

Pattern Recognition –HW#5

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 machine learning.
- Understanding idea of feature selection in machine learning.

Step1:

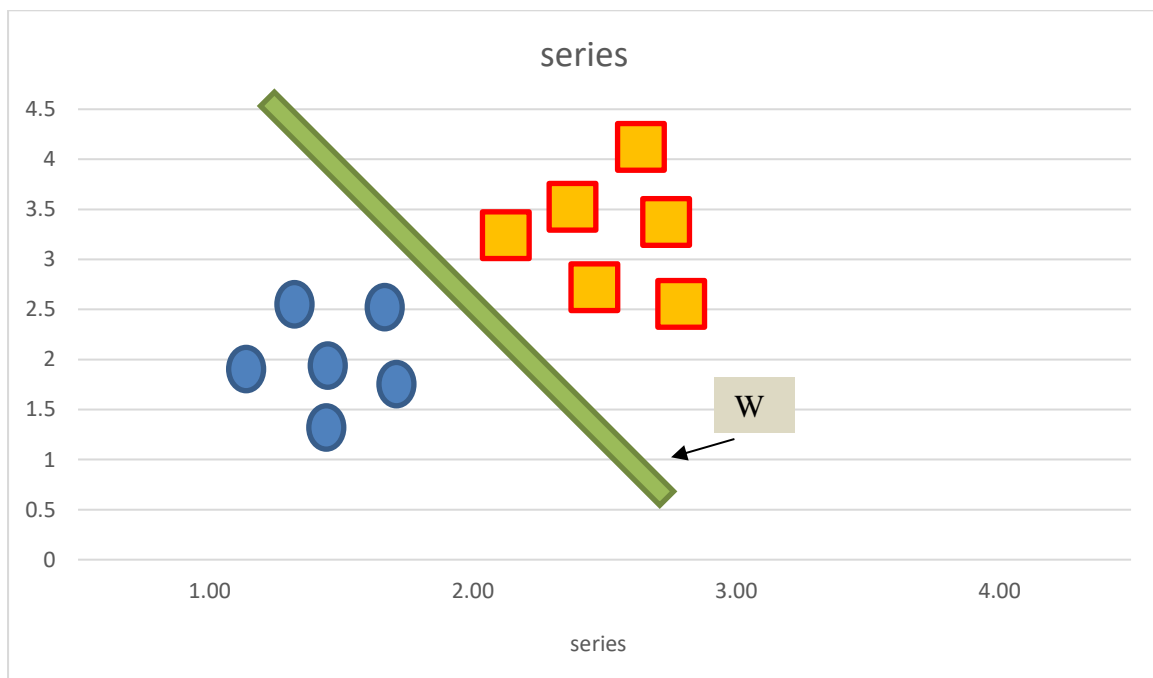


Fig. 1: A simple visualization of SVM.

In this HW, you are expected to make an experiment with SVM classifier, explained in class. Recall the previous lab, 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, we will project a given test sample (test vector) onto each hyperplane, as represented with w in the Fig. 1. After observing the obtained scores, we can make decision that the best matched class is coincided with best similarity score. It means that the higher similarity score refers to predicted target class of processed sample.

However, in this study, we can use only the any SVM code for training and predictions.

Step1: Feature Extraction

Download data from the link:

<https://drive.google.com/file/d/1brWPjht-ZkutXkPobutFhU8EtzAaMTry/view?usp=sharing>

By running the given code, you can observe that an image is in the 128x128x3 format. If it is converted to vector, then it becomes a vector with the format of 1x49152 size. Instead of using the whole image data (128x128x3 size.), we have to extract some meaningful features in image. In this study, we will extract 512 HOG features.

You can use ***sklearn*** to extract hog features

- Orientations = 8
- pixels_per_cell = 16 .
- cells_per_block = 8
- block_norm = L2

After feature extraction the training data (1457x49152) will be represented as (1457x512).

Step2: Feature Selection

The main motivation behind the feature selection is the dimension reduction. It means that we can reduce the feature size by eliminating the redundant ones. After feature selection, we will choose the more meaningful and rich features.

You can use any python library for feature selection process. Also, you can choose filter or wrapper approach. You are expected to reduce the size of feature vector from 1x512 to 1x256 format. It means that half of features will be removed after feature selection process is completed.

Don't use mlxtend library.

You are expected to fill the required fields in given code template.

You have to use the given template, namely HW#_sent.py

Submit the Assignment

Ex: No_Name_Surname_HW#.zip

Hint

You can look at the implementations available in internet.

Note in English:

You will be graded over 5 points if you send your friend's code or any snippet code available on internet. You have to consider following 3 rules in case of uploading the hw.

- 1) Send a screen shot that shows your code worked successfully without error.
- 2) Send a screen shot related to generated results in spyder
- 3) Only send **py extension file**, other file format will be graded over 5 points. You have to use **spyder** when implementing any homework.

Screen shots must be related to spyder.

If one of the rule above is ignored, then your homework will be graded only over 50 points.

Take care yourself.

Note in Turkish:

Vereceğim ödevlerde internetteki **hazır kodu** benimle paylaşmanız veya kendi **arkadaşınızın kodunu** paylaşmanız. Direk bakılmadan **5 puan** olarak değerlendirilecektir.

Ödevleri yüklerken aşağıdaki 3 şeyi dikkate alarak yükleme yapınız.

- 1) Ödevin **hatasız** çalıştığına dair ekran görüntüsü
- 2) Ödevin ürettiği **sonuçlara** ait ekran görüntüsü
- 3) Ödevi yüklerken sadece **py uzantılı** dosyayı yükleyiniz. Bunun dışındaki formatta gönderenler sadece **5 puan** alır. Ödevleri **spyderda** yazmanız gerekmektedir.

Ekran görüntüsü spydera ait olmalı.

Bu üç maddeden biri eksik olanların ödevleri 50 üzerinden değerlendirmeye alınacaktır

