Homework 3

After you finish Homework 1 and 2, you will already know the basic pipeline of large-scale multimedia analysis. In Homework 1, you use audio features for event detection while in Homework 2 you use visual features for event detection. In Homework 3, you are required to use multiple features from both audio and visual domains for event detection.

The task is a bit different this time. Take SIFT feature as an example. For each event, your training set will be all the positive examples and the first 200 negative examples from the features obtained by:
perl select_files.pl filelists/sift_spbof.list vids_MED10_dev_evl.id > YOUR OWN WORKING DIRECTORY/MED10.sift_spbof.list
To get the training set, you follow a procedure similar to Step 4) in Stage 1 of Homework 2. The difference is that you select all the positive examples and the first 200 negative examples as your development (training) set.
Note that you are supposed to use several provided features to build your classification model this time. You can even add any other features that you think are helpful for improving the accuracy. You may choose early fusion or late fusion or any other fusion strategies you want.

After you have the classification model, apply it on the evaluation (testing) set and return the predictions. We will give you the evaluation after you finish Homework 1 and 2.

Your score for this homework will be given based on your AP on the evaluation set in comparison with your classmates. The higher AP you get, the higher score you have for this homework.