Age & Gender Classification

Patrick Wahrmann
Christoph Körner

Age & Gender Classification for *The Profiler*

IMDb & Wiki Dataset

IMDb

460,723 images

Wikipedia

62,328 images

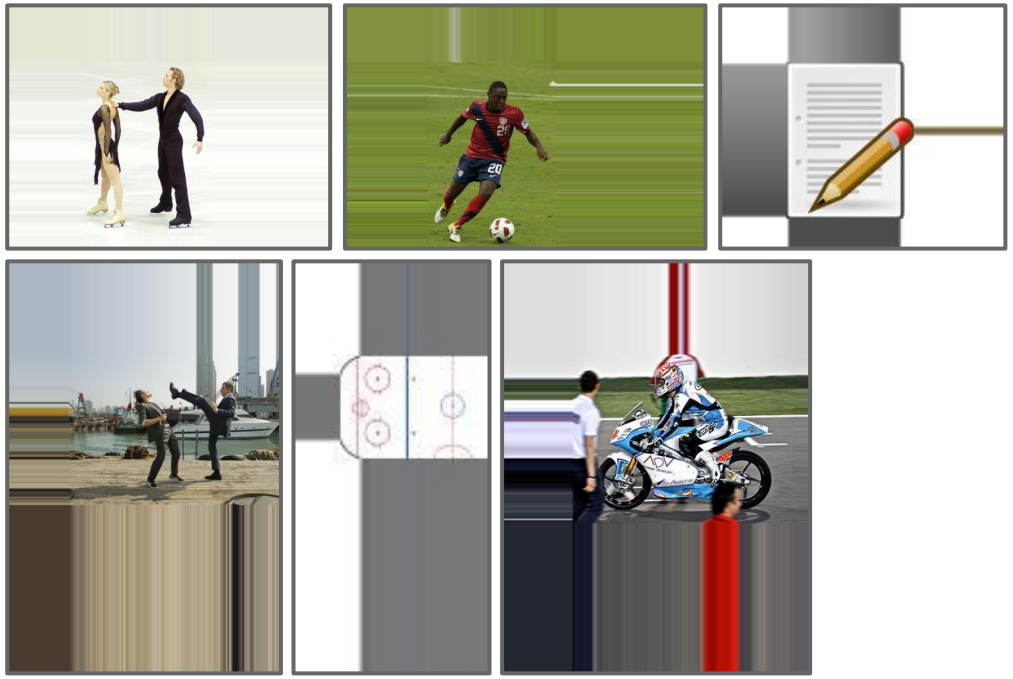
Labels

gender: 0, 1

age: date of birth and photo taken

IMDb & Wiki Dataset - Preprocessing: Cleaning

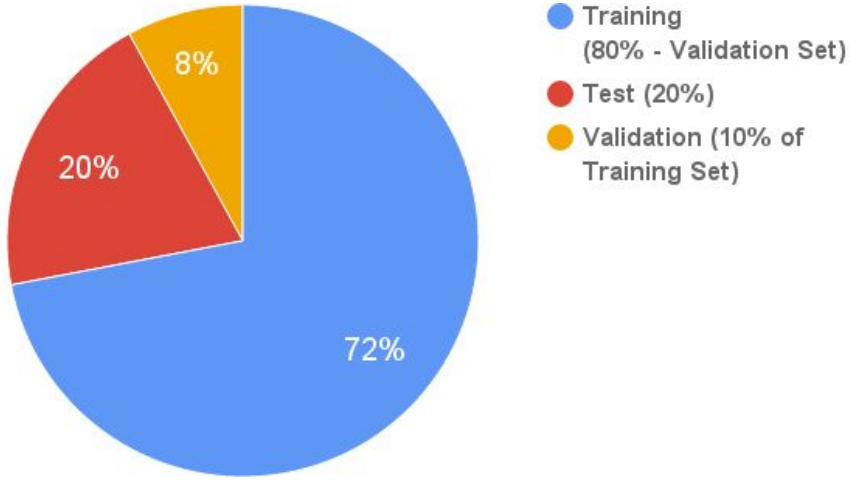
- Remove 1x1 pixel images
- Remove non-square images
- Resize images to fixed dimensions (224x224px RGB)

