(SUPER) Cyan team Challenge

"You shall not pass, young elf!"

Mihaela Sorostinean, Laurent Cetinsoy, Nawel Medjkoune, Herilalaina Rakotoarison, Ahmed MAZARI, Matthieu Ré, NGO HO Anh Khoa

Outline

- Introduction and Motivation
- Materials and Methods
  - The dataset
  - Data preprocessing and features extraction
- Evaluation Metrics
- Results and Discussion
- Conclusion

Introduction

Age estimation

Estimating the specific age of a person given his/her face image

Age interval= 1 year

Sometimes: age groups (baby, young, adult, senior)

Humans are very good at it

What makes it difficult ?

People age differently: genes, health, environment, living style

Understanding of how human vision works to get precise estimation


YOU SHALL NOT PASS

Applications

Objective of the challenge

Binary classification: predict age group of a person from facial image

Two age categories: **minor/ major**

Threshold: 30 yrs

WIKI Dataset

Materials and Methods

Preprocessing

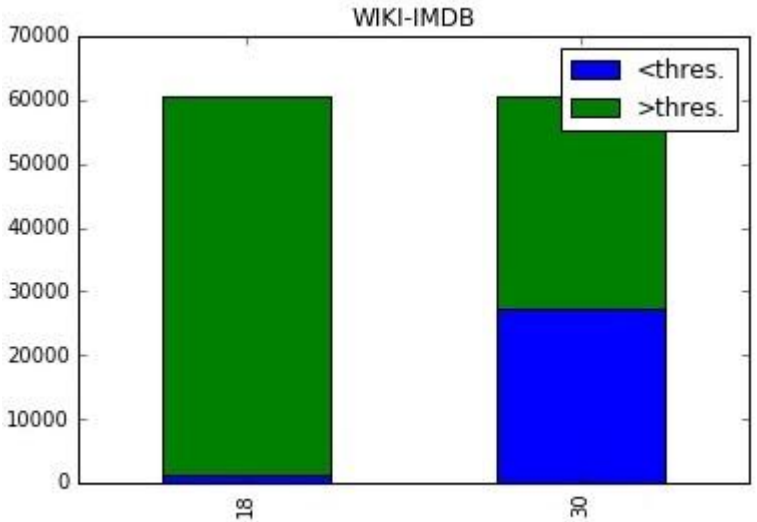
1. Input image 2. Face detection 3. Cropped face

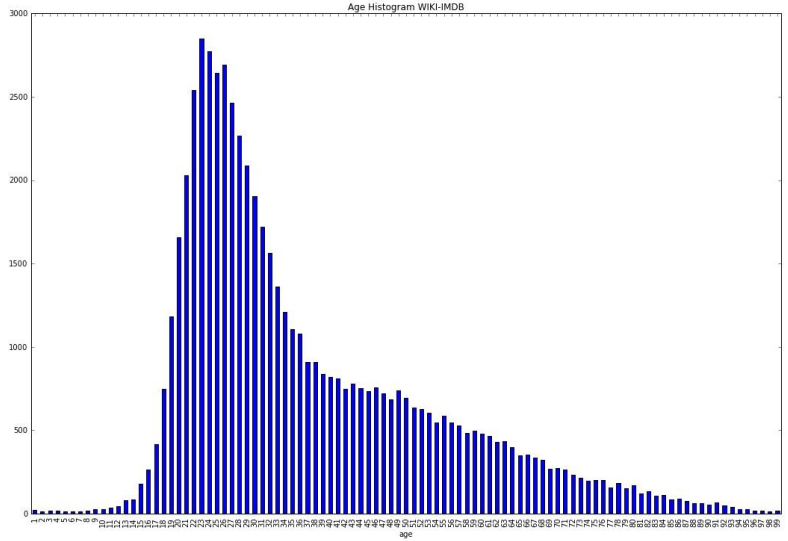
- Detector Mathias et al
  → Face without bells and whistles
- Rotated version
  → 90°, -90°, 180° ...

