Getting Started

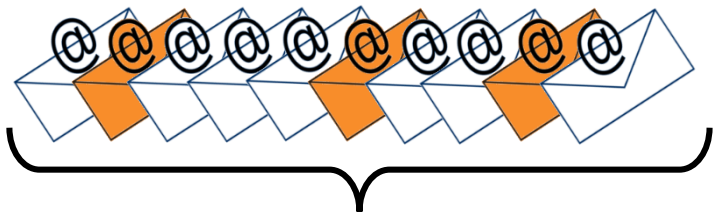
CS771: Introduction to Machine Learning

Purushottam Kar

Please enrol on Piazza

http://tinyurl.com/ml17-18adf

Recap



Training Data

ML Algorithm

Test Data

Model

Output

Recap

Input driven ML

Batch Learning

Active Learning

Robust Learning

Semi-supervised Learning

Supervised Learning

Online Learning

Unsupervised Learning

Reinforcement Learning

Output driven ML

Regression

Subject: [**SPAM**] Free movie tickets every month
X-Barracuda-Spam-Score: 4.89

Topic Modelling

Ranking

Multi-classification

Primary Social Promotions Updates Forums

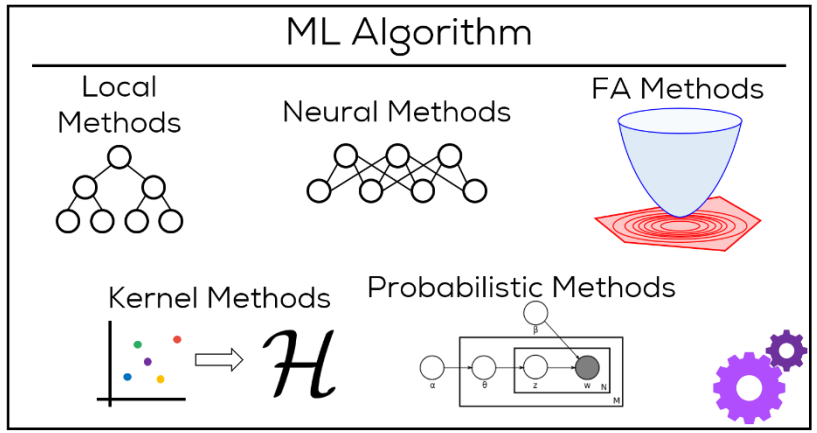
Tagging

URGENT, OFFICIAL, TAX

Binary Classification

Subject: [**SPAM**] Free movie tickets every month
X-Barracuda-Spam-Status: Yes

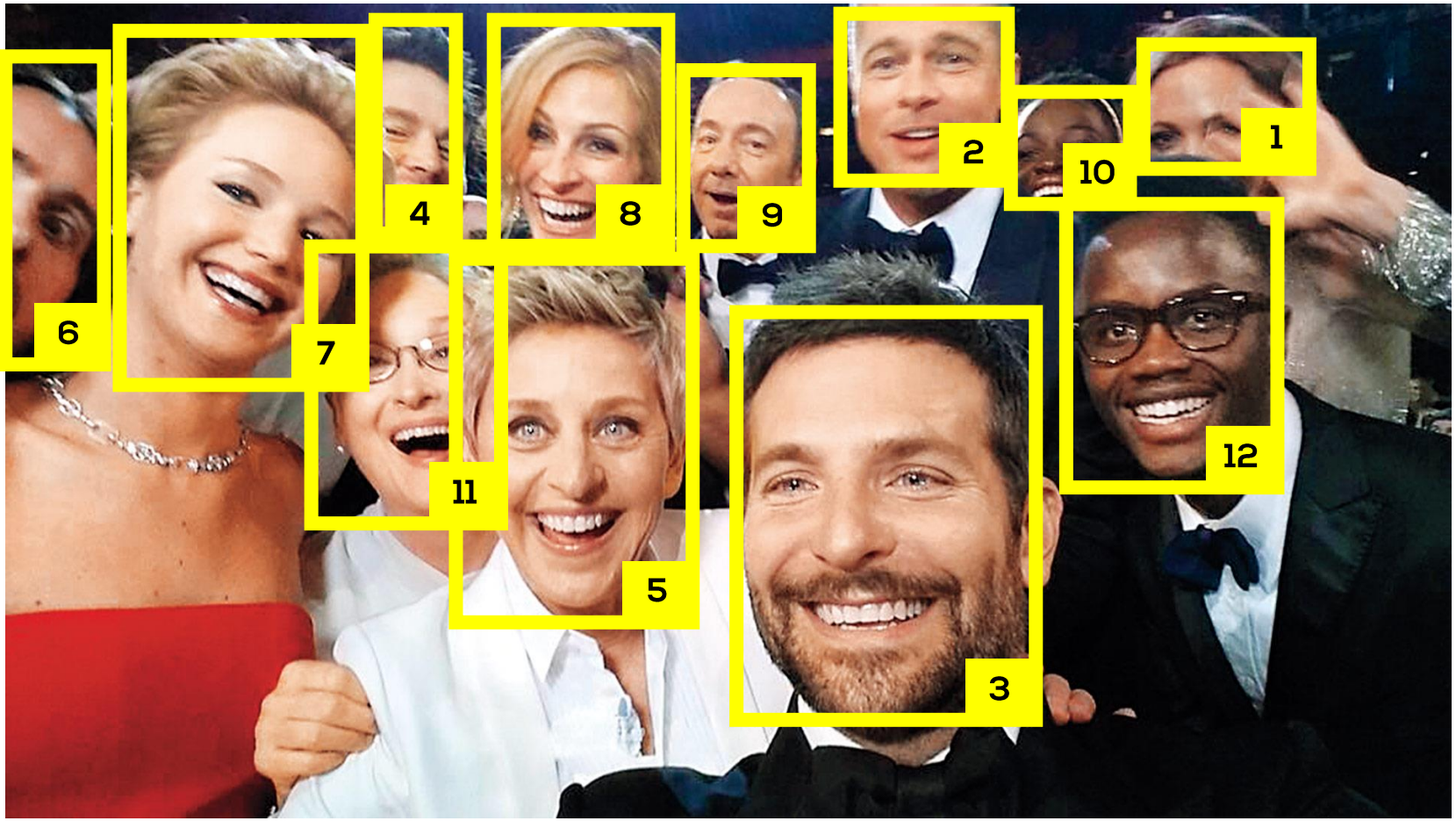
Process driven ML



ML Primitives

A first look

Image Tagging



Celebrity Names

1. Angelina
2. Brad
3. Bradley
4. Channing
5. Ellen
6. Jared
7. Jennifer
8. Julia
9. Kevin
10. Lupita
11. Meryl
12. Peter

