April 10, 2014

1 Classifier competence as a function of image

../figs/longitude0.pdf
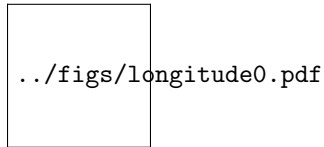
The following figure investigates the ability of a naive bayes classifier to classify workers who were primed under different treatments. Specifically, we consider the binary classification task of classifying workers who drawn from either the cultural-images or ingredients-images treatments.

We train the classifier to perform a binary classification between workers from the two treatments, providing it 80 randomly selected workers from both treatments as a training set, and using 20 workers to test. We train the classifier using only the labels that workers provided for one of the test images, and provide the classifier's performance results for when using different images as the basis of classification.

For each such test of the classifier, we look at two cases, one in which we treat each text input associated to an image in the HIT as as distinguished source of featuers (tokens), and one in which no distinction is made between the 5 text inputs associated to an image. We denote the former as a classifier test with "position-based features", and the latter with "position-less featuers". So, in a test of the classifier with position-based features, an entry of "soup" into the first text input is regarded distinct from an entry of "soup" into the second text input, whereas in a test with position-less features these two token occurences are encoded as one and the same feature.