- Crop to 40% of its margin
- Scale to 224 x 224

The dataset

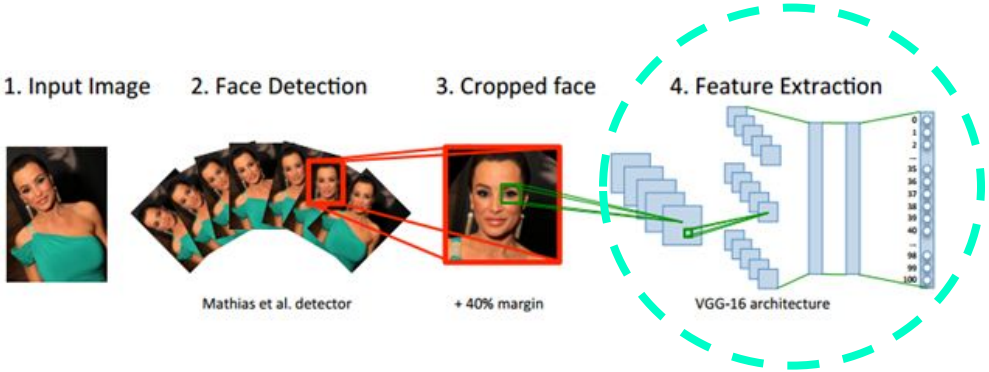
- The WIKI Database => 62,328 images of celebrities => age and gender label => cleaned



