Assignment: Week 5
Due date: 17 Feb 2024 (Before Class)

1. Questions in Lec5_LogReg.ipynb:
    a. Implement stochastic and mini-batch gradient descent for logistic regression in the notebook. Compare precision, recall, and F1-score keeping all the other parameters unchanged.
    b. Take more than one feature (sepal_length, sepal_width, petal_length, petal_width) and try the above.
2. Questions in Lec5_Naive_Bayes.ipynb:
    a. Modify the code to include more than one feature and check the results.
    b. Use one of the kernel density estimation (KDE) methods instead of a normal distribution to model the various distributions of the Naive-Bayes classifier. Link to sklearn KDE.
3. The *k*-fold cross-validation (KFCV) method, discussed in the previous class was employed for *model selection.* To elaborate, we obtained the optimal "hyper-parameters" by judicious use of the training data. This method has another crucial use: *model assessment*, i.e., estimation of **test error** (e.g. RMSE) again by just using the training data!

    Consider the polynomial regression problem (5-HO-3-polynomial_regression_sklearn.ipynb). For a fixed regularization parameter (say alpha=1e-8), plot the mean RMSE from KFCV ($RMSE_{KFCV}$) and the test error ($RMSE_{test}$) as a function of polynomial degree (n). Record your observations. Reference: Chapter 5 of Introduction to Statistical Learning

4. Leave-one-out cross-validation (LOOCV) is a special case of k-fold validation when k=n, where *n* is the number of samples in the training data.
    The LOOCV error for linear/polynomial regressions reduces to the following expression:

    $CV_n = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - \widehat{y_i}}{1-h_i}\right)^2$, where $h_i = \frac{1}{n} + \frac{\left(x_i - \bar{x}\right)^2}{\sum_{j=1}^{n}\left(x_j - \bar{x}\right)^2}$, called the leverage of x_i and other symbols have their usual meanings.
    a. What is the range of h_i?
    b. Implement a function to calculate this error that accepts x, y, and yhat as arguments.
    c. Apply it in deciding the order of the polynomial (degree) and regularization constant (alpha) for polynomial regression of the dataset in the notebook.

    Reference: Chapter 5 of Introduction to Statistical Learning