

Natural Computing, Assignment 6

Dennis Verheijden - s4455770 Pauline Lauron - s1016609
Joost Besseling - s4796799

May 9, 2018

Note

After a lot of struggling to get the source code up and running using Matlab, compiling the binaries necessary, we managed to get it to run.

However, we quickly stumbled upon a multitude of errors. We believe they were due to incompatible code, where the code was written for <MATLAB 2018a. For example, the function `svmtrain` was used, which has been deprecated.

We tried to circumvent this issue by using an older version of MATLAB, 2015b. Here the functions were not deprecated, however we did not manage to successfully compile the necessary binaries. As a last resort, we copied the source code for these deprecated functions, which ultimately worked. However, we then got presented with the error that we could not predict multi-class problems with this function. Which is weird, because the original implementation was built to handle multi-class problems.

We also had to fix some errors where the input for certain functions were incorrect, due to typing errors. Such as for the function `idivide`.

The latter could be fixed, such that we could successfully run Ad-REM. However we did not get CORAL to run, due to many deprecations.

We then tried to implement CORAL ourselves in python. We implemented it as described in the provided paper, but the results are not as expected. Where, if we did not use CORAL but instead use both raw datasets, the results were better than when we applied CORAL.

1 Using Zero Target Data

1.1 Ad-REM

VGG: 0.8931

SURF: 0.4283

This difference might be accounted to the number of features that are created for each method. The VGG creates 4096 features, most having some meaningful values (i.e. not-sparse). Whereas SURF outputs only 800 sparse integer-values.

1.2 CORAL

VGG: —

SURF: —

2 Experiment

2.1 SURF

For the experiment we randomly selected a percentage p of the labeled target-data for each class. Each experiment was repeated 10 times. The results may be seen in figure 1. What we can see from this boxplot is that using more labeled data increases the accuracy. Which is to be expected, following the “more data is better” ideology.

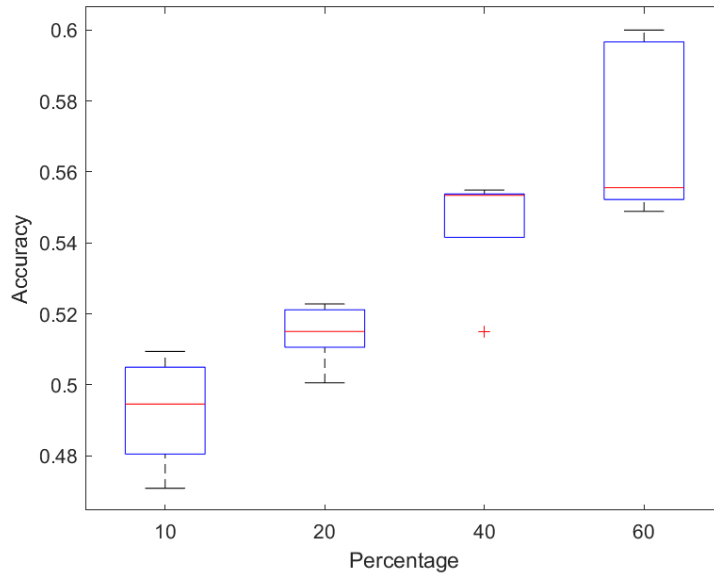


Figure 1: Boxplot of the results for Ad-REM on the SURF dataset.

2.2 VGG

We also do the same experimentation with the VGG dataset. We can conclude the same that using more labeled data increases the accuracy. Which is to be expected, following the “more data is better” ideology.

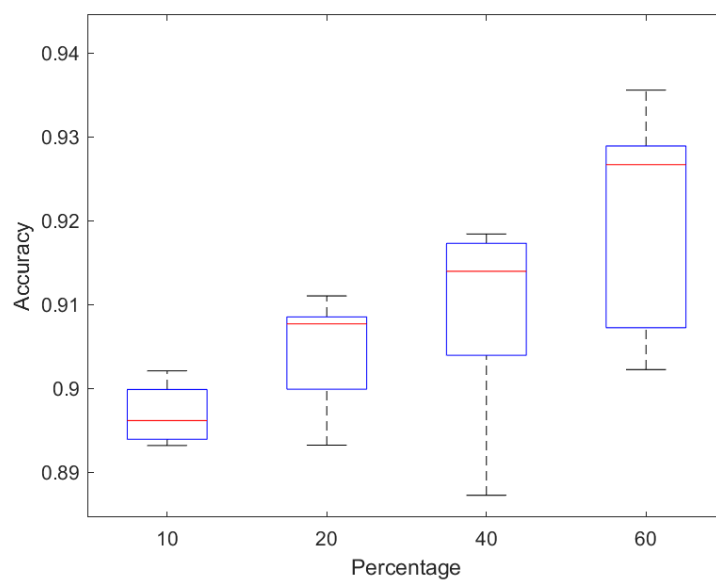


Figure 2: Boxplot of the results for Ad-REM on the VGG dataset.