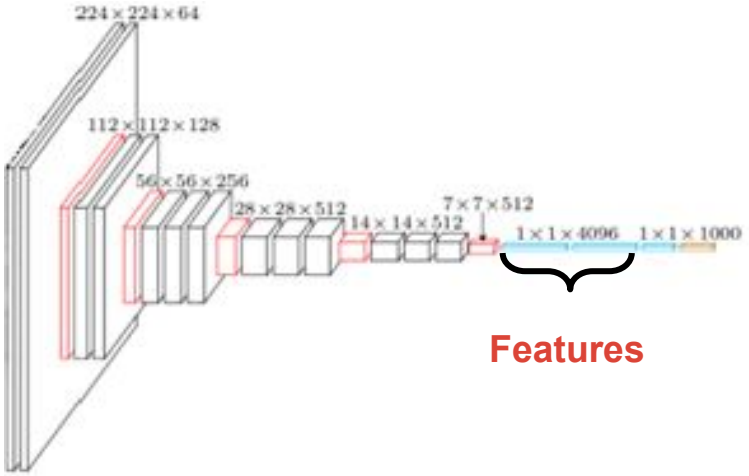


Data preprocessing and exploration

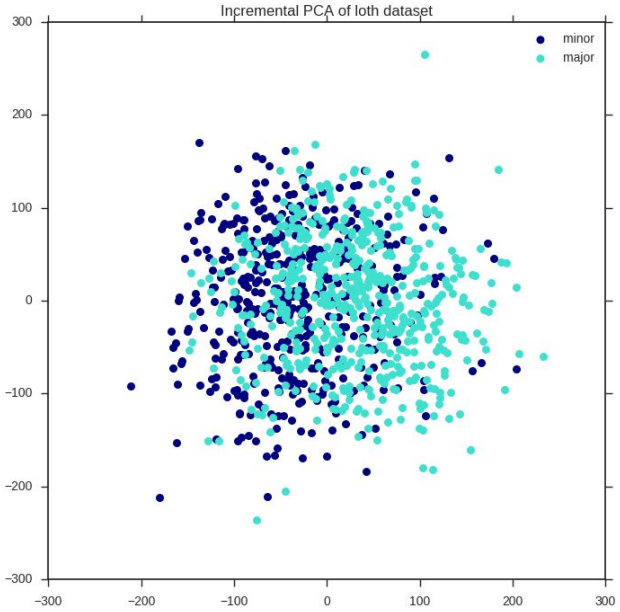
Feature extraction