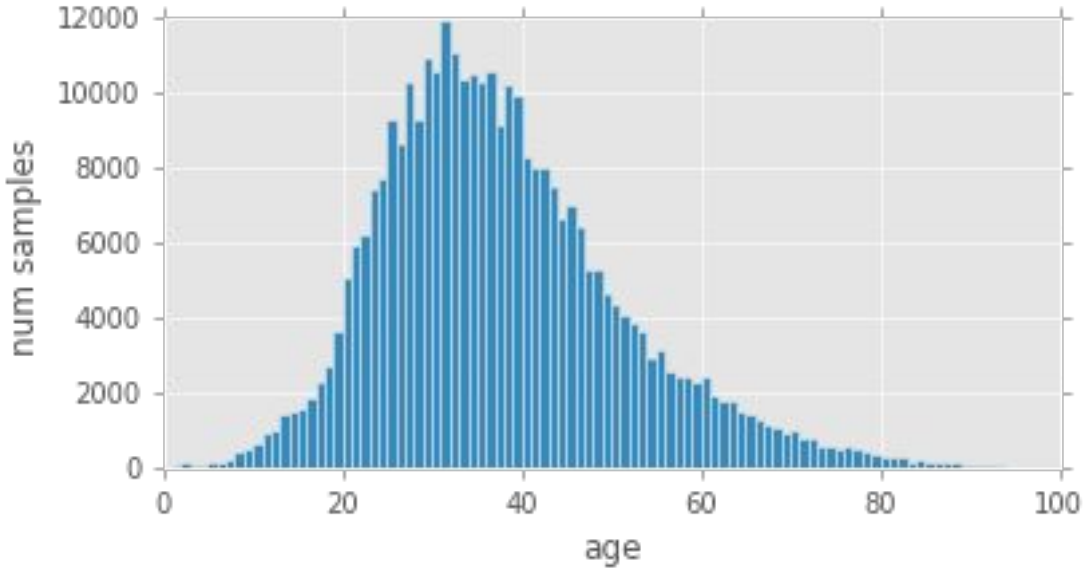


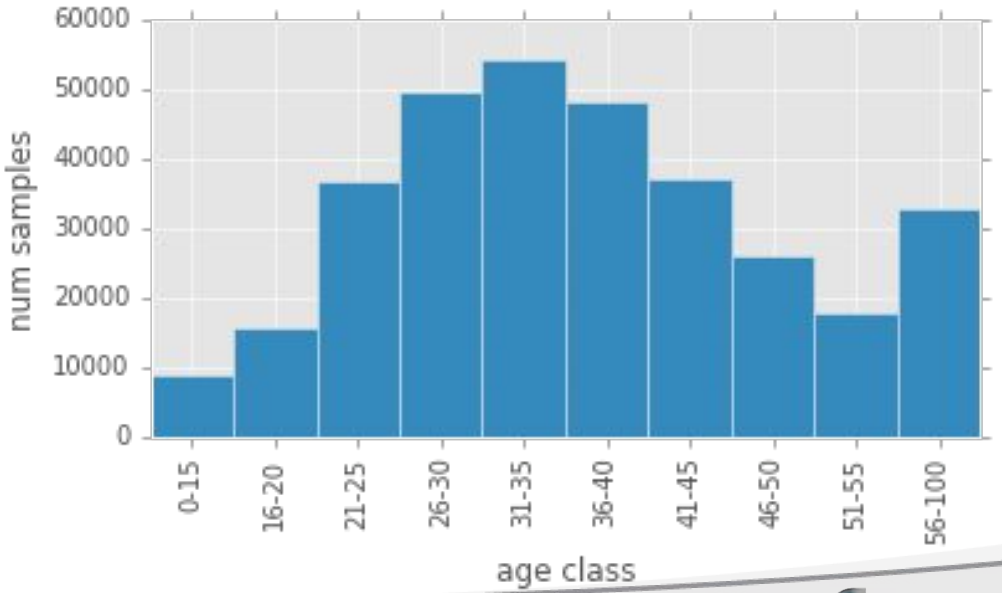
	WIKI	IMDb
total	62328	460723
image dimensions 1x1	8070	0
image non-square	24124	163452
age > 100	44	107
age < 1	57	244
total after cleaning	**30033**	**296920**

