Pattern Recognition –HW#3

About the Assignment

The main aim of the assignment is to learn SVM classification and feature extraction. Contributions of this lab are;

- Ability to analyze the separable condition of features.
- Ability to analyze the non-separable condition of features.
- Understanding idea of feature extraction in the machine learning.
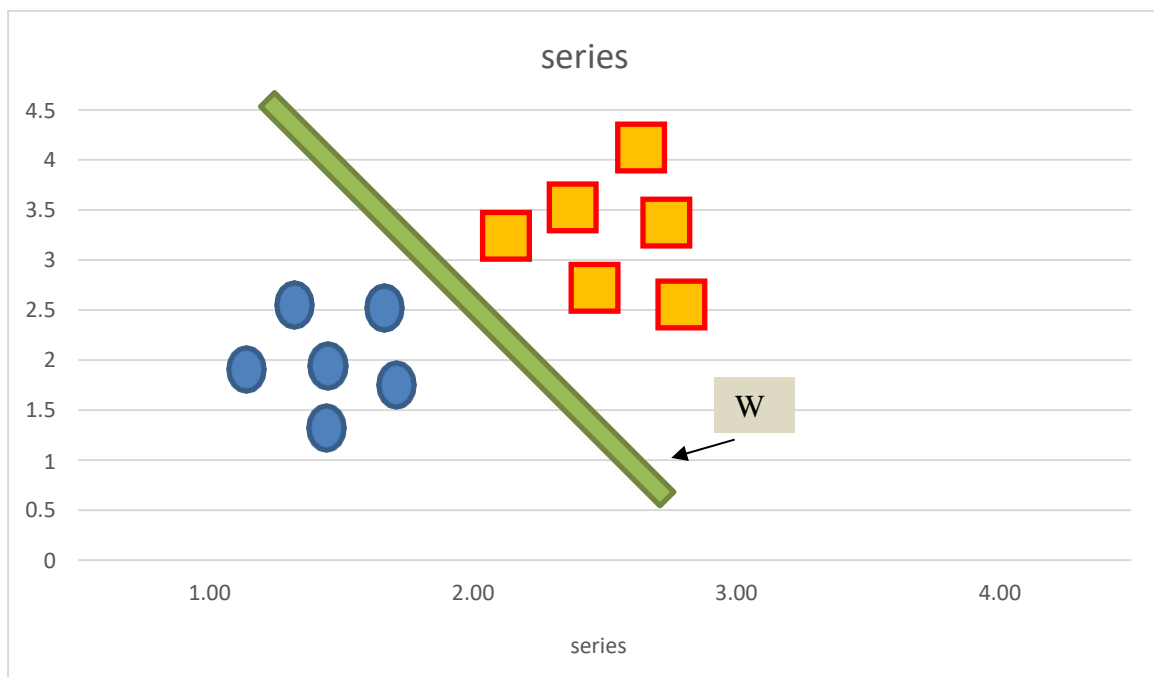
Step1:

Fig. 1: A simple visualization of SVM.

In this HW, you are expected to make an experiment with SVM classifier, explained in class. Recall the previous homework, in like manner, the experiment will be conducted on the real life problem, called a classification problem among multi categories. The experiment is about Caltech-101 datasets, there are 15 classes, and each one contains different number of samples. The aim is to use SVM classifier, one-against-all methodology, in order to find different hyperplanes that is capable to separate classes.

According to one-against-all methodology, we have to find 15 hyperplanes in case of training stage, since there are 15 classes. In case of test case, the SVM will project a given test sample (test vector) onto each hyperplane, as represented with w in the Fig. 1. After observing the obtained scores, the SVM can make decision that the best-matched class is associated with best similarity score. It means that the higher similarity score refers to predicted target class of processed sample.

Step1: Feature Extraction

Read images in RGB format. Then convert image to HSV format. An image must be in the 224x224x3 format. Instead of using the whole image data (224x224x3 size), we have to extract some meaningful features in image. In this study, we will use following feature extraction method.

- skimage.feature.multiscale_basic_features
- **from skimage.color import rgb2hsv**

- https://scikit-image.org/docs/dev/api/skimage.feature.html#skimage.feature.multiscale_basic_features

You can see that an image will be represented only with ndarray features. The dimension of feature vector per sample is 1xn. It means that training data (1457 samples) will be represented as (1457xn). Let's call the X matrix as training matrix and y is label vector, which keeps the class name of samples. The size of X matrix is (1457xn) and the size of y vector is (1457x1). You are expected to fill the X matrix with features and y vector with class label per each sample.

Step2:
You are expected to train with SVM classifier. For this purpose, you can use the following SVM class.

- https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html

Step3:
You are expected to test with SVM classifier. Create confusion matrix and show the accuracy for test samples.

Submit the Assignment
Ex: No_Name_Surname_HW#.zip