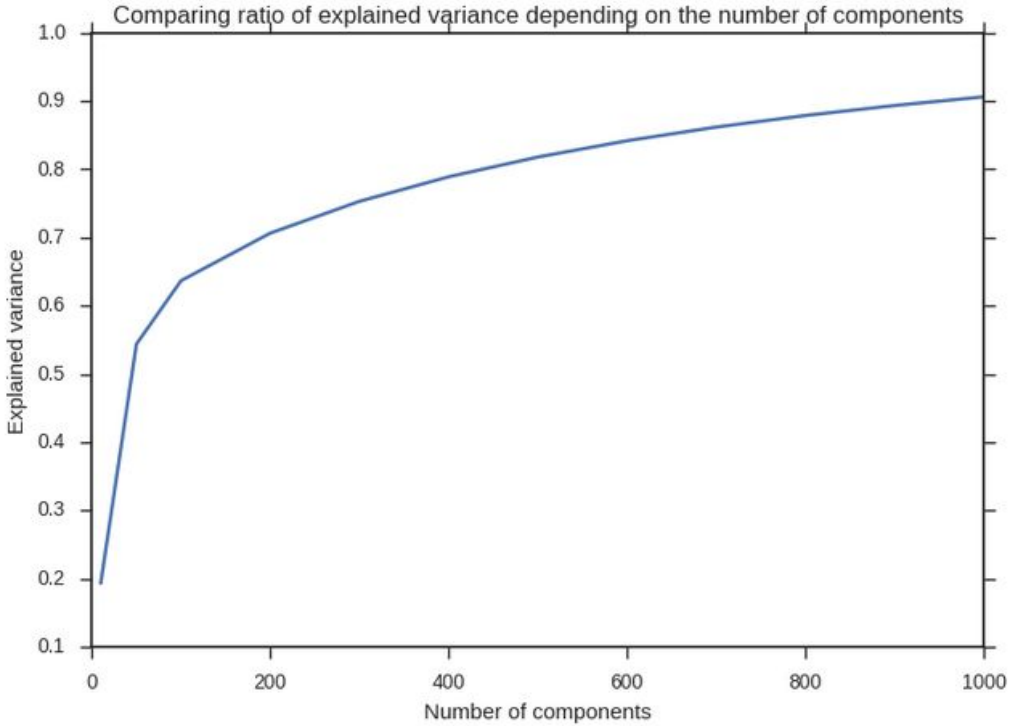


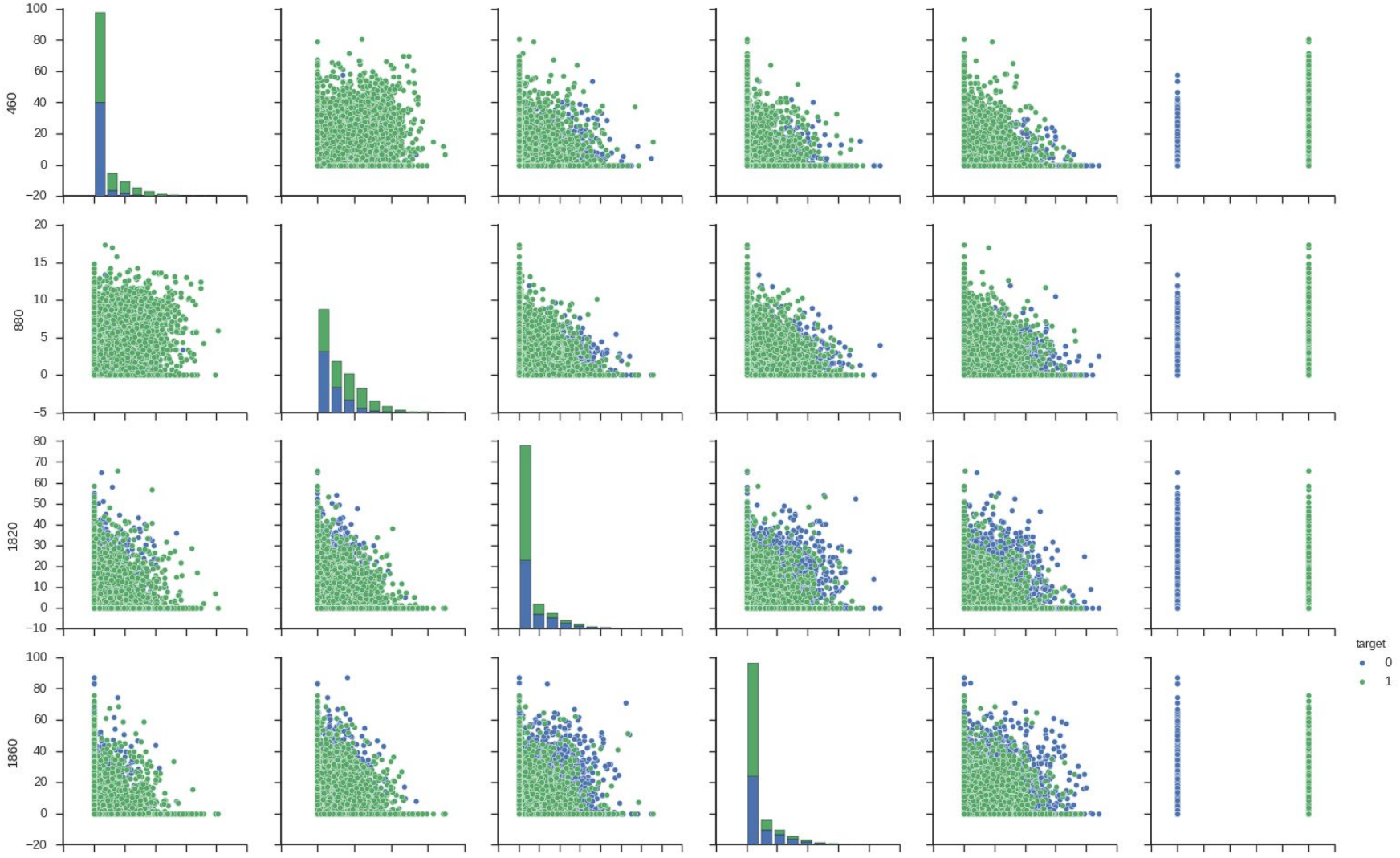
- CNN pre-trained (imageNet)
- Architecture VGG-16