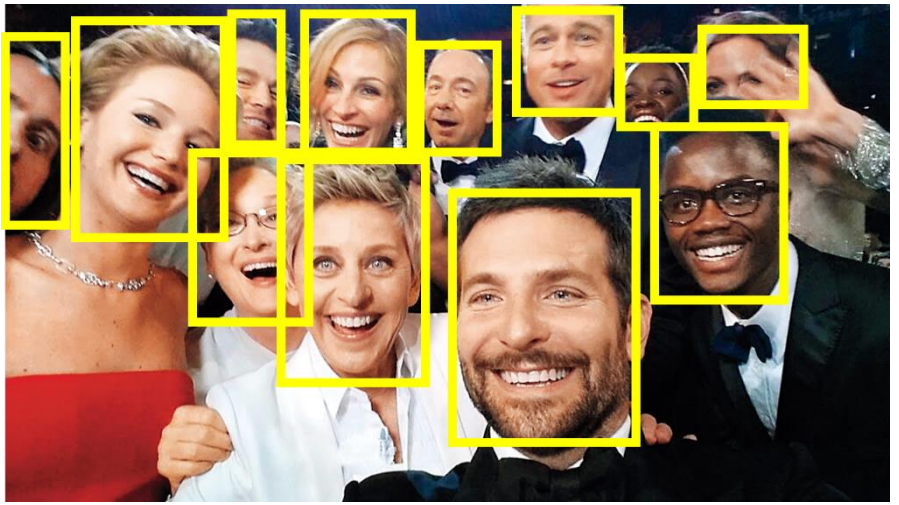
Image Tagging as Multi-label Classification

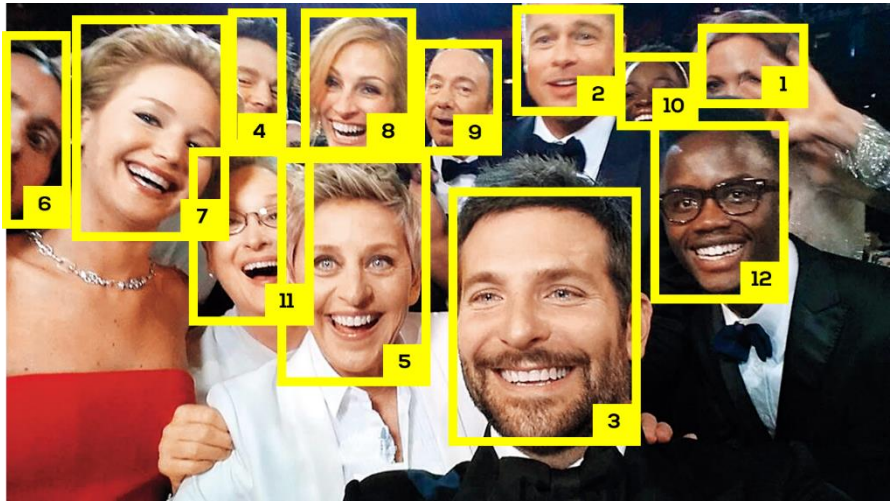
Celebrity Names

1. Asin
2. Angelina
3. Aamir
4. Brad
5. Bradley
6. Channing
7. Deepika
8. Dhanush
9. Ellen
10. Hansika
11. Hrithik
12. Ileana
13. Jared
14. Jennifer
15. Julia
16. Kajal
17. Katrina
18. Kevin
19. Lupita
20. Meryl
21. Mohanlal
22. Nayantara
23. Peter
24. Prabhas
25. Rajnikanth
26. Shahrukh
27. Suriya
28. Sonam
29. Taapsee
30. Vikram