IMDb & Wiki Dataset - Preprocessing: Splitting

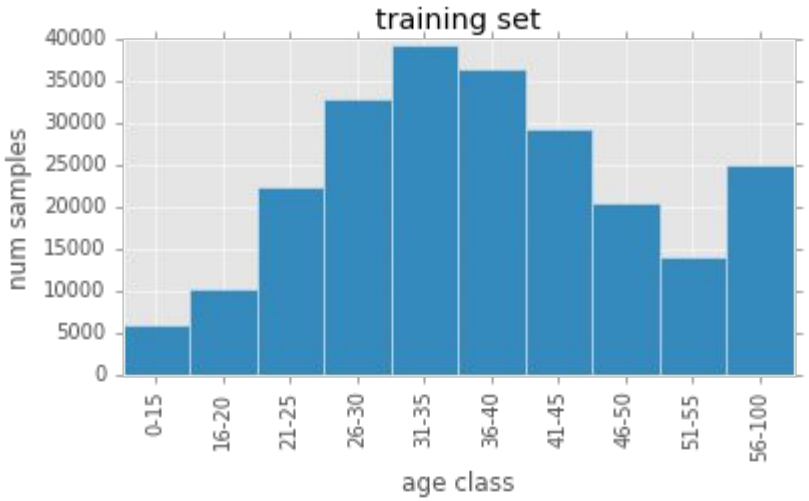
- Shuffle cleaned dataset
- Partition into age classes
- Split into test, training and validation set