PCA

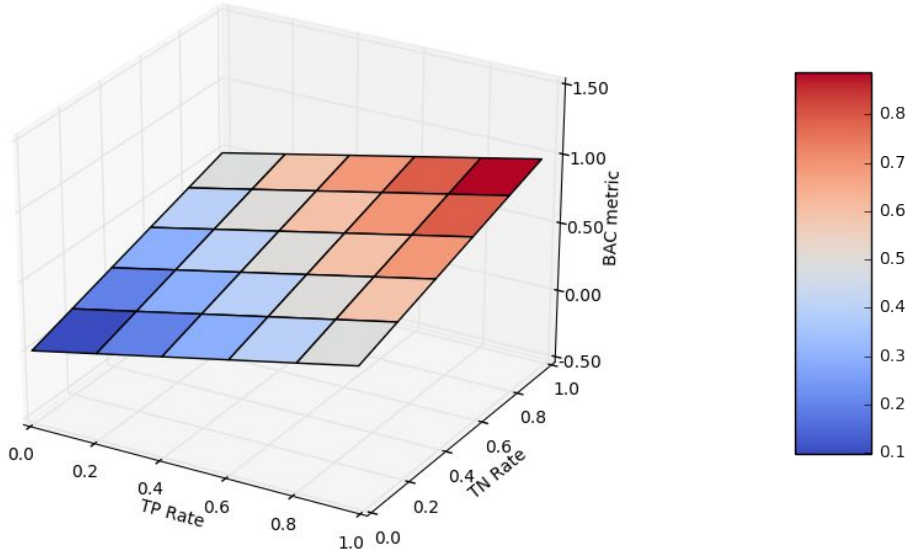

Incremental PCA of loth dataset
minor
major
300
200
100
0
−100
−200
−300
−300
−200
−100
0
100
200
300


Comparing ratio of explained variance depending on the number of components
Explained variance
Number of components
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
200
400
600
800
1000


