Homework 4:
Variational Inference

Instructions: You may discuss this assignment with other students in the class, but you must submit your own answers to the questions below. Include an honor pledge with your submission. Submit on-line and in pdf, providing a pdf of your Python notebook. This homework is worth 100 points and the point totals for each question are shown in parentheses.

1. (33) Use the EM implementation for Gaussian mixture models (GMM) in sklearn to segment the image Haiti_Image.tif into four classes. Then answer the following questions:

   - Segment this image using K-means using 4 classes by selecting appropriate parameters in the GMM sklearn function. Compare the classification results obtained with the GMM to those obtained from K-means (KM) segmentation of the same image. Show the results of the two approaches to segmentation as images with the separate classes shown in different colors. Comment on the similarities and differences in the two segmentations.
   - Write the equation for the posterior prediction of the probability of the class label for each data point in the image.
   - Show four separate images, one for each class, that color code the probabilities of a class label for each data point in the image. Comment on the insights provided by this probabilistic view of the segmentation. Very briefly describe example of how this probabilistic perspective might be used by governmental or non-profit agencies concerned with improving the environment or the living conditions in Haiti.

2. (50) Modify the Neural Network Variational Inference notebook (https://www.kaggle.com/billbasener/pyr variation-inference-neural-network) by finding a different nonlinear 2-dimensional data and training the network and analyzing the results provided for the make moons data. You can use data provided in sklearn.datasets or any 2D data you can find or create. You can modify the network (number of neurons or layers) as needed.

   - Include and discuss the probability results for your data.
   - Include and discuss the uncertainty results derived from the standard deviation in the posterior predictive.
   - include and discuss the plot of the ELBOW, and what that tells about convergence to maximizing the likelihood.
   - Include and discuss the traceplot, focusing on what this tells you about your network.

References

[1] Theodoridas, Sergios *Machine Learning: A Bayesian and Optimization-Perspective*, Elsevier, 2015.