HW 1 Report

Younkyung Jwa

Task 2 : SVM & MLP Classifier

1. With the default settings, which classifier performs better? Why?

   MLP is better than SVM. In my case, the Kaggle test set result is 0.32085 for SVM and 0.34224 for MLP. SVM works well when the number of labels is 2, but it doesn't work well for more than that. Therefore, MLP scored better in Task 2 problem with more labels than two

2. Which default parameter do you want to tune? Why? How do you change it and what's the difference?

Task 3 : More Powerful SoundNet features

1. How do you generate a single-vector representation for each video using SoundNet features?

   In the first step, I change mp4 to mp3 using ffmpeg. Next, by using pretrained SoundNet model and forward function, I extract features to .csv files.

2. You extract features from which layer? Why you choose this layer? What's the performance on the public test set?

   I picked layer 6. In Tabel 6 in the paper, conv6 has higher accuracy than conv7 and conv8. With the same MLP architecture, conv6's validation accuracy is 0.4978 and conv7's validation accuracy is 0.4578.

Task 4 : Improve your model

1. Introduce how do you improve the model. What's the performance? Why it works? If you have extra time and GPU resource, what's your next steps.
2. What's the bottleneck of the current system?