Beauty and Bias
A Critical Analysis of Beauty Standards in Machine Learning

Michael Sam, Isabella Bossa

Background

Why analyze beauty standards in ML?

Relevance and increase in popularity of Computer Vision

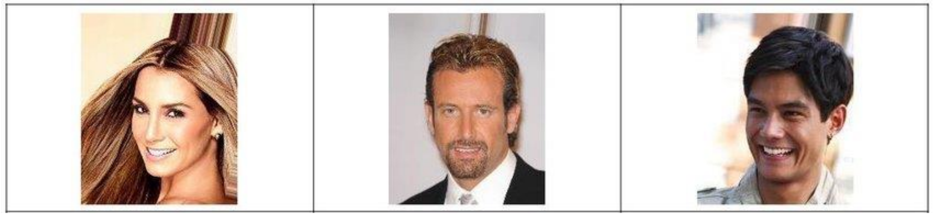
Facial recognition dataset:
CelebFaces Attributes (CelebA)

Features

Index	Definition	Index	Definition	Index	Definition	Index	Definition
1	5o'ClockShadow	11	Blurry	21	Male	31	Sideburns
2	ArchedEyebrows	12	BrownHair	22	MouthSlightlyOpen	32	Smiling
3	Attractive	13	BushyEyebrows	23	Mustache	33	StraightHair
4	BagsUnderEyes	14	Chubby	24	NarrowEyes	34	WavyHair
5	Bald	15	DoubleChin	25	NoBeard	35	WearingEarrings
6	Bangs	16	Eyeglasses	26	OvalFace	36	WearingHat
7	BigLips	17	Goatee	27	PaleSkin	37	WearingLipstick
8	BigNose	18	GrayHair	28	PointyNose	38	WearingNecklace
9	BlackHair	19	HeavyMakeup	29	RecedingHairline	39	WearingNecktie
10	BlondHair	20	HighCheekbones	30	RosyCheeks	40	Young

Attributes in the dataset

Objectivity and subjectivity



Attractive

Our main concerns

No detail about criteria used to determine the annotations

No documentation on how annotations were made

Subjectivity of most feature annotations

What is “big nose”, “narrow eyes”?

Attractiveness as a hard Yes/No label

Shouldn’t this depend on the culture, region, person, etc.?

Exploratory data analysis

What visual patterns can we find?

EDA Results

202599 observations, 40 Yes/No features

No missing data

Response variable **apparently** well balanced: 51.3% deemed “Attractive”

Highest Positive Correlation	
Attribute	Correlation
Wearing Lipstick	0.480
Heavy Makeup	0.477
Young	0.388
Arched Eyebrows	0.250
Pointy Nose	0.228

Highest Negative Correlation	
Attribute	Correlation
Male	-0.395
Big Nose	-0.277
Chubby	-0.237
Eyeglasses	-0.223
Double Chin	-0.201

“Correlation does not imply causation”

Models & results

What drives attractiveness?

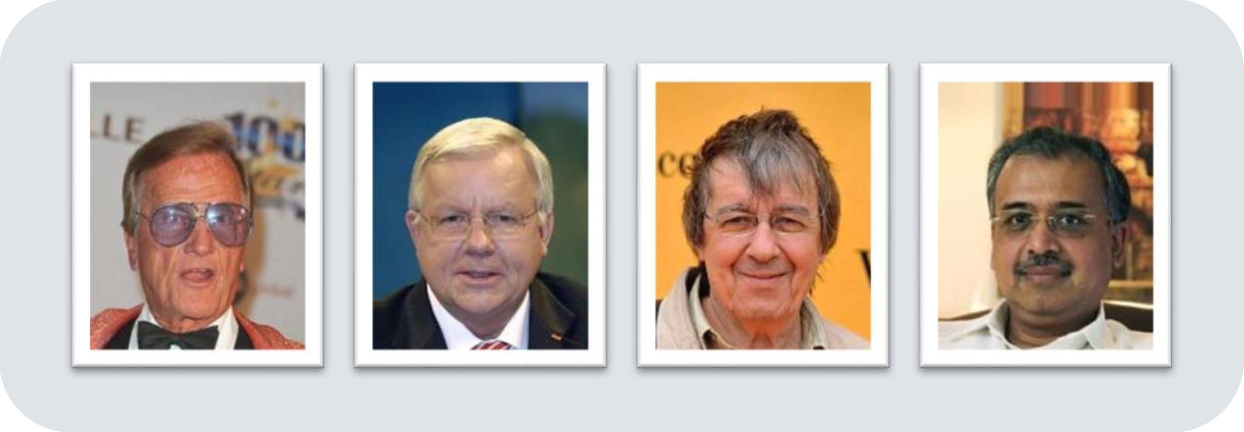
Logistic Regression

Largest Positive Predictors	
Attribute	Change in Odds
Young	4.51
Wearing Lipstick	2.65
Heavy Makeup	2.48
Smiling	2.44
Pale Skin	2.35

Largest Negative Predictors	
Attribute	Change in Odds
Chubby	0.17
Blurry	0.20
Gray Hair	0.24
Bald	0.25
Double Chin	0.34
Eyeglasses	0.34

Critical Analysis

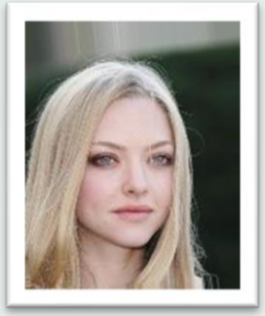
What do our results mean?



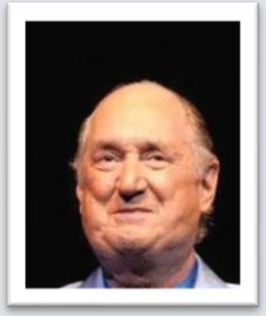
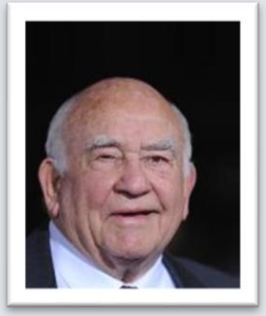
Correlation Results

- Positive correlation
  - Western features (arched eyebrows, pointy noise).
  - Young.
  - Stylish (heavy makeup, wearing lipstick).
- Negative correlation
  - Not young.





Logistic Regression Results

- Positive predictors
  - Western features (pale skin).
  - Young.
  - Stylish (wearing lipstick, heavy makeup).
  - Ethnicity.
- Negative predictors
  - Not young.
  - Men.
  - **452** men versus **4** women who are ~~bald~~, chubby, with a double chin and grey hair and wearing glasses.

Gender Biases

Is this a reflection of the entertainment industry?

Gender	Attractive	Count	Proportion
Male	No	60,855	**72.1%**
	Yes	23,576	27.9%
Female	No	37,911	32.1%
	Yes	80,254	**67.9%**

Age Biases

Do we seem "less attractive" to society as we age?

Age	Attractive	Count	Proportion
Young	No	59,975	38.3%
	Yes	96,759	**61.7%**
Not Young	No	38,791	**84.6%**
	Yes	7,074	15.4%

Lack of Diversity

"Cultural diversity" versus CelebA

Cultural Biases

Discussion

"CelebA has large diversities, large quantities, and rich annotations."

Underrepresentation of large segments of population.

Perpetuation and amplification of stereotypes and societal biases.

Some images are wrongly labeled.

The dataset is employed for face attribute recognition, face detection, landmark (or facial part) localization, and face editing and synthesis.

THANK YOU