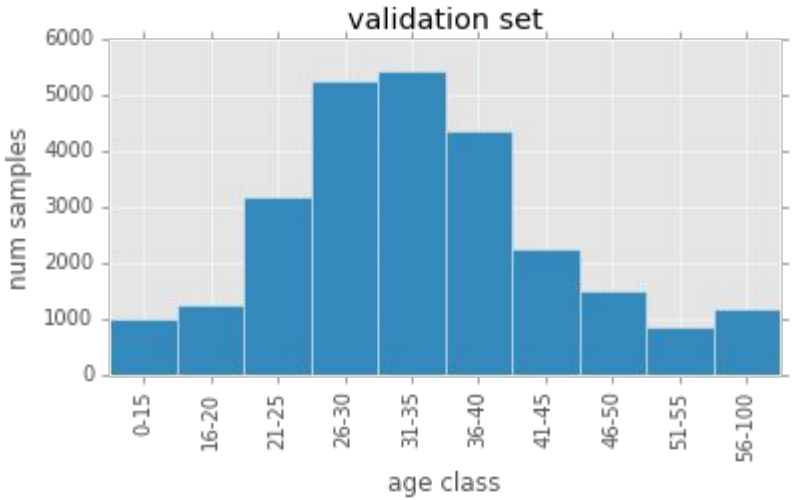


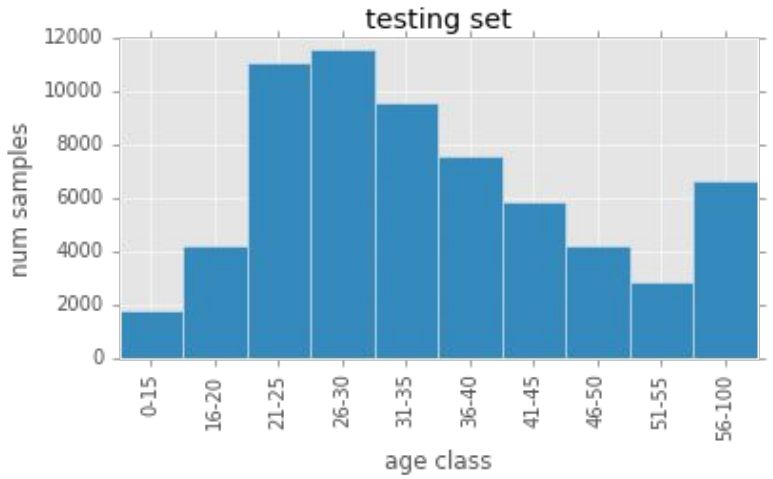


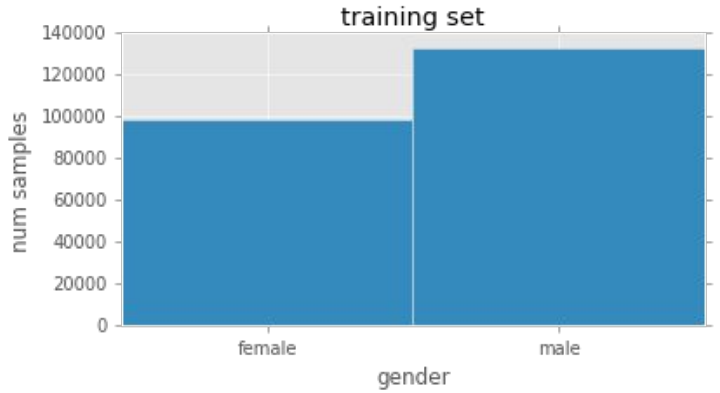


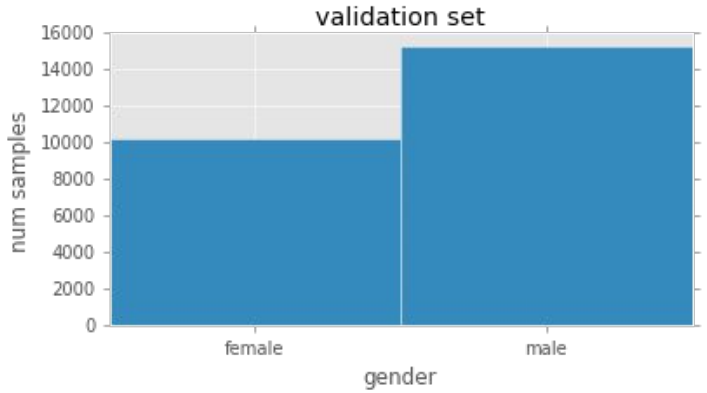
IMDb & Wiki Dataset - Preprocessing: Age Class and Gender

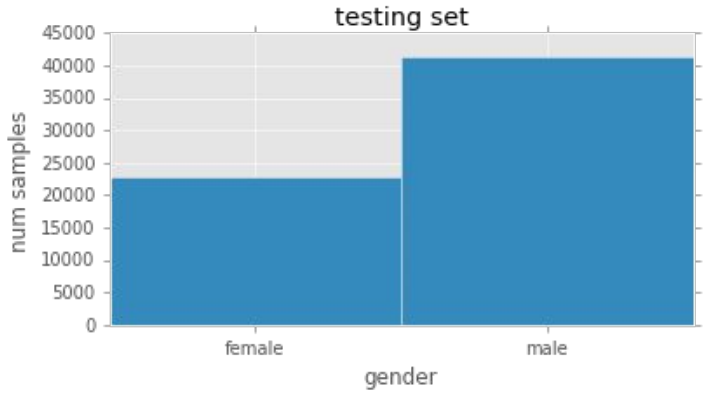




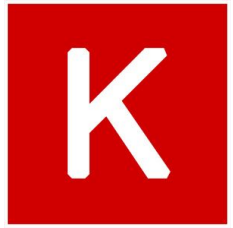








Deep Learning Model



- Deep Learning (Keras on Theano)
  - VGG-16 (easy customizable)
  - SGD with Nesterov Momentum
- Loss function
  - cross-entropy for both age and gender classes
  - ordinal loss for age classes (e.g. MSE or weighted kappa)

Training on Amazon EC2 (p2.xlarge)

- Hardware Specs (0.9 $/h)
  - 4 vCPUs @ 2.7 GHz, 61 GB RAM
  - 1 NVIDIA K80 GPU, 2496 cores, 12 GB Memory
  - Deep Learning AMI (CUDA, cuDNN, Theano, Keras, etc.)
- Configuration
  - Dataset stored in S3
  - Data processed in EC2

Schedule

- ~~Data acquisition~~
- ~~Data preprocessing~~
- ~~Model selection and pipeline~~
- Model tuning
- Model training
- Model evaluation

Thank you!

- Patrick Wahrmann
- Msc. Student in Visual Computing @ TU Wien
- patrick.wahrmann@student.tuwien.ac.at

- Christoph Körner
- Msc. Student in Visual Computing @ TU Wien
- christoph.koerner@tuwien.ac.at