Predictive Analytics: An Efficient Model for Emotion Detection in Face Images

Project 3 – Group 8 – Tim Schleicher

Presentation in the course GR5243 “Applied Data Science”

Columbia University | 10/31/2019

After an intense process of model construction and selection, we now achieve over 50% training accuracy

Award ceremony

The models have severe differences in their efficiency as well as their predictive performance

Award ceremony based on test set performance

Baseline Model
Accuracy: < 45%

- Training feature construction: 1.550s
- Testing feature construction: 0.132s
- Model training: 307.863s
- Model testing: 9.674s

Linear Discriminant Analysis on Principal Components
Accuracy: 51.2%

- Training feature construction: 0.230s
- Testing feature construction: 0.057s
- Model training: 10.953s
- Model testing: 0.004s

Convolutional Neural Net
Accuracy: < 20%

- Not enough images per class to successfully deploy a CNN
- Either use pre-trained CNN or other model

SOURCE: Own analysis on 16GB RAM MacBook

Overview of project steps

Feature Selection & Dimension Reduction

- PCA, Forward Selection and Random Forest deployed
- PCA proves best to reduce dimensions (50 top components with 60% proportion of variance used)

Feature Construction

- Deletion of right side features
- Deletion of duplicated horizontal lines
- 44 * 43 – 34 * 9 = 1586 points

Model Selection

- CNN, SVM, LDA, QDA, AdaBoost, Gradient Boosting and XGBoost
- LDA is best model in CV accuracy and efficiency

