CS/ECE/ME532 Assignment 9

1. **Face Emotion Classification with a three layer neural network**. In this problem we return to the face emotion data studied previously. You may find it very helpful to use code from an activity (or libraries such as Keras and Tensorflow).

   a) Build a classifier using a full connected three layer neural network with logistic activation functions. Your network should

   - take a vector $\boldsymbol{x} \in \mathbb{R}^{10}$ as input (nine features plus a constant offset),
   - have a single, fully connected hidden layer with 32 neurons
   - output a scalar $\widehat{y}$.

   Note that since the logistic activation function is always positive, your decision should be as follows: $\widehat{y} > 0.5$ corresponds to a 'happy' face, while $\widehat{y} \leq 0.5$ is not happy.

   b) Train your classifier using stochastic gradient descent (start with a step size of $\alpha = 0.05$) and create a plot with the number of epochs on the horizontal axis, and training accuracy on the vertical axis. Does your classifier achieve 0% training error? If so, how many epoch does it take for your classifier to achieve perfect classification on the training set?

   c) Find a more realistic estimate of the accuracy of your classifier by using 8-fold cross validation. Can you achieve perfect test accuracy?

2. **Face Emotion Classification with Kernel Classifier.** In this problem you will apply a kernel classifier to the face emotion dataset. You may find it very helpful to use code from an activity.

   a) Build a kernel classifier using

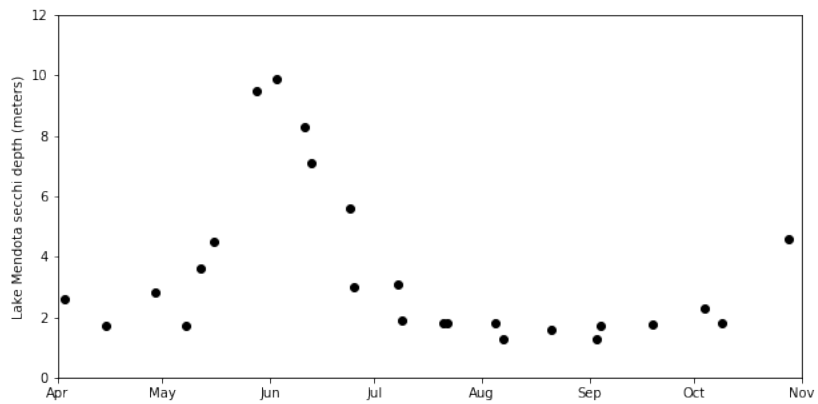
   - the squared error loss function
   - an ℓ_2 regularizer with $\lambda = 0.5$.
   - the Guassian Kernel $K(\boldsymbol{u}, \boldsymbol{v}) = \exp(-||\boldsymbol{u} - \boldsymbol{v}||^2/(2\sigma^2)$.

   b) Train your classifier choosing for different values of σ and create a plot with σ on the horizontal axis and accuracy on the vertical axis and comment on the plot. Does your classifier achieve 0% training error?

   c) Find a more realistic estimate of the accuracy of your classifier by using 8-fold cross validation. Can you achieve perfect test accuracy?

3. Kernel Regression, Lake Mendota Clarity. The *Secchi depth* is a measure of water clarity obtained by lowering a black and white disk off the shady side of a boat and recording the depth at which the disk is no longer visible.

A dataset obtained from the University of Wisconsin's Limnology department contains Secchi disk readings (in meters) on Lake Mendota from 2019 and 2020. A Secchi depth of less than 2 meters is consider poor clarity, while a Secchi depth greater than 6 meters is consider very clear. Lake Mendota can have very clear water in late spring when native zooplankton *daphnia pulicaria* consume large amounts of algae and phytoplankton (for more details, see https://blog.limnology.wisc.edu/2019/06/12/whats-behind-this-extended-phase-of-crazy-clear-water-in-lake-mendota/).



a) Use kernel ridge regression with a Gaussian kernel to fit the measurements. You may find it useful to use code from an activity. Use regularization parameter $\lambda = 0.01$ and scale parameter $\sigma = 10$. Plot the resulting fit, and comment on the results. Do these parameters overfit or underfit the data? Adjust the regularization parameter to find a visually better fit.

b) Describe how you could use k-fold cross validation to systematically find a good value of σ and λ.