Attractiveness Prediction Results

Amanda Song, Chad Atalla

Feb 19th

1 PCA result

Everyone has 200 rating points for 200 faces. We computed every face's 29 features as listed in the Chicago face database paper and get a $200 * 29$ feature matrix. We first apply PCA to the feature matrix and plot the eigenvalues. Fig. 1 shows that the first two principal components explained 95% of the total variance, the first 6 principal components explained 99% of the variance.
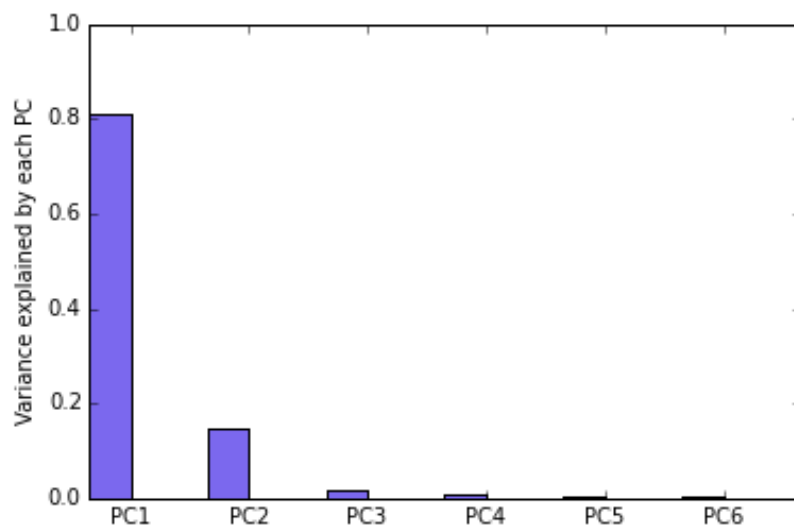

Figure 1: Variance explained by each PC component

2 Linear Regression

3 Correlation, Variance explained, Mean Squared Error

We fit a linear regression model for every subject using the first 6 principal components and calculate the correlation between predicted score and the actual ratings. We also calculate the explained variance captured by the linear model, and the mean squared error of the model. We repeated the same procedure to all 1548 raters and calculate the average correlation, variance score and mean squared error of the group. Fig. 3
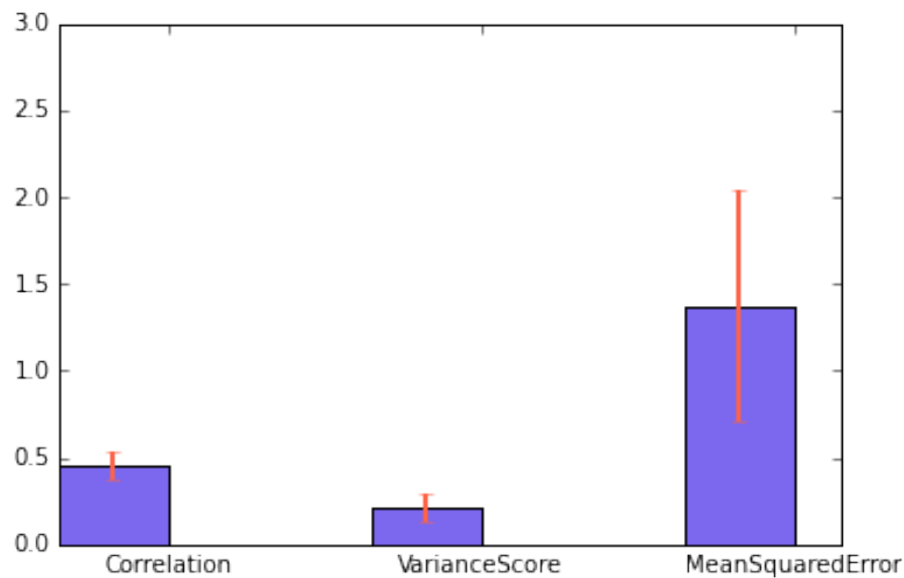

Figure 2: Average Correlation, Variance Score and MSE from PCA prediction

We then calculate the mean and the standard deviation of coefficient for each PC.

A curious phenomenon is that the coefficient for PC 4,5,6 are much higher than PC 1,2,3. It is suggesting that the feature dimensions that capture less variance is more informative in predicting attractiveness. What does that mean?
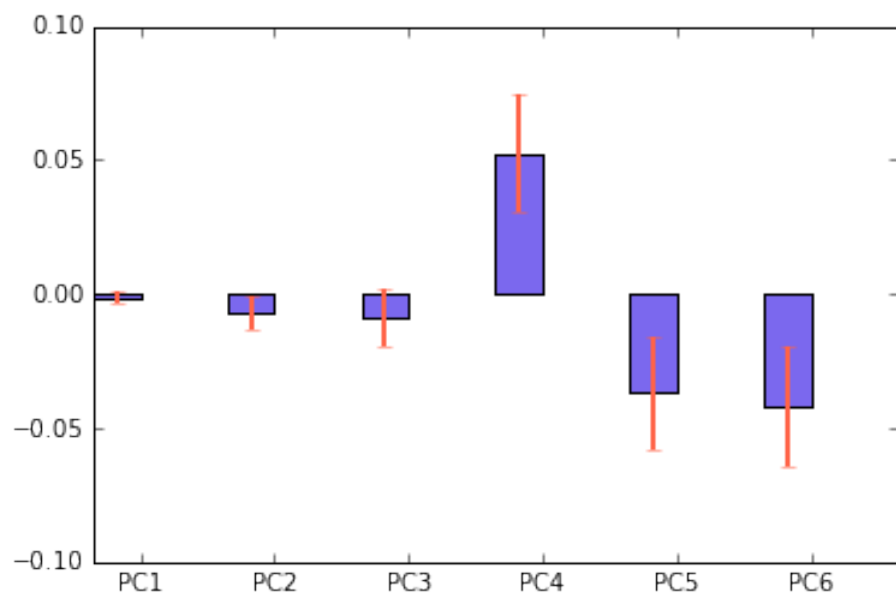
Fig 4

3.1 Project it back

4 Statistical test
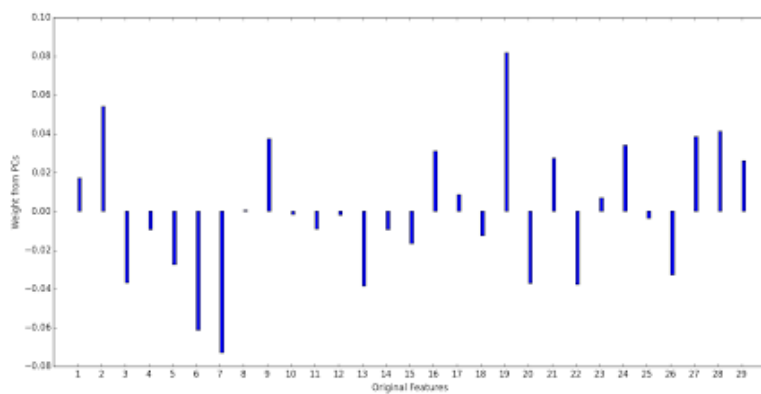
Figure 3: Mean and std for each PC's coefficient

Figure 4: Project back