460
880
1820
1860
target
0
1

Evaluation Metrics

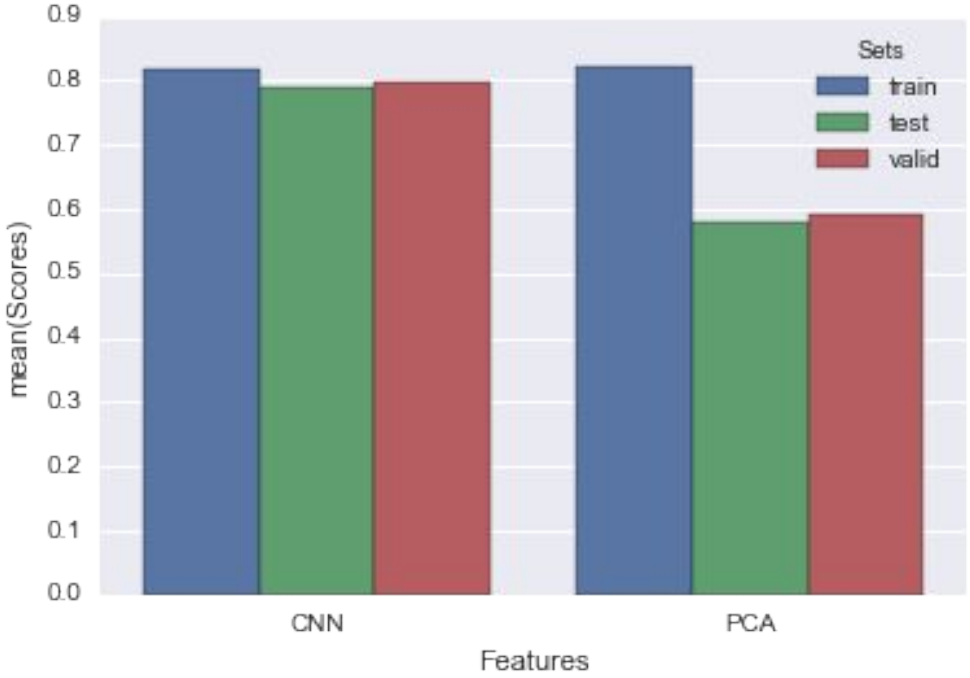
BAC Metric



$$BAC = (TP/P + TN/N)/2$$

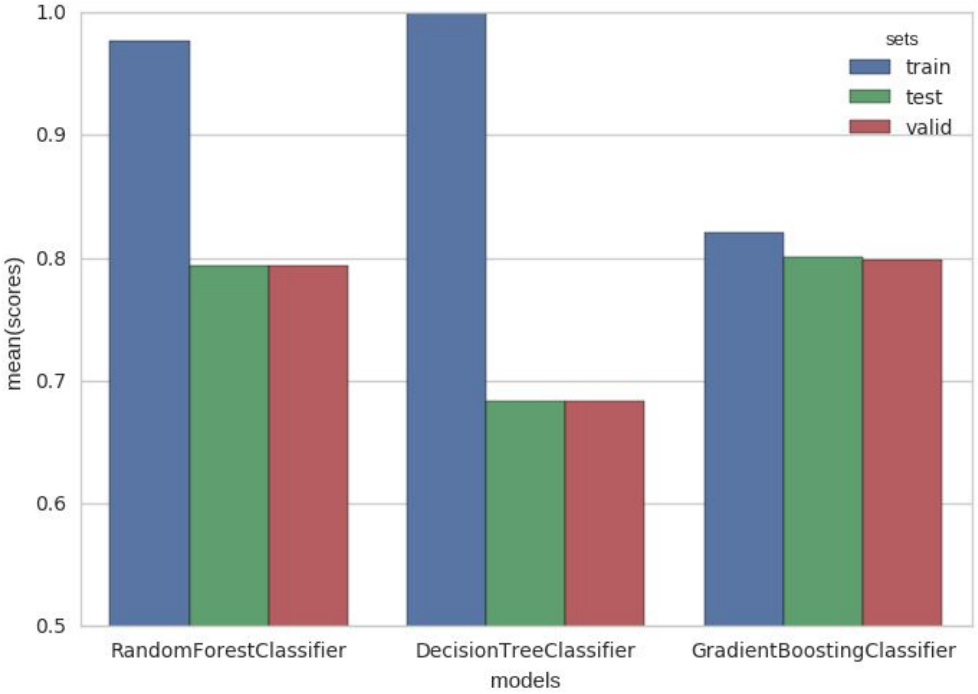
Results and discussion

Results and discussion



- PCA vs CNN

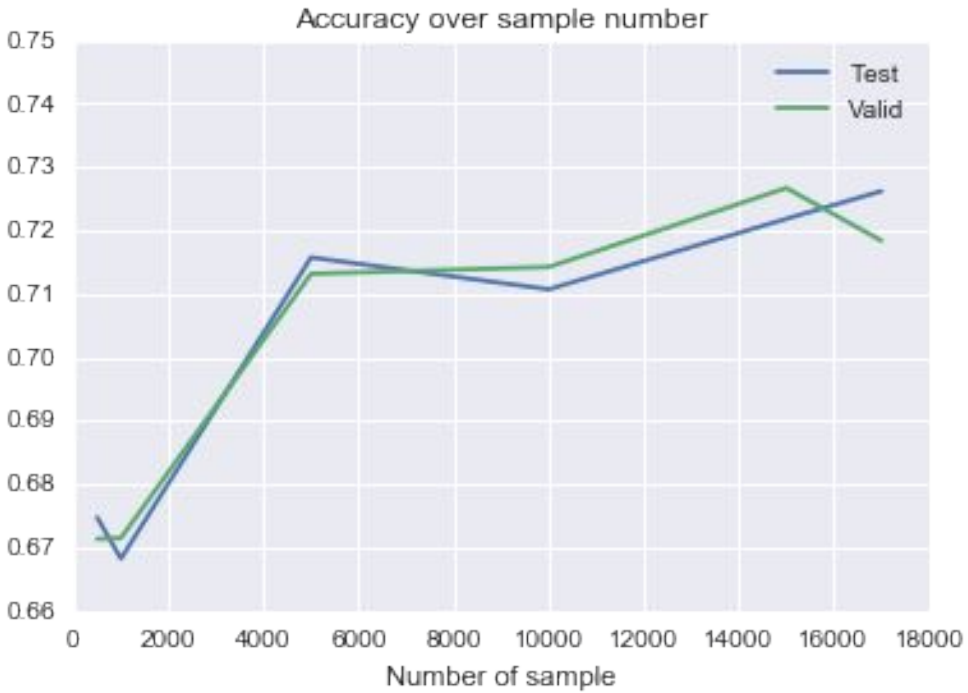
Results and discussion

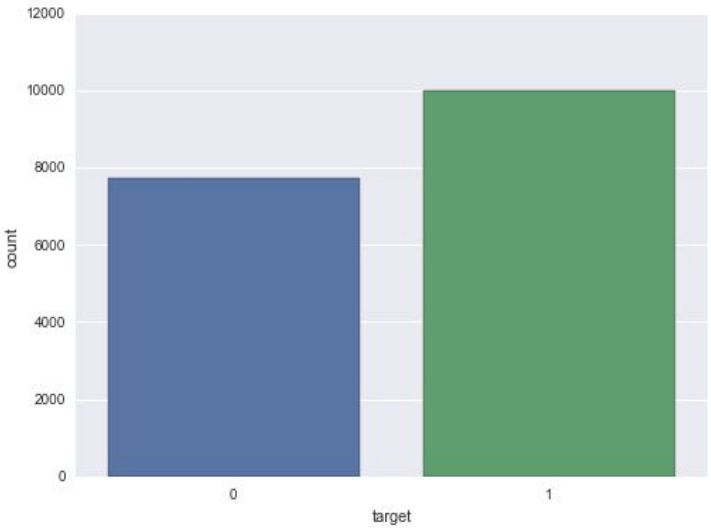


Algorithm ?