Image Tagging

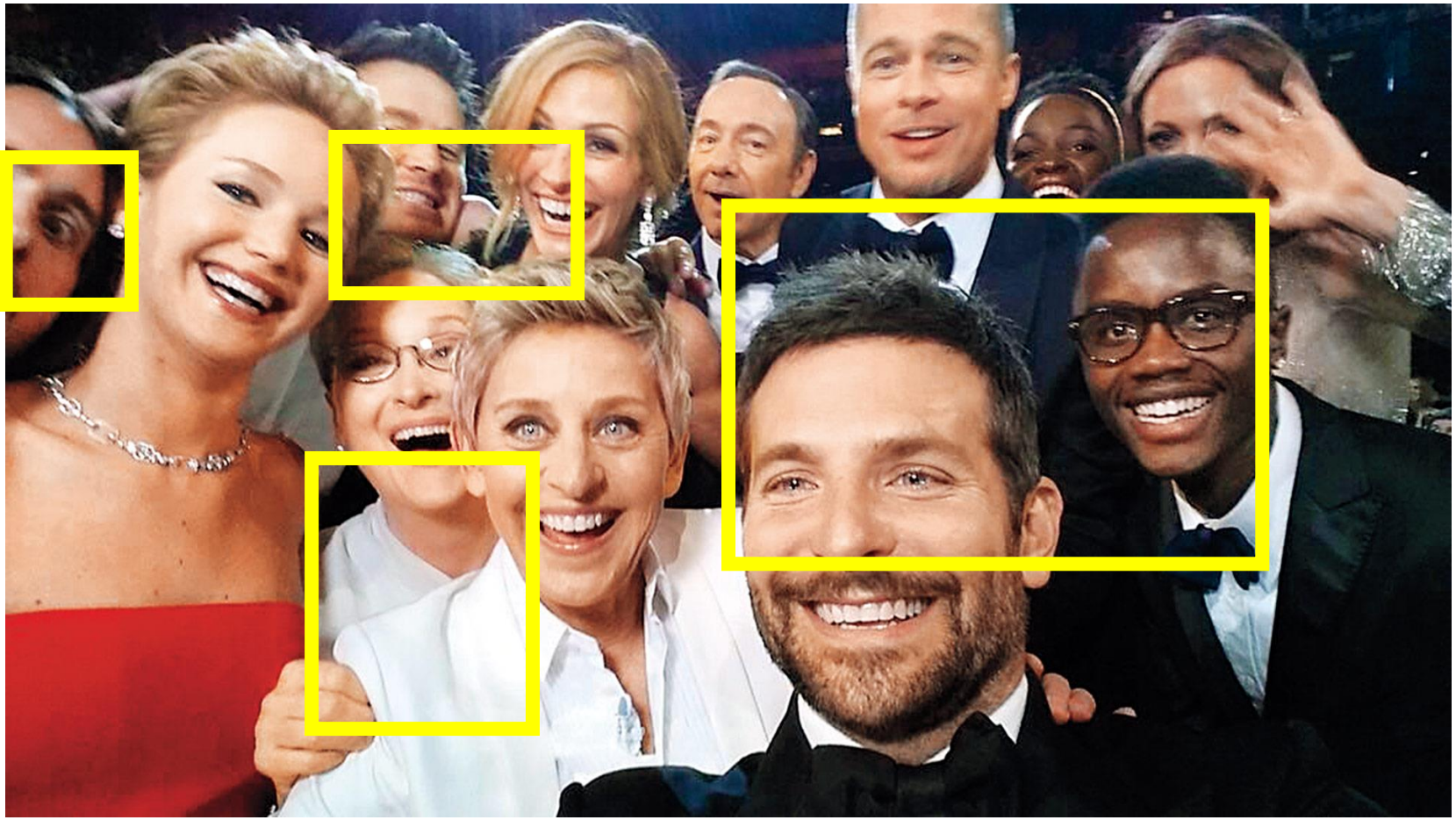
Face Detection



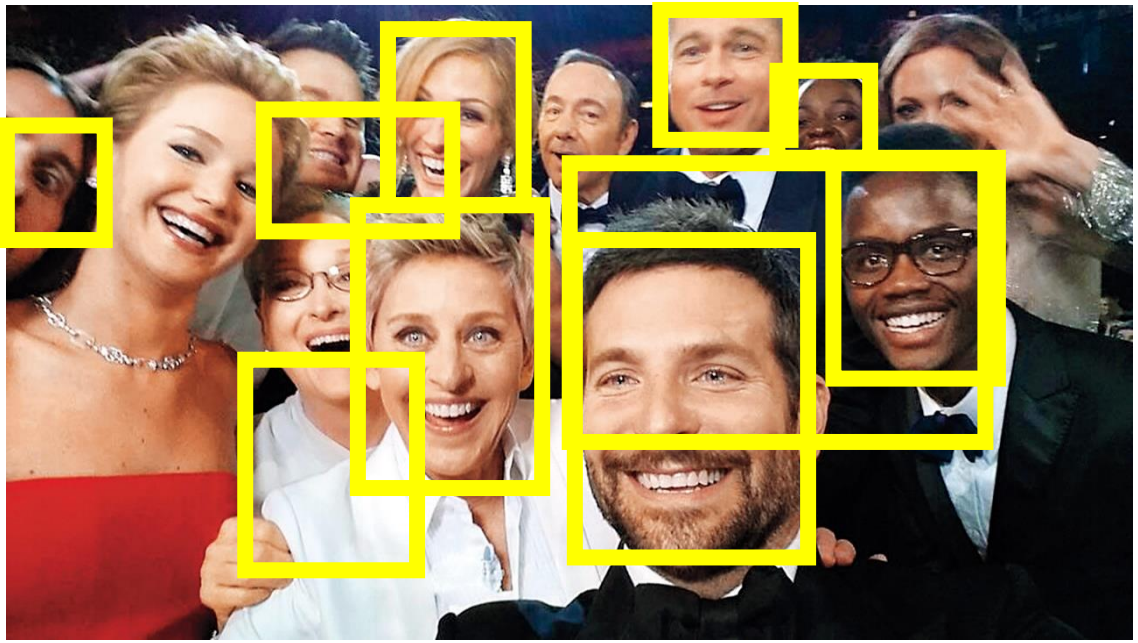


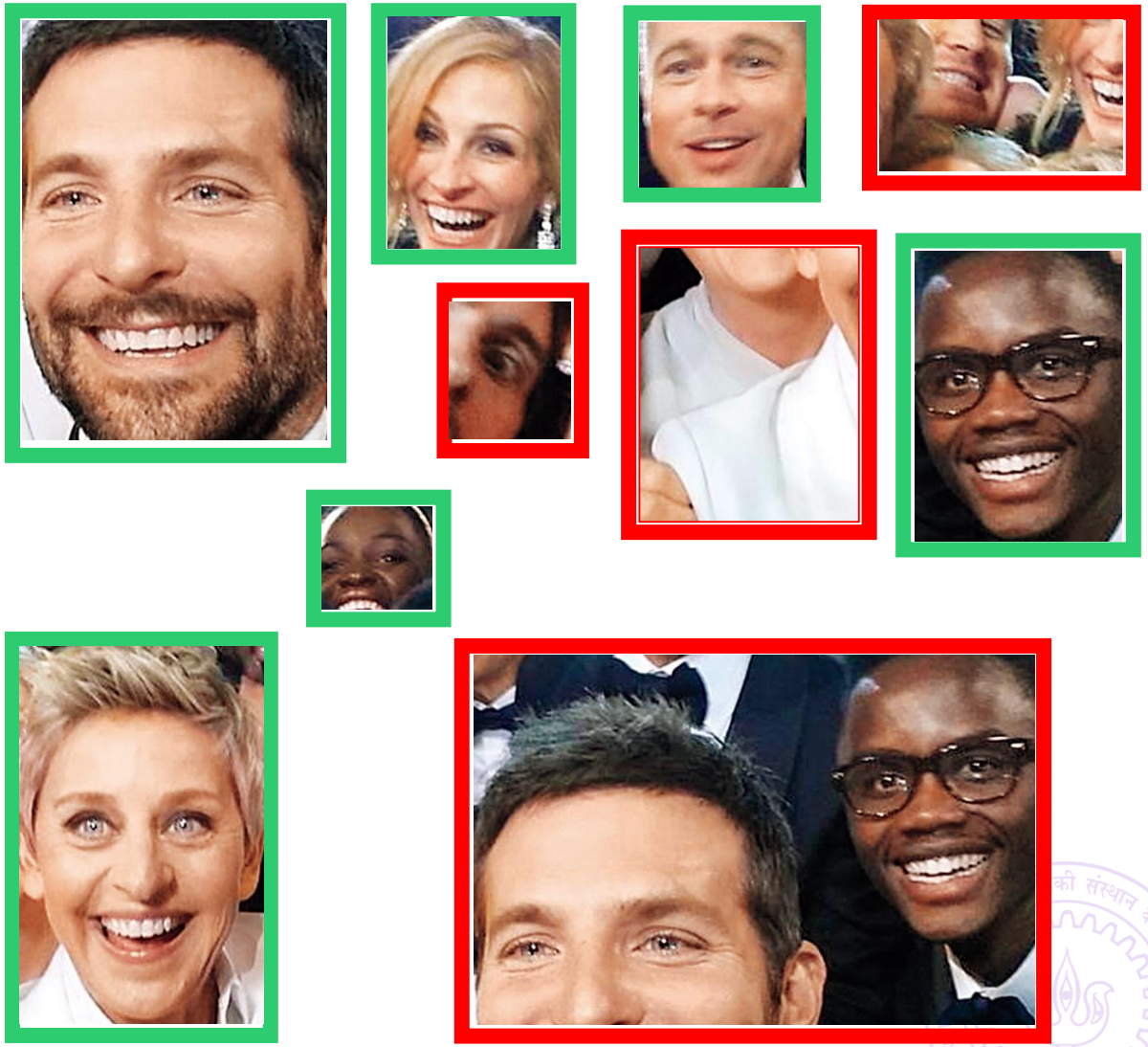
Face Tagging

Face Detection



