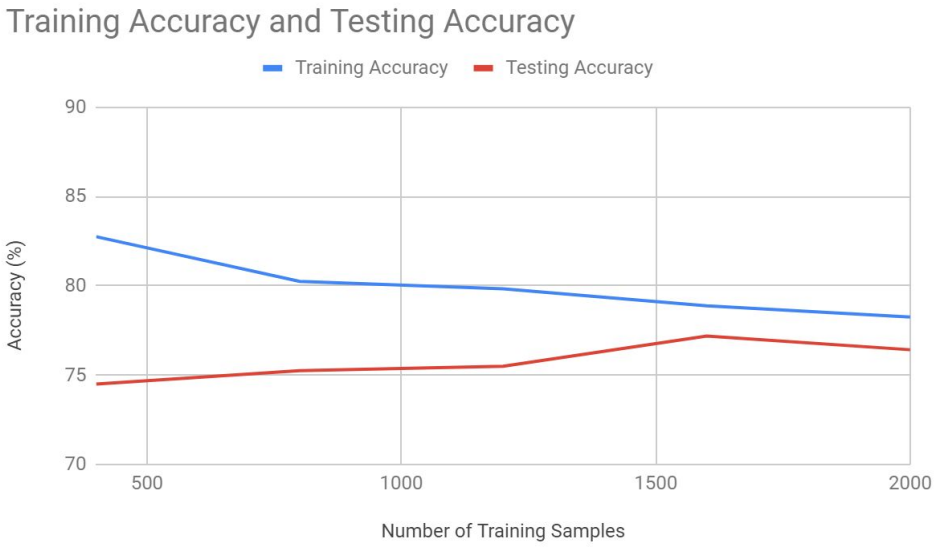
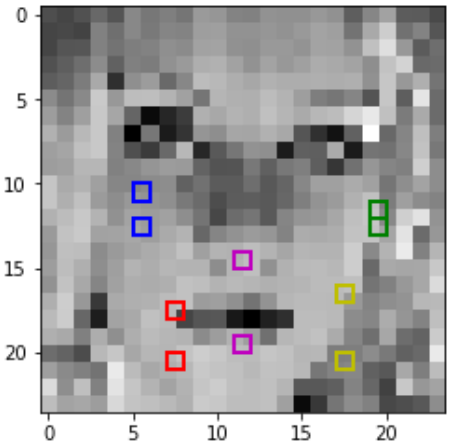
Caleb Ralphs
3/18/2019

Homework 1

n	Training Accuracy	Testing Accuracy
400	82.75	74.50
800	80.25	75.25
1200	79.83	75.50
1600	78.88	77.19
2000	78.25	76.42



The training accuracy and the testing accuracy seem to have an inverse relationship. This would appear to be the case because as we increase the number of samples in the training model, the chance of overfitting decreases. This means that as we train on a larger data set, we increase the chances of having more generalized predicting features which would, in change, reduce the amount of overfitting to the training data. As the size of the training data increases, it becomes more difficult to create features that generalize to the data. This is why we see a negative relationship between training accuracy and number of training samples. This explains why the testing accuracy increases as the number of samples that the model is trained on increases. There are more generalized features that we predict the test labels on, resulting in higher testing accuracy (i.e pixel pair permutations that generally are better at telling whether someone is smiling or not).



Note: the colors on the image are the pairs of the features chosen for the regression model. Rather than having all the r1c1 blue and r2c2 red, I think seeing the actual pairs would be more effective.