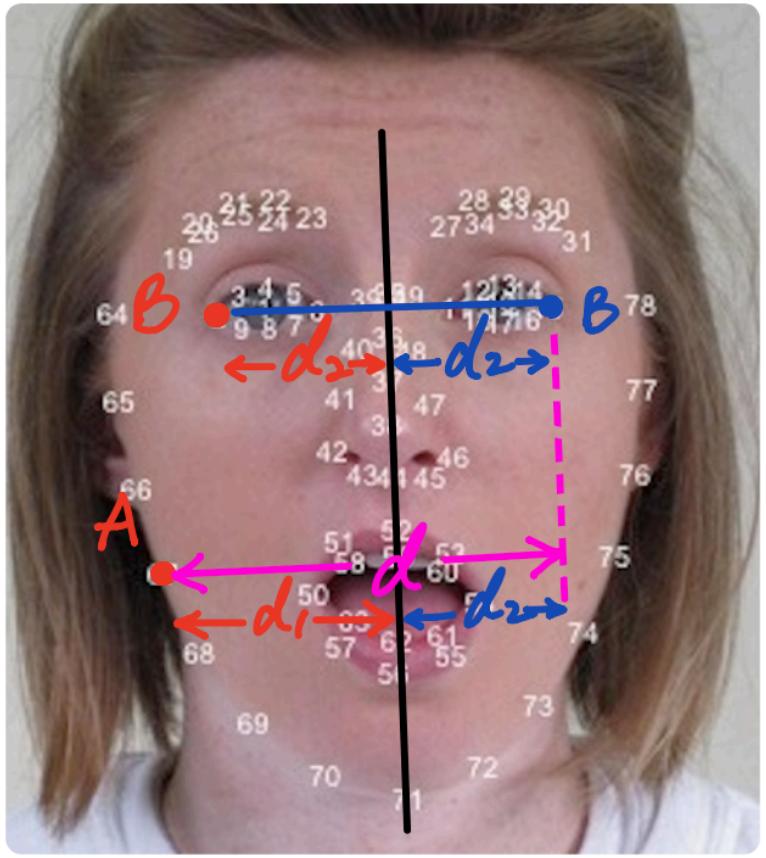
Predictions on Test Data

Model Training

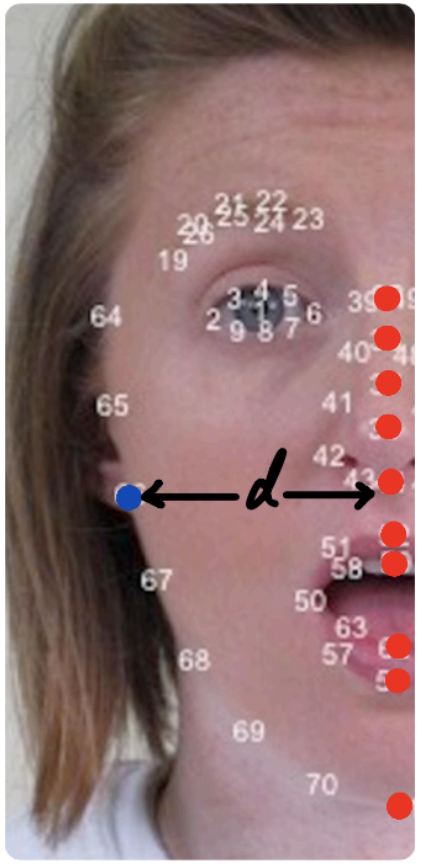
SOURCE: Own depiction

For our winning model, we first deleted irrelevant features for emotion detection

Process steps of feature construction



1. **Deletion of right side features**



2. **Deletion of duplicate horizontal lines**

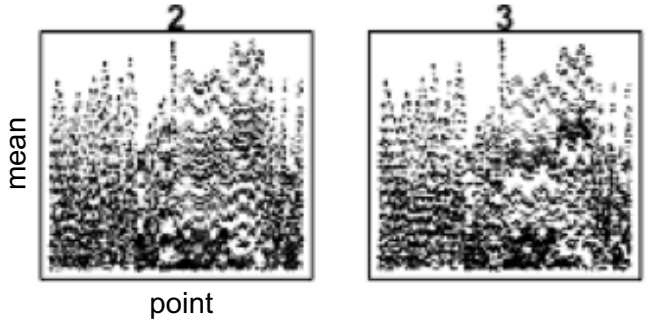
Assumption: Emotions are symmetrically represented in the face

SOURCE: Own analysis

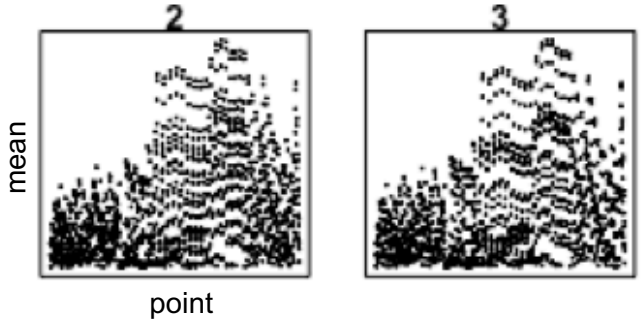
Feature selection and dimensionality reduction are essential for our winning model

Comparison between representations of happy (2) and sad (3) with different features/components