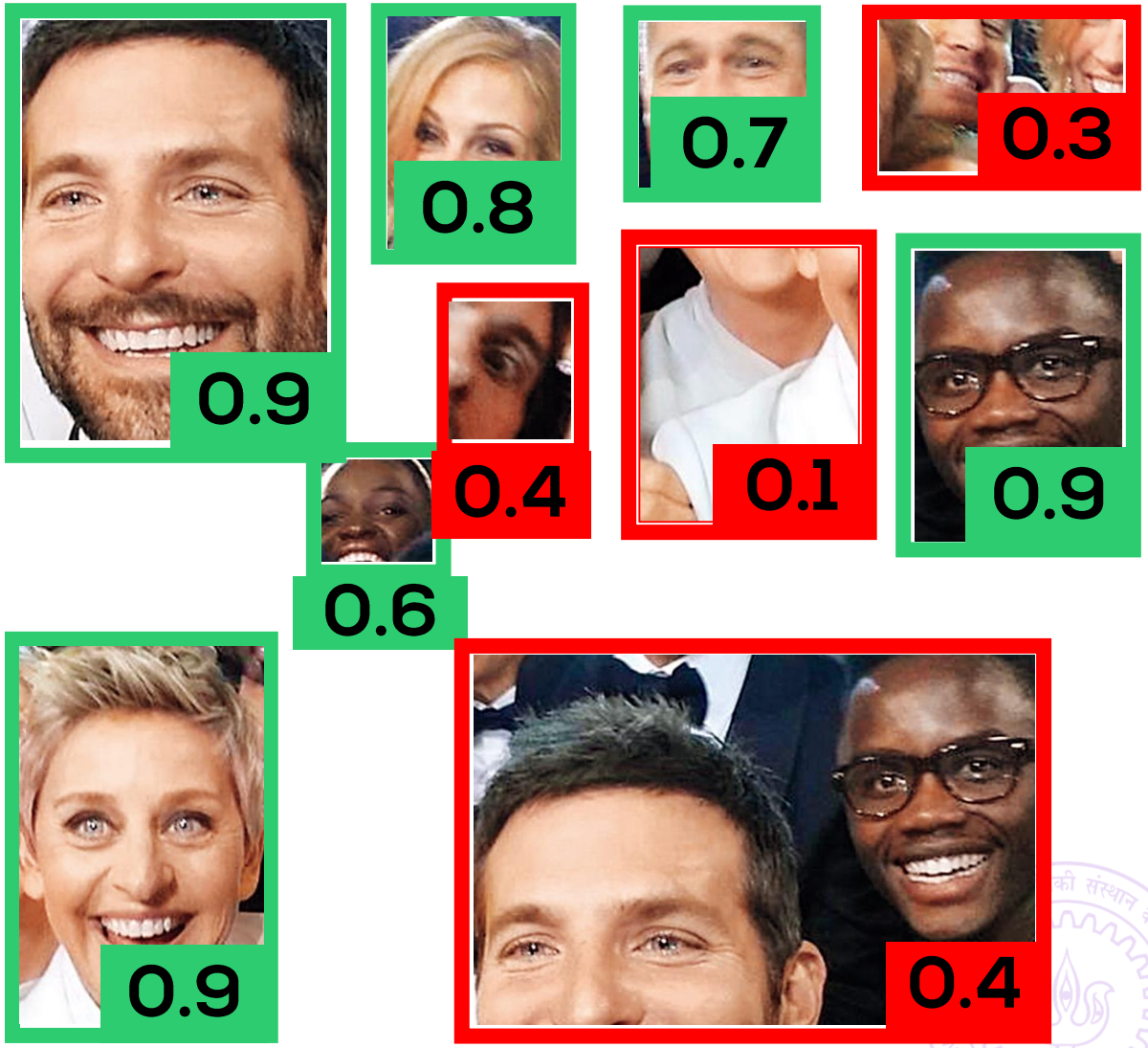
Face Detection as Binary Classification





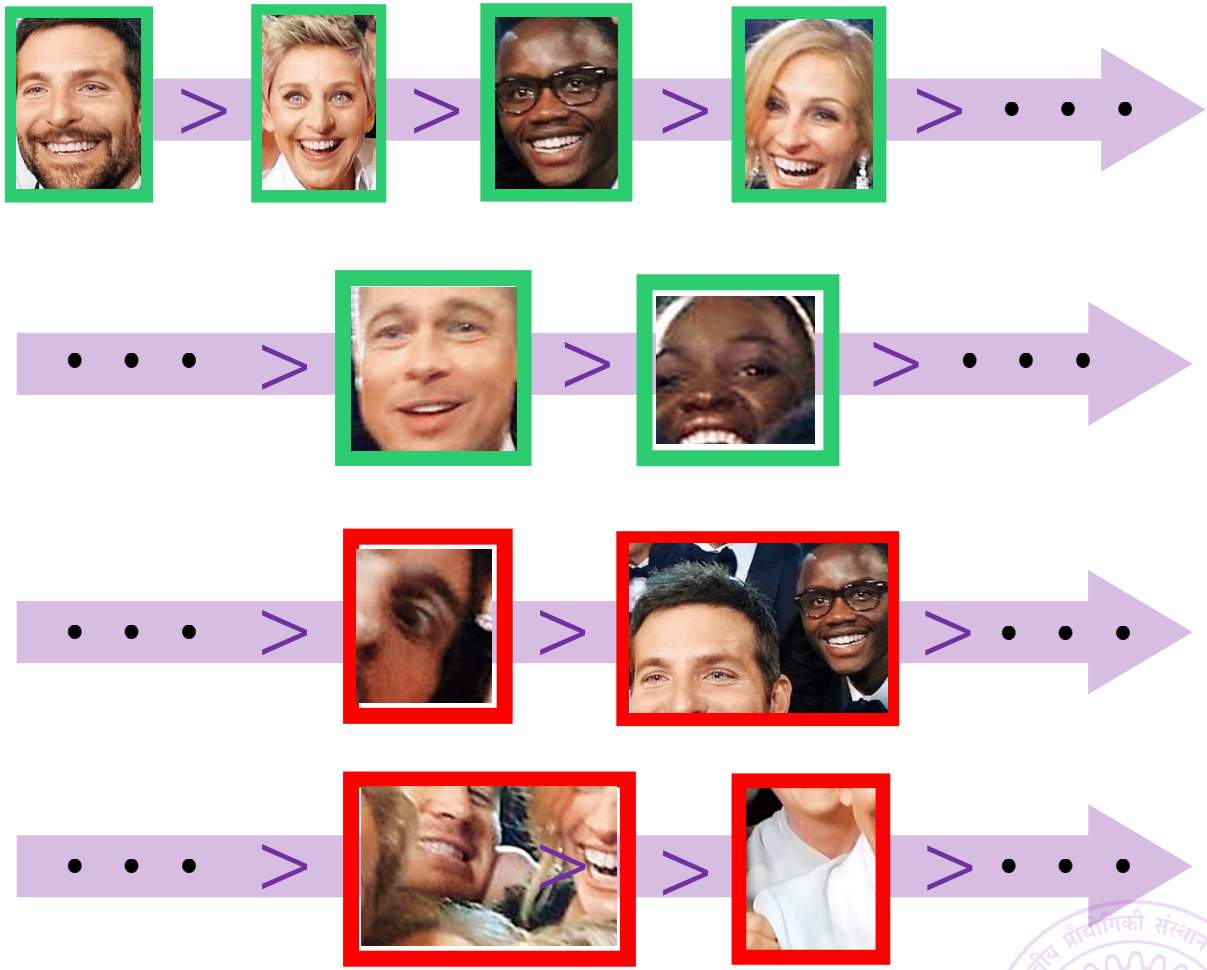
Given any bounding box, predict whether it contains a single face or not!

