Problem 1 SVM

(a) Use datasets.make_moons in sklearn package to generate a dataset with 200 samples with random noise. Visualize positive and negative instances with different colors. Split the dataset into training set and test set by 7:3.

(b) Use SVM model in sklearn to fit the training data with "RBF" kernel. Try parameter C with different values, i.e. [0.05, 0.1, 1, 2, 5, 10, 50, 100]. Visualize the decision boundary and support vectors along with the plot in part (a). What effects does parameter C have on the model?

(c) We have learned in class that kernel functions allow for learning in high-dimensional feature spaces without explicit mapping into feature space. Sklearn provides common kernels like 'linear', 'poly', 'rbf', 'sigmoid'. Train the models using different kernels and visualize decision boundaries. Evaluate the models with test accuracy. Which kernel gives the best performance on the test dataset?

Problem 2 (Artificial Neural Networks)

The Artificial Neural Network Model

(1) Consider the following two activation functions:

The sigmoid function $f_1(x) = \frac{1}{1+e^{-x}}$ and the hyperbolic tangent function $f_2(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$
Show that:

$$f_1'(x) = f_1(x)(1 - f_1(x))$$

$$f_2'(x) = 1 - f_2^2(x)$$

(2) Learn how to use the MLP (multi-layer perceptron) model in sklearn package by reading corresponding sections from sklearn documentations (https://scikit-learn.org/stable/).

(3) Design your MLP **model** and apply the MLP classifier from sklearn package to the ***digits*** data set. Find the best number of hidden nodes that achieves the best classification result.