treat each feature separately
     - find maximum likelihood estimate of the parameters (μ, σ^2) of different classes using training set,
     - using these parameters, predict the classes of the instances in the test set, and calculate the error
4. (1pt) First submission bonus.