Face Detection as Regression



Given any bounding box, predict the likelihood score of it containing a single face

Face Detection as Ranking



Given a picture, rank all its bounding boxes in order of likelihood of containing a face

Face Tagging as Multi-classification

Celebrity Names

1. Angelina
2. Brad
3. Bradley
4. Channing
5. Ellen
6. Jared
7. Jennifer
8. Julia
9. Kevin
10. Lupita
11. Meryl
12. Peter

Face Tagging as Multi-classification

Celebrity Names

1. Angelina
2. Brad
3. Bradley
4. Channing
5. Ellen
6. Jared
7. Jennifer
8. Julia
9. Kevin
10. Lupita
11. Meryl
12. Peter

Face Tagging as Multi-classification

Celebrity Names

1. Angelina
2. Brad
3. Bradley
4. Channing
5. Ellen
6. Jared
7. Jennifer
8. Julia
9. Kevin
10. Lupita
11. Meryl
12. Peter

Face Tagging as Multi-classification

Celebrity Names

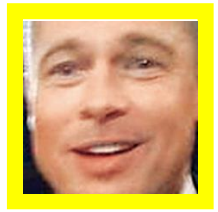
1. Angelina
2. Brad
3. Bradley
4. Channing
5. Ellen
6. Jared
7. Jennifer
8. Julia
9. Kevin
10. Lupita
11. Meryl
12. Peter

Face Tagging as Multi-classification

Celebrity Names

1. Angelina
2. Brad
3. Bradley
4. Channing
5. Ellen
6. Jared
7. Jennifer
8. Julia
9. Kevin
10. Lupita
11. Meryl
12. Peter

Face Tagging as Multi-classification



Celebrity Names

1. Angelina
2. Brad
3. Bradley
4. Channing
5. Ellen
6. Jared
7. Jennifer
8. Julia
9. Kevin
10. Lupita
11. Meryl
12. Peter

Face Tagging as Regression

Celebrity Names

1.	Angelina	0.01
2.	Brad	0.01
3.	Bradley	0.01
4.	Channing	0.01
5.	Ellen	0.99
6.	Jared	0.01
7.	Jennifer	0.25
8.	Julia	0.34
9.	Kevin	0.01
10.	Lupita	0.01
11.	Meryl	0.45
12.	Peter	0.01

Face Tagging as Regression

Celebrity Names

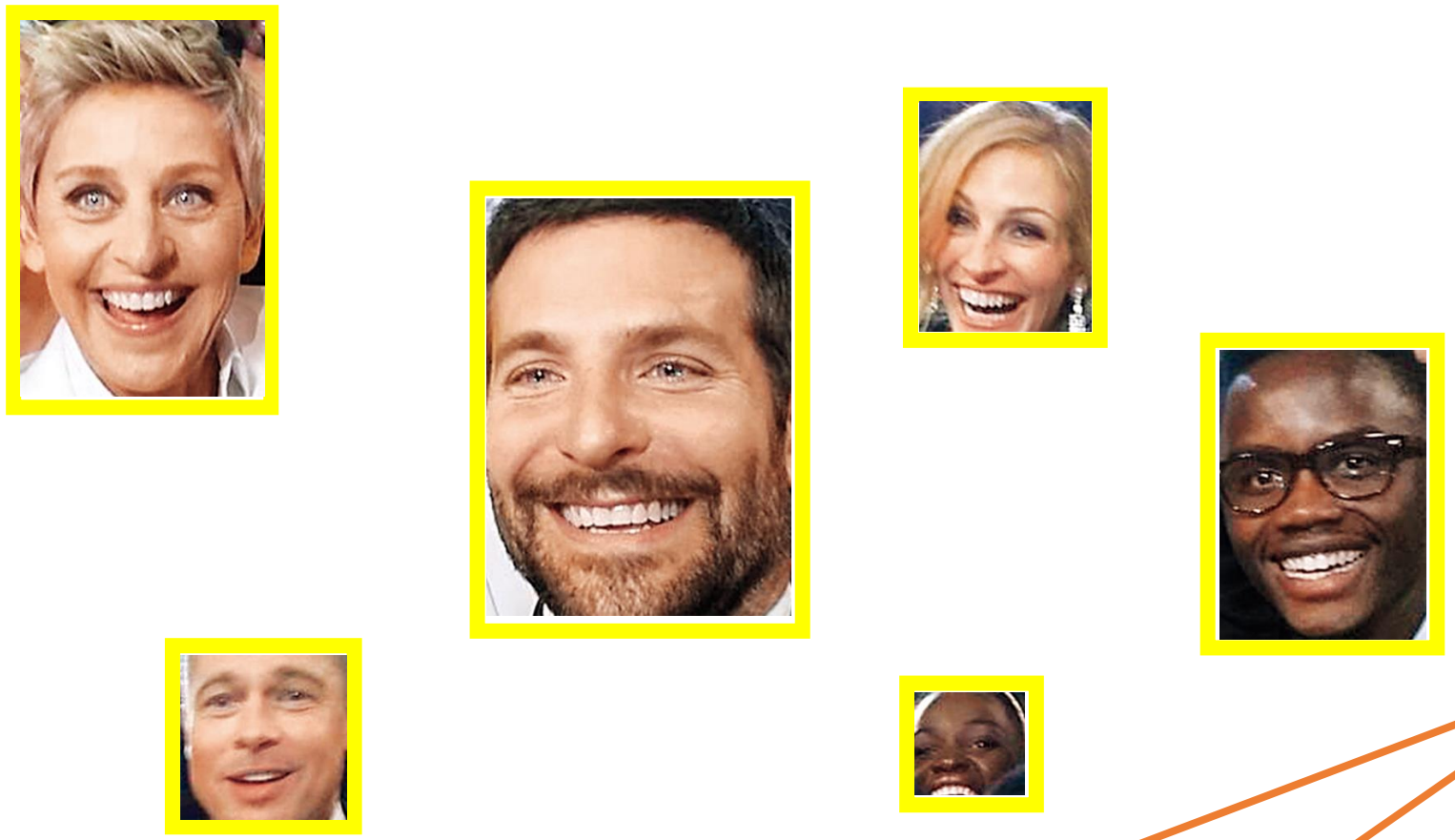
1. Angelina	0.01
2. Brad	0.25
3. Bradley	0.99
4. Channing	0.01
5. Ellen	0.01
6. Jared	0.01
7. Jennifer	0.01
8. Julia	0.01
9. Kevin	0.11
10. Lupita	0.01
11. Meryl	0.01
12. Peter	0.01

