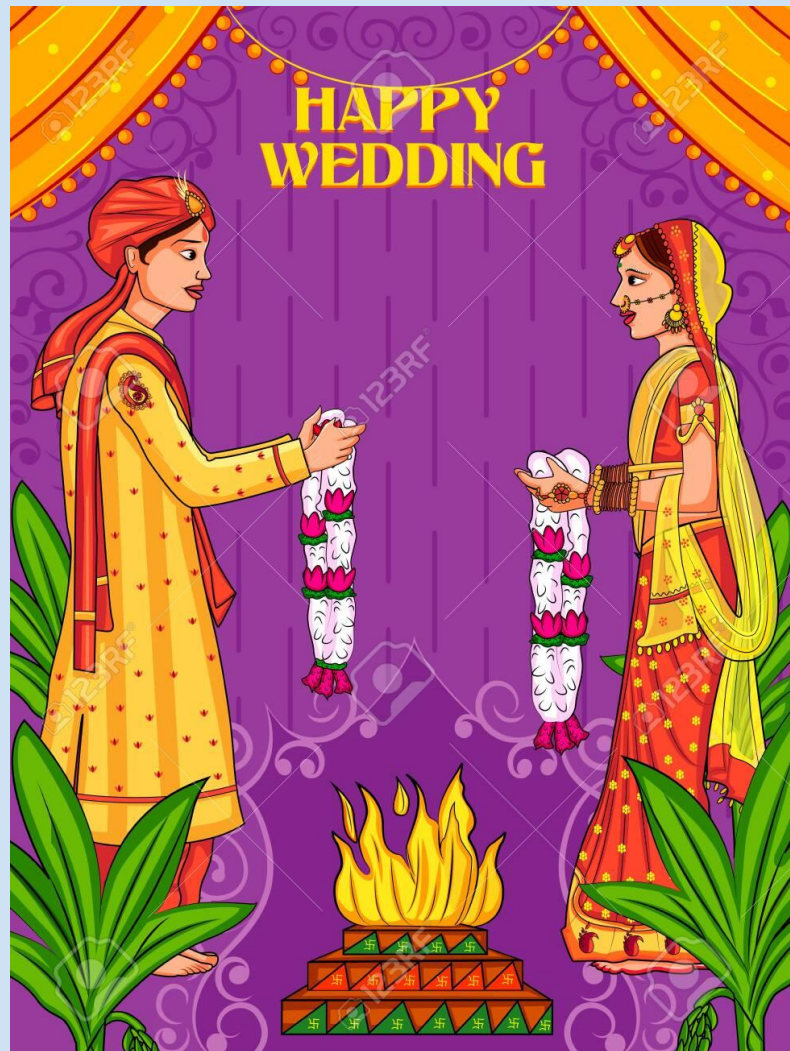


# **Image Processing & Face Recognition**

**Aditya & Youssef  
BUS 212A, Prof. Namini**

# Why did we choose this



# Problem Statement

“The problem is Image Classification, more specifically, Face Classification”



WE DID IT ON OUR OWN

Classification	Regression
Classify faces to their owner in a supervised setting.  Dataset for 115 images for 5 Celebrities. (Kaggle)	Predict the age of a person from a picture of the face.  Dataset of 10,000 images with their age. (Kaggle)

# Data:

1. 115 Images for 5 Celebrities:



Ben Affleck



Elton John



Jerry Seinfeld



Madonna



Mindy Kaling

# Data Cont'

10,000 images for people aged from 1 - 100



1000.png  
Size 42.79 KB



1018.png  
Size 32.71 KB



1030.png  
Size 39.78 KB



1062.png  
Size 43.08 KB



1068.png  
Size 54.05 KB

Age = 1



1075.png



1077.png



1081.png



1087.png



1089.png

Age = 2



1385.png  
Size 41.47 KB



141.png  
Size 35.58 KB



1428.png  
Size 46.28 KB



143.png  
Size 40.54 KB



1491.png  
Size 51.69 KB

Age = 9



1751.png  
Size 44.95 KB



1978.png  
Size 46.79 KB



1995.png  
Size 52.67 KB



2082.png  
Size 51.92 KB



2119.png  
Size 41.49 KB

Age = 10

•  
•

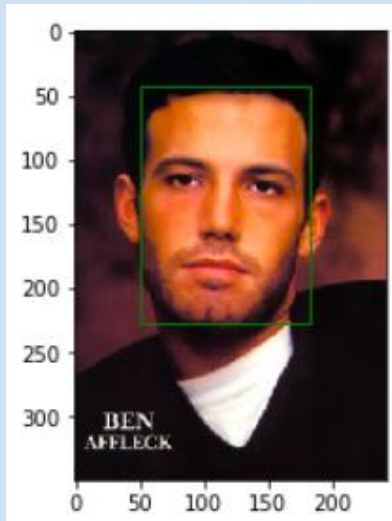


# Data Description & Preparation

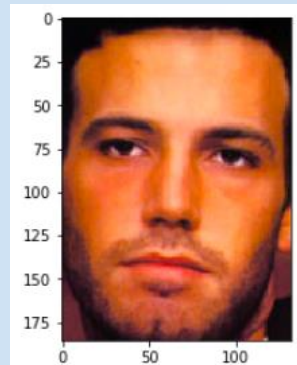
Preparation steps (The same for both datasets)



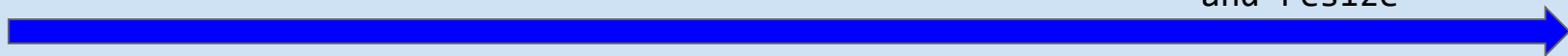
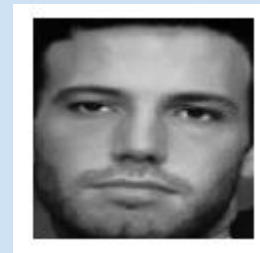
Face Detection, MTCNN



Face Cropping



Convert colors  
and resize



# Data Preparation Cont'

Final product after flattening them:

0	1	2	3	4	5	6	7	8	9	...	9991	9992	9993	9994	9995	9996	9997	9998	9999	target
97	97	77	59	59	71	71	60	60	72	...	167	167	142	142	41	41	133	140	140	elton_john
177	177	177	213	213	213	247	247	242	242	...	184	193	193	117	117	117	36	36	36	jerry_seinfeld
228	230	230	194	114	79	79	79	57	67	...	76	75	75	76	76	79	66	66	77	ben_afflek
46	49	44	44	38	32	33	35	35	25	...	163	163	128	12	8	24	24	13	6	jerry_seinfeld
38	44	49	54	64	64	100	104	131	158	...	1	0	0	5	5	12	19	22	24	jerry_seinfeld
135	135	135	162	162	162	164	164	164	98	...	0	0	0	0	0	0	0	0	0	elton_john
3	0	0	2	2	2	2	1	1	1	...	191	175	189	196	189	156	3	1	3	ben_afflek
20	20	56	56	58	73	73	67	67	82	...	8	8	9	9	8	7	7	9	9	mindy_kaling
24	24	24	24	12	12	12	12	12	17	...	10	10	10	10	10	13	13	13	13	mindy_kaling
103	127	142	165	188	186	186	178	197	210	...	11	11	10	10	10	10	13	25	35	madonna

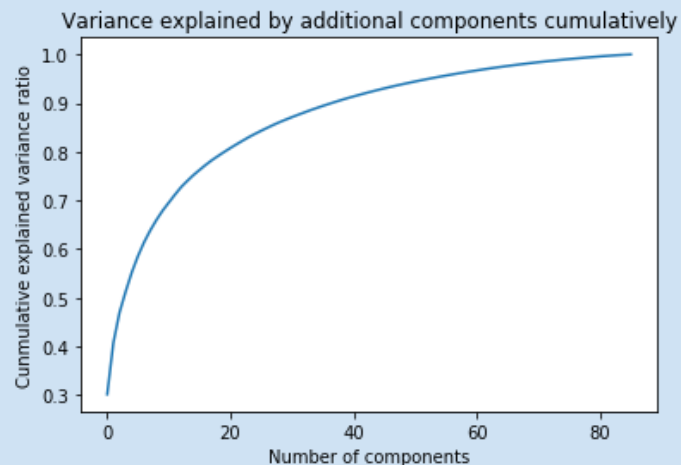
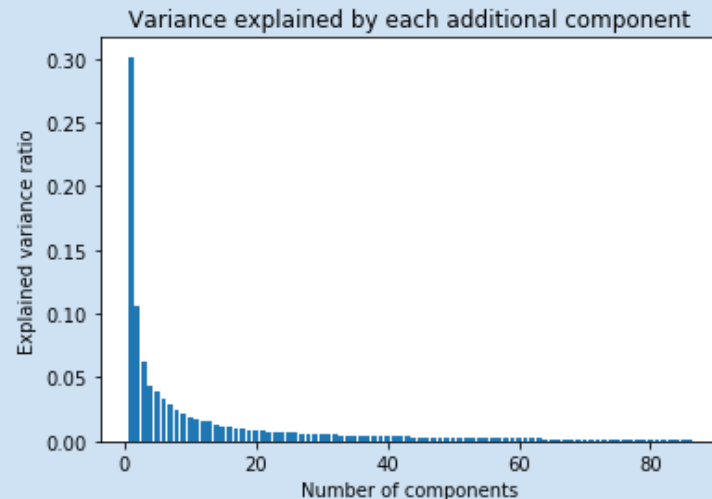
# Original Faces



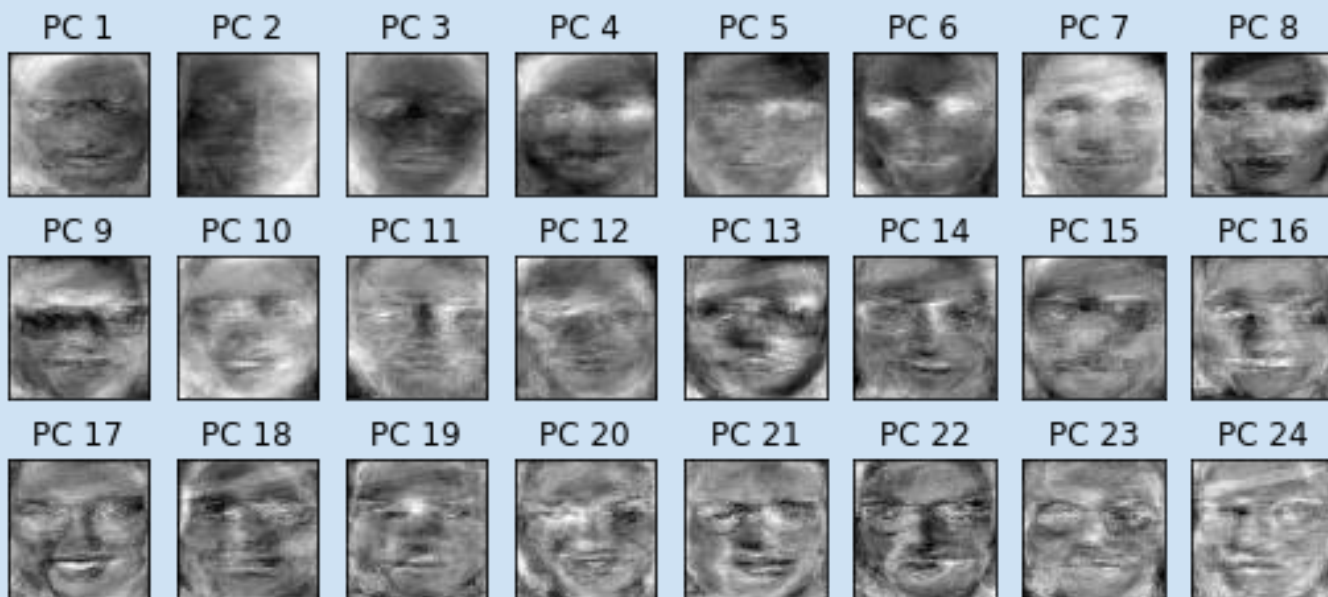


# Dimensionality Reduction

## Principal Component Analysis



# Eigen Faces



<u>Classification Methods</u>	<u>Accuracy</u>
SVC with PCA	72.76%
SVC without PCA	73.45%
KNN	64.55%
Logistic Regression	75.79%
Linear Discriminant Analysis	76.00%
Random Forest	62.58%
Gaussian Naive Bayes	61.72%
Ensemble	74.55%

**With Monte-Carlo Simulation  
& Cross Validation**

# Regression Algorithms

To predict the age, we regressed the images on the age:

Linear Regression

Ridge

Lasso

Support Vector Regression

Random Forest

<u>Regression</u>	<u>MSE</u>	<u>R<sup>2</sup></u>
Support Vector Regression	288.6	0.56
Ridge Regression	417.9	0.21
Lasso Regression	277.7	0.47
Random Forest	321	0.39
Linear Regression	421.9	0.19
Ensemble	345.4	0.36

**With Monte-Carlo Simulation  
& Cross Validation**

# Conclusion

We classified faces to their owners with ~75% accuracy.

We regressed images to their age, not as good.

Insights:

1. Working with images is a lot harder than other data.
2. Running time is usually in hours 😞.
3. There are so many moving parts in dealing with images
4. Deep learning is probably better than classical ML algorithms.



# Thank You

Aditya & Youssef  
BUS 212A, Prof. Namini