- Random Forest
- Decision Tree
- **Gradient Boosting**

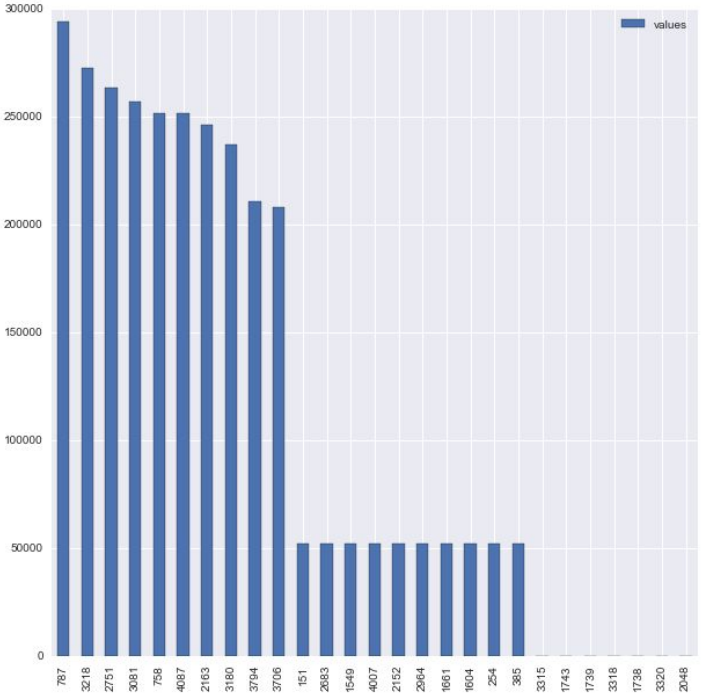
Results and discussion





Role of Sample Size

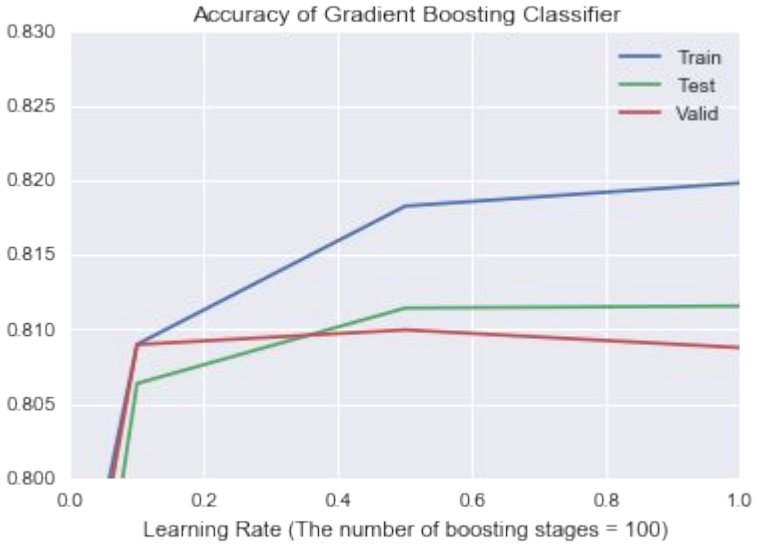
Results and discussion

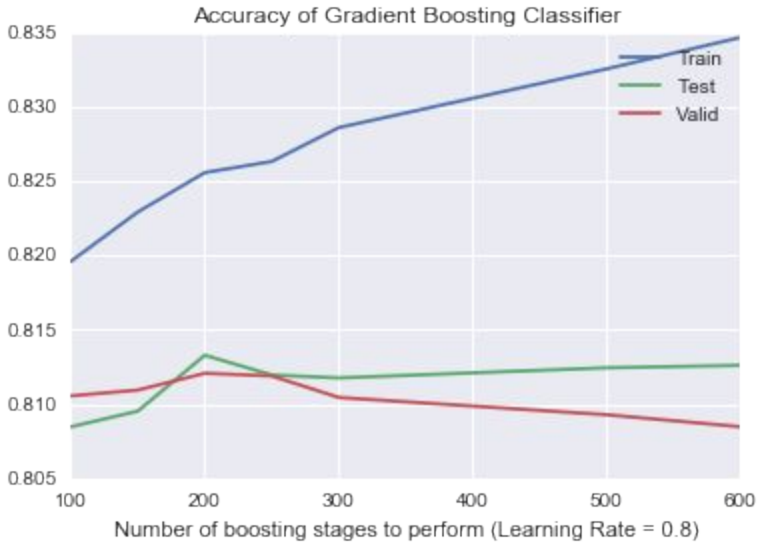


Role of Features

Results and discussion

BAC Score = 0.81 ± 0.0038





Parameters : **Learning Rate (Scale Value)** and Number of boosting stages

Conclusion

Conclusion

Age detection task

Binary classification

Major / Minor categories

Features extraction:
pre-trained CNN

Gradient Boosting Classifier

Reduction of the size of the featureset