Face Tagging as Regression

Celebrity Names

1. Angelina	0.65
2. Brad	0.01
3. Bradley	0.01
4. Channing	0.01
5. Ellen	0.44
6. Jared	0.01
7. Jennifer	0.01
8. Julia	0.99
9. Kevin	0.01
10. Lupita	0.01
11. Meryl	0.22
12. Peter	0.01

Face Tagging as Regression

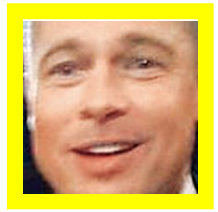


Celebrity Names

1.	Angelina	0.65
2.	Brad	0.01
3.	Bradley	0.01
4.	Channing	0.01
5.	Ellen	0.44
6.	Jared	0.01
7.	Jennifer	0.01
8.	Julia	0.99
9.	Kevin	0.01
10.	Lupita	0.01
11.	Meryl	0.22
12.	Peter	0.01

Note: a separate regression problem for each celebrity

Face Tagging as Regression



… or else, a multi-regression /vector regression problem

Celebrity Names

1. Angelina	0.65
2. Brad	0.01
3. Bradley	0.01
4. Channing	0.01
5. Ellen	0.44
6. Jared	0.01
7. Jennifer	0.01
8. Julia	0.99
9. Kevin	0.01
10. Lupita	0.01
11. Meryl	0.22
12. Peter	0.01

Face Tagging as Ranking

Celebrity Names

5. Ellen
8. Julia
11. Meryl
1. Angelina
4. Channing
6. Jared
7. Jennifer
2. Brad
9. Kevin
10. Lupita
3. Bradley
12. Peter

Face Tagging as Ranking

Celebrity Names

3. Bradley
2. Brad
9. Kevin
6. Jared
12. Peter
4. Channing
7. Jennifer
8. Julia
11. Meryl
10. Lupita
1. Angelina
5. Ellen

Face Tagging as Ranking

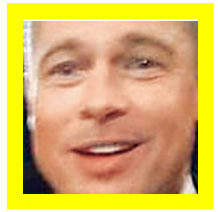
Celebrity Names

8. Julia
11. Meryl
10. Lupita
5. Ellen
1. Angelina
7. Jennifer
3. Bradley
2. Brad
9. Kevin
12. Peter
6. Jared
4. Channing

Face Tagging as Ranking





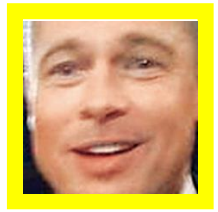



Celebrity Names

8. Julia
11. Meryl
10. Lupita
5. Ellen
1. Angelina
7. Jennifer
3. Bradley
2. Brad
9. Kevin
12. Peter
6. Jared
4. Channing

Note: can use regression to solve the ranking problem!

Face Tagging as Ranking



Celebrity Names

8. Julia
11. Meryl
10. Lupita
5. Ellen
1. Angelina
7. Jennifer
3. Bradley
2. Brad
9. Kevin
12. Peter
6. Jared
4. Channing

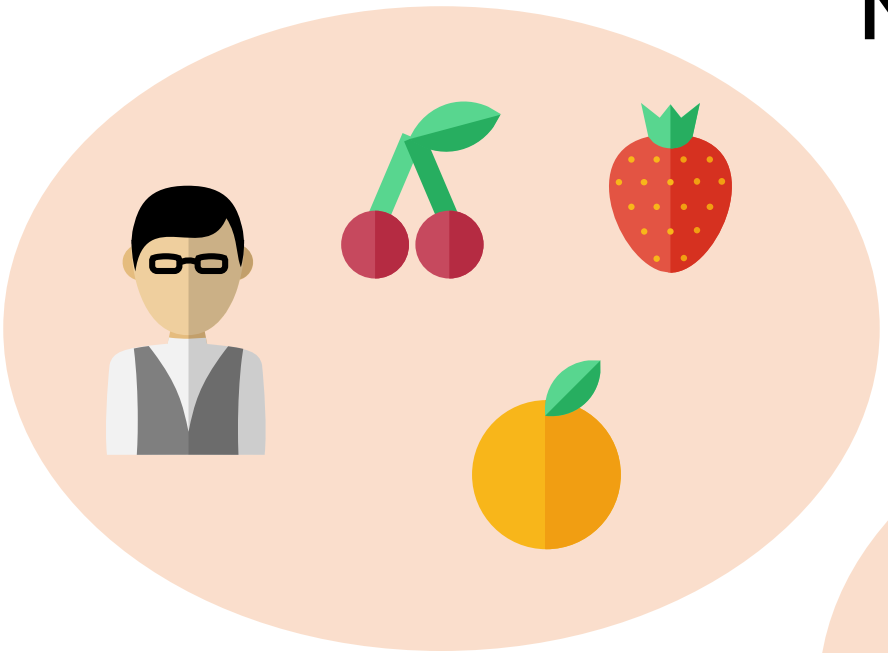
… but need not … ranking this way is not scalable