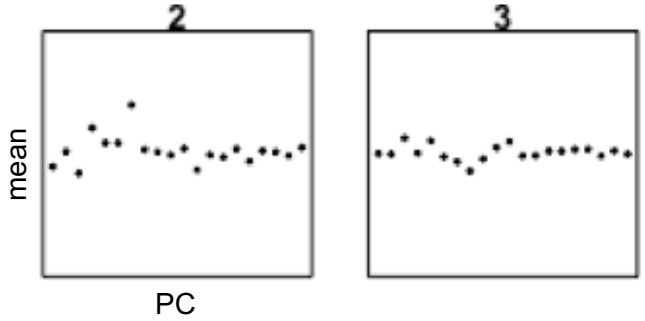
6006 fiducial points



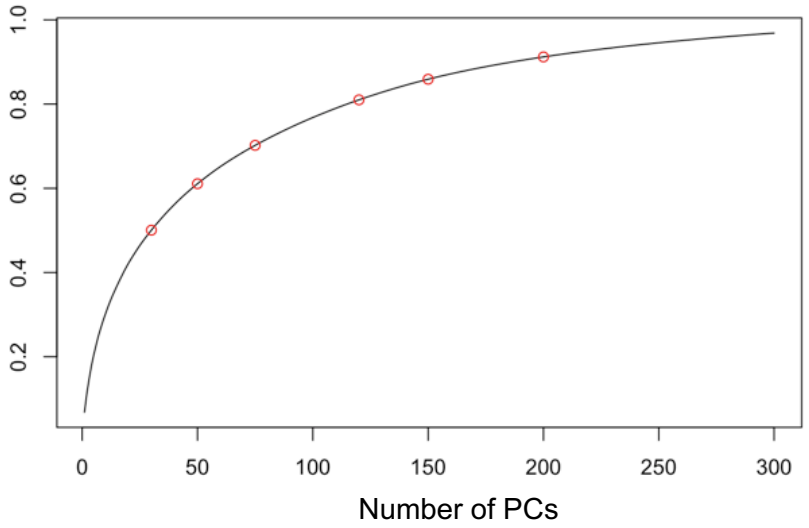
1586 fiducial points



20 top PCs



Proportion of variance



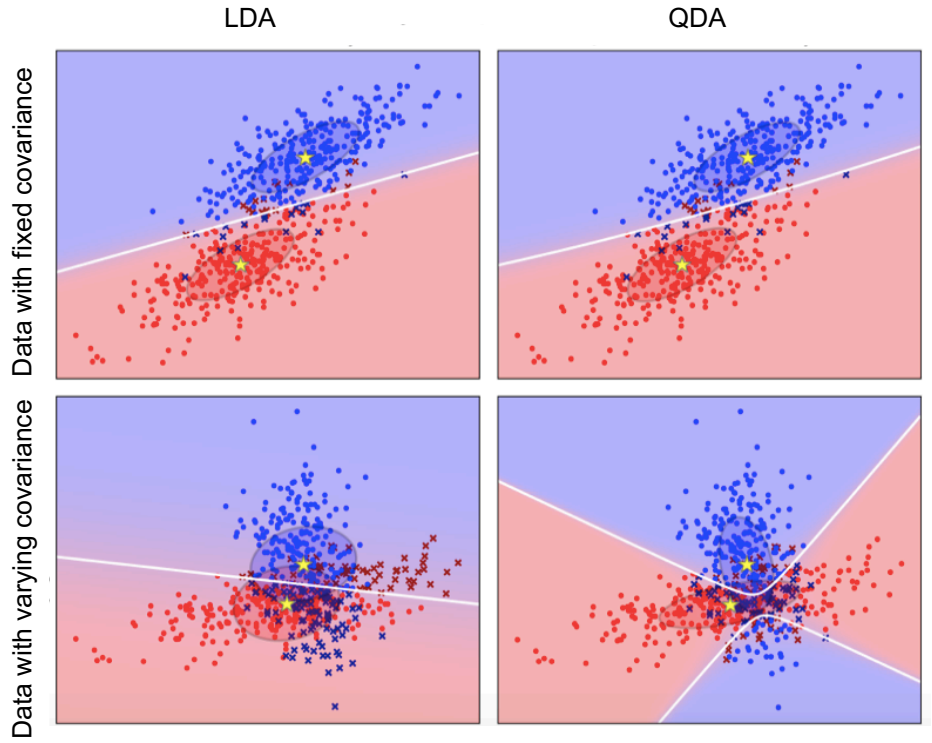
It is much easier to distinguish the emotions with the PCs than by drawing on the (reduced) fiducial points

SOURCE: Own analysis

Linear Discriminant Analysis performs best on the top 50 PCs and is at the heart of our model

Why does LDA perform so well?

- LDA consists of statistical properties for each emotion of our data
- Means and covariance matrix (latter assumed same for all emotions)
- LDA has closed-form solutions that can easily be computed; predictions obtained with Bayes
- Better than QDA in this case since QDA overfits due to different covariance matrices



LDA outperforms QDA as well as GBM! We can use LDA since we have enough unbiased data for each emotion to assume covariance matrix to stay the same

SOURCE: Scikit Learn

Thank you for your attention!

Stefano Longo, Yang Meng, Quiyu Rang, Tim Schleicher, Nuanjun Zhao