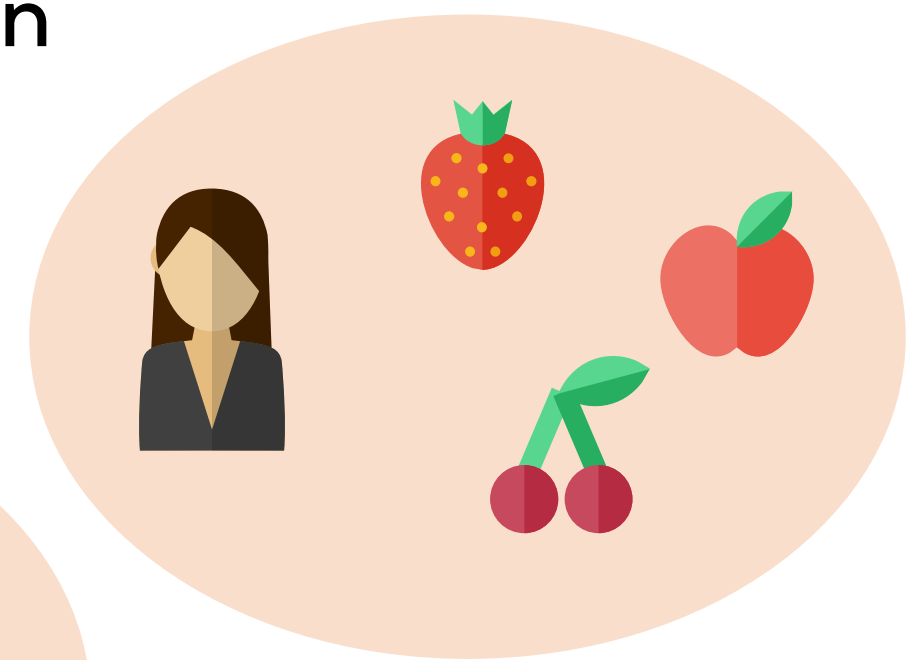
Exercise: Recommendation Systems

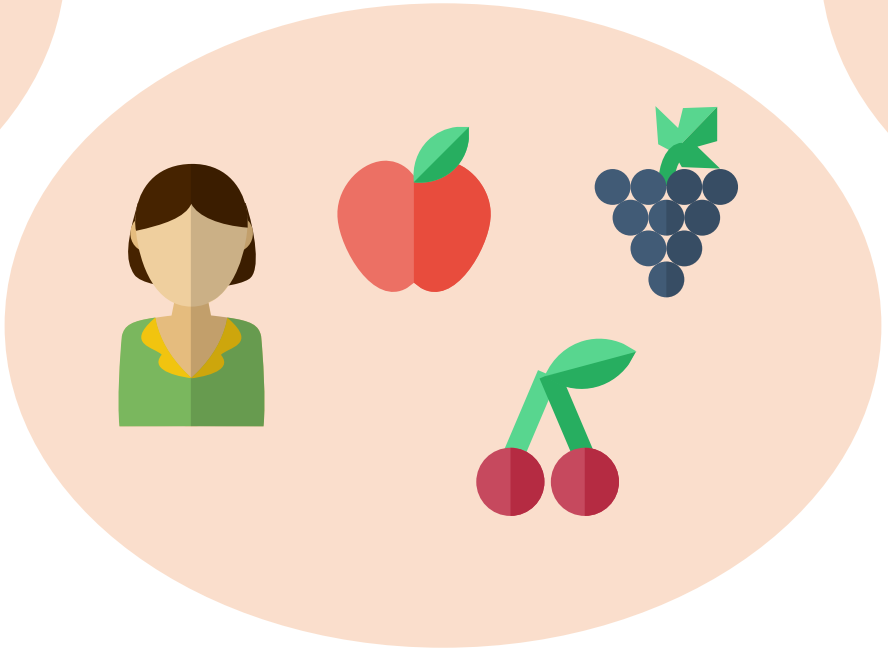
Multi-Classification

Regression

Ranking







Multi-Label Classification

Binary Classification

Hint: you may have to solve several classification or regression problems for every instance

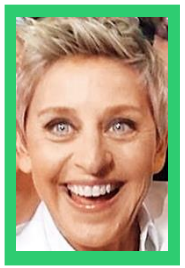
ML Workflows

Revisited

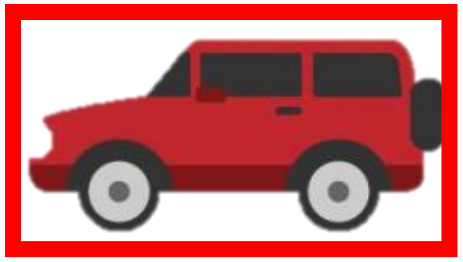
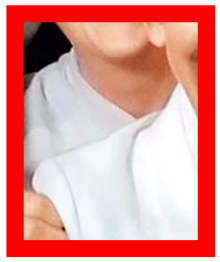
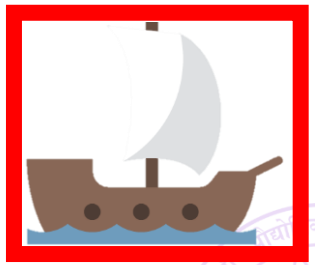
Supervised Batch Binary Classification

Yes

No


Training Data

freepik.com, dribble.com

Supervised Batch Binary Classification

Yes

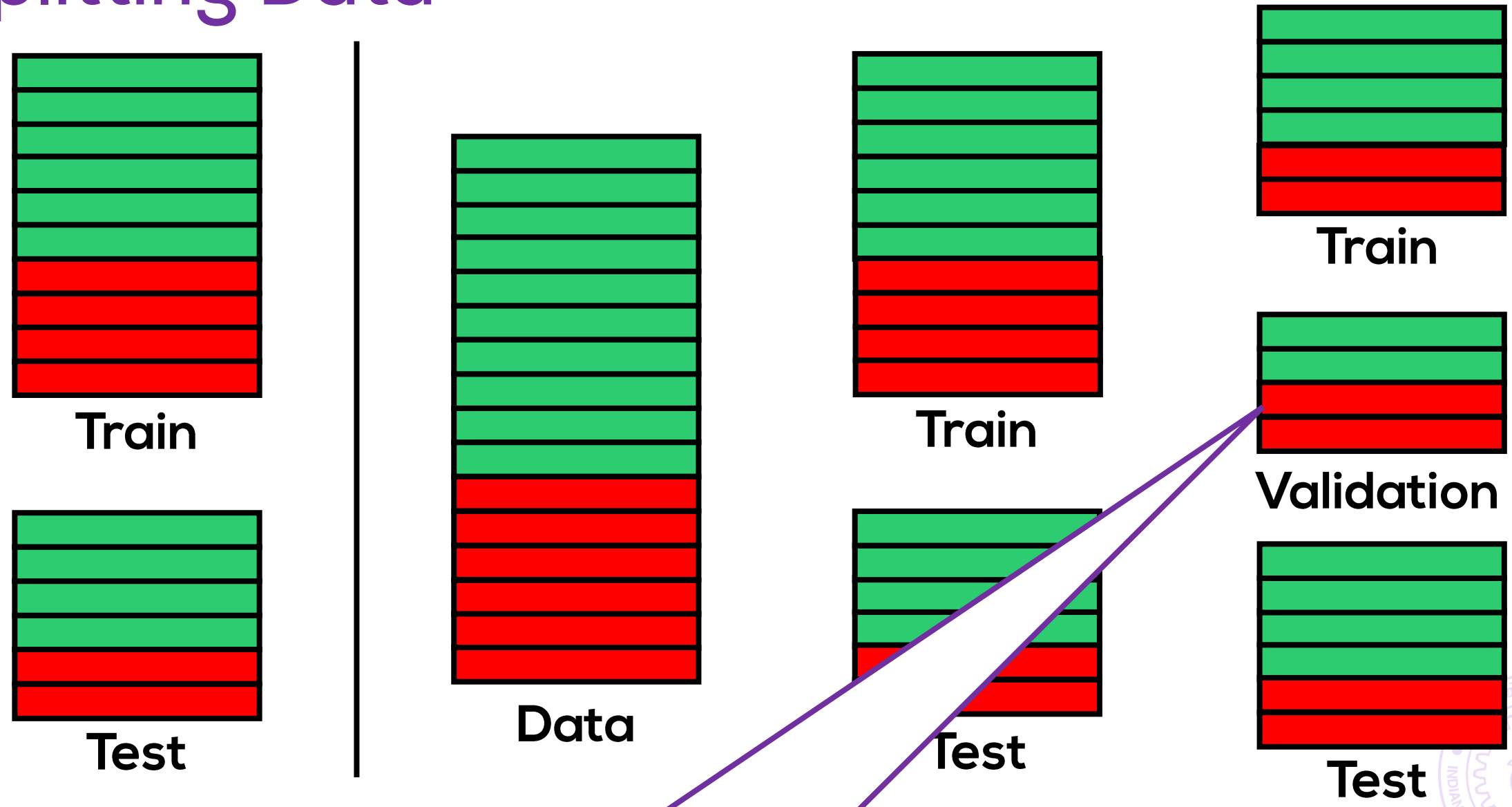
No

One-class Classification!

Too much variety in Red class

freepik.com, dribble.com

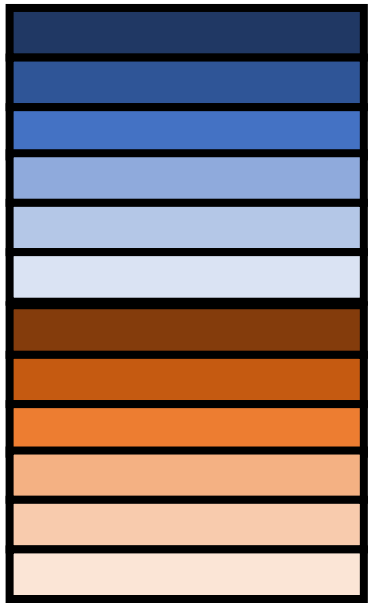
Splitting Data



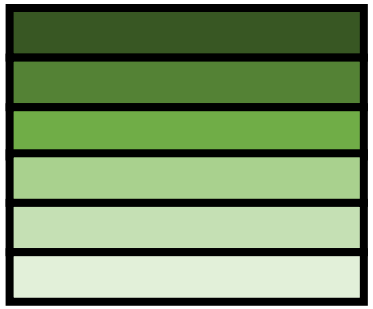
"Held-out" validation set

Multiple Splits

Split 1

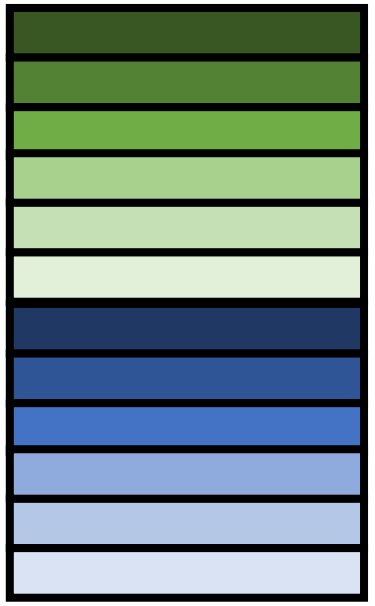


Train



Test

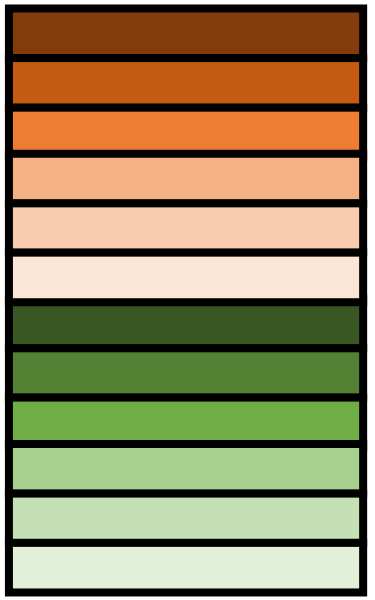
Split 2



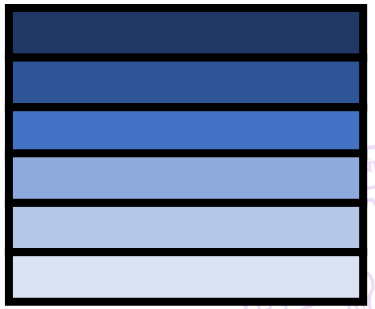
Train

Test

Split 3



Train

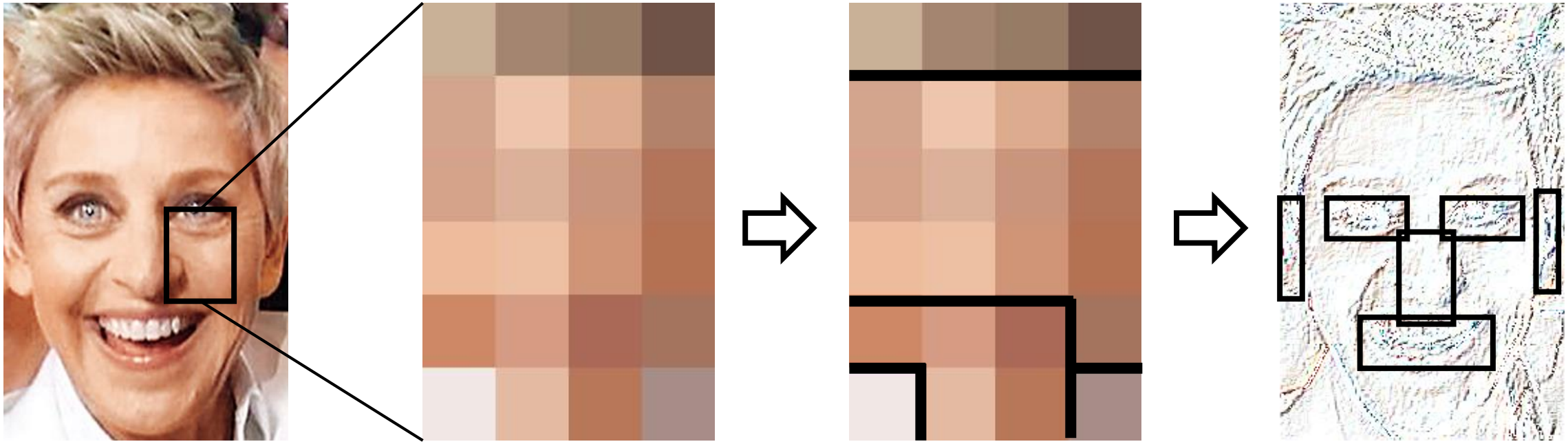


Test

Fantastic Features

... and how to find them

What is a feature?



Raw/ Low-level features

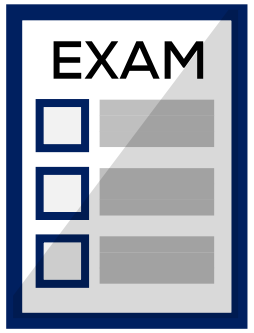
Derived/ High-level features

What is raw for you may have been derived by someone else

Types of Features

- Numerical features (pixel value, temperature)
- Categorical features (income bracket, blood type)
- Structured features (graph, tree, list)
- Relational features (neighbourhood, similarity)
- Bagged features (count statistics of other features)
  - Bag of words, bag of edges
- Pooled features (max, average of other features)
  - Popular in neural networks
- Missing and latent/hidden features

Exert caution with features



- Tricks, mnemonics lessen cognitive load, increase speed
- Easy questions can be solved in one step with a mnemonic!
- Too many mnemonics can confuse you at time of exam

- Derived features make learning easier, faster at test
- What you are trying to predict is just another (latent) feature!
- Too many